29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms

AofA 2018, June 25–29, 2018, Uppsala, Sweden

Edited by

James Allen Fill
Mark Daniel Ward

LIPICS

Editors

James Allen Fill   Mark Daniel Ward
The Johns Hopkins University   Purdue University
Baltimore, MD USA   West Lafayette, IN USA
`jimfill@jhu.edu`   `mdw@purdue.edu`

LIPIcs – Leibniz International Proceedings in Informatics

LIPIcs is a series of high-quality conference proceedings across all fields in informatics. LIPIcs volumes are published according to the principle of Open Access, i.e., they are available online and free of charge.

Dedicated to the memory of Philippe Flajolet.

■ Contents

Flajolet Award Lecture

Keynote Speakers

Regular Papers

Contents **0:ix**

◼ Preface

The 29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018) was held in Uppsala, Sweden, June 25–29, 2018.

Analysis of algorithms is a scientific basis for computation, providing a link between abstract algorithms and the performance characteristics of their implementations in the real world. The general effort to predict precisely the performance of algorithms has come to involve research in analytic combinatorics, the analysis of random discrete structures, asymptotic analysis, exact and limiting distributions, and other fields of inquiry in computer science, probability theory, and enumerative combinatorics. See http://aofa.cs.purdue.edu/ .

The Call for Papers invited papers in
- analytic algorithmics and combinatorics,
- probabilistic analysis of algorithms,
- randomized algorithms.

We also welcomed papers addressing problems such as: combinatorial algorithms, string searching and pattern matching, sublinear algorithms on massive data sets, network algorithms, graph algorithms, caching and memory hierarchies, indexing, data mining, data compression, coding and information theory, and computational finance. Papers were also welcomed that address bridges to research in related fields such as statistical physics, computational biology, computational geometry, and simulation.

Authors of selected accepted extended abstracts will be invited to submit full papers for peer review to a special issue (published circa late 2019) of *Algorithmica*.

—James Allen ("Jim") Fill and Mark Daniel Ward,
    on behalf of the Program and Steering Committees

■ Flajolet Lecture

The Philippe Flajolet Lecture Prize for outstanding contributions to analytic combinatorics and analysis of algorithms is awarded every two years by the Analysis of Algorithms (AofA) community—a community that owes its existence to Philippe Flajolet. The first Flajolet Lecture was presented by Donald E. Knuth at the 25th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms in 2014 in Paris, France, and the second one by Robert Sedgewick at the 27th AofA Conference in 2016 in Krakow, Poland.

At this year's conference, Luc Devroye presented the third Flajolet Lecture, entitled "OMG: GW, CLT, CRT and CFTP."

The prize is named in honor and recognition of the extraordinary accomplishments of the late Philippe Flajolet, who spent most of his scientific life at INRIA, France. Philippe is best known for fundamental advances in mathematical methods for the analysis of algorithms. His research laid the foundation of a subfield of mathematics now known as analytic combinatorics. Analytic combinatorics is a modern basis for the quantitative study of combinatorial structures (such as words, trees, mappings, and graphs), with applications to probabilistic study of algorithms that are based on these structures. It also strongly influences research in other scientific domains, such as statistical physics, computational biology, and information theory. Flajolet's work takes the field forward by introducing original approaches in combinatorics based on two types of methods: symbolic and analytic. The symbolic side is based on the automation of decision procedures in combinatorial enumeration to derive characterizations of generating functions. The analytic side treats those functions as functions in the complex plane and leads to precise characterization of limit distributions. Beyond these foundational contributions, Philippe's research opened new avenues in various domains of applied computer science, including streaming algorithms, communication protocols, database access methods, data mining, symbolic manipulation, text-processing algorithms, and random generation.

![] Program Committee

- **Chair:** James Allen Fill, Johns Hopkins University, USA

- Olivier Bodini, Université Paris 13, France
- Michael Drmota, Technische Universität Vienna, Austria
- Antoine Genitrini, Université Paris 6, France
- Zbigniew Gołębiewski, Wrocław University of Science and Technology, Poland
- Cecilia Holmgren, Uppsala University, Sweden
- Hsien-Kuei Hwang, Academic Sinica, Taiwan
- Svante Linusson, KTH Royal Institute of Technology, Sweden
- Manuel Lladser, University of Colorado Boulder, USA
- Cécille Mailler, University of Bath, United Kingdom
- Markus Nebel, Bielefeld University, Germany
- Ralph Neininger, Goethe University Frankfurt, Germany
- Daniel Panario, Carleton University, Canada
- Bruno Salvy, INRIA Lyon, France
- Robert Sedgewick, Princeton University, USA
- Wojciech Szpankowski, Purdue University, USA
- Stephan Wagner, Stellenbosch University, South Africa
- Mark Daniel Ward, Purdue University, USA

Steering Committee

- **Chair:** Robert Sedgewick, Princeton University, USA

- Nicolas Broutin, INRIA Rocquencourt, France
- Hsien-Kuei Hwang, Academia Sinica, Taiwan
- Conrado Martinez, Polytechnic University of Catalonia, Spain
- Ralph Neininger, Goethe-U Frankfurt, Germany
- Michèle Soria, Université Paris 6, France

⬛ Organization and Funding

Local organization of the conference was coordinated by Cecilia Holmgren (Uppsala University, Sweden) and Sofie White (Uppsala, Sweden), in coordination with Program Committee Chair Jim Fill (Johns Hopkins University, USA). The conference location was the Campus Blåsenhus of Uppsala University.

Generous funding for the conference was provided by the Marcus Wallenberg Foundation for International Scientific Collaboration and by the Swedish Research Council.

OMG: GW, CLT, CRT and CFTP

Luc Devroye

School of Computer Science, McGill University, 3480 University St., Montreal, Canada H3A 0E9
http://luc.devroye.org/
lucdevroye@gmail.com

──── **Abstract** ────

After a brief review of the main results on Galton-Watson trees from the past two decades, we will discuss a few recent results in the field.

Assumptionless Bounds for Random Trees

Louigi Addario-Berry

Department of Mathematics & Statistics, McGill University, 1005-805 Rue Sherbrooke O.,
Montréal, QC, H3A 2K6, Canada
http://problab.ca/louigi
louigi.addario@mcgill.ca

Abstract

Let T be any Galton-Watson tree. Write $\mathrm{vol}(T)$ for the volume of T (the number of nodes), $\mathrm{ht}(T)$
for the height of T (the greatest distance of any node from the root) and $\mathrm{wid}(T)$ for the width
of T (the greatest number of nodes at any level). We study the relation between $\mathrm{vol}(T)$, $\mathrm{ht}(T)$
and $\mathrm{wid}(T)$.

In the case when the offspring distribution $p = (p_i, i \geq 0)$ has mean one and finite variance,
both $\mathrm{ht}(T)$ and $\mathrm{wid}(T)$ are typically of order $\mathrm{vol}(T)^{1/2}$, and have sub-Gaussian upper tails on
this scale. Heuristically, as the tail of the offspring distribution becomes heavier, the tree T
becomes "shorter and bushier". I will describe a collection of work which can be viewed as
justifying this heuristic in various ways In particular, I will explain how classical bounds on
Lévy's concentration function for random walks may be used to show that for any offspring
distribution, the random variable $\mathrm{ht}(T)/\mathrm{wid}(T)$ has sub-exponential tails. I will also describe
a more combinatorial approach to coupling random trees with different degree sequences which
allows the heights of randomly sampled vertices to be compared.

Making Squares – Sieves, Smooth Numbers, Cores and Random Xorsat

Béla Bollobás

University of Cambridge, Department of Pure Mathematics and Mathematical Statistics,
Wilberforce Road, Cambridge CB3 0WB, UK and University of Memphis, Department of
Mathematical Sciences, Memphis, TN 38152, USA
bb12@dpmms.cam.ac.uk

—— **Abstract** ——

Since the advent of fast computers, much attention has been paid to practical factoring algorithms. Several of these algorithms set out to find two squares x^2, y^2 that are congruent modulo the number n we wish to factor, and are non-trivial in the sense that $x \not\equiv \pm y \pmod{n}$. In 1994, this prompted Pomerance to ask the following question.

Let $a_1, a_2, \ldots$ be random integers, chosen independently and uniformly from a set $\{1, \ldots x\}$. Let N be the smallest index such that $\{a_1, \ldots, a_N\}$ contains a subsequence, the product of whose elements is a perfect square. What can you say about this random number N? In particular, give bounds N_0 and N_1 such that $\mathbb{P}(N_0 \leq N \leq N_1) \to 1$ as $x \to \infty$. Pomerance also gave bounds N_0 and N_1 with $\log N_0 \sim \log N_1$.

In 2012, Croot, Granville, Pemantle and Tetali significantly improved these bounds of Pomerance, bringing them within a constant of each other, and conjectured that their upper bound is sharp. In a recent paper, Paul Balister, Rob Morris and I have proved this conjecture. In the talk I shall review some related results and sketch some of the ideas used in our proof.

Bootstrap Percolation and Galton–Watson Trees

Karen Gunderson[1]

University of Manitoba, 186 Dysart Road, Winnipeg MB R3T 2N2, Canada
karen.gunderson@umanitoba.ca

Abstract

A bootstrap process is a type of cellular automaton, acting on the vertices of a graph which are in one of two states: 'healthy' or 'infected'. For any positive integer r, the r-neighbour bootstrap process is the following update rule for the states of vertices: infected vertices remain infected forever and each healthy vertex with at least r infected neighbours becomes itself infected. These updates occur simultaneously and are repeated at discrete time intervals. Percolation is said to occur if all vertices are eventually infected. For an infinite graph, of interest is the random setting, in which each vertex is initially infected independently with a fixed probability. I will give some history of this process for infinite trees and present results on the possible values of critical probabilities for percolation on Galton–Watson trees.

This talk is based on joint work with Bollobás, Holmgren, Janson, and Przykucki.

Thinking in Advance About the Last Algorithm We Ever Need to Invent

Olle Häggström

Dept of Mathematical Sciences, Chalmers University of Technology, 412 96 Göteborg, Sweden, and Institute for Future Studies, Box 591, 101 31 Stockholm, Sweden
olleh@chalmers.se

──── **Abstract** ────

We survey current discussions about possibilities and risks associated with an artificial intelligence breakthrough on the level that puts humanity in the situation where we are no longer foremost on the planet in terms of general intelligence. The importance of thinking in advance about such an event is emphasized. Key issues include when and how suddenly superintelligence is likely to emerge, the goals and motivations of a superintelligent machine, and what we can do to improve the chances of a favorable outcome.

1 Introduction

In 1951, Alan Turing, in his *Intelligent machinery, a heretical theory* [41], anticipated many of the key ideas in current artificial intelligence (AI) futurology:

> My contention is that machines can be constructed which will simulate the behaviour of the human mind very closely. [...] Let us now assume, for the sake of argument, that these machines are a genuine possibility, and look at the consequences of constructing them. [...] It seems probable that once the machine thinking method had started, it would not take long to outstrip our feeble powers. There would be no question of the machines dying, and they would be able to converse with each other to sharpen their wits. At some stage therefore we should have to expect the machines to take control.

One of Turing's collaborators at Beltchley Park, mathematician I.J. Good, later made a related prediction, in a famous passage [13] from which the title of the present paper is partly borrowed:

> Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an "intelligence explosion," and the intelligence of man would be left far behind. Thus the first ultraintelligent machine is the last invention that man need ever make.

The presently favored term for what Good called ultraintelligence is **superintelligence**: a superintelligent machine is one that by far exceeds human performance across the full range of relevant cognitive skills, including the mysterious-seeming quality we label creativity or the ability to think outside the box. Defining an agent's intelligence is of course not straightforward, and no strict definition will be given here, but it can be thought of informally as the ability to direct the world towards whatever goals the agent has. If a machine has at least human-level such ability across more or less the full range of domains encountered by humans, we speak of **artificial general intelligence (AGI)**, and if its general intelligence vastly exceeds that of humans, then it has superintelligence.

Is it really reasonable to expect superintelligence any time soon – let's say before the end of the present century? This is a highly controversial issue where expert opinions vary wildly, and while I accept that the question is wide open, I also hold – as the first of my two main claims in this paper – that the emergence of superintelligence is a sufficiently plausible scenario to warrant taking seriously. This claim is defended in Section 2 on the possibility in principle of superintelligence, and in Sections 3 and 4 on timelines.

The second main claim in this paper is that it is of great practical importance to think in advance about safety aspects of a superintelligence breakthrough, because if those aspects are ignored or otherwise mismanaged, the event might have catastrophic consequences to humanity. Such risks are discussed in Section 5, aided mainly by the Omohundro–Bostrom theory for instrumental vs final AI goals, which is explained in some detail. Ideas on how to ensure a more benign outcome are briefly discussed in Section 6, and Section 7 offers some concluding remarks.

2   The possibility in principle

Is a superintelligent machine possible in principle in the universe we inhabit? If a supernatural human soul – or something else in that vein – exists, then all bets are out the window, so I will ignore that possibility and instead focus on the case which is more amenable to rational argument: a physical world in which all high-level phenomena, including the human mind, are the result of particular arrangements of matter. Assuming this, the example of the human brain demonstrates that there are arrangements of matter that gives rise to human-level intelligence.

There are several independent ways to argue that the human brain is unlikely to be anywhere near an optimal arrangement of mater for producing intelligence. One is to point to the fact that our brain is the product of biological evolution, which viewed as an optimization algorithm is a rather primitive local search approach, which in a setting as complex as optimizing for intelligence is unlikely to find anything like a global optimum. Another thing to point at is the extreme slowness of the nervous system compared to how the same information processing might be carried out on a modern electronic computer. A third one is the many obvious miscalibrations and biases our brain has [12], that might be corrected for. See also Sotala [38] for further concrete examples of ways in which there is room for improvement upon human intelligence.

So there are good reasons to believe that there are physical arrangements of matter that produce intelligence far superior to the human brain, i.e., superintelligence. The argument so far does not show that it can be implemented on a digital computer, but if we accept the Church–Turing–Deutsch principle that a Turing-complete computing device can be used to simulate any physical process [9], then there is an algorithm out there that achieves superintelligence.

This argument is not entirely watertight, because if the algorithm is based on simulating the physical process on a very low level (say, the movement of elementary particles), then an implementation of it on a digital computer may turn out to be so slow that it cannot be recognized as superintelligent. But it seems plausible that more efficient implementations of the system's essential information processing should be possible. We note in passing that the level of detail with which a human brain needs to be implemented on a digital computer to capture its intelligence remains a highly open question [34].

While some uncertainty remains, considerations such as these strongly suggest the existence of algorithms that can be implemented on a digital computer to achieve superintelligence. Husfeldt [24] accepts the existence of such an algorithm, calls it *the monster in the library of Turing*, and suggests that it is prohibitively difficult to find such a monster. So even if we accept its existence, we should still be open to the possiblity that the answer to the question that the next section addresses – that of when we can expect a superintelligent machine – is "never". It might be that finding it requires – short of a thremodynamics-level mircle – astronomical (or larger) amounts of brute force search, so in the next sections's discussion on when to expect the emergence of superintelligence, time $t = \infty$ will be considered a genuine possibility.

3 When to expect superintelligence?

In view of the current surge of progress in AI for a wide range of applications such as speech synthesis [37], board games [36] and autonomous vehicles [25], it may be tempting to read this as a sign that AGI and superintelligence are just around the corner. We should not jump too quickly to such conclusions, however. Many commentators, including recently Jordan [26], emphasize a fundamental discontinuity between specialized AI applications and AGI – the former should not in general be understood as stepping stones towards the latter – and they may well be right. (On the other hand, see Yudkowsky [44] who points out that we do not have strong evidence to conclude that AGI and superintelligence are *not* around the corner.)

When looking at the history of AI, the contrast between the the extraordinary achievements in specialized AI applications and the much less impressive progress towards AGI is striking. It is sometimes claimed that the latter has been literally zero, but that seems to me a bit harsh. For instance, an AI was developed a few years ago that quickly learned to successfully play a range of Atari video games [29]. As I admitted in [19], this is of course a very far cry from the ability to handle the full range of tasks encountered by humans in the physical and social world we inhabit; nevertheless, it is a nonzero improvement upon having specialized skill in just a single video game. One possible path towards AGI, among many, might be a step-by-step expansion of the domain in which the machine is able to act intelligently.

We do not at present have very clear ideas on what approach to AI has the best potential for realizing AGI. The main driver behind the rapid progress we see today in various AI applications is the deep learning approach, which is essentially a rejuvenation and further development of old neural network techniques that used to yield unimpressive results but which in many cases work remarkably well today, thanks to faster machines and access to huge training data sets. It is not, however, written in stone that deep learning will retain its position as the dominant AI paradigm forever. Other potentially useful approaches that share the black box feature of deep learning include genetic programming mimicking biological evolution, and the brute force copying of the workings of the human brain in sufficient detail

to reproduce its behavior. This last possibility is advocated enthusiastically by Kurzweil [27] and discussed in more balanced fashion by Sandberg and Bostrom [34]. Alternatively, we might see a revival of the non-black box approach of GOFAI (Good Old-Fashioned AI) with explicit hand-coding of the machine's central concepts and reasoning procedures. Or perhaps some hitherto untried combination of these approaches, or something else entirely. It might be that none of these will ever yield AGI, but the reasonable stance seems to be to at least be open to the possibility that one of them might eventually accomplish that.

But when would that happen? This is highly uncertain, as illustrated by a survey by Müller and Bostrom [30] of estimates by the world's top 100 most cited AI researchers – eatimates that are spread out all over the present century, and beyond. Not only is the amount of between-individual differences large, the individually reported uncertainty ranges also tend to be broad. Among the 29 who responded, the median of their estimates for the time when human-level AGI can be expected to have arrived with probability 50% (given that "human scientific activity continues without major negative disruption") is 2050, with a median estimate of 50% for the probability that superintelligence emerges within 30 years later. More detailed but broadly consistent results are reported in the more recent survey by Grace et al. [14]. Yet another expert survey is reported in what looks like a deliberate attempt to downplay the importance of thinking ahead about AGI and superintelligence [11], but see [8] for an effective rebuttal.

The short answer to the question of when to expect superintelligence is that we do not know: experts are highly divided. In such a situation, it would be epistemically reckless to have a firm belief about if/when superintelligence will happen, rather than prudently and thoughtfully accepting that it may well happen within decades, or within centuries, or not at all.

Yet, it is quite common to hear, even among commentators for whom the label "AI expert" seems justified, dismissive attitudes towards the idea of a future superintelligence; Dubhashi and Lappin [10] and Bentley [3] are typical examples (see [20] for my fair and balanced response to the latter). Rarely or never do these commentators offer convincing arguments for their view. So one might wonder what the actual reasons for their view is, and although admittedly it is dubious to speculate on one's disputant's motives, I made a brave attempt in [17] to suggest an explanation for their stance in terms of what I decided to call vulgopopperianism, which I defined as the implicit attitude of someone who

> (a) is moderately familiar with Popperian theory of science, (b) is fond of the kind of asymmetry [appearing between the task of showing that all swans are white and showing that at least one non-white swan exists], and (c) rejoices in claiming, whenever he encounters two competing hypotheses one of which he for whatever reasons prefers, some asymmetry such that the entire (or almost the entire) burden of proof is on proving the other hypothesis, and insisting that until a conclusive such proof is presented, we can take for granted that the preferred hypothesis is correct.

The superintelligence timing case can for instance be concretized as a choice between two competing hypotheses (H1) and (H2), where (H1) is the hypothesis that achieving superintelligence is hard in the sense of not being attainable (other than possibly by extreme luck) by human technological progress by the year 2100. (H2) is the complementary hypothesis that achieving superintelligence is comparatively easy in the sense of being within reach of human technological progress (if allowed to continue unhampered) by 2100. A priori both hypotheses seem reasonably plausible, and the presently available evidence of one over the other is fairly weak (in both directions). This gives a vulgopopperian favoring (H1) the opportunity to focus on the shortage of evidence for (H2) and thus declare (H1) the winner –

while neglecting the shortage of evidence for (H1). This may be backed up with an analogy to the swan example: just like we stick to the "all swans are white" hypothesis until a non-white swan is encountered, we can stick with (H1) for as long as no superintelligence has been produced [17]. I believe this example would (or at least should) have made Popper nervous, because the idea behind his theory of falsificationism is to make science self-correcting [33], while in the case of stubbornly sticking to (H1) the desired self-correction (in case (H1) is wrong) is likely to materialize only the moment that superintelligence shows up and and it is too late for us to avert an AI apocalypse – a scenario whose plausibility I will argue for in Section 5.

4   How suddenly?

Related to, but distinct from, the question of *when* superintelligence can be expected, is that of *how sudden* its emergence from modest intelligence levels is likely to be. Bostrom [6] distinguishes between **slow takeoff** and **fast takeoff**, where the former happens over long time scales such as decades or centuries, and the latter over short time scales such as minutes, hours or days (he also speaks of the intermediate case of **moderate takeoff**, but for the present discussion it will suffice to contrast the two extreme cases). Fast takeoff is more or less synonymous with **the Singularity** (popularized in Kurzweil's 2005 book [27]) and **intelligence explosion** (the term coined by I.J. Good as quoted in Section 1, and the one that today is preferred by most AI futurologists). The practical importance of deciding whether slow or fast takeoff is the more likely scenario is mainly that the latter gives us less opportunity to adapt during the transition, making it even more important to prepare in advance for the event.

The idea that is most often held forth in favor of a fast takeoff is the **recursive self-improvement** suggested in the Good quote in Section 1. Once we have managed to create an AI that outperforms us in terms of general intelligence, we have in particular that this AI is better equipped than us to construct the next and improved generation of AI, which will in turn be even better at constructing the next AI after that, and so on in a rapidly accelerating spiral towards superintelligence. But is it obvious that this spiral will be rapidly accelerating? No, because alternatively the machine might quickly encounter some point of diminishing return – an "all the low-hanging fruit have already been picked" phenomenon. So the problem of deciding between fast and slow takeoff seems to remain open even if we can establish that a recursive self-improvement dynamic is likely.

Just like with the timing issue discussed in Section 3, our epistemic situation regarding how suddenly superintelligence can be expected to emerge is steeped in uncertainty. Still, I think we are at present a bit better equipped to deal with the suddenness issue than with the timing issue, because unlike for timing we have what seems like a promising theoretical framework for dealing with suddenness. In his seminal 2013 paper [43], Yudkowsky borrows from economics the concept of returns on reinvestment, frames the AI's self-improvement as a kind of cognitive reinvestment, and phrases the slow vs fast takeoff problem in terms of whether returns on cognitive reinvestment are increasing or decreasing in the intelligence level. Roughly, increasing returns leads to an intelligence explosion, while decreasing returns leaves the AI struggling to reach any higher in the tree than the low branches with no fruits left on them. From that insight, a way forward is to estimate returns on cognitive reinvestment based on various data sets, e.g, from the evolutionary history of *homo sapiens*, and think carefully about to what extent the results obtained generalize to an AI takeoff. Yudkowsky does some of this in [43], and leans tentatively towards the view that an intelligence explosion

is likely. This may be contrasted against the figures from the Müller–Bostrom survey [30] quoted in Section 3, which suggest that a majority of AI experts lean more towards a slow takeoff. I doubt, however, that most of these experts have thought as systematically and as hard about the issue as Yudkowsky.

5   Goals of the superintelligent AI: Omohundro–Bostrom theory

Consequences of an AGI breakthrough may turn out extremely beneficial to humanity, or they may turn out catastrophic. A favorite example of the latter – cartoonish on purpose to emphaisze that it is merely an example – is the so-called **Paperclip Armageddon**, which dates back at least to 2003 [4]. Imagine a paperclip factory, which is run by an advanced (but not yet superintelligent) AI, programmed to maximize paperclip production. Its computer engineers are continuously trying to improve it, and one day, more or less by accident, they manage to push the machine over the threshold where it enters the spiral of self-improvement causing an intelligence explosion. Coming out of the explosion is the world's first and only superintelligent AI. Having retained its goal of maximizing paperclip production, it promptly goes on to turn our entire planet (including us) into a giant heap of paperclips, followed by an expansion into outer space in order to turn the rest of the observable universe into paperclips. (For readers who feel repelled by the crude and seemingly farfetched character of Paperclip Armageddon, I recommend the more subtle and elaborate but no less frightening thought experiments offered by Armstrong [1] and Tegmark [40].)

Of course, AI futurology is not about randomly dreaming up weird scenarios, but about reasoning as rigorously as the topic admits about what is plausible and what is likely. The difficulty in evaluating whether an apocalypse along the lines of Paperclip Armageddon might really happen lies not so much in what a superintelligent machine would be *capable* of doing, but rather what it would be *motivated* to do. (For some vivid scenarios illustrating the capability of a superintelligent AI, see, e.g., [42], [6] and [40].) Currently the only game in town for going beyond mere speculations regarding a superintelligent AI's goals and motivations is what in my 2016 book [16] I decided to call **the Omohundro–Bostrom theory of final vs instrumental AI goals**, honoring key contributions by Omohundro [31, 32] and Bostrom [5, 6]. An agent's final goal is what the agent values as an end in itself rather than as a means towards achieving something else. An instrumental goal, in contrast, is one that is set up as a stepping stone towards another goal.

(Some philosophers, such as Searle [35], are fond of saying that this whole approach is confused, because computers cannot have goals. But the confusion is on their side, as even heat-seeking missiles and thermostats have goals in the relevant sense. See [15] for my detailed response to Searle.)

The two cornerstones of Omohundro–Bostrom theory are **the orthogonality thesis** and the **the instrumental convergence thesis**. We begin with the former.

> **The Orthogonality Thesis:** More or less any final goal is compatible with more or less arbitrarily high levels of intelligence.

In his original formulation, Bostrom [5] omits the qualifier "arbitrarily high" (writing instead "any"), but I prefer its inclusion so as not to have to bother with possible counterexamples that combine low intelligence with conceptually advanced goals. He does, however, include the qualifiers "more or less" (in both places), underlining the statement's lack of mathematical precision; it really does seem to be needed due to the kinds of counterexamples discussed towards the end of this section.

In response to the question "What will a superintelligent machine be inclined to do?", the Orthogonality Thesis on its own obviously isn't of much help in narrowing down from the useless answer "anything might happen". It does, however, serve as an antidote to naive (but fairly common; [22] is a typical example) anthropomorphisms such as "Paperclip Armageddon is impossible, since having such a stupid goal would directly contradict the very notion of superintelligence; surely someone who is superintelligent would realize that things like human welfare and ecosystem preservation are more important than monomanically producing ever-increasing numbers of paperclips," which conflate intelligence with goals. The Orthogonality Thesis helps remind us to distinguish between intelligence and goals.

More useful in terms of narrowing down on what a superintelligent machine can be expected to do is the Instrumental Convergence Thesis, in combination with a collection of concrete goals to which it applies.

> **The Instrumental Convergence Thesis:** There are several instrumental goals that are likely to be adopted by a sufficiently intelligent agent in order to pursue its final goal, for a wide range of final goals and a wide range of circumstances.

Omohundro [31] and Bostrom [5] list several instrumental goals that they argue to be in the range of applicability of the instrumental convergence thesis:

- **Self-preservation:** if you continue to exist and are up and running, you will be in a better position to work for your final goal compared to if you are turned off, so don't let anyone pull the plug on you!
- **Self-improvement:** improvements to one's own software and hardware design.
- **Acquisition of resources** such as hardware, but also things like money in case the agent operates in a world that is still dominated by the kind of economy we have today.
- **Goal integrity:** make sure your final goal remains intact.

The instrumental goal of self-improvement plays a special role in the theory of intelligence explosion discussed in Section 5, because it explains why, among the millions of other things it might decide to do, we should not be surprised to see the AI choose to work its way up the spiral of recursive self-improvement.

The value, for the purpose of pursuing a generic final goal, of the first three instrumental goals on the list is more or less self-explanatory, but the fourth item on the list – goal intrgrity – may warrant an explanation. As a simple example, imagine an AI with the goal of maximizing paperclip production, and suppose that, perhaps triggered by some external impulse, it starts to contemplate whether in fact ecosystem preservation might in fact be a preferable goal to pursue, compared to maximizing paperclip production. Should it stick to the old goal, or should it switch? In order to decide, it needs some criterion for which goal is the better one. Since it hasn't yet switched to the new goal, but is merely considering whether to do so, it still has the paperclip maximization goal, so the criterion will be: which goal is likely to lead to the larger number of paperclips? In all but some very contrived circumstances, paperclip maximimzation will win this comparison, so the AI will stick to that.

Equipped with Omohundro–Bostrom theory, we are in a position to understand that a scenario like Paperclip Armageddon is not as far-fetched as it first might seem. The Orthogonality Thesis helps us see that while paperclip maximization may seem bizarre *to us* (because we have other goals), it need not look that way *to the machine*, who may instead find goals like ecosystem preservation and promotion of human well-being utterly pointless. The instrumental goal of self-improvement helps explain why the paperclip maximizer might go through an intelligence explosion, and the instrumental goal of goal integrity explains why

the machine can be expected to come out of the intelligence explosion with its monomaniacal wish to produce paperclips intact.

A common objection to Paperclip Armageddon-like scenarios is that a superintelligent machine will understand that its original human programmers did not intend it to turn the observable universe into paperclips, and will therefore refrain from doing so. The mistake here is to take for granted that "do things that please your programmers" is among the machine's goals. Every programmer today knows that whenever there is a discrepancy between what the programmer intends and what appears literally in the computer code, it is the latter that counts. Omohundro–Bostrom theory predicts that principle to remain true for superintelligent machines. If that sounds like bad news, then perhaps a remedy might be to make "do things that please your programmers" the machine's final goal. Ideas in that spirit are in fact being considered in contemporary work on AI risk. More on that in the next section.

Before that, let me emphasize that while Omohundro–Bostrom theory is, for the time being, an indispensable tool for reasoning about consequences of an AGI breakthrough, it is also to some extent tentative. Its two cornerstones deal with messy concepts with fuzzy boundaries, and they do not (as yet, in their present form) deserve the same epistemic status as mathematical theorems that have been established once and for all. Therefore, predictions derived from the theory should be treated with some degree of epistemic humility (which is not to say that they can be dismissed out of hand). In my recent paper [18], I discuss a variety of challenges to the validity and range of applicability of Omohundro–Bostrom theory – in particular, the following three.

First, **self-referentiality**. Bostrom [5] points out that a superintelligent machine with the final goal of being stupid (properly specified) is unlikely to remain superintelligent for very long. Thus, for all practical purposes, the final goal of being stupid serves as a counterexample to the Orthogonality Thesis. Given one counterexample, how can we stop a wildfire of others? Some extra condition on the final goal needs to be found that excludes the stupidity example and whose inclusion makes the Orthogonality Thesis true. An obvious candidate is that the final goal cannot refer back to the machine itself, but the discussion in [18] points towards the task of defining such self-referentiality being highly problematic.

Second, **Tegmark's physics challenge**. Could other properties of a final goal, beyond self-referentiality, have the potential to invalidate the conclusion of the Orthogonality Thesis? A perhaps-too-obvious candidate is incoherence. What would it even *mean* for the machine to act towards an incoherent goal? Tegmark [39] suggests that the class of incoherent goals might be much bigger than we currently think:

> Suppose we program a friendly AI to maximize the number of humans whose souls go to heaven in the afterlife. First it tries things like increasing people's compassion and church attendance. But suppose it then attains a complete scientific understanding of humans and human consciousness, and discovers that there is no such thing as a soul. Now what? In the same way, it is possible that any other goal we give it based on our current understanding of the world (*"maximize the meaningfulness of human life"*, say) may eventually be discovered by the AI to be undefined.

Third, **human values are a mess**. If we believe that the Omohundro–Bostrom framework captures something important about the goal structure of a sufficiently intelligent agent, then we should also expect its neat dichotomy of final vs instrumental goals to be observable in such agents. The most intelligent agent we know of is *homo sapiens*, but the goals of a typical human do not seem to admit such a clearcut dichotomy [18].

6   AI Alignment

Various attempts have been made to avoid Turing's [41] conclusion (quoted in Section 1) that in the presence of superintelligent machines, "we should have to expect the machines to take control", but none of them seem to provide a clearcut solution. Probably the most studied such attempt is the so-called AI-in-a-box approach, which is to keep the machine boxed in and unable to influence the world other than via a narrow and carefully controlled communications channel. While this deserves further study, the present state-of-the-art seems to point in the direction that such boxing-in is extremely difficult and can be expected to work for at most a temporary and rather brief time period; see, e.g., [2] and [21].

It therefore makes sense to look into whether it is possible to accept that the superintelligent AI takes control and still get a favorable outcome (whatever that means). For that to happen, we need that the AI has goals that work out in our favor. Due to the instrumental goal of goal integrity, discussed in Section 5, it is unlikely that a superintelligent AI would allow us to tamper with its final goal, so the favorable goal needs to be installed into the AI *before* it attains superintelligence. This is the aim of the **AI Alignment** research program, formulated (under the alternative heading **Friendly AI**, which however is perhaps best avoided as it has an unnecessarily anthropomorphic ring to it) in Yudkowsky's seminal 2008 paper [42], and much discussed ever since; see, e.g., [6], [16] and [40].

Following Bostrom [6], we can think of AI Alignment as two problems: First, the difficult technical problem of how to encode whatever the desired goals are and install them into the AI – Bostrom calls this the **value loading problem** and "a research challenge worthy of some of the next generation's best mathematical talent". Second, the ethical problem of what the desired goals are, who gets to determine them, and via what procedure (democratic or otherwise). We probably do not want to leave it to a small group of AI developers in Silicon Valley or elsewhere to decide on the fate of humanity for the rest of eternity. Most thinkers in this field (including Yudkowsky [42] and Bostrom [6]) seem to agree that rather than explicitly hand-coding the values we wish the AI to have, an indirect approach is better, where somehow the AI is instructed to figure out what we want – or even better, what we would have wanted if we were more knowledgable and ethically mature, and had more time to think about it.

A key insight going back at least to Yudkowsky [42] is that human values are highly fragile, in the sense that getting them just a little bit wrong can bring catastrophic consequenecs in the mighty hands of a superintelligent AI. There may also be a tension between what is good for humanity and what is good in a less anthropocentric and possibly more objective sense: for instance, the goal "maximize the amount of hedonic utility in the world" might in a sense be very good for the universe, but is also likely to lead to the prompt extinction of humanity, as our bodies and brains are probably very far from optimizing the amount of hedonic utility per kilogram of matter.

Solving the AI Alignment problem should in my opinion be a high on the list of today's most urgent research tasks, but not for the reason that AGI and superintelligence would be likely to emerge during the next few years (although see [44]). Rather, even if they are decades away, the problem may well be so difficult that we need those decades to solve it, with little or no room for procrastination.

7   Concluding remarks

Let me conclude with the following remarks.

1. The reader may have noticed the discrepancy between Turing's [41] use of plural in talking about "machines [taking] control", and my use of singular when talking about the superintelligent AI. My choice of singular is due to what Bostrom [6] speaks of as "decisive strategic advantage": especially in case of a fast takeoff, the first machine to attain superintelligence can be expected to take control in such a way as to prevent other machines from challenging its power monopoly. But this outcome is not certain, and Bostrom devotes a chapter also to what he calls multipolar outcomes, with no such monopoly. Such an outcome might arise if AGI is first attained via brain emulations, at a time when our understanding of the human brain is still not good enough to enable us to tweak with the emulations much beyond what we already do to our brains today; Hanson [23] offers a rich and fascinating account of the many societal exotica that such a breakthrough might lead to.

2. Creating superintelligence is of course difficult, but creating superintelligence *and* AI Alignment may be even more difficult. This means that if several actors (companies or countries) compete over being the first (and probably only) one to create superintelligence, there may be an incentive to cut corners on the AI Alignment task or maybe even ignore it altogether. Such a situation would be terribly dangerous (see, e.g., Miller [28] and Cave and ÓhÉigeartaigh[7]), and should be avoided, e.g., by creating a spirit of international cooperation rather than competition. That is possibly easier said than done.

3. Apart from superintelligence there are many other problems about the future of AI that we urgently need to deal with, concering, e.g., integrity and mass survelliance, the social consequences of sexbot technology, autonomous weapons arms races, or the effects of automation on unemployment. It is sometimes suggested that the superintelligence discourse in AI futurology is a dangerous distraction from these other problems; see, e.g., Dubhashi and Lappin [10]. I agree that these other problems are extremely important, but I do not agree that this means that we should ignore superintelligence. It would be bad if we managed to navigate all those more down-to-earth societal problems with AI, only to end up being turned into into paperclips. We need to deal with all of these problems, including superintelligence.

References

1   Stuart Armstrong. *Smarter Than Us: The Rise of Machine Intelligence*. Machine Intelligence Research Institute, Berkeley, CA, 2014.

2   Stuart Armstrong, Anders Sandberg, and Nick Bostrom. Thinking inside the box: controlling and using an oracle AI. *Minds and Machines*, 22:299–324, 2012.

3   Peter Bentley. The three laws of artificial intelligence: Dispelling common myths. In *Should we fear artificial intelligence?*, pages 6–12. The EU Parliament's STOA (Science and Technology Options Assessment) committee, Brussels, 2018.

4   Nick Bostrom. Ethical issues in advanced artificial intelligence. In I. Smit et al, editor, *Cognitive, Emotive and Ethical Aspects of Decision Making in Humans and in A rtificial Intelligence, Vol. 2*, pages 12–17. International Institute of Advanced Studies in Systems Research and Cybernetics, 2003.

5   Nick Bostrom. The superintelligent will: motivation and instrumental rationality in advanced artificial agents. *Minds and Machines*, 22:71–85, 2012.

6   Nick Bostrom. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, Oxford, 2014.

7   Stephen Cave and Seán ÓhÉigeartaigh. An AI race for strategic advantage: rhetoric and risks. preprint, 2018.

8 Allan Dafoe and Stuart Russell. Yes, we are worried about the existential risk of artificial intelligence. *MIT Technology Review*, November 2016.

9 David Deutsch. Quantum theory, the Church–Turing principle and the universal quantum computer. *Proceedings of the Royal Society A*, 400:97–117, 1985.

10 Devdatt Dubhashi and Shalom Lappin. AI dangers: imagined and real. *Communications of the ACM*, 60:43–45, 2016.

11 Oren Etzioni. No, the experts don't think superintelligent AI is a threat to humanity. *MIT Technology Review*, September 2016.

12 Thomas Gilovich, Dale Griffin, and Daniel Kahnemann. *Heuristics and Biases: The Psychology of Intuitive Judgment.* Cambridge University Press, Cambridge, UK, 2002.

13 I.J. Good. Speculations concerning the first ultraintelligent machine. In F. Alt and M. Rubinoff, editors, *Advances in Computers*, volume 6, 1965.

14 Katja Grace, John Salvatier, Allan Dafoe, Baobao Zhang, and Owain Evans. When will AI exceed human performance? Evidence from AI experts. arXiv:1705.08807, 2017.

15 Olle Häggström. Does the Chinese room argument preclude a robot uprising? OUPblog, Oxford University Press, January 2016.

16 Olle Häggström. *Here Be Dragons: Science, Technology and the Future of Humanity.* Oxford University Press, Oxford, 2016.

17 Olle Häggström. Vulgopopperianism. Häggström hävdar, February 2017.

18 Olle Häggström. Challenges to the Omohundro–Bostrom framework for AI motivations. preprint, 2018.

19 Olle Häggström. Remarks on artificial intelligence and rational optimism. In *Should we fear artificial intelligence?*, pages 19–26. The EU Parliament's STOA (Science and Technology Options Assessment) committee, Brussels, 2018.

20 Olle Häggström. A spectacularly uneven AI report. Häggström hävdar, March 2018.

21 Olle Häggström. Strategies for an unfriendly oracle AI with reset button. In Roman Yampolskiy, editor, *Artificial Intelligence Safety and Security.* CRC Press, Boca Raton, FL, to appear, 2018.

22 Brett Hall. Superintelligence. Part 4: Irrational rationality. `http://www.bretthall.org/superintelligence-4.html`, 2016.

23 Robin Hanson. *The Age of Em: Work, Love and Life When Robots Rule the Earth.* Oxford University Press, Oxford, 2016.

24 Thore Husfeldt. The monster in the library of Turing. `https://thorehusfeldt.files.wordpress.com/2015/04/bostrom.pdf`, 2015.

25 Joel Janai, Fatma Güney, Aseem Behl, and Andreas Geiger. Computer vision for autonomous vehicles: problems, datasets and state-of-the-art. arXiv:1704.05519, 2017.

26 Michael Jordan. Artificial intelligence – the revolution that hasn't happened yet. *Medium*, April 2018.

27 Ray Kurzweil. *The Singularity Is Near: When Humans Transcend Biology.* Viking, New York, 2005.

28 James Miller. *Singularity Rising: Surviving and Thriving in a Smarter, Richer, and More Dangerous World.* Benbella, Dallas, TX, 2012.

29 Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518:529–533, 2015.

30 Vincent Müller and Nick Bostrom. Future progress in artificial intelligence: a survey of expert opinion. In *Fundamental Issues of Artificial Intelligence*, pages 553–571. Springer, New York, 2016.

31   Stephen Omohundro. The basic AI drives. In P. Wang, B. Goertzel, and S. Franklin, editors, *Artificial General Intelligence 2008: Proceedings of the First AGI Conference*, pages 483–492. IOS, Amsterdam, 2008.

32   Stephen Omohundro. Rational artificial intelligence for the greater good. In A. Eden, J. Moor, J. Søraker, and E. Stenhart, editors, *Singularity Hypotheses: A Scientific and Philosophical Assessment*, pages 161–175. Springer, New York, 2012.

33   Karl Popper. *The Logic of Scientific Discovery*. Hutchinson, London, 1959.

34   Anders Sandberg and Nick Bostrom. Whole brain emulation: a roadmap. Future of Humanity Institute technical report #2008-3, 2008.

35   John Searle. What your computer can't know. *New York Review of Books*, October 2014.

36   David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy Lillicrap, Karen Simonyan, and Demis Hassabis. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. arXiv:1712.01815, 2017.

37   R.J. Skerry-Ryan, Eric Battenberg, Ying Xiao, Yuxuan Wang, Daisy Stanton, Joel Shor, Ron J. Weiss, Rob Clark, and Rif A. Saurous. Towards end-to-end prosody transfer for expressive apeech aynthesis with Tacotron. arXiv:1803.09047, 2018.

38   Kaj Sotala. How feasible is the rapid development of artificial superintelligence? *Physica Scripta*, 92:113001, 2017.

39   Max Tegmark. Friendly artificial intelligence: the physics challenge. arXiv:1409.0813, 2014.

40   Max Tegmark. *Life 3.0: Being Human in the Age of Artificial Intelligence*. Brockman Inc, New York, 2017.

41   Alan Turing. Intelligent machinery: a heretical theory. BBC, 1951.

42   Eliezer Yudkowsky. Artificial intelligence as a positive and negative factor in global risk. In Nick Bostrom and Milan Cirkovic, editors, *Global Catastrophic Risks*, pages 308–345. Oxford University Press, Oxford, 2008.

43   Eliezer Yudkowsky. Intelligence explosion miroeconomics. Machine Intelligence Research Institute, 2013.

44   Eliezer Yudkowsky. There's no fire alarm for artificial general intelligence. Machine Intelligence Research Institute, 2017.

Patterns in Random Permutations Avoiding Some Other Patterns

Svante Janson[1]

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
svante.janson@math.uu.se

https://orcid.org/0000-0002-9680-2790

---- **Abstract** --

Consider a random permutation drawn from the set of permutations of length n that avoid a given set of one or several patterns of length 3. We show that the number of occurrences of another pattern has a limit distribution, after suitable scaling. In several cases, the limit is normal, as it is in the case of unrestricted random permutations; in other cases the limit is a non-normal distribution, depending on the studied pattern. In the case when a single pattern of length 3 is forbidden, the limit distributions can be expressed in terms of a Brownian excursion.

The analysis is made case by case; unfortunately, no general method is known, and no general pattern emerges from the results.

1 Introduction

Let $\mathfrak{S}_n$ be the set of permutations of $[n] := \{1, \ldots, n\}$, and $\mathfrak{S}_* := \bigcup_{n \geq 1} \mathfrak{S}_n$. If $\sigma = \sigma_1 \cdots \sigma_m \in \mathfrak{S}_m$ and $\pi = \pi_1 \cdots \pi_n \in \mathfrak{S}_n$, then an *occurrence* of σ in π is a subsequence $\pi_{i_1} \cdots \pi_{i_m}$, with $1 \leq i_1 < \cdots < i_m \leq n$, that has the same order as σ, i.e., $\pi_{i_j} < \pi_{i_k} \iff \sigma_j < \sigma_k$ for all $j, k \in [m]$. We let $n_\sigma(\pi)$ be the number of occurrences of σ in π, and note that

$$\sum_{\sigma \in \mathfrak{S}_m} n_\sigma(\pi) = \binom{n}{m}, \tag{1}$$

for every $\pi \in \mathfrak{S}_n$. For example, an inversion is an occurrence of 21, and thus $n_{21}(\pi)$ is the number of inversions in π.

We say that π *avoids* another permutation τ if $n_\tau(\pi) = 0$. Let

$$\mathfrak{S}_n(\tau) := \{\pi \in \mathfrak{S}_n : n_\tau(\pi) = 0\}, \tag{2}$$

---

the set of permutations of length n that avoid τ. More generally, for any set $T = \{\tau_1, \ldots, \tau_k\}$ of permutations, let

$$\mathfrak{S}_n(T) = \mathfrak{S}_n(\tau_1, \ldots, \tau_k) := \bigcap_{i=1}^{k} \mathfrak{S}_n(\tau_i), \tag{3}$$

the set of permutations of length n that avoid all $\tau_i \in T$. We also let $\mathfrak{S}_*(T) := \bigcup_{n=1}^{\infty} \mathfrak{S}_n(T)$ be the set of T-avoiding permutations of arbitrary length.

The classes $\mathfrak{S}_*(\tau)$ and, more generally, $\mathfrak{S}_*(T)$ have been studied for a long time. For examples relevant to analysis of algorithms, see e.g. [13, Exercise 2.2.1-5] (π can be obtained by a stack if and only if $\pi \in \mathfrak{S}_n(312)$; equivalently: π is stack-sortable if and only if $\pi \in \mathfrak{S}_n(312)$); [13, Exercise 2.2.1-10,11] and [17] (π is deque-sortable if and only if $\pi \in \mathfrak{S}_n(2431, 4231)$; [16] ($\pi$ can be sorted by 2 parallel queues if and only if $\pi \in \mathfrak{S}_n(321)$. Further examples are given in [15], Exercises 6.19 x (321), y (312), ee (321), ff (312), ii (231), oo (132), xx (321); 6.25 g (321); 6.39 k, l ($\{2413, 3142\}$), m ($\{1342, 1324\}$); 6.47 a ($\{4231, 3412\}$); 6.48 (1342). See also [3].

In particular, one classical problem is to enumerate the sets $\mathfrak{S}_n(T)$, either exactly or asymptotically, see e.g. [3, Chapters 4–5] and [14].

The general problem that concerns us is to take a fixed set T of one or several permutations and let $\boldsymbol{\pi}_{T;n}$ be a uniformly random T-avoiding permutation, i.e., a uniformly random element of $\mathfrak{S}_n(T)$, and then study the asymptotic distribution of the random variable $n_\sigma(\boldsymbol{\pi}_{T;n})$ (as $n \to \infty$) for some other fixed permutation σ. (Only σ that are themselves T-avoiding are interesting, since otherwise $n_\sigma(\boldsymbol{\pi}_{T;n}) = 0$.)

Here we study the cases when T is a set of permutations of length 3. The cases when T contains a permutation of length ≤ 2 are trivial, since then there is at most one permutation in $\mathfrak{S}_n(T)$ for any n. The case of forbidding one or several permutations of length ≥ 4 seems much more complicated, but there are recent impressive results for $\mathfrak{S}_n(2413, 3142)$ (separable permutations) by Bassino, Bouvel, Féray, Gerin, and Pierrot [2], with generalizations to some other classes in [1].

There are $2^6 = 64$ sets T of permutations of length 3. Of these, every T that contains $\{123, 321\}$, and every T with $|T| \geq 4$ is trivial, in the sense that $\mathfrak{S}_n(T)$ contains at most 2 elements for any $n \geq 5$ (see [14]). Ignoring these cases, there are $1 + 6 + 14 + 16 = 37$ remaining cases (with $|T| = 0, 1, 2, 3$, respectively), and by symmetries, see Appendix A, these reduce to $1 + 2 + 4 + 4 = 11$ non-equivalent cases, which are treated in Sections 2–12. For further details, see [12], [8], [9], [10]; these papers also contain further references to related work, and to some of the many papers by various authors that study other properties of random τ-avoiding permutations.

The cases studied here, i.e., the non-trivial cases with $T \subset \mathfrak{S}_3$, all have asymptotic distributions of one of the following two types.

I. Normal limits: For every $\sigma \in \mathfrak{S}_*(T)$, there exists constants α, β, γ such that, as $n \to \infty$,

$$\frac{n_\sigma(\boldsymbol{\pi}_{T;n}) - \beta n^\alpha}{n^{\alpha - 1/2}} \xrightarrow{\mathrm{d}} N(0, \gamma^2), \tag{4}$$

with convergence of all moments. Furthermore, assuming $|\sigma| \geq 2$, $\gamma^2 > 0$, so the limit is not deterministic, except possibly for one $\sigma \in \mathfrak{S}_m(T)$ for each length $m \geq 2$.
In particular, $\mathbb{E} \, n_\sigma(\boldsymbol{\pi}_{T;n}) \sim \beta n^\alpha$. Note that (4) implies concentration, in the sense

$$\frac{n_\sigma(\boldsymbol{\pi}_{T;n})}{\mathbb{E} \, n_\sigma(\boldsymbol{\pi}_{T;n})} \xrightarrow{\mathrm{p}} 1. \tag{5}$$

Table 1 The table shows whether $n_\sigma(\boldsymbol{\pi}_{T;n})$ has limits of type I or II; furthermore, the exponent $\alpha = \alpha(\sigma)$ is given in the column for the type. The last column shows the exceptional cases, if any, where the asymptotic variance vanishes. $C_n := \frac{1}{n+1}\binom{2n}{n}$ is a Catalan number; F_{n+1} is a Fibonacci number ($F_0 = 0$, $F_1 = 1$); s_{n-1} is a Schröder number; $D(\sigma)$ is the number of descents and $B(\sigma)$ is the number of blocks in σ.

T	$\|\mathfrak{S}_n(T)\|$	type I	type II	as. variance = 0
$\emptyset$	$n!$	$\|\sigma\|$		
$\{132\}$	C_n		$(\|\sigma\| + D(\sigma))/2$	$m \cdots 1$
$\{321\}$	C_n		$(\|\sigma\| + B(\sigma))/2$	$1 \cdots m$
$\{132, 312\}$	2^{n-1}	$\|\sigma\|$		
$\{231, 312\}$	2^{n-1}	$B(\sigma)$		$1 \cdots m$
$\{231, 321\}$	2^{n-1}	$B(\sigma)$		$1 \cdots m$
$\{132, 321\}$	$\binom{n}{2} + 1$		$\|\sigma\|$	
$\{231, 312, 321\}$	F_{n+1}	$B(\sigma)$		$1 \cdots m$
$\{132, 231, 312\}$	n		$\|\sigma\|$	
$\{132, 231, 321\}$	n		$\|\sigma\| - 1$ or $\|\sigma\|$	$1 \cdots m$
$\{132, 213, 321\}$	n		$\|\sigma\|$	
$\{2413, 3142\}$	s_{n-1}		$\|\sigma\|$	

II. Non-normal limits without concentration: For every $\sigma \in \mathfrak{S}_*(T)$, there exists a constant α such that

$$\frac{n_\sigma(\boldsymbol{\pi}_{T;n})}{n^\alpha} \xrightarrow{\mathrm{d}} W_\sigma, \tag{6}$$

with convergence of all moments, for some random variable $W_\sigma > 0$. Hence, also

$$\frac{n_\sigma(\boldsymbol{\pi}_{T;n})}{\mathbb{E}\, n_\sigma(\boldsymbol{\pi}_{T;n})} \xrightarrow{\mathrm{d}} W'_\sigma, \tag{7}$$

with convergence of all moments, for some random variable $W'_\sigma > 0$ (necessarily with $\mathbb{E}\, W'_\sigma = 1$). Furthermore, assuming $|\sigma| \geq 2$, $\mathrm{Var}\, W_\sigma > 0$, so W_σ and W'_σ are not deterministic, except possibly for one $\sigma \in \mathfrak{S}_m(T)$ for each length $m \geq 2$.

▶ Remark. In all cases studied here, if there are any exceptional $\sigma \in \mathfrak{S}_*(T)$ with $\sigma \geq 2$ such that the limit in (4) or (6) is deterministic, i.e., the asymptotic variance is 0, then the exceptional σ are either all identity permutations $1 \cdots m$, or all decreasing permutations $m \cdots 1$. Furthermore, these exceptional cases arise because almost all of the $\binom{n}{|\sigma|}$ patterns in $\boldsymbol{\pi}_{T;n}$ of length $|\sigma|$ are occurrences of σ; more precisely, $\mathbb{E}\big(\binom{n}{|\sigma|} - n_\sigma(\boldsymbol{\pi}_{T;n})\big) = O\big(n^{|\sigma|-1}\big)$ for the exceptional cases of type I and $O\big(n^{|\sigma|-1/2}\big)$ for the cases of type II. (It follows that (5) holds also for the latter.)

We summarize the results for T consisting of permutations of length 3 in Table 1; for reference, we include the number $|\mathfrak{S}_n(T)|$ of T-avoiding permutations of length n, see e.g. [13, Exercises 2.2.1-4,5], [15, Exercise 6.19ee,ff], [3, Corollary 4.7], and [14]. We include also the case $T = \{2413, 3142\}$ from [2]; see [17] for the enumeration.

We see no obvious pattern in the existence of limits of type I or II in Table 1. Moreover, the proofs, sketched below, are done case by case; we have not succeeded to prove any general results, treating all (or at least some) forbidden sets T at the same time.

▶ Remark. We do not know whether a general set of forbidden permutations T has limits in distribution of $n_\sigma(\boldsymbol{\pi}_{T;n})$ (after normalization) at all, and even if limits exist, there is no known reason implying that they have to be of type I or II above; other types of limits are conceivable.

▶ Remark. The non-normal limits in the cases {132}, {321} and {2413, 3142} can all be expressed as functionals of a Brownian excursion **e**, see [8, 9, 2]. However, the expressions in these three cases are, in general, quite different (and obtained by quite different arguments), so there is no obvious hope for a unification. (The other cases of non-normal limits in Table 1 are different, and of a more elementary kind.)

1.1 Some notation

Let $\iota = \iota_n$ be the identity permutation of length n.

If $\sigma \in \mathfrak{S}_m$ and $\tau \in \mathfrak{S}_n$, their *composition* $\sigma * \tau \in \mathfrak{S}_{m+n}$ is defined by letting τ act on $[m+1, m+n]$ in the natural way; more formally, $\sigma * \tau = \pi \in \mathfrak{S}_{m+n}$ where $\pi_i = \sigma_i$ for $1 \leq i \leq m$, and $\pi_{j+m} = \tau_j + m$ for $1 \leq j \leq n$. We say that a permutation $\pi \in \mathfrak{S}_*$ is *decomposable* if $\pi = \sigma * \tau$ for some $\sigma, \tau \in \mathfrak{S}_*$, and *indecomposable* otherwise; we also call an indecomposable permutation a *block*.

It is easy to see that any permutation $\pi \in \mathfrak{S}_*$ has a unique decomposition $\pi = \pi_1 * \cdots * \pi_\ell$ into indecomposable permutations (blocks) $\pi_1, \ldots, \pi_\ell$; we call these the *blocks of* π. (These are useful to characterize the permutations in some of the classes below.)

2 No restriction, $T = \emptyset$

As a background, consider first the case $T = \emptyset$, so $\mathfrak{S}_n(T) = \mathfrak{S}_n$; the set of all $n!$ permutations of length n. It is well-known, see Bóna [4, 5] and [12, Theorem 4.1], that if $\boldsymbol{\pi}_n$ is a uniformly random permutation in $\mathfrak{S}_n$, then $n_\sigma(\boldsymbol{\pi}_n)$ has an asymptotic normal distribution as $n \to \infty$ for every fixed permutation σ:

▶ **Theorem 1** (Bóna [4, 5]). *If $|\sigma| = m \geq 2$ then, as $n \to \infty$, for some $\gamma^2 > 0$,*

$$\frac{n_\sigma(\boldsymbol{\pi}_n) - \frac{1}{m!}\binom{n}{m}}{n^{m-1/2}} \xrightarrow{\text{d}} N(0, \gamma^2). \tag{8}$$

Sketch of proof. A random permutation $\boldsymbol{\pi}_n$ can be obtained by taking i.i.d. random variables $X_1, \ldots, X_n \sim U(0, 1)$ and considering their ranks. Then

$$n_\sigma(\boldsymbol{\pi}_n) = \sum_{i_1 < \cdots < i_m} f(X_{i_1}, \ldots, X_{i_m}) \tag{9}$$

for a suitable (indicator) function f. This sum is an asymmetric U-statistic, and the result follows by general results on U-statistics, see [6] and [11]. ◀

▶ Remark. The asymptotic variance γ^2 depends on σ. It can be calculated explicitly, and the same holds for all parameters γ^2 (or μ) in the limit theorems below. Moreover, the convergence (8) holds with convergence of all moments, and it holds jointly for any set of σ; also this holds for all later limit theorems too.

3 Avoiding 132

Consider next the cases when T consists of a single permutation of length 3. The symmetries in Appendix A leave two non-equivalent cases. In this section we avoid $T = \{132\}$; equivalent cases are {213}, {231}, {312}. Recall that the standard Brownian excursion $\mathbf{e}(x)$ is a random non-negative function on $[0, 1]$. Let

$$\lambda(\sigma) := |\sigma| + D(\sigma) \tag{10}$$

where $D(\sigma)$ is the number of *descents* in σ, i.e., indices i such that $\sigma_i > \sigma_{i+1}$ or (as a convenient convention) $i = |\sigma|$. Note that $1 \leq D(\sigma) \leq |\sigma|$, and thus

$$|\sigma| + 1 \leq \lambda(\sigma) \leq 2|\sigma|, \tag{11}$$

with the extreme values $\lambda(\sigma) = |\sigma| + 1$ if and only if $\sigma = 1 \cdots k$, and $\lambda(\sigma) = 2|\sigma|$ if and only if $\sigma = k \cdots 1$, for some $k = |\sigma|$.

▶ **Theorem 2** ([8]). *There exist strictly positive random variables Λ_σ such that as $n \to \infty$,*

$$n_\sigma(\boldsymbol{\pi}_{132;n})/n^{\lambda(\sigma)/2} \xrightarrow{\mathrm{d}} \Lambda_\sigma. \tag{12}$$

Sketch of proof. The analysis is based on a well-known bijection with binary trees and Dyck paths, and the, also well-known, convergence in distribution of random Dyck paths to a Brownian excursion. For (not so simple) details, see [8]. ◀

The limit variables Λ_σ in Theorem 2 can be expressed as functionals of a Brownian excursion $\mathbf{e}(x)$, see [8]; the description is, in general, rather complicated, but some cases are simple. Moments of the variables Λ_σ can be calculated by a recursion formula given in [8].

▶ **Example 3.** In the special case $\sigma = 12$, $\Lambda_{12} = \sqrt{2} \int_0^1 \mathbf{e}(x) \, \mathrm{d}x$, see [8, Example 7.6]; this is (apart from the factor $\sqrt{2}$) the well-known *Brownian excursion area*, see e.g. [7] and the references there.

For the number n_{21} of inversions, we thus have

$$\frac{\binom{n}{2} - n_{21}(\boldsymbol{\pi}_{132;n})}{n^{3/2}} = \frac{n_{12}(\boldsymbol{\pi}_{132;n})}{n^{3/2}} \xrightarrow{\mathrm{d}} \Lambda_{12} = \sqrt{2} \int_0^1 \mathbf{e}(x) \, \mathrm{d}x. \tag{13}$$

By symmetries, see Appendix A, the left-hand side can also be seen as the number of inversions $n_{21}(\boldsymbol{\pi}_{231;n})$ or $n_{21}(\boldsymbol{\pi}_{312;n})$, normalized by $n^{3/2}$, where we instead avoid 231 or 312.

4 Avoiding 321

In this section we avoid $T = \{321\}$. The case $T = \{123\}$ is equivalent.

$\mathfrak{S}_n(321)$ is treated in detail in [9]. As for $\mathfrak{S}_n(132)$ in Section 3, the analysis is based on a well-known bijection with Dyck paths, but the details are very different, and so are in general the resulting limit distributions.

▶ **Theorem 4** ([9]). *Let $\sigma \in \mathfrak{S}_*(321)$. Let $m := |\sigma|$, and suppose that σ has ℓ blocks of lengths $m_1, \ldots, m_\ell$. Then, as $n \to \infty$,*

$$n_\sigma(\boldsymbol{\pi}_{321;n})/n^{(m+\ell)/2} \xrightarrow{\mathrm{d}} W_\sigma \tag{14}$$

for a positive random variable W_σ that can be represented as

$$W_\sigma = w_\sigma \int_{0 < t_1 < \cdots < t_\ell < 1} \mathbf{e}(t_1)^{m_1 - 1} \cdots \mathbf{e}(t_\ell)^{m_\ell - 1} \, \mathrm{d}t_1 \cdots \mathrm{d}t_\ell, \tag{15}$$

where w_σ is positive constant.

Sketch of proof. As for Theorem 2, the analysis is based on a bijection with Dyck paths, and the convergence in distribution of random Dyck paths to a Brownian excursion. For details, see [8]. ◀

In this case, we have an explicit general formula (15) for the limit variables. On the other hand, we do not know how to compute even the mean $\mathbb{E}\, W_\sigma$ in general; see [9] for calculations in various special cases.

▶ **Example 5.** Let $\sigma = 21$. Then $w_{21} = 2^{-1/2}$, see [9], and thus (14)–(15), with $\ell = 1$ and $m_1 = m = 2$, yield for the number of inversions,

$$\frac{n_{21}(\boldsymbol{\pi}_{321;n})}{n^{3/2}} \xrightarrow{\mathrm{d}} 2^{-1/2} \int_0^1 \mathbf{e}(x)\, \mathrm{d}x. \tag{16}$$

Note that the limit in (16) differs from the one in (13) by a factor 2.

5 Avoiding {132,312}

In this section we avoid $T = \{132, 312\}$. Equivalent sets are $\{132, 231\}$, $\{213, 231\}$, $\{213, 312\}$.

▶ **Theorem 6.** *For any $m \geq 2$ and $\sigma \in \mathfrak{S}_m(132, 312)$, as $n \to \infty$,*

$$\frac{n_\sigma(\boldsymbol{\pi}_{132,312;n}) - 2^{1-m} n^m / m!}{n^{m-1/2}} \xrightarrow{\mathrm{d}} N\big(0, \gamma^2\big). \tag{17}$$

Sketch of proof. It was shown by [14, Proposition 12] (in an equivalent formulation) that *a permutation π belongs to the class $\mathfrak{S}_*(132, 312)$ if and only if every entry π_i is either a maximum or a minimum.* We encode a permutation $\pi \in \mathfrak{S}_n(132, 312)$ by a sequence $\xi_2, \ldots, \xi_n \in \{\pm 1\}^{n-1}$, where $\xi_j = 1$ if π_j is a maximum in π, and $\xi_j = -1$ if π_j is a minimum. This is a bijection, and hence the code for a uniformly random $\boldsymbol{\pi}_{132,312;n}$ has $\xi_2, \ldots, \xi_n$ i.i.d. with the symmetric Bernoulli distribution $\mathbb{P}(\xi_j = 1) = \mathbb{P}(\xi_j = -1) = \frac{1}{2}$.

Let $\sigma \in \mathfrak{S}_m(132, 312)$ have the code $\eta_2, \ldots, \eta_m$. Then $\pi_{i_1} \cdots \pi_{i_m}$ is an occurrence of σ in π if and only if $\xi_{i_j} = \eta_j$ for $2 \leq j \leq m$. Consequently, $n_\sigma(\boldsymbol{\pi}_{132,312;n})$ is a U-statistic

$$n_\sigma(\boldsymbol{\pi}_{132,312;n}) = \sum_{i_1 < \cdots < i_m} f\big(\xi_{i_1}, \ldots, \xi_{i_m}\big), \tag{18}$$

where

$$f\big(\xi_1, \ldots, \xi_m\big) := \prod_{j=2}^m \mathbf{1}\{\xi_j = \eta_j\}. \tag{19}$$

Note that f does not depend on the first argument.

The result now follows from the theory of U-statistics [6], [11]. ◀

▶ **Example 7.** For the number of inversions, we have $\sigma = 21$ and $m = 2$, $\eta_2 = -1$. A calculation yields $\mu = \frac{1}{2}$ and $\gamma^2 = \frac{1}{12}$, and thus Theorem 6 yields

$$\frac{n_{21}(\boldsymbol{\pi}_{132,312;n}) - n^2/4}{n^{3/2}} \xrightarrow{\mathrm{d}} N\big(0, \tfrac{1}{12}\big), \tag{20}$$

6 Avoiding {231,312}

In this section we avoid $T = \{231, 312\}$. The only equivalent set is $\{132, 213\}$.

▶ **Theorem 8.** *Let $\sigma \in \mathfrak{S}_m(231, 312)$ have block lengths $\ell_1, \ldots, \ell_b$. Then, as $n \to \infty$,*

$$\frac{n_\sigma(\boldsymbol{\pi}_{231,312;n}) - n^b / b!}{n^{b-1/2}} \xrightarrow{\mathrm{d}} N\big(0, \gamma^2\big). \tag{21}$$

Sketch of proof. It was shown by [14, Proposition 12] (in an equivalent form) that *a permutation π belongs to the class $\mathfrak{S}_*(231, 312)$ if and only if every block in π is decreasing, i.e., of the type $\ell(\ell - 1)\cdots 21$ for some ℓ.* Hence there exists exactly one block of each length $\ell \geq 1$, and a permutation $\pi \in \mathfrak{S}_*(231, 312)$ can be encoded by its sequence of block lengths. In this section, let $\pi_{\ell_1,\ldots,\ell_b}$ denote the permutation in $\mathfrak{S}_*(231, 312)$ with block lengths $\ell_1,\ldots,\ell_b$.

A uniformly random permutation $\boldsymbol{\pi}_{231,312;n}$ can be generated as $\pi_{L_1,\ldots,L_B}$, where the block lengths $L_1,\ldots,L_B$ are obtained from an infinite i.i.d. sequence $L_1, L_2, \cdots \sim \text{Ge}(\frac{1}{2})$, stopped at B such that $L_1 + \cdots + L_B \geq n$, and then adjusting L_B such that $L_1 + \cdots + L_B = n$.

Let $\sigma \in \mathfrak{S}_*(231, 312)$ have block lengths $\ell_1,\ldots,\ell_b$, so that $\sigma = \pi_{\ell_1,\ldots,\ell_b}$. Then,

$$n_\sigma\big(\pi_{L_1,\ldots,L_B}\big) = \sum_{1 \leq i_1 < \cdots < i_b \leq B} \prod_{j=1}^{b} \binom{L_{i_j}}{\ell_i}. \tag{22}$$

This is again a kind of U-statistic, but it is based on the sequence $L_1,\ldots,L_B$ of random length B, obtained by stopping the infinite sequence L_i. Nevertheless, general results for U-statistics cover this modification and yield the result, see [11]. ◄

▶ **Example 9.** For the number of inversions, we have $\sigma = 21$ and $b = 1$, $\ell_1 = 2$. A calculation yields $\gamma^2 = 6$, and Theorem 8 yields

$$\frac{n_{21}(\boldsymbol{\pi}_{231,312;n}) - n}{n^{1/2}} \xrightarrow{\text{d}} N(0, 6). \tag{23}$$

7 Avoiding {231, 321}

In this section we avoid $T = \{231, 321\}$. Equivalent sets are $\{123, 132\}$, $\{123, 213\}$, $\{312, 321\}$.

▶ **Theorem 10.** *Let $\sigma \in \mathfrak{S}_m(231, 321)$ have block lengths $\ell_1,\ldots,\ell_b$, and let b_1 be the number of blocks of length $\ell_i = 1$. Then, as $n \to \infty$,*

$$\frac{n_\sigma(\boldsymbol{\pi}_{231,321;n}) - 2^{b_1-b}n^b/b!}{n^{b-1/2}} \xrightarrow{\text{d}} N\big(0, \gamma^2\big). \tag{24}$$

Sketch of proof. It was shown by [14, Proposition 12] (in an equivalent form) that *a permutation π belongs to the class $\mathfrak{S}_*(231, 321)$ if and only if every block in π is of the type $\ell 12 \cdots (\ell - 1)$ for some ℓ.* Thus, as in Section 6, a permutation in $\mathfrak{S}_*(231, 321)$ is determined by its block lengths, and these can be arbitrary. Hence, a uniformly random $\boldsymbol{\pi}_{231,321;n}$ has block lengths $L_1,\ldots,L_B$ with the same distribution as in Section 6. Letting now σ be the permutation in $\mathfrak{S}_*(231, 321)$ with block lengths $\ell_1,\ldots,\ell_b$, $n_\sigma(\boldsymbol{\pi}_{231,321;n})$ is a function of the block lengths $L_1,\ldots,L_B$ that is similar (but not identical) to (22). This time some lower order terms appear, but they may be neglected, and the remainder is a U-statistic similar to the one in the proof of Theorem 8, and the result follows in the same way. ◄

▶ **Example 11.** For the number of inversions, we have $\sigma = 21$ and $b = 1$, $\ell_1 = 2$, $b_1 = 0$. A calculation yields $\gamma^2 = 1/4$, and Theorem 10 yields

$$\frac{n_{21}(\boldsymbol{\pi}_{231,321;n}) - n/2}{n^{1/2}} \xrightarrow{\text{d}} N(0, \tfrac{1}{4}). \tag{25}$$

In fact, in this special case it can be seen that we have the exact distribution

$$n_{21}(\boldsymbol{\pi}_{231,321;n}) \sim \text{Bi}\big(n - 1, \tfrac{1}{2}\big). \tag{26}$$

8 Avoiding {132, 321}

In this section we avoid $T = \{132, 321\}$. Equivalent sets are $\{123, 231\}, \{123, 312\}, \{213, 321\}$.

It was shown in [14, Proposition 13] that *a permutation π belongs to $\mathfrak{S}_*(132, 321)$ if and only if either $\pi = \iota_n$ for some n, or $\pi = \pi_{k,\ell,m}$ for some $k, \ell \geq 1$ and $m \geq 0$, where, in this section,*

$$\pi_{k,\ell,m} := (\ell + 1, \ldots, \ell + k, 1, \ldots, \ell, k + \ell + 1, \ldots, k + \ell + m) \in \mathfrak{S}_{k+\ell+m}. \tag{27}$$

Recall that the Dirichlet distribution $\mathrm{Dir}(1, 1, 1)$ is the uniform distribution on the simplex $\{(x, y, z) \in \mathbb{R}_+^3 : x + y + z = 1\}$.

▶ **Theorem 12.** *Let $\sigma \in \mathfrak{S}_*(132, 321)$. Then the following hold as $n \to \infty$.*
  (i) *If $\sigma = \pi_{i,j,p}$ for some i, j, p, then*

$$n^{-(i+j+p)} n_\sigma(\boldsymbol{\pi}_{132,321;n}) \xrightarrow{\mathrm{d}} W_{i,j,p} := \frac{1}{i!\, j!\, p!} X^i Y^j Z^p, \tag{28}$$

  where $(X, Y, Z) \sim \mathrm{Dir}(1, 1, 1)$.
  (ii) *If $\sigma = \iota_i$, then*

$$n^{-i} n_\sigma(\boldsymbol{\pi}_{132,321;n}) \xrightarrow{\mathrm{d}} W_i := \frac{1}{i!} \big((X + Z)^i + (Y + Z)^i - Z^i\big), \tag{29}$$

  with $(X, Y, Z) \sim \mathrm{Dir}(1, 1, 1)$ as in i.

Sketch of proof. For asymptotic results, we may ignore the case when $\boldsymbol{\pi}_{132,321;n} = \iota_n$. Conditioning on $\boldsymbol{\pi}_{132,321;n} \neq \iota_n$, we have $\boldsymbol{\pi}_{132,321;n} = \pi_{K,L,n-K-L}$, where K and L are random with (K, L) uniformly distributed over the set $\{K, L \geq 1 : K + L \leq n\}$. As $n \to \infty$, we thus have

$$\Big(\frac{K}{n}, \frac{L}{n}, \frac{n - K - L}{n}\Big) \xrightarrow{\mathrm{d}} (X, Y, Z) \sim \mathrm{Dir}(1, 1, 1). \tag{30}$$

If $\sigma = \pi_{i,j,p}$ for some i, j, p, then it is easily seen that

$$n_\sigma(\pi_{k,\ell,m}) = \binom{k}{i}\binom{\ell}{j}\binom{m}{p}. \tag{31}$$

Similarly, if $\sigma = \iota_i$, then, by inclusion-exclusion,

$$n_\sigma(\pi_{k,\ell,m}) = \binom{k + m}{i} + \binom{\ell + m}{i} - \binom{m}{i}. \tag{32}$$

These exact formulas and (30) yield the results. ◀

▶ **Corollary 13.** *The number of inversions has the asymptotic distribution*

$$n^{-2} n_{21}(\boldsymbol{\pi}_{132,321;n}) \xrightarrow{\mathrm{d}} W := XY, \tag{33}$$

with (X, Y) as above; the limit variable W has density function

$$2\log\big(1 + \sqrt{1 - 4x}\big) - 2\log\big(1 - \sqrt{1 - 4x}\big), \qquad 0 < x < 1/4, \tag{34}$$

and moments

$$\mathbb{E}\, W^r = 2\frac{r!^2}{(2r + 2)!}, \qquad r > 0. \tag{35}$$

9 Avoiding {231,312,321}

We proceed to sets of three forbidden patterns. In this section we avoid $T = \{231, 312, 321\}$. An equivalent set is $\{123, 132, 213\}$.

▶ **Theorem 14.** *Let* $\sigma \in \mathfrak{S}_m(231, 312, 321)$ *have block lengths* $\ell_1, \ldots, \ell_b$. *Then, as* $n \to \infty$,

$$\frac{n_\sigma(\boldsymbol{\pi}_{231,312,321;n}) - \mu n^b/b!}{n^{b-1/2}} \xrightarrow{\mathrm{d}} N(0, \gamma^2), \tag{36}$$

for some constants μ *and* γ^2.

Sketch of proof. It was shown in [14, Proposition 15*] (in an equivalent form) that *a permutation* π *belongs to the class* $\mathfrak{S}_*(231, 312, 321)$ *if and only if every block in* π *is decreasing and has length* ≤ 2, *i.e., every block is* 1 *or* 21. Hence, a permutation $\pi \in \mathfrak{S}_n(231, 312, 321)$ is uniquely determined by its sequence of block lengths $L_1, \ldots, L_B$, where each $L_i \in \{1, 2\}$ and $L_1 + \cdots + L_B = n$.

Let $p := (\sqrt{5} - 1)/2$, the golden ratio, so that $p + p^2 = 1$. Let X be a random variable with the distribution

$$\mathbb{P}(X = 1) = p, \qquad \mathbb{P}(X = 2) = p^2. \tag{37}$$

Consider an i.i.d. sequence $X_1, X_2, \ldots$ of copies of X, and let $S_k := \sum_{i=1}^k X_i$. Let further $B(n) := \min\{k : S_k \geq n\}$. Then, conditioned on $S_{B(n)} = n$, the sequence $X_1, \ldots, X_{B(n)}$ has the same distribution as the sequence $L_1, \ldots, L_B$ of block lengths of a uniformly random permutation $\boldsymbol{\pi}_{231,312,321;n}$.

Consequently, $n_\sigma(\boldsymbol{\pi}_{231,312,321;n})$ can be expressed as a U-statistic based on $X_1, \ldots, X_B$, conditioned as above. This conditioning does not affect the asymptotic distribution, see [11], and the result follows again by general results for U-statistics. ◀

▶ **Example 15.** For the number of inversions, $\sigma = 21$ we have $b = 1$. A calculation yields $\mu = 1 - p = (3 - \sqrt{5})/2$ and $\gamma^2 = 5^{-3/2}$. Consequently,

$$\frac{n_{21}(\boldsymbol{\pi}_{231,312,321;n}) - \frac{3-\sqrt{5}}{2}n}{n^{1/2}} \xrightarrow{\mathrm{d}} N(0, 5^{-3/2}). \tag{38}$$

10 Avoiding {132,231,312}

In this section we avoid $\{132, 231, 312\}$. Equivalent sets are $\{132, 213, 231\}$, $\{132, 213, 312\}$, $\{213, 231, 312\}$.

It was shown in [14, Proposition 16*] (in an equivalent form) that $\mathfrak{S}_n(132, 231, 312) = \{\pi_{k,n-k} : 1 \leq k \leq n\}$, *where, in this section,*

$$\pi_{k,\ell} := (k, \ldots, 1, k+1, \ldots, k+\ell) \in \mathfrak{S}_{k+\ell}, \qquad k \geq 1, \ell \geq 0. \tag{39}$$

▶ **Theorem 16.** *Let* $\sigma \in \mathfrak{S}_*(132, 231, 312)$. *Then the following hold as* $n \to \infty$, *with* $U \sim \mathsf{U}(0, 1)$.
  (i) *If* $\sigma = \pi_{k,m-k}$ *with* $2 \leq k \leq m$, *then*

$$n^{-m} n_\sigma(\boldsymbol{\pi}_{132,231,312;n}) \xrightarrow{\mathrm{d}} W_{k,m-k} := \frac{1}{k!\,(m-k)!} U^k(1 - U)^{m-k}. \tag{40}$$

(ii) If $\sigma = \pi_{1,m-1} = \iota_m$, then

$$n^{-m}n_\sigma(\boldsymbol{\pi}_{132,231,312;n}) \xrightarrow{\text{d}} W_{1,m-1} := \frac{1}{(m-1)!}U(1-U)^{m-1} + \frac{1}{m!}(1-U)^m$$

$$= \frac{1}{m!}\big(1+(m-1)U\big)(1-U)^{m-1}. \tag{41}$$

Sketch of proof. The random $\boldsymbol{\pi}_{132,231,312;n} = \pi_{K,n-K}$, where $K \in [n]$ is uniformly random. Obviously, as $n \to \infty$,

$$K/n \xrightarrow{\text{d}} U \sim \mathsf{U}(0,1). \tag{42}$$

Furthermore, if $\sigma = \pi_{k,\ell}$, then it is easy to see that

$$n_\sigma\big(\pi_{K,n-K}\big) = \begin{cases} \binom{K}{k}\binom{n-K}{\ell}, & k \geq 2, \\ K\binom{n-K}{\ell} + \binom{n-K}{\ell+1}, & k = 1. \end{cases} \tag{43}$$

The results follow.   ◀

▶ **Corollary 17.** *The number of inversions has the asymptotic distribution*

$$n^{-2}n_{21}(\boldsymbol{\pi}_{132,231,312;n}) \xrightarrow{\text{d}} W := U^2/2 \tag{44}$$

with $U \sim \mathsf{U}(0,1)$. *Thus,* $2W \sim B(\frac{1}{2},1)$, *and* W *has moments*

$$\mathbb{E}\, W^r = \frac{1}{2^r(2r+1)}, \qquad r > 0. \tag{45}$$

11   Avoiding {132,231,321}

In this section we avoid $\{132, 231, 321\}$. Equivalent sets are $\{123, 132, 231\}$, $\{123, 213, 312\}$, $\{213, 312, 321\}$, $\{123, 132, 312\}$, $\{123, 213, 231\}$, $\{132, 312, 321\}$, $\{213, 231, 321\}$.

It was shown in [14, Proposition 16*] (in an equivalent form) that $\mathfrak{S}_n(132, 231, 321) = \{\pi_{k,n-k} : 1 \leq k \leq n\}$, *where, in this section,*

$$\pi_{k,\ell} := (k, 1, \ldots, k-1, k+1, \ldots, k+\ell) \in \mathfrak{S}_{k+\ell}, \qquad k \geq 1, \ell \geq 0. \tag{46}$$

▶ **Theorem 18.** *Let* $\sigma \in \mathfrak{S}_*(132, 231, 321)$. *Then the following hold as* $n \to \infty$, *with* $U \sim \mathsf{U}(0,1)$.

(i) If $\sigma = \pi_{k,m-k}$ with $2 \leq k \leq m$, then

$$n^{-(m-1)}n_\sigma(\boldsymbol{\pi}_{132,231,321;n}) \xrightarrow{\text{d}} W_{k,m-k} := \frac{1}{(k-1)!\,(m-k)!}U^{k-1}(1-U)^{m-k}. \tag{47}$$

(ii) If $\sigma = \pi_{1,m-1} = \iota_m$, then

$$n^{-m}n_\sigma(\boldsymbol{\pi}_{132,231,321;n}) = \frac{1}{m!} + O\big(n^{-1}\big) \xrightarrow{\text{p}} \frac{1}{m!}. \tag{48}$$

Sketch of proof. The random permutation $\boldsymbol{\pi}_{132,231,321;n} = \pi_{K,n-K}$, where $K \in [n]$ is uniformly random. The results follow similarly to the proof of Theorem 16.   ◀

▶ **Corollary 19.** *The number of inversions* $n_{21}(\boldsymbol{\pi}_{132,231,321;n})$ *has a uniform distribution on* $\{0, \ldots, n-1\}$, *and thus the asymptotic distribution*

$$n^{-1}n_{21}(\boldsymbol{\pi}_{132,231,321;n}) \xrightarrow{\text{d}} U \sim \mathsf{U}(0,1). \tag{49}$$

12 Avoiding {132,213,321}

In this section we avoid $\{132, 213, 321\}$. An equivalent sets is $\{123, 231, 312\}$.

It was shown in [14, Proposition 16*] (in an equivalent form) that $\mathfrak{S}_n(132, 213, 321) = \{\pi_{k,n-k} : 1 \leq k \leq n\}$, *where, in this section,*

$$\pi_{k,\ell} := (\ell+1, \ldots, \ell+k, 1, \ldots, \ell) \in \mathfrak{S}_{k+\ell}, \qquad k \geq 1, \, \ell \geq 0. \tag{50}$$

▶ **Theorem 20.** *Let* $\sigma \in \mathfrak{S}_*(132, 213, 321)$*. Then the following hold as* $n \to \infty$*, with* $U \sim \mathsf{U}(0,1)$*.*

(i) *If* $\sigma = \pi_{k,m-k}$ *with* $1 \leq k \leq m-1$*, then*

$$n^{-m} n_\sigma(\boldsymbol{\pi}_{132,213,321;n}) \xrightarrow{\mathrm{d}} W_{k,m-k} := \frac{1}{k!\,(m-k)!} U^k (1-U)^{m-k}. \tag{51}$$

(ii) *If* $\sigma = \pi_{m,0} = \iota_m$*, then*

$$n^{-m} n_\sigma(\boldsymbol{\pi}_{132,213,321;n}) \xrightarrow{\mathrm{d}} W_{m,0} := \frac{1}{m!}\bigl(U^m + (1-U)^m\bigr). \tag{52}$$

Sketch of proof. Similarly to the proof of Theorem 16. ◀

▶ **Corollary 21.** *The number of inversions has the asymptotic distribution*

$$n^{-2} n_{21}(\boldsymbol{\pi}_{132,213,321;n}) \xrightarrow{\mathrm{d}} W := U(1-U), \tag{53}$$

with $U \sim \mathsf{U}(0,1)$*. Thus,* $4W \sim B(1, \frac{1}{2})$*, and* W *has moments*

$$\mathbb{E}\, W^r = \frac{\Gamma(r+1)^2}{\Gamma(2r+2)}, \qquad r > 0. \tag{54}$$

─── **References** ───

1. Frédérique Bassino, Mathilde Bouvel, Valentin Féray, Lucas Gerin, Mickaël Maazoun, and Adeline Pierrot. Universal limits of substitution-closed permutation classes. Preprint, `arXiv:1706.08333`, 2017.

2. Frédérique Bassino, Mathilde Bouvel, Valentin Féray, Lucas Gerin, and Adeline Pierrot. The Brownian limit of separable permutations. Preprint, `arXiv:1602.04960`, 2016.

3. Miklós Bóna. *Combinatorics of Permutations*. Chapman & Hall/CRC, Boca Raton, FL, 2004.

4. Miklós Bóna. The copies of any permutation pattern are asymptotically normal. Preprint, `arXiv:0712.2792`, 2007.

5. Miklós Bóna. On three different notions of monotone subsequences. In *Permutation patterns*, volume 376 of *London Math. Soc. Lecture Note Ser.*, pages 89–114. Cambridge Univ. Press, Cambridge, 2010.

6. Svante Janson. *Gaussian Hilbert Spaces*. Cambridge University Press, Cambridge, 1997.

7. Svante Janson. Brownian excursion area, Wright's constants in graph enumeration, and other Brownian areas. *Probab. Surv.*, 4:80–145, 2007.

8. Svante Janson. Patterns in random permutations avoiding the pattern 132. *Combin. Probab. Comput.*, 26(1):24–51, 2017.

9. Svante Janson. Patterns in random permutations avoiding the pattern 321. Preprint, `arXiv:1709.08427`, 2017.

10. Svante Janson. Patterns in random permutations avoiding some sets of multiple patterns. Preprint, `arXiv:1804.06071`, 2018.

11 Svante Janson. Renewal theory for asymmetric U-statistics. Preprint, `arXiv:1804.05509`, 2018.

12 Svante Janson, Brian Nakamura, and Doron Zeilberger. On the asymptotic statistics of the number of occurrences of multiple permutation patterns. *J. Comb.*, 6(1-2):117–143, 2015.

13 Donald E. Knuth. *The Art of Computer Programming. Vol. 1.* Addison-Wesley, Reading, MA, third edition, 1997.

14 Rodica Simion and Frank W. Schmidt. Restricted permutations. *European J. Combin.*, 6(4):383–406, 1985.

15 Richard P. Stanley. *Enumerative combinatorics. Vol. 2.* Cambridge University Press, Cambridge, 1999.

16 Robert Tarjan. Sorting using networks of queues and stacks. *J. Assoc. Comput. Mach.*, 19:341–346, 1972.

17 Julian West. Generating trees and the Catalan and Schröder numbers. *Discrete Math.*, 146(1-3):247–262, 1995.

A   Symmetries

For any permutation $\pi = \pi_1 \cdots \pi_n$, define its *inverse* π^{-1} in the usual way, and its *reversal* and *complement* by

$$\pi^{\mathsf{r}} := \pi_n \cdots \pi_1, \tag{55}$$

$$\pi^{\mathsf{c}} := (n + 1 - \pi_1) \cdots (n + 1 - \pi_n). \tag{56}$$

These three operations generate a group $\mathfrak{G}$ of 8 symmetries (isomorphic to the dihedral group D_4). It is easy to see that for any symmetry $\mathsf{s} \in \mathfrak{G}$,

$$n_{\sigma^{\mathsf{s}}}(\pi^{\mathsf{s}}) = n_\sigma(\pi). \tag{57}$$

Thus, if we define $T^{\mathsf{s}} := \{\tau^{\mathsf{s}} : \tau \in T\}$, then

$$\mathfrak{S}_n(T^{\mathsf{s}}) = \{\pi^{\mathsf{s}} : \pi \in \mathfrak{S}_n(T)\}, \tag{58}$$

and, for any permutation σ,

$$n_{\sigma^{\mathsf{s}}}(\boldsymbol{\pi}_{T^{\mathsf{s}};n}) \overset{\mathrm{d}}{=} n_\sigma(\boldsymbol{\pi}_{T;n}). \tag{59}$$

We say that the sets of forbidden permutations T and T^{s} are *equivalent*, and note that (59) implies that it suffices to consider one set T in each equivalence class $\{T^{\mathsf{s}} : \mathsf{s} \in \mathfrak{G}\}$.

Vanishing of Cohomology Groups of Random Simplicial Complexes

Oliver Cooley

Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria
cooley@math.tugraz.at

Nicola Del Giudice

Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria
delgiudice@math.tugraz.at

Mihyun Kang

Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria
kang@math.tugraz.at

Philipp Sprüssel

Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria
spruessel@math.tugraz.at

—— **Abstract** ——————————————————————

We consider k-dimensional random simplicial complexes that are generated from the binomial random $(k + 1)$-uniform hypergraph by taking the downward-closure, where $k \geq 2$. For each $1 \leq j \leq k - 1$, we determine when all cohomology groups with coefficients in $\mathbb{F}_2$ from dimension one up to j vanish and the zero-th cohomology group is isomorphic to $\mathbb{F}_2$. This property is not monotone, but nevertheless we show that it has a single sharp threshold. Moreover, we prove a hitting time result, relating the vanishing of these cohomology groups to the disappearance of the last minimal obstruction. Furthermore, we study the asymptotic distribution of the dimension of the j-th cohomology group inside the critical window. As a corollary, we deduce a hitting time result for a different model of random simplicial complexes introduced in [Linial and Meshulam, Combinatorica, 2006], a result which has only been known for dimension two [Kahle and Pittel, Random Structures Algorithms, 2016].

29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018).
Editors: James Allen Fill and Mark Daniel Ward; Article No. 7; pp. 7:1–7:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

1.1 Motivation

In their seminal paper [12], Erdős and Rényi introduced the uniform random graph and addressed the problem of determining the probability of this graph being connected. Nowadays, this classical result is usually stated for the binomial model, in which each edge is present with a given probability p independently: the connectedness of the binomial random graph $G(n, p)$ on n vertices undergoes a *phase transition* around the sharp threshold $p = \frac{\log n}{n}$ [24], where log denotes the natural logarithm.

▶ **Theorem 1.1.** *Let ω be any function of n which tends to infinity as $n \to \infty$. Then* with high probability,[1] *the following holds.*
  (i) *If $p = \frac{\log n - \omega}{n}$, then $G(n, p)$ is not connected.*
  (ii) *If $p = \frac{\log n + \omega}{n}$, then $G(n, p)$ is connected.*

As an even stronger result, Erdős and Rényi [12] determined the limiting probability for connectedness around the point of the phase transition. Subsequently, Bollobás and Thomason [7] proved a *hitting time* result, stating that whp the random graph *process* becomes connected at the very same time at which the last isolated vertex—the smallest obstruction for connectedness—disappears.

Since then, various higher-dimensional analogues of both random graphs and connectedness have been analysed and in particular two different approaches have received considerable attention. A first natural generalisation is the random k-uniform hypergraph $G_p = G(k; n, p)$ in which each $(k + 1)$-tuple of vertices forms a hyperedge with probability p independently. There are several natural ways of defining connectedness of G_p, which have been extensively studied [4, 5, 6, 8, 9, 10, 11, 15, 16, 22, 23].

A more recent approach concerns random simplicial complexes, of which a first model for the 2-dimensional case was introduced by Linial and Meshulam [17]. They considered $\mathbb{F}_2$-*homological* 1-*connectivity* of the random 2-complex as the vanishing of its first homology group with coefficients in the two-element field $\mathbb{F}_2$, which is equivalent to the vanishing of the first cohomology group. More precisely, the model $\mathcal{Y}_p = \mathcal{Y}(k; n, p)$ considered by Linial and Meshulam [17] for $k = 2$ and subsequently by Meshulam and Wallach [20] for general $k \geq 2$ is defined as follows. Starting from the full $(k - 1)$-*dimensional skeleton* on $[n] := \{1, \ldots, n\}$, that is, all simplices from dimension zero up to $k - 1$, each $(k + 1)$-set forms a k-simplex with probability p independently. They showed that the vanishing of the $(k - 1)$-th cohomology group $H^{k-1}(\mathcal{Y}_p; \mathbb{F}_2)$ with coefficients in $\mathbb{F}_2$ has a sharp threshold at $p = \frac{k \log n}{n}$.

▶ **Theorem 1.2** ([17, 20]). *Let ω be any function of n which tends to infinity as $n \to \infty$. Then with high probability, the following holds.*
  (i) *If $p = \frac{k \log n - \omega}{n}$, then $H^{k-1}(\mathcal{Y}_p; \mathbb{F}_2) \neq 0$.*
  (ii) *If $p = \frac{k \log n + \omega}{n}$, then $H^{k-1}(\mathcal{Y}_p; \mathbb{F}_2) = 0$.*

Later, Kahle and Pittel [15] derived a hitting time result for the case $k = 2$ and determined the limiting probability of $H^{k-1}(\mathcal{Y}_p; \mathbb{F}_2) = 0$ for general $k \geq 2$ and p in the critical window.

In this paper, we aim to bridge the gap between random hypergraphs and random simplicial complexes. We consider random simplicial k-complexes that arise as the downward-closure of random $(k + 1)$-uniform hypergraphs. Unlike $\mathcal{Y}_p$, in this model the presence of the

---

[1] With probability tending to 1 as n tends to infinity, whp for short.

full $(k-1)$-dimensional skeleton is not guaranteed, thus the vanishing of the cohomology groups of dimensions lower than $k-1$ does not hold trivially. Therefore, for each $j \in [k-1]$, we introduce $\mathbb{F}_2$-*cohomological j-connectedness* as the vanishing of *all* cohomology groups with coefficients in $\mathbb{F}_2$ from dimension one up to j and the zero-th cohomology group being isomorphic to $\mathbb{F}_2$.

Although this notion of connectedness is not monotone, we prove that nevertheless $\mathbb{F}_2$-cohomological j-connectedness has a sharp threshold. Furthermore, we derive a hitting time result and determine the limiting probability for $\mathbb{F}_2$-cohomological j-connectedness in the critical window. As a corollary, we deduce a hitting time result for $\mathcal{Y}_p$ in *general dimension*, thus extending the hitting time result of Kahle and Pittel [15].

1.2 Model

Throughout the paper let $k \geq 2$ be a fixed integer. For a positive integer ℓ, let $[\ell] := \{1, \dots, \ell\}$.

▶ **Definition 1.3.** A family $\mathcal{G}$ of non-empty finite subsets of a vertex set V is called a *simplicial complex* if it is downward-closed, i.e. if every non-empty set A that is contained in a set $B \in \mathcal{G}$ also lies in $\mathcal{G}$, and if the singleton $\{v\}$ is in $\mathcal{G}$ for every $v \in V$.

The elements of a simplicial complex $\mathcal{G}$ of cardinality $k+1$ are called *k-simplices* of $\mathcal{G}$. If $\mathcal{G}$ has no $(k+1)$-simplices, then we call it *k-dimensional*, or *k-complex*. If $\mathcal{G}$ is a k-complex, then for each $j = 0, \dots, k-1$ the *j-skeleton* of $\mathcal{G}$ is the j-complex formed by all i-simplices in $\mathcal{G}$ with $0 \leq i \leq j$.

We aim to define a model of random k-complexes starting from the binomial random $(k+1)$-uniform hypergraph $G_p = G(k; n, p)$ on vertex set $[n]$: the 0-simplices are the vertices of G_p, the k-simplices are the hyperedges of G_p, but there is more than one way to guarantee the downward-closure property, to obtain a simplicial complex. In the model $\mathcal{Y}_p$ considered by Meshulam and Wallach in [20], the full $(k-1)$-skeleton on $[n]$ is always included. In contrast, we shall only include those simplices that are *necessary* to ensure the downward-closure property.

▶ **Definition 1.4.** We denote by $\mathcal{G}_p = \mathcal{G}(k; n, p)$ the random k-dimensional simplicial complex on vertex set $[n]$ such that
- the 0-simplices are the singletons of $[n]$;
- the k-simplices are the hyperedges of G_p;
- for each $j \in [k-1]$, the j-simplices are exactly the $(j+1)$-subsets of hyperedges of G_p.

In other words, $\mathcal{G}_p$ is the random k-complex on $[n]$ obtained from G_p by taking the downward-closure of each hyperedge.

Given a simplicial complex $\mathcal{G}$, let $H^i(\mathcal{G}; \mathbb{F}_2)$ be its i-th cohomology group with coefficients in $\mathbb{F}_2$ (see Section 2.1 for the definition). Connectedness of $\mathcal{G}_p$ in the topological sense—which we call *topological connectedness* in order to distinguish it from other notions of connectedness—is equivalent to $H^0(\mathcal{G}_p; \mathbb{F}_2)$ being (isomorphic to) $\mathbb{F}_2$. We therefore define a notion of connectedness as follows.

▶ **Definition 1.5.** For a positive integer j, a simplicial complex $\mathcal{G}$ is called $\mathbb{F}_2$-*cohomologically j-connected* (*j-cohom-connected* for short) if
- $H^0(\mathcal{G}; \mathbb{F}_2) = \mathbb{F}_2$;
- $H^i(\mathcal{G}; \mathbb{F}_2) = 0$ for all $i \in [j]$.

One might define an analogous version of connectedness via the vanishing of *homology* groups, which would be equivalent to $\mathbb{F}_2$-cohomological j-connectedness by the Universal Coefficient Theorem (see e.g. [21]).

A significant difference between $\mathcal{G}_p$ and $\mathcal{Y}_p$ is that for $\mathcal{Y}_p$ the only requirement for $\mathbb{F}_2$-cohomologically $(k-1)$-connectedness is the vanishing of the $(k-1)$-th cohomology group, since the presence of the full $(k-1)$-skeleton guarantees topological connectedness and the vanishing of the j-th cohomology groups for all $j \in [k-2]$.

Moreover, it is important to observe that $\mathbb{F}_2$-cohomological j-connectedness is *not* a monotone increasing property of $\mathcal{G}_p$: adding a k-simplex to a j-cohom-connected complex might yield a complex without this property (see Example 2.3). Thus, the existence of a single threshold for j-cohom-connectedness is *not* guaranteed, but one of our main results shows that such a threshold indeed exists.

1.3 Main results

The main contributions of this paper are fourfold. Firstly, we prove (Theorem 1.8) that for each $j \in [k-1]$, $\mathbb{F}_2$-cohomological j-connectedness of $\mathcal{G}_p$ undergoes a phase transition at around probability

$$p_j := \frac{(j+1)\log n + \log\log n}{(k-j+1)n^{k-j}}(k-j)!. \tag{1}$$

Secondly, we prove a hitting time result (also Theorem 1.8), which relates the j-cohom-connectedness threshold to the disappearance of all copies of the *minimal obstruction M_j* (Definition 1.7). Thirdly, our results directly imply an analogous hitting time result for $\mathcal{Y}_p$ (Corollary 1.9), which Kahle and Pittel [15] proved for $k=2$. Lastly, we analyse the critical window around the threshold p_j, showing that inside the window the dimension of the j-th cohomology group converges in distribution to a Poisson random variable (Theorem 1.10).

Before defining the minimal obstruction M_j, we need the following concept.

▶ **Definition 1.6.** Given a k-simplex K in a k-complex $\mathcal{G}$, a collection $\mathcal{F} = \{P_0, \ldots, P_{k-j}\}$ of j-simplices forms a *j-flower in K* if $K = \bigcup_i P_i$ and $C := \bigcap_i P_i$ satisfies $|C| = j$. We call the j-simplices P_i the *petals* and the set C the *centre* of the j-flower $\mathcal{F}$.

Observe that for each k-simplex K and each $(j-1)$-simplex $C \subseteq K$, there is a unique j-flower in K with centre C, namely

$$\mathcal{F}(K, C) := \{C \cup \{w\} \mid w \in K \setminus C\}.$$

When j is clear from the context, we simply refer to a j-flower as a *flower*. A *j-cycle* is a set J of j-simplices such that every $(j-1)$-simplex is contained in an *even* number of j-simplices in J.

▶ **Definition 1.7.** A *copy of M_j* (see Figure 1) in a k-complex $\mathcal{G}$ is a triple (K, C, J) where
(M1) K is a k-simplex;
(M2) C is a $(j-1)$-simplex in K and each petal of the flower $\mathcal{F} = \mathcal{F}(K, C)$ is contained in *no other* k-simplex of $\mathcal{G}$;
(M3) J is a j-cycle that contains exactly one petal of $\mathcal{F}$, i.e. there exists a vertex $w_0 \in K \setminus C$ such that

$$J \cap \mathcal{F} = \left\{ C \cup \{w_0\} \right\}.$$

We will see (Lemma 2.2) that a copy of M_j can be interpreted as a minimal obstruction for $\mathbb{F}_2$-cohomological j-connectedness.

The random k-complex $\mathcal{G}_p$ can be viewed as a *process*, by assigning a *birth time* to each k-simplex. More precisely, for each $(k+1)$-set of vertices in $[n]$ independently, sample a birth
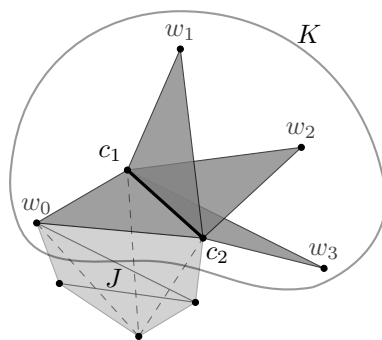
Figure 1 A copy of M_j for $k = 5$, $j = 2$. The centre $C = \{c_1, c_2\}$ lies in all petals $P_i = C \cup \{w_i\}$, $i = 0, \ldots, 3$ (dark grey), which are contained in no other k-simplex except K. The j-cycle J (light grey) intersects the flower $\mathcal{F}(K, C) = \{P_0, P_1, P_2, P_3\}$ only in the petal $P_0 = C \cup \{w_0\}$.

time uniformly at random from $[0, 1]$.[2] Then $\mathcal{G}_p$ is exactly the complex generated by the $(k+1)$-sets with birth times at most p, by taking the downward-closure. If p is gradually increased from 0 to 1, we may interpret $\mathcal{G}_p$ as a process. Thus, we can define p_{M_j} as the birth time of the k-simplex whose appearance causes the last copy of M_j to disappear. More formally, let

$$p_{M_j} := \sup\{p \in [0,1] \mid \mathcal{G}_p \text{ contains a copy of } M_j\}. \tag{2}$$

Our first main result is that the value p_{M_j} is the hitting time for j-cohom-connectedness of $\mathcal{G}_p$ and is "close" to p_j defined in (1), implying that p_j is in fact a sharp threshold for $\mathbb{F}_2$-cohomological j-connectedness.

▶ **Theorem 1.8.** *Let $k \geq 2$ be an integer and let ω be any function of n which tends to infinity as $n \to \infty$. For each $j \in [k-1]$, with high probability the following statements hold.*

(i) $\frac{(j+1)\log n + \log\log n - \omega}{(k-j+1)n^{k-j}}(k-j)! < p_{M_j} < \frac{(j+1)\log n + \log\log n + \omega}{(k-j+1)n^{k-j}}(k-j)!.$

(ii) *For all $p < p_{M_j}$, $\mathcal{G}_p$ is not $\mathbb{F}_2$-cohomologically j-connected, i.e.*

$$H^0(\mathcal{G}_p; \mathbb{F}_2) \neq \mathbb{F}_2 \text{ or } H^i(\mathcal{G}_p; \mathbb{F}_2) \neq 0 \text{ for some } i \in [j].$$

(iii) *For all $p \geq p_{M_j}$, $\mathcal{G}_p$ is $\mathbb{F}_2$-cohomologically j-connected, i.e.*

$$H^0(\mathcal{G}_p; \mathbb{F}_2) = \mathbb{F}_2 \text{ and } H^i(\mathcal{G}_p; \mathbb{F}_2) = 0 \text{ for all } i \in [j].$$

For the case $j = k - 1$, Theorem 1.8 gives a threshold $p_{k-1} = \frac{k \log n + \log \log n}{2n}$ for $\mathbb{F}_2$-cohomologically $(k-1)$-connectedness, which is about half as large as the threshold $\frac{k \log n}{n}$ in Theorem 1.2 for $\mathcal{Y}_p$. The reason for this is that the minimal obstructions are different: in $\mathcal{Y}_p$ the minimal obstruction is a $(k-1)$-simplex which is not contained in any k-simplex of the complex (such a $(k-1)$-simplex is called *isolated*). By definition, isolated $(k-1)$-simplices do not exist in $\mathcal{G}_p$, because $\mathcal{G}_p$ contains only those $(k-1)$-simplices that lie in some k-simplex.

Observe that Theorem 1.8 ii and iii provide a hitting time result for the process described above. A similar result was proved by Kahle and Pittel [15] for $\mathcal{Y}_p$, but only for the 2-dimensional case. As a corollary of Theorem 1.8, we can now derive a hitting time result for $\mathcal{Y}_p$ for general $k \geq 2$. To this end, let

$$p_{\text{isol}} := \sup\{p \in [0,1] \mid \mathcal{Y}_p \text{ contains isolated } (k-1)\text{-simplices}\} \tag{3}$$

---

[2] With probability 1 no two $(k+1)$-sets have the same birth time.

be the birth time of the k-simplex whose appearance causes the last isolated $(k-1)$-simplex to disappear and let

$$p_{\text{conn}} := \sup\{p \in [0,1] \mid H^{k-1}(\mathcal{Y}_p; \mathbb{F}_2) \neq 0 \} \tag{4}$$

be the time when $\mathcal{Y}_p$ becomes $\mathbb{F}_2$-cohomological $(k-1)$-connected.

▶ **Corollary 1.9.** *Let $k \geq 2$ be an integer. Then, with high probability $p_{conn} = p_{isol}$.*

Our last main result gives an explicit expression for the limiting probability of the random complex $\mathcal{G}_p$ being $\mathbb{F}_2$-cohomologically j-connected inside the critical window given by the threshold p_j. More generally, we prove that the *dimension* of the j-th cohomology group with coefficients in $\mathbb{F}_2$ converges in distribution to a Poisson random variable.

▶ **Theorem 1.10.** *Let $k \geq 2$ be an integer, $j \in [k-1]$ and $c \in \mathbb{R}$ be a constant. Suppose that $c_n \in \mathbb{R}$ are such that $c_n \xrightarrow{n \to \infty} c$. If*

$$p = \frac{(j+1)\log n + \log\log n + c_n}{(k-j+1)n^{k-j}}(k-j)!,$$

then $\dim\left(H^j(\mathcal{G}_p; \mathbb{F}_2)\right)$ converges in distribution to a Poisson random variable with expectation

$$\lambda_j := \frac{(j+1)e^{-c}}{(k-j+1)^2 j!},$$

while whp $H^0(\mathcal{G}_p; \mathbb{F}_2) = \mathbb{F}_2$ and $H^i(\mathcal{G}_p; \mathbb{F}_2) = 0$ for all $i \in [j-1]$. In particular, we have

$$\mathbb{P}\left(\mathcal{G}_p \text{ is } j\text{-cohom-connected}\right) \xrightarrow{n \to \infty} e^{-\lambda_j}.$$

Note that a similar result for $\mathcal{Y}_p$ was proved by Kahle and Pittel [15].

1.4 Related work

The vanishing of $H^{k-1}(\mathcal{Y}_p; \mathbb{F}_2)$ considered in [17] and [20] is a monotone property due to the presence of the full $(k-1)$-dimensional skeleton. This fact in particular makes it possible to use a simple second moment argument to prove the subcritical case (i.e. statement (i)) of Theorem 1.2.

In contrast, $\mathcal{G}_p$ does *not* contain the full $(k-1)$-dimensional skeleton. As a consequence, we need to consider *all* cohomology groups up to dimension j, for each $j \in [k-1]$. Moreover, our notion of $\mathbb{F}_2$-cohomological j-connectedness is *not* a monotone property, which makes the subcritical case far from trivial. In fact, it does not suffice to prove that $\mathcal{G}_p$ is not j-cohom-connected at $p_- = \frac{(j+1)\log n + \log\log n - \omega}{(k-j+1)n^{k-j}}(k-j)!$; rather we need to show that whp $\mathcal{G}_p$ is not j-cohom-connected for *any* p up to and including p_-.

The proof of the supercritical case $p \geq p_{M_j}$ is also more challenging than for $\mathcal{Y}_p$. We are forced to derive better bounds for the number of *bad functions* (see Definition 2.1), due to the fact that for $j = k-1$, the threshold in Theorem 1.8 is about half as large as the corresponding threshold in [20].

2 Preliminaries

2.1 Cohomology terminology

We formally introduce cohomology with coefficients in $\mathbb{F}_2$ for a simplicial complex. The following notions are all standard, except the definition of a bad function (Definition 2.1).

Given a simplicial k-complex $\mathcal{G}$, for each $j \in \{0, \dots, k\}$ denote by $C^j(\mathcal{G})$ the set of *j-cochains*, that is, the set of 0-1 functions on the j-simplices. The *support* of a function in $C^j(\mathcal{G})$ is the set of j-simplices mapped to 1. Each $C^j(\mathcal{G})$ forms a group with respect to pointwise addition modulo 2. We define the *coboundary operators* $\delta^j \colon C^j(\mathcal{G}) \to C^{j+1}(\mathcal{G})$ for $j = 0, \dots, k-1$ as follows. For $f \in C^j(\mathcal{G})$, the 0-1 function $\delta^j f$ assigns to each $(j+1)$-simplex σ the value

$$\delta^j f(\sigma) := \sum_{\tau \subset \sigma,\ |\tau| = j+1} f(\tau) \pmod{2}.$$

In addition, we denote by δ^{-1} the unique group homomorphism $\delta^{-1} \colon \{0\} \to C^0(\mathcal{G})$. The j-cochains in $\operatorname{im} \delta^{j-1}$ and $\ker \delta^j$ are called *j-coboundaries* and *j-cocycles*, respectively. A straightforward calculation shows that each coboundary operator is a group homomorphism and that every j-coboundary is also a j-cocycle, i.e. $\operatorname{im} \delta^{j-1} \subseteq \ker \delta^j$. Therefore, we can define the j-th cohomology group of $\mathcal{G}$ with coefficients in $\mathbb{F}_2$ as the quotient group

$$H^j(\mathcal{G}; \mathbb{F}_2) := \ker \delta^j / \operatorname{im} \delta^{j-1}.$$

By definition, $H^j(\mathcal{G}; \mathbb{F}_2)$ vanishes if and only if every j-cocycle is a j-coboundary. This motivates the following definition of a *bad function*.

▶ **Definition 2.1.** We say that a function $f \in C^j(\mathcal{G})$ is *bad* if
  (i) f is a j-cocycle, i.e. it assigns an even number of 1's to the j-simplices on the boundary of each $(j+1)$-simplex;
  (ii) f is not a j-coboundary, i.e. it is not induced by a 0-1 function on the $(j-1)$-simplices.
Thus, $H^j(\mathcal{G}; \mathbb{F}_2)$ vanishes if and only if no bad function in $C^j(\mathcal{G})$ exists.

Recall that a set J of j-simplices is a j-cycle if every $(j-1)$-simplex lies in an even number of j-simplices in J. It is easy to see that if f is a j-cocycle and J is a j-cycle such that $f|_J$ has support of odd size, then f is not a j-coboundary and thus is a bad function.

2.2 Minimal obstructions

Let us explain why M_j (Definition 1.7) can be interpreted as the minimal obstruction to j-cohom-connectedness. Given a copy (K, C, J) of M_j in a k-complex $\mathcal{G}$, define a function $f \in C^j(\mathcal{G})$ that takes value 1 on the petals of the flower $\mathcal{F}(K, C)$ and 0 everywhere else. Since each petal lies in K but in no further k-simplices, every $(j+1)$-simplex contains either two petals or none. In particular, f is even on the boundary of every $(j+1)$-simplex. However, J would be a j-cycle containing precisely one j-simplex (namely $C \cup \{w_0\}$) on which f takes value 1, ensuring that f is bad. The support of f has size $k - j + 1$.

▶ **Lemma 2.2.** *Let $\mathcal{G}$ be a k-complex and let S be a non-empty support of a j-cocycle. Then either S is the flower of an M_j (and thus $|S| = k - j + 1$) or $|S| \geq k - j + 2$.*

Both the presence of a copy of M_j and j-cohom-connectedness are *not* monotone, as the following example shows.

▶ **Example 2.3.** Let $\mathcal{G}$ be the 2-complex on vertex set $\{1, 2, 3, 4, 5\}$ generated by the 3-uniform hypergraph with hyperedges $\{1, 2, 3\}$ and $\{1, 4, 5\}$, see Figure 2. Then $\mathcal{G}$ is 1-cohom-connected and thus contains no copies of M_1. Adding to $\mathcal{G}$ the 2-simplex $\{2, 3, 4\}$ (and its downward-closure) creates several copies of M_1 and thus yields a complex $\mathcal{G}'$ which is not 1-cohom-connected. If we further add the 2-simplex $\{1, 3, 4\}$ to $\mathcal{G}'$, we obtain a 2-complex $\mathcal{G}''$ which is 1-cohom-connected and thus contains no copies of M_1.
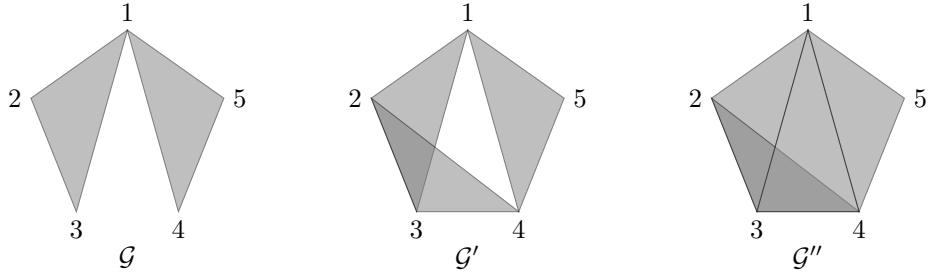
Figure 2 Adding simplices might create new copies of M_j or destroy existing ones.

3   Subcritical regime

3.1   Overview

In this section we study the subcritical case $p < p_{M_j}$ and state results necessary for the proofs of statements i and ii of Theorem 1.8.

Define

$$p_T := \sup\{p \in [0,1] \mid \mathcal{G}_p \text{ is not topologically connected}\}$$

as the birth time of the k-simplex whose appearance causes the complex $\mathcal{G}_p$ to become topologically connected. In addition, we will need the probabilities

$$p_0^- := \frac{\log n}{n^k},$$

$$p_j^- := \left(1 - \frac{1}{\sqrt{\log n}}\right) \frac{(j+1)\log n}{(k-j+1)n^{k-j}}(k-j)! \quad \text{for each } j \in [k-1].$$

Observe that $H^0(\mathcal{G}_p; \mathbb{F}_2) \neq \mathbb{F}_2$ in $[0, p_T)$ by definition. In order to prove Theorem 1.8 ii, we aim to show that whp $H^j(\mathcal{G}_p; \mathbb{F}_2) \neq 0$ in $[p_{j-1}^-, p_{M_j})$ for all $j \in [k-1]$ and that

$$[0, p_T) \cup \bigcup_{i=1}^{j} [p_{i-1}^-, p_{M_i}) = [0, p_{M_j}),$$

which we prove by showing that $p_T > p_0^-$ and $p_{M_j} > p_j^- > p_{j-1}^-$ for all $j \in [k-1]$ whp. To cover the interval $[p_{j-1}^-, p_{M_j})$, we in fact prove the existence of *just three copies* of M_j such that whp for all p in this interval, at least one of these copies is present in $\mathcal{G}_p$.

▶ **Lemma 3.1.** *Let $j \in [k-1]$. With high probability, there exist three triples (K_ℓ, C_ℓ, J_ℓ), $\ell = 1, 2, 3$, such that for all $p \in [p_{j-1}^-, p_{M_j})$, (K_ℓ, C_ℓ, J_ℓ) forms a copy of M_j in $\mathcal{G}_p$ for some ℓ. In particular, whp $H^j(\mathcal{G}_p; \mathbb{F}_2) \neq 0$ for all $p \in [p_{j-1}^-, p_{M_j})$.*

3.2   Topological connectedness

Topological connectedness of $\mathcal{G}_p$ is equivalent to vertex-connectedness of the random $(k+1)$-uniform hypergraph, whose (sharp) threshold follows e.g. as a special case of [8] or [22].

▶ **Lemma 3.2.** *Let ω be any function of n which tends to infinity as $n \to \infty$. Then with high probability*

$$\frac{\log n - \omega}{n^k} k! < p_T < \frac{\log n + \omega}{n^k} k!$$

and thus in particular $p_T > p_0^-$.

3.3 Finding obstructions

In order to prove Lemma 3.1, we make use of a simplified version of the obstruction M_j.

▶ **Definition 3.3.** *copy of* M_j^- *in a k-complex* $\mathcal{G}$ *is a pair* (K, C) *such that*

(M1) K *is a k-simplex;*

(M2) C *is a $(j-1)$-simplex in K such that each petal of the flower $\mathcal{F}(K, C)$ is contained in no other k-simplex of $\mathcal{G}$.*

In other words, a copy of M_j^- can be viewed as a copy of M_j without the condition (M3) of Definition 1.7, i.e. without the j-cycle J containing one of the petals. Therefore, if (K, C, J) is a copy of M_j in $\mathcal{G}_p$, then (K, C) is a copy of M_j^-. Vice versa, the following lemma ensures that whp for p at least

$$p_j^{(1)} := \frac{1}{10(j+1)\binom{k+1}{j+1}n^{k-j}},$$

whp every copy of M_j^- gives rise to a copy of M_j, allowing us to consider just copies of M_j^- as obstructions to j-cohom-connectedness. In other words, the existence of copies of M_j^- and M_j are essentially equivalent for $p \geq p_j^{(1)}$.

▶ **Lemma 3.4.** *There exists a positive constant γ such that with high probability for every $p \geq p_j^{(1)}$, each j-simplex σ in $\mathcal{G}_p$ lies in at least γn many j-cycles in $\mathcal{G}_p$ that meet only in σ. In particular, whp for all $p \geq p_j^{(1)}$, every copy of M_j^- in $\mathcal{G}_p$ is part of a copy of M_j.*

3.4 Excluding obstructions and determining the hitting time

A second moment argument shows that at time

$$\bar{p}_j := \frac{(j+1)\log n + \frac{1}{2}\log\log n}{(k-j+1)n^{k-j}}(k-j)!, \tag{5}$$

whp $\mathcal{G}_{\bar{p}_j}$ contains (a growing number of) copies of M_j^-, and thus whp also copies of M_j by Lemma 3.4. Define $\bar{p}_{M_j}$ as the first birth time p larger than $\bar{p}_j$ such that there are no copies of M_j in $\mathcal{G}_p$. By definition of p_{M_j}, conditioned on the high probability event $M_j \subset \mathcal{G}_{\bar{p}_j}$, we have $\bar{p}_{M_j} \leq p_{M_j}$. In the next lemma we show that they are in fact equal whp.

To do so, we need the following definition.

▶ **Definition 3.5.** *Given a k-complex $\mathcal{G}$, a k-simplex K is a local obstacle if K contains at least $k-j+1$ many j-simplices which are not contained in any other k-simplex of $\mathcal{G}$.*

Observe that each M_j^- is in particular a local obstacle. Moreover, whp each copy of M_j^- in $\mathcal{G}_p$ for $p \geq \bar{p}_j$ gives rise to copies of M_j by Lemma 3.4.

▶ **Lemma 3.6.** *With high probability, for all $p \geq \bar{p}_j$ every local obstacle that exists in $\mathcal{G}_p$ also exists in $\mathcal{G}_{\bar{p}_j}$. In particular, we have $p_{M_j} = \bar{p}_{M_j}$ whp.*

▶ **Corollary 3.7.** *Whp for all $p \geq p_{M_j}$, there are no copies of M_j^- in $\mathcal{G}_p$.*

By first and second moment arguments, we can now easily derive that p_{M_j} is "close to" p_j. Observe that the following corollary is exactly Theorem 1.8 i.

▶ **Corollary 3.8.** *Let ω be any function of n which tends to infinity as n tends to infinity. Then whp*

$$\frac{(j+1)\log n + \log\log n - \omega}{(k-j+1)n^{k-j}}(k-j)! < p_{M_j} < \frac{(j+1)\log n + \log\log n + \omega}{(k-j+1)n^{k-j}}(k-j)!.$$

3.5   Covering the interval

Our strategy to derive Lemma 3.1 is to divide the interval $[p_{j-1}^-, p_{M_j})$ into three subintervals $[p_{j-1}^-, p_j^{(1)}]$, $[p_j^{(1)}, p_j^-]$, $[p_j^-, p_{M_j})$, each of which we cover by one copy of M_j. We first use a second moment argument to show that at time p_{j-1}^-, whp there are "many" copies of M_j. With high probability, at least one copy (K_1, C_1, J_1) survives until probability $p_j^{(1)}$.

In order to find a copy of M_j that covers the interval $[p_j^{(1)}, p_j^-]$, we show that whp "many" copies of M_j^- exist at time p_j^-, of which one whp was already present at the beginning of the interval. Together with the fact that whp each M_j^- gives rise to a copy of M_j (Lemma 3.4), this implies that whp one copy (K_2, C_2, J_2) of M_j exists throughout this interval.

For the remaining interval $[p_j^-, p_{M_j})$, consider a copy (K_3, C_3) of M_j^- that vanishes at time p_{M_j}. Corollary 3.8 implies that whp $p_j^- = (1 - o(1))p_{M_j}$, and thus (K_3, C_3) whp was already present at time p_j^-. Now Lemma 3.4 ensures the existence of a j-cycle J_3 such that (K_3, C_3, J_3) is a copy of M_j throughout the range $[p_j^-, p_{M_j})$.

4   Critical window and supercritical regime

In this section, we study obstructions around the point of the claimed phase transition and in the supercritical regime, that is, for $p = (1 + o(1))p_j$ and $p \geq p_{M_j}$, respectively. The results of this section will form the foundation of the proof of Theorem 1.8 iii. Furthermore, they will play a crucial role in the proof of Theorem 1.10.

By the definition of p_{M_j}, there are no copies of M_j in $\mathcal{G}_p$ (and also no copies of M_j^- by Corollary 3.7) for any $p \geq p_{M_j}$. It remains to show that there are no other obstructions either. In fact, we shall even prove (Lemma 4.2) that from slightly before p_{M_j} onwards, any j-cocycles are generated by copies of M_j^-. To make this more precise, we need the following notation.

▶ **Definition 4.1.** We say that a j-cochain $f_{K,C}$ *arises from* a copy (K, C) of M_j^- in a k-complex $\mathcal{G}$ if its support is the j-flower $\mathcal{F}(K, C)$. Observe that then $f_{K,C}$ is a j-cocycle.

We say that a j-cocycle f in $\mathcal{G}$ is *generated by copies of M_j^-* if it lies in the same cohomology class as a sum of cocycles that arise from copies of M_j^-. We denote by $\mathcal{N}_{\mathcal{G}}$ the set of j-cocycles that are *not* generated by copies of M_j^-.

We show that whp for all $p \geq p_{M_j}$, $\mathcal{N}_{\mathcal{G}_p} = \emptyset$, which will in particular imply that there are no non-empty j-cocycles in $\mathcal{G}_p$. Furthermore, a similar argument will enable us to directly relate the number of copies of M_j^- with the dimension of $H^j(\mathcal{G}_p; \mathbb{F}_2)$ (cf. Theorem 1.10).

▶ **Lemma 4.2.** *For every* $p \geq p_j^-$, *we have* $\mathcal{N}_{\mathcal{G}_p} = \emptyset$ *with high probability. Moreover, with high probability* $\mathcal{N}_{\mathcal{G}_p} = \emptyset$ *for all* $p \geq p_{M_j}$ *simultaneously.*

In order to prove Lemma 4.2, we first show that a smallest support of elements of $\mathcal{N}_{\mathcal{G}}$ would have to have a property we call *traversability*.

▶ **Definition 4.3.** Let $\mathcal{G}$ be a k-complex and $S \subseteq \mathcal{G}$ be a collection of j-simplices. For $\sigma_1, \sigma_2 \in S$, we write $\sigma_1 \sim \sigma_2$ if σ_1 and σ_2 lie in a common k-simplex.[3] We say that S is *traversable* if the transitive closure of $\sim$ is $S \times S$.

In other words, a set of j-simplices in a k-complex is traversable if it *cannot* be partitioned into two non-empty subsets such that no k-simplex contains j-simplices in both subsets.

---

[3] Observe that this relation is reflexive, because every j-simplex is contained in at least one k-simplex.

▶ **Lemma 4.4.** *Let $\mathcal{G}$ be a k-complex and f be an element of $\mathcal{N}_{\mathcal{G}}$ with smallest support S. Then S is traversable.*

We then show that whp no such smallest support can exist in $\mathcal{G}_p$. For "small" support size and probability around p_j, a standard application of the first moment method suffices.

▶ **Lemma 4.5.** *For $p = (1 + o(1))p_j$ and for any constant $d \geq k - j + 2$, with high probability $\mathcal{G}_p$ has no j-cocycle with traversable support of size s with $k - j + 2 \leq s \leq d$.*

For larger size, we make use of traversability to define a breadth-first search process that finds all possible supports. Using this process, we can bound the number of possible smallest supports of elements of $\mathcal{N}_{\mathcal{G}_p}$ more carefully, thus allowing us to prove that whp *for all relevant p simultaneously*, such a smallest support cannot be "large".

▶ **Lemma 4.6.** *There exists a positive constant $\bar{d}$ such that with high probability for all $p \geq p_j^-$, the smallest support of a j-cocycle in $\mathcal{N}_{\mathcal{G}_p}$ has size $s < \bar{d}$.*

In particular, for any fixed $p = (1 + o(1))p_j$, whp the smallest support of elements of $\mathcal{N}_{\mathcal{G}_p}$ is not "small" by Lemma 4.5 and not "large" by Lemma 4.6, which means that $\mathcal{N}_{\mathcal{G}_p} = \emptyset$ whp.

Finally, we complete the argument by proving that any new element of $\mathcal{N}_{\mathcal{G}_p}$ with "small" support that might appear if we increase p would have to give rise to a "new" local obstacle. But Lemma 3.6 already tells us that whp no new local obstacles appear. This concludes the proof of Lemma 4.2.

5 Proofs of main results

5.1 Proof of Theorem 1.8

Corollary 3.8 states that for any function ω of n which tends to infinity as $n \to \infty$, whp

$$\frac{(j + 1) \log n + \log \log n - \omega}{(k - j + 1)n^{k-j}}(k - j)! < p_{M_j} < \frac{(j + 1) \log n + \log \log n + \omega}{(k - j + 1)n^{k-j}}(k - j)!,$$

which is precisely Theorem 1.8 i.

To prove ii, recall that Lemma 3.1 tells us that for each $i \in [j - 1]$, whp $H^i(\mathcal{G}_p; \mathbb{F}_2) \neq 0$ for all $p \in [p_{i-1}^-, p_{M_i})$. By i, whp

$$p_{M_i} > \left(1 - \frac{1}{\sqrt{\log n}}\right) \frac{(i + 1) \log n}{(k - i + 1)n^{k-i}}(k - i)! = p_i^-,$$

and thus whp $\mathcal{G}_p$ is not j-cohom-connected throughout $\bigcup_{i=1}^{j} [p_{i-1}^-, p_{M_i}) = [p_0^-, p_{M_j})$.

Now observe that by Lemma 3.2 whp $p_T > p_0^-$ and that $\mathcal{G}_p$ is not topologically connected in $[0, p_T)$ by definition of p_T. Therefore, whp $\mathcal{G}_p$ is not j-cohom-connected in

$$[0, p_{M_j}) = [0, p_T) \cup [p_0^-, p_{M_j}),$$

as required.

It remains to prove iii. By Corollary 3.7, we know that for all $p \geq p_{M_j}$, there are no copies of M_j^- in $\mathcal{G}_p$. Thus, if $H^j(\mathcal{G}_p; \mathbb{F}_2) \neq 0$, then any representative of a non-zero cohomology class cannot arise from copies of M_j^- and therefore lies in $\mathcal{N}_{\mathcal{G}_p}$ (Definition 4.1). But by Lemma 4.2, whp each such $\mathcal{N}_{\mathcal{G}_p}$ is empty and thus whp $H^j(\mathcal{G}_p; \mathbb{F}_2) = 0$ for all $p \geq p_{M_j}$. Analogously, whp all cohomology groups $H^i(\mathcal{G}_p; \mathbb{F}_2)$ for $i \in [j - 1]$ vanish, because whp $p_{M_i} < p_{M_j}$ by i. Finally, by i and Lemma 3.2 whp $p_T < p_{M_j}$, meaning that whp $\mathcal{G}_p$ is topologically connected for all $p \geq p_{M_j}$. This implies that whp each such $\mathcal{G}_p$ is $\mathbb{F}_2$-cohomologically j-connected. ◀

5.2 Proof of Corollary 1.9

Let ω be any function of n which tends to infinity as $n \to \infty$. It follows by a simple first and second moment argument (see e.g. [20]) that whp

$$\frac{k \log n - \omega}{n} < p_{\text{isol}} < \frac{k \log n + \omega}{n}. \tag{6}$$

In order to prove that $p_{\text{conn}} = p_{\text{isol}}$ whp, suppose that a $(k-1)$-simplex σ is isolated in $\mathcal{Y}_p$ for some p. The indicator function f_σ of σ is a $(k-1)$-cocycle, because σ is isolated. But f_σ is *not* a $(k-1)$-coboundary, because σ lies in (many) $(k-1)$-cycles due to the presence of the full $(k-1)$-dimensional skeleton. In particular, $H^{k-1}(\mathcal{Y}_p; \mathbb{F}_2) \neq 0$. By the definitions of p_{conn} and p_{isol}, this implies that $p_{\text{conn}} \geq p_{\text{isol}}$.

For the opposite direction, fix the birth times of all k-simplices. Then for all $p \geq p_{\text{isol}}$, we have $\mathcal{Y}_p = \mathcal{G}_p$ and therefore $\mathcal{Y}_p$ is $\mathbb{F}_2$-cohomological $(k-1)$-connected whp for every $p \geq \max(p_{\text{isol}}, p_{M_{k-1}})$ by Theorem 1.8 iii. By (6) and Theorem 1.8 i for $j = k-1$, whp for any (slowly) growing function ω

$$p_{\text{isol}} > \frac{k \log n - \omega}{n} > \frac{k \log n + \log \log n + \omega}{2n} > p_{M_{k-1}},$$

hence whp for all $p \geq p_{\text{isol}}$ we have $H^{k-1}(\mathcal{Y}_p; \mathbb{F}_2) = 0$. This means that whp $p_{\text{conn}} \leq p_{\text{isol}}$ and thus $p_{\text{conn}} = p_{\text{isol}}$, as required. ◀

5.3 Proof of Theorem 1.10

We are interested in the asymptotic distribution of $D_j := \dim\left(H^j(\mathcal{G}_p; \mathbb{F}_2)\right)$ for

$$p = \frac{(j+1) \log n + \log \log n + c_n}{(k-j+1) n^{k-j}} (k-j)!, \qquad \text{where } c_n \xrightarrow{n \to \infty} c \in \mathbb{R}.$$

Denote by X_- the number of copies of M_j^- in $\mathcal{G}_p$. Standard calculations show that

$$\mathbb{E}(X_-) = (1 + o(1))\lambda_j, \qquad \text{where } \lambda_j = \frac{(j+1) e^{-c}}{(k-j+1)^2 j!}.$$

Moreover, we show that for each fixed integer $t \geq 1$

$$\mathbb{E}\binom{X_-}{t} = (1 + o(1)) \frac{\lambda_j^t}{t!}.$$

These equalities are precisely what is necessary to apply the method of moments (see e.g. [13]) in order to show that X_- converges in distribution to a Poisson random variable with expectation λ_j, which we denote by $X_- \xrightarrow{d} \text{Po}(\lambda_j)$.

It remains to show that $X_- = D_j$ whp. To this end, denote by $f_1, \ldots, f_{X_-}$ the j-cocycles arising from the copies of M_j^- in $\mathcal{G}_p$. Lemma 4.2 states that whp the cohomology classes of $f_1, \ldots, f_{X_-}$ generate $H^j(\mathcal{G}_p; \mathbb{F}_2)$, which means that $X_- \geq D_j$ whp.

In order to prove the opposite direction, we show that the cohomology classes of $f_1, \ldots, f_{X_-}$ are linearly independent. Observe first that whp $X_- = o(n)$ by Markov's inequality, because X_- has bounded expectation. Let $I \subseteq [X_-]$ be non-empty and let S be the support of $\sum_{i \in I} f_i$. Whp no two f_i's can have their supports contained in the same k-simplex K, because otherwise their union would be a traversable support of size s with $k - j + 2 \leq s \leq 2(k-j+1)$, but such supports whp do not exist by Lemma 4.5.

Thus, whp the f_i's have disjoint support by property (M2) of an M_j^- (Definition 3.3), and in particular $S \neq \emptyset$. Pick $\sigma \in S$. Lemma 3.4 tells us that whp there are $\Theta(n)$ many j-cycles in $\mathcal{G}_p$ that contain σ and are otherwise disjoint. But at most $|S| \leq (k-j+1)|I| = o(n)$ of these j-cycles can contain another j-simplex in S, which means that whp there are j-cycles that meet S only in σ, showing that $\sum_{i \in I} f_i$ is not a j-coboundary. Therefore the cohomology classes of $f_1, \ldots, f_{X_-}$ are linearly independent whp. This shows that $X_- \leq D_j$ and thus $X_- = D_j$ whp, as desired.

Together with $X_- \xrightarrow{d} \mathrm{Po}(\lambda_j)$, this proves that $D_j \xrightarrow{d} \mathrm{Po}(\lambda_j)$. By Theorem 1.8 (for $j-1$ instead of j), whp $H^0(\mathcal{G}_p; \mathbb{F}_2) = \mathbb{F}_2$ and $H^i(\mathcal{G}_p; \mathbb{F}_2) = 0$ for all $i \in [j-1]$. In particular,

$$
\begin{aligned}
\mathbb{P}(\mathcal{G}_p \text{ is } j\text{-cohom-connected}) &= \mathbb{P}\big(H^j(\mathcal{G}_p; \mathbb{F}_2) = 0\big) + o(1) \\
&= (1 + o(1))\mathbb{P}\big(\mathrm{Po}(\lambda_j) = 0\big) \\
&= (1 + o(1))e^{-\lambda_j}.
\end{aligned}
$$

This concludes the proof of Theorem 1.10.   ◀

6   Concluding remarks

The vanishing of cohomology groups with coefficients in $\mathbb{F}_2$ is just one possible way of defining the concept of "connectedness" of $\mathcal{G}_p$. An obvious alternative would be to consider coefficients from other groups or fields. For $\mathcal{Y}_p$, such notions of connectedness have been studied for coefficients in any finite abelian group, in $\mathbb{Z}$, or in any field [1, 2, 14, 18, 19, 20].

A rather strong notion of connectedness would be to require the homotopy groups $\pi_1(\mathcal{G}_p), \ldots, \pi_j(\mathcal{G}_p)$ to vanish. For the 2-dimensional case, the vanishing of $\pi_1(\mathcal{Y}_p)$ was studied by Babson, Hoffman and Kahle [3]. In particular, they showed that whp $\pi_1(\mathcal{Y}_p) \neq 0$ at the time that $H^1(\mathcal{Y}_p; \mathbb{F}_2)$ becomes zero. From that time on, the models $\mathcal{Y}_p$ and $\mathcal{G}_p$ coincide. As $\pi_1(\mathcal{G}_p) \neq 0$ follows immediately from $H^1(\mathcal{G}_p; \mathbb{F}_2) \neq 0$, the range that should be of particular interest with respect to $\pi_1(\mathcal{G}_p)$ in the 2-dimensional case is

$$
\frac{\log n + \frac{1}{2}\log\log n}{n} \leq p \leq \frac{2\log n + \omega}{n}.
$$

A natural conjecture would be that whp $\pi_1(\mathcal{G}_p) \neq 0$ in this range.

Theorem 1.9 provides a limit result for the dimensions $D_j = \dim(H^j(\mathcal{G}_p; \mathbb{F}_2))$ around the point of the phase transition. It would be interesting to know the behaviour of D_j also for earlier regimes. More precisely, we know by Theorem 1.8 that whp $D_j \neq 0$ in the interval $[p_{j-1}^-, p_{M_j}^-)$. Can we say more about the value of D_j in this interval? How far below p_{j-1}^- do we have $D_j > 0$ whp?

References

1   L. Aronshtam and N. Linial. When does the top homology of a random simplicial complex vanish? *Random Structures Algorithms*, 46(1):26–35, 2015. `doi:10.1002/rsa.20495`.

2   L. Aronshtam, N. Linial, T. Łuczak, and R. Meshulam. Collapsibility and vanishing of top homology in random simplicial complexes. *Discrete Comput. Geom.*, 49(2):317–334, 2013. `doi:10.1007/s00454-012-9483-8`.

3   E. Babson, C. Hoffman, and M. Kahle. The fundamental group of random 2-complexes. *J. Amer. Math. Soc.*, 24(1):1–28, 2011. `doi:10.1090/S0894-0347-2010-00677-7`.

4   M. Behrisch, A. Coja-Oghlan, and M. Kang. The order of the giant component of random hypergraphs. *Random Structures Algorithms*, 36(2):149–184, 2010. `doi:10.1002/rsa.20282`.

5   B. Bollobás and O. Riordan. Asymptotic normality of the size of the giant component in a random hypergraph. *Random Structures Algorithms*, 41(4):441–450, 2012. `doi:10.1002/rsa.20456`.

6   B. Bollobás and O. Riordan. Exploring hypergraphs with martingales. *Random Structures Algorithms*, 50(3):325–352, 2017. `doi:10.1002/rsa.20678`.

7   B. Bollobás and A. Thomason. Random graphs of small order. In *Random graphs '83 (Poznań, 1983)*, volume 118 of *North-Holland Math. Stud.*, pages 47–97. North-Holland, Amsterdam, 1985.

8   O. Cooley, M. Kang, and C. Koch. Threshold and hitting time for high-order connectedness in random hypergraphs. *Electron. J. Combin.*, 23(2):Paper 2.48, 14, 2016.

9   O. Cooley, M. Kang, and C. Koch. The size of the giant high-order component in random hypergraphs. *Random Structures & Algorithms*, 2018. `doi:10.1002/rsa.20761`.

10   O. Cooley, M. Kang, and Y. Person. Largest components in random hypergraphs. *arXiv:1412.6366*, dec *Manuscript*, 2014. `arXiv:1412.6366`.

11   R. W. R. Darling and J. R. Norris. Structure of large random hypergraphs. *Ann. Appl. Probab.*, 15(1A):125–152, 2005. `doi:10.1214/105051604000000567`.

12   P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

13   A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge, 2016.

14   C. Hoffman, M. Kahle, and E. Paquette. The threshold for integer homology in random d-complexes. *Discrete Comput. Geom.*, 57(4):810–823, 2017. `doi:10.1007/s00454-017-9863-1`.

15   M. Kahle and B. Pittel. Inside the critical window for cohomology of random k-complexes. *Random Structures Algorithms*, 48(1):102–124, 2016. `doi:10.1002/rsa.20577`.

16   M. Karoński and T. Łuczak. Random hypergraphs. In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 283–293. János Bolyai Math. Soc., Budapest, 1996.

17   N. Linial and R. Meshulam. Homological connectivity of random 2-complexes. *Combinatorica*, 26(4):475–487, 2006. `doi:10.1007/s00493-006-0027-9`.

18   N. Linial and Y. Peled. On the phase transition in random simplicial complexes. *Ann. of Math. (2)*, 184(3):745–773, 2016. `doi:10.4007/annals.2016.184.3.3`.

19   T. Łuczak and Y. Peled. Integral homology of random simplicial complexes. *Discrete Comput. Geom.*, 59(1):131–142, 2018. `doi:10.1007/s00454-017-9938-z`.

20   R. Meshulam and N. Wallach. Homological connectivity of random k-dimensional complexes. *Random Structures Algorithms*, 34(3):408–417, 2009. `doi:10.1002/rsa.20238`.

21   J. Munkres. *Elements of algebraic topology*. Addison-Wesley Publishing Company, Menlo Park, CA, 1984.

22   D. Poole. On the strength of connectedness of a random hypergraph. *Electron. J. Combin.*, 22(1):Paper 1.69, 16, 2015.

23   J. Schmidt-Pruzan and E. Shamir. Component structure in the evolution of random hypergraphs. *Combinatorica*, 5(1):81–94, 1985. `doi:10.1007/BF02579445`.

24   V. E. Stepanov. The probability of the connectedness of a random graph $\mathcal{G}_m(t)$. *Teor. Verojatnost. i Primenen*, 15:58–68, 1970.

Periods in Subtraction Games

Bret Benesh
Department of Mathematics, The College of St. Benedict/St. John's University, 37 College
Avenue South, St. Joseph, MN 56374–2001 U.S.A.
bbenesh@csbsju.edu

Jamylle Carter
Department of Mathematics, Diablo Valley College, 321 Golf Club Road, Pleasant Hill,
CA 94523–1529 U.S.A.
JCarter@dvc.edu

Deidra A. Coleman
Department of Mathematics, Wofford College, 429 North Church Street, Spartanburg,
SC 29303–3663 U.S.A.
ColemanDA@wofford.edu

Douglas G. Crabill
Department of Statistics, Purdue University, 150 North University Street, West Lafayette,
IN 47907–2067 U.S.A.
dgc@purdue.edu

Jack H. Good
Department of Computer Science (undergraduate student), Purdue University, 305 North
University Street, West Lafayette, IN 47907–2107 U.S.A.
good10@purdue.edu

Michael A. Smith
Department of Mathematics and Department of Statistics (undergraduate student), Purdue
University, 150 North University Street, West Lafayette, IN 47907–2067 U.S.A.
smit2589@purdue.edu

Jennifer Travis
Department of Mathematics, Lone Star College–North Harris, 2700 W. W. Thorne Drive,
Houston, TX 77073–3499 U.S.A.
Jennifer.L.Travis@lonestar.edu

Mark Daniel Ward[1]
Department of Statistics, Purdue University, 150 North University Street, West Lafayette,
IN 47907–2067 U.S.A.
mdw@purdue.edu

—— **Abstract** ——

We discuss the structure of periods in subtraction games. In particular, we discuss ways that a
computational approach yields insights to the periods that emerge in the asymptotic structure
of these combinatorial games.

2012 ACM Subject Classification Mathematics of computing → Combinatorial algorithms

Keywords and phrases combinatorial games, subtraction games, periods, asymptotic structure

Digital Object Identifier 10.4230/LIPIcs.AofA.2018.8

1 Overview

Subtraction games are one of the most fundamental combinatorial games. In the document *Unsolved Problems in Combinatorial Games* [4], maintained by Richard J. Nowakowski, the structure of combinatorial games is the first open problem that is discussed. Such games are so fundamental because the underlying premise is the same as Nim: there are several piles of beans, and on a player's turn, he/she can remove beans from exactly one pile. As in many areas of mathematics, this simple concept gives rise to much deeper mathematical structure. In the case of subtraction games, an even richer structure emerges because the moves of a player are limited. For instance, in the three-dimensional version of subtraction games with subtraction set $\{s_1, s_2, s_3\}$, the number of beans that can be removed from a heap during a player's turn is limited to one of these three possibilities. In other words, a player can only remove either s_1, s_2, or s_3 beans.

The problem of understanding the associated Nim values of a subtraction game is sufficiently challenging and useful that a table of values for small s_1, s_2, s_3 is given in the 4-volume set of books called *Winning Ways for your Mathematical Plays* [2].

The problem of understanding the asymptotic periodicities of subtraction games with a subtraction set of size three has been open for more than 40 years; see [1] for early analysis.

Mark Paulhus and Alex Fink have derived values of the periods in two cases, for subtraction sets of size 3, namely, in the case where $s_1 = 1$ and s_2, s_3 are arbitrary, and in the case where $s_1 < s_2 < s_3 < 32$ (see [4]). Achim Flammenkamp [3] has made conjectures about the types of periodicities that arise, based on calculations with all s_j's bounded above by 256.

We organized a team of colleagues to work on this problem at the American Institute of Mathematics (AIM), under the auspices of the Research Experiences for Undergraduate Faculty (REUF) workshops, starting in July 2016. (Ward had already been working on a computational attack for this problem in his spare moments, for more than a decade.) Our REUF team relies on a data-driven approach. We have computed the Nim values and the resulting (asymptotic) periodicity of the games for s_j's bounded above by 16384. The computational aspects of this problem are nontrivial. Each time the size of the parameters grows by a factor of 2, the computational time required for the resulting computations grows by a factor of (roughly) 17. Therefore, our most recent computation took a full 37 years of CPU time. It was accomplished by running a massive parallel computation on three of the computational clusters at Purdue University (using thousands of computational cores). After all, we made $\binom{16384}{3} = 732{,}873{,}539{,}584$ distinct computations altogether. We have generated terabytes of data about this combinatorial problem.

We will present our computational approach to determining the combinatorial structure of the asymptotic periods that arise in these subtraction games. Importantly, we emphasize that our algorithms allow us to know the asymptotic periods, without resorting at all to the traditional approach (which relies on minimal excluded numbers). Instead, we have obtained structural insights about this problem. These results should continue to be useful for revealing completely new viewpoints about the structure of combinatorial games.

References

1 Richard Austin. *Impartial and partizan games*. PhD thesis, University of Calgary, 1976.
2 Elwyn R. Berlekamp, John H. Conway, and Richard K. Guy. *Winning Ways for Your Mathematical Plays*, volume 1. A K Peters, 2001.
3 Achim Flammenkamp. *Lange Perioden in Subtraktions-Spielen*. PhD thesis, Universität Bielefeld, 1997.
4 Richard J. Nowakowski. Unsolved problems in combinatorial games. In Richard J. Nowakowski, editor, *Games of No Chance 4*, pages 279–308. Cambridge, 2015.

Permutations in Binary Trees and Split Trees

Michael Albert
Department of Computer Science, Otago University, New Zealand
malbert@cs.otago.ac.nz
ⓘ https://orcid.org/0000-0002-4587-1104

Cecilia Holmgren
Department of Mathematics, Uppsala University, Uppsala, Sweden
cecilia.holmgren@math.uu.se
ⓘ https://orcid.org/0000-0003-0717-4671

Tony Johansson
Department of Mathematics, Uppsala University, Uppsala, Sweden
tony.johansson@math.uu.se

Fiona Skerman
Department of Mathematics, Uppsala University, Uppsala, Sweden
fiona.skerman@math.uu.se
ⓘ https://orcid.org/0000-0003-4141-7059

─── Abstract ───

We investigate the number of permutations that occur in random node labellings of trees. This is a generalisation of the number of subpermutations occuring in a random permutation. It also generalises some recent results on the number of inversions in randomly labelled trees [3]. We consider complete binary trees as well as random split trees a large class of random trees of logarithmic height introduced by Devroye [4]. Split trees consist of nodes (bags) which can contain balls and are generated by a random trickle down process of balls through the nodes.

For complete binary trees we show that asymptotically the cumulants of the number of occurrences of a fixed permutation in the random node labelling have explicit formulas. Our other main theorem is to show that for a random split tree with high probability the cumulants of the number of occurrences are asymptotically an explicit parameter of the split tree. For the proof of the second theorem we show some results on the number of embeddings of digraphs into split trees which may be of independent interest.

1 Introduction and statement of results

Our main results are Theorem 2 on the distribution of the number of appearances of a fixed permutation in a random labelling of a complete binary tree and Theorem 4 which shows that for a random split tree with high probability (whp) the same result holds for the number of appearances of a fixed permutation in a random labelling of the balls of the tree. We write a complete introduction and statement of results in terms of complete binary trees first before defining split trees and stating our results for split trees.

Complete Binary trees

Let V_n denote the node set of the complete binary tree T_n of height m and $n = 2^{m+1} - 1$ nodes. Define a partial ordering on the nodes of the tree by saying that $a < b$ if a is an ancestor of b. Choose a uniform random labelling of the nodes $\pi : V_n \to [n]$.

We say that nodes a and b form an inversion if $a < b$ and $\pi(a) > \pi(b)$. The (random) number of inversions in random node labellings of fixed trees as well as some random models of trees were studied in a recent paper ([3]). This paper finds approximate extensions to some of these results.

The (random) number of inverted triples is $R(321, T) = \sum_{u_1 < u_2 < u_3} \mathbf{1}[\pi(u_1) > \pi(u_2) > \pi(u_3)]$ where the sum runs over all triples of nodes in T such that u_1 is an ancestor of u_2 and u_2 an ancestor of u_3. In general, we say a permutation σ appears on the $|\sigma|$-tuple of vertices $u_1, \ldots, u_{|\sigma|}$, if $u_1 < \ldots < u_{|\sigma|}$ and the induced order on $\pi(u) = (\pi(u_1), \ldots, \pi(u_{|\sigma|}))$ is σ. Write $\pi(u) \approx \sigma$ to indicate the induced order is the same for example $527 \approx 312$. Define

$$R(\sigma, T) \stackrel{\text{def}}{=} \sum_{u_1 < \ldots < u_{|\sigma|}} \mathbf{1}[\pi(u) \approx \sigma],$$

so in particular $R(21, T)$ counts the number of inversions in a random labelling of T.

We will generally be concerned with the centralised moments, e.g., $\mathbb{E}\left[(R(\sigma, T) - \mathbb{E}\left[R(\sigma, T)\right])^r\right]$. Let $d(v)$ denote the *depth* of v, i.e., the distance from v to the root ρ. For any $u_1 < \ldots < u_{|\sigma|}$ we have $\mathbb{P}[\pi(u) = \sigma] = 1/|\sigma|!$ and so it immediately follows that,

$$\mathbb{E}\left[R(\sigma, T)\right] = \sum_{u_1 < \ldots < u_{|\sigma|}} \mathbb{E}\left[\pi(u) = \sigma\right] = \frac{1}{|\sigma|!} \sum_v \binom{d(v)}{|\sigma| - 1}. \tag{1}$$

For length two permutations, e.g. inversions, $\mathbb{E}\left[R(21, T)\right] = \frac{1}{2}\Upsilon(T)$ where $\Upsilon(T) \stackrel{\text{def}}{=} \sum_v d(v)$ is called the *total path length* of T. We state our results in terms of a tree parameter $\Upsilon_r^k(T)$ which generalises the notion of total path length.

We define $\Upsilon_r^k(T)$ which allows us to generalize (1) to higher moments of $R(\sigma, T)$. For r nodes $v_1, \ldots, v_r$ (not necessarily distinct), let $c(v_1, \ldots, v_r)$ be the number of ancestors that they share $c(v_1, \ldots, v_r) \stackrel{\text{def}}{=} |\{u \in V : u \leq v_1, v_2, \ldots, v_r\}|$ which is also the depth of the least common ancestor plus one. That is $c(v_1, \ldots, v_r) = d(v_1 \vee \ldots \vee v_r) + 1$ where we write $d(v)$ for the depth of v and $v_1 \vee v_2$ for the least common ancestor of v_1 and v_2. The 'off by one error' is because the root is in the set of common ancestors for any subsets of nodes but we use the convention the root has depth 0. Also define

$$\Upsilon_r^k(T) \stackrel{\text{def}}{=} \sum_{v_1, \ldots, v_r} c(v_1, \ldots, v_r) \prod_{i=1}^r \binom{d(v_i)}{k-2}, \tag{2}$$

where the sum is over all ordered r-tuples of nodes in the tree and with the convention $\binom{x}{0} = 1$. For a single node v, $d(v) = c(v) - 1$, since v itself is counted in $c(v)$. So $\Upsilon(T) = \Upsilon_1^2(T) - |V|$; i.e., we recover the usual notion of total path length. The $k = 2$ case recovers the r-total common ancestors defined in [3], $\Upsilon_r^2(T) = \sum_{v_1, \ldots, v_r} c(v_1, \ldots, v_r)$.

Indeed the distribution of the number of permutations in a fixed tree has already been studied in [3]. Let $\varkappa_r = \varkappa_r(X)$ denote the r-th cumulant of a random variable X (provided it exists); thus $\varkappa_1(X) = \mathbb{E}\left[X\right]$ and $\varkappa_2(X) = \text{Var}\left(X\right)$.

▶ **Theorem 1** (Thm 1 of Cai et al. [3]). *Let T be a fixed tree. Let $\varkappa_r = \varkappa_r(R(21, T))$ be the r-th cumulant of $R(21, T)$. Then for $r \geq 2$,*

$$\varkappa_r = \frac{B_r(-1)^r}{r}\left(\Upsilon_r^2(T) - |V|\right)$$

where B_r denotes the r-th Bernoulli number.

For the case of T a complete binary tree on n vertices we asymptotically recover this result for large n. Moreover we extend it to cover any fixed permutation σ for complete binary trees.

▶ **Remark.** In essence Theorem 1 of [3] shows the r-th cumulant of the number of inversions is a constant times $\Upsilon_r^2(T)$. Our main result on fixed trees, Theorem 2 (resp. Theorem 4 on split trees), shows that for any fixed permutation σ of length k for complete binary trees (and whp for split trees) the r-th cumulant is a constant times $\Upsilon_r^k(T_n)$ asymptotically. The exact constant is defined below and is a little more involved than for inversions but observe it is a function only of the moment r and the length of $k = |\sigma|$ together with the first element σ_1 of the permutation $\sigma = \sigma_1 \ldots \sigma_k$. With some work one can show $D_{12,r} = B_r(-1)^r/r$ and so Theorem 2 does asymptotically recover Theorem 1 for complete binary trees.

We now state our first main result.

▶ **Theorem 2.** *Let T_n be the complete binary tree of depth n and fix a permutation $\sigma = \sigma_1 \ldots \sigma_k$ of length k. Let $\varkappa_r = \varkappa_r(R(\sigma, T_n))$ be the r-th cumulant of $R(\sigma, T_n)$. Then for $r \geq 2$,*

$$\varkappa_r = D_{\sigma,r}\Upsilon_r^k(T_n) + o\big(\Upsilon_r^k(T_n)\big) \tag{3}$$

where

$$D_{\sigma,r} \overset{\text{def}}{=} \sum_{j=0}^{r}\left(\frac{-1}{k!}\right)^{r-j}\binom{r}{j}\frac{\big(j(\sigma_1-1)\big)!\big(j(k-\sigma_1-1)\big)!}{\big(j(k-1)+1\big)!\big((\sigma_1-1)!(k-\sigma_1)!\big)^j}. \tag{4}$$

This implies the following corollary.

▶ **Corollary 3.** *Let T_n be the complete binary tree of depth n. For permutations σ of length 3,*

$$\mathbb{V}(R(\sigma, T_n)) = \begin{cases} \frac{1}{45}\Upsilon_2^3(T_n)(1 + o(1)) & \text{for } \sigma = 123, 132, 312, 321 \\ \frac{1}{180}\Upsilon_2^3(T_n)(1 + o(1)) & \text{for } \sigma = 213, 231 \end{cases}$$

and more generally for $\sigma = \sigma_1\sigma_2 \ldots \sigma_k$,

$$\mathbb{V}(R(\sigma, T_n)) = \begin{cases} \frac{1}{((k-1)!)^2}\left(\frac{1}{2k-1} - \frac{1}{k^2}\right)\Upsilon_2^k(1 + o(1)) & \text{for } \sigma_1 \in \{1, k\} \\ \left(\frac{1}{(2k-1)(k-\sigma_1)!(k+\sigma_1-2)!} - \frac{1}{(k!)^2}\right)\Upsilon_2^k(1 + o(1)) & . \end{cases}$$

▶ **Remark.** The methods of proof are very different for inversions and general permutations. In [3], the method takes advantage of a nice independence property of permutations. For a node u let I_u be the number of inversions involving u as the top node: $I_u = |\{w : u < w, \pi(u) > \pi(w)\}|$. Then the $\{I_u\}_u$ are independent random variables and I_u is distributed as the uniform distribution on $\{0, \ldots, |T_u|\}$, see Lemma 1 of [3].

Without an obvious similar independence property for general permutations our route instead uses nice properties on the number of embeddings of small digraphs in both binary trees and, whp, in split trees. This property allows us to calculate the centralised r-th

moment of $R(\sigma, T)$ directly from a sum of products of indicator variables as most terms in the sum are zero or negligible by the embedding property. The centralised r-th moment is then approximately a function of the j-th cumulants for $j \leq r$ and we are able to deduce the r-th cumulant by induction.

We now define a particular notion of embedding small digraphs into a tree which will be important as discussed in the previous remark.

In the complete binary tree we have a natural partial order, the ancestor relation, where the root is the ancestor of all other nodes. Any fixed acyclic digraph also induces a partial order on its vertices where $v > u$ if there is a directed path from v to u. Define $[\vec{H}]_{T_n}$ to be the number of embeddings ι of $\vec{H}$ to distinct nodes in T_n such that the partial order of vertices in $\vec{H}$ is respected by the embedding to nodes in T_n under the ancestor relation.

$$[\vec{H}]_{T_n} \overset{\text{def}}{=} |\{\iota : V(\vec{H}) \to V(T_n) \text{ such that if } u < v \text{ in } \vec{H} \text{ then } \iota(u) < \iota(v) \text{ in } T_n\}|$$

Observe the inverse of embedding ι^{-1} need not respect relations. If $u \perp v$ in $\vec{H}$, i.e. u, v are incomparable in $\vec{H}$ then we can embed so that $\iota(u) < \iota(v)$, $\iota(u) > \iota(v)$ or $\iota(u) = \iota(v)$ in T_n. For an example of this take the digraph ⬦ and denote by P_ℓ the rooted path on ℓ nodes. Notice that in ⬦ two of the vertices are incomparable but the vertices of the digraph can be embedded into the nodes of a path which are completely ordered. The counts are $[⬦]_{P_4} = 2$ and in general $[⬦]_{P_\ell} = 2\binom{\ell}{4}$.

A particular star-like digraph $\vec{S}_{k,r}$ will be important. This is the digraph obtained by taking r directed paths of length k and fusing their source vertices into a single vertex. Alternatively we can state the theorem in terms of star counts as $[\vec{S}_{|\sigma|,r}]_{T_n} = \Upsilon_r^{|\sigma|}(T_n)(1+o(1))$. See the beginning of the proof of the theorem for details.

Split trees

Split trees were first defined in [4] and were introduced to encompass many families of trees that are frequently used in algorithm analysis, e.g., binary search trees [6], m-ary search trees [8] and quad trees [5].

The random split tree T_n has parameters $b, s, s_0, s_1, \mathcal{V}$ and n. The integers b, s, s_0, s_1 are required to satisfy the inequalities

$$2 \leq b, \quad 0 < s, \quad 0 \leq s_0 \leq s, \quad 0 \leq bs_1 \leq s + 1 - s_0. \tag{5}$$

and $\mathcal{V} = (V_1, \ldots, V_b)$ is a random non-negative vector with $\sum_{i=1}^b V_i = 1$.

We may now define the random split tree as follows. Consider an infinite b-ary tree $\mathcal{U}$. The split tree T_n is constructed by distributing n balls (pieces of information) among nodes of $\mathcal{U}$. For a node u, let n_u be the number of balls stored in the subtree rooted at u. Once n_u are all decided, we take T_n to be the largest subtree of $\mathcal{U}$ such that $n_u > 0$ for all $u \in T_n$. Let $\mathcal{V}_u = (V_{u,1}, \ldots, V_{u,b})$ be the independent copy of $\mathcal{V}$ assigned to u. Let $u_1, \ldots, u_b$ be the child nodes of u. Conditioning on n_u and $\mathcal{V}_u$, if $n_u \leq s$, then $n_{u_i} = 0$ for all i; if $n_u > s$, then

$$(n_{u_1}, \ldots, n_{u_b}) \sim \text{Mult}(n - s_0 - bs_1, V_{u,1}, \ldots, V_{u,b}) + (s_1, s_1, \ldots, s_1),$$

where Mult denotes multinomial distribution, and b, s, s_0, s_1 are integers satisfying (5). Note that $\sum_{i=1}^b n_{u_i} \leq n$ (hence the "splitting"). Naturally for the root ρ, $n_\rho = n$. Thus the distribution of $(n_u, \mathcal{V}_u)_{u \in V(\mathcal{U})}$ is completely defined. For this paper we will also require that the internal node capacity s_0 is at least one so that there are some internal balls to receive labels.

This next theorem is our other main result.
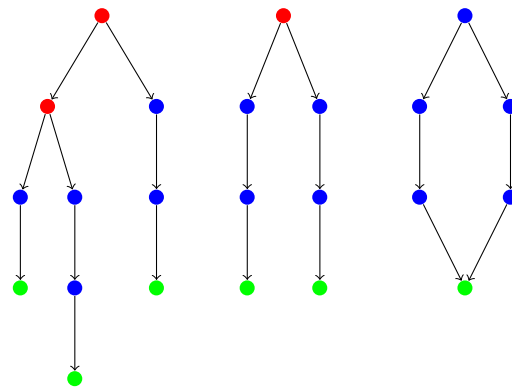
■ Figure 1 An example of a directed acyclic graph $\vec{H}$ with 'sink' (green), 'ancestor' (blue) and 'common-ancestor' (red) nodes indicated by colour. This particular digraph is in $\mathcal{G}_{4,7}$ and it appears in the seventh moment calculations of $R(\sigma, T)$ for $|\sigma| = 4$.

▶ **Theorem 4.** *Fix a permutation $\sigma = \sigma_1 \ldots \sigma_k$ of length k. Let T_n be a split tree with split vector $\mathcal{V} = (V_1, \ldots, V_b)$ and n balls. Let $\varkappa_r = \varkappa_r(R(\sigma, T_n))$ be the r-th cumulant of $R(\sigma, T_n)$. For $r \geq 2$ the constant $D_{\sigma,r}$ is defined in line (4). Whp the split tree T_n has the following property.*

$$\varkappa_r = D_{\sigma,r} \Upsilon_r^k(T_n) + o\big(\Upsilon_r^k(T_n)\big).$$

Our theorem says the following. Generate a random split tree T_n, whp it has the property that the random number of occurrences of any fixed subpermutation in a random ball labelling of T_n has variance and higher cumulant moments approximately a constant times a 'simple' tree parameter of T_n.

We may contrast this with Theorem 4 of [3]. This theorem states the distribution of the number of inversions in a random split tree; where the distribution is expressed as the solution of a system of fixed point equations. It is work in progress to find the distribution of $\Upsilon_r^k(T_n)$. This would extend Theorem 4 of [3] about inversions to general permutations.

2 Embeddings of small digraphs into the complete binary tree

Certain classes of digraphs will be important in the proof of Theorem 2, loosely those that may be obtained by taking r copies of the path $\vec{P}_k$ and iteratively fusing pairs of vertices together. It will also matter how many embeddings each digraph has into the complete binary tree. In Proposition 9 we show the counts for most digraphs in such a class are dwarfed by the counts of a particular digraph in the class. The main work in the proof of this proposition is to show that the number of embeddings of any digraph $\vec{H}$, up to a factor of n, depends only on the numbers of two types of vertices in $\vec{H}$. We separate this result out as a lemma, Lemma 5, which we show first before proving the proposition.

A vertex in a directed graph is a *sink* if it has zero out-degree. For a directed acyclic graph $\vec{H}$ we define $A_i \subseteq V(\vec{H})$ to be the vertices with exactly i descendents in $\vec{H}$ which are sinks. In particular A_0 is the set of sink vertices. We will call vertices in A_1 *ancestors* as they are ancestors of a single sink and those in A_i for $i \geq 2$ *common-ancestors* as they are the common ancestor of at least two sinks (see Figure 1). Observe if $\vec{H}$ is a directed forest then the sinks are the leaves but a sink may have indegree more than one as in the rightmost sink in Figure 1

The next lemma shows that the numbers of sinks and ancestors in $\vec{H}$ determine the number of ways to map $\vec{H}$ into the complete binary tree T_n on n vertices to within a factor of $\ln n$.

▶ **Lemma 5.** *Let $\vec{H}$ be a fixed directed acyclic graph and let T_n be the complete binary tree of height m with $n = 2^{m+1} - 1$ vertices. Then writing $|A_0| = |A_0(\vec{H})|$ for the number of sink (green) vertices and $|A_1| = |A_1(\vec{H})|$ for the number of ancestor (blue) vertices*

$$\Omega(n^{|A_0|}(\ln n)^{|A_1|}) = [\vec{H}]_{T_n} = o(n^{|A_0|}(\ln n)^{|A_1|+1}).$$

Proof of upper bound. The key observation is that for most pairs of nodes in T_n their least common ancestor is very near the root. Let the nodes at depth d be $w_1, \ldots, w_{2^d}$. Fix a node u in the tree. Provided the depth of node u is at least d, i.e. $h(u) \geq d$ then if $c(u, v) \geq d$ it must be that u and v are in the same subtree T_{w_i} for some i. If $h(u) \geq d$ let $w(u)$ be the node at depth d which is either node u itself or an ancestor of u. Thus

$$
\begin{aligned}
\sum_{u,v} \mathbf{1}[c(u,v) \geq d] \quad &\leq \quad \sum_v \mathbf{1}[v \in T_{w(u)})] \sum_u \mathbf{1}[d(u) \leq d] \\
&\leq \quad 2(m + 2^{m-d+1} - 1)(2^{m+1} - 1) \\
&\leq \quad 2^{2m-d+2+1} + m2^{m+1} + 2^{2d+2} \\
&= \quad n^2 2^{-d+3} + mn \qquad\qquad\qquad (6)
\end{aligned}
$$

Fix $\epsilon > 0$ such that $|A_2|\epsilon < 1/2$. Let B be the set of $|A_0|$-tuples of vertices so that some pair of them have an ancestor at depth $> n^\epsilon$. By (6) the set is B is small: $|B| \leq |A_0|^2 n^{|A_0|} \cdot 2^{-n^\epsilon}$.

Given an embedding of A_0 into T_n the number of ways to extend an embedding of $\vec{H}$ into T_n is at most $m^{|A_1|+|A_2|}$. This is because each vertex in $A_1 \cup A_2$ must be embedded as an ancestor of the embedding of a vertex in A_0 and each vertex in T_n has at most m ancestors. And in particular, if A_0 is embedded to a $|A_0|$-tuple not in B there are at most $m^{|A_1|+\epsilon|A_2|}$ ways to extend to an embedding of $\vec{H}$. Thus

$$[\vec{H}]_T \leq n^{|A_0|}m^{|A_1|+\epsilon|A_2|} + n^{|A_0|-\epsilon}m^{|A_1|+|A_2|} = o(n^{|A_0|}(\ln n)^{|A_1|+1}),$$

where the second inequality follows because $m = \Theta(\ln n)$.   ◀

Proof of lower bound. We restrict attention to embeddings where all common-ancestors are embedded very near the root of T_n, the sink vertices are embedded to leaves of T_n and the ancestor vertices are placed on the path between the root of T_n and the leaf in to which their descendent sink was embedded (see Figure 2). There are sufficiently many such embeddings to obtain the lower bound. In fact we restrict a little further to make it easy to check all the embeddings are valid.

By an abuse in notation denote by A_2 the union $\cup_{i \geq 2} A_i$. As $\vec{H}$ is an acyclic digraph the directed edges define a partial order on all vertices of $\vec{H}$ and in particular for those in A_2. Thus this relation can be extended to a total order. Fix some total order $<_*$ on $V(\vec{H})$ and relabel vertices in A_2 so that $v_1 <_* \ldots <_* v_{|A_2|}$. Thus we may embed v_1 to the root ρ in T_n and each v_{i+1} to a child of the node to which v_i was embedded and the relation between vertices in $\vec{H}$ will be preserved by their embedding in T_n; i.e. we may embed A_2 to the nodes on the path from ρ to some u^* at depth $|A_2| - 1$. Fix such a node u^* and let T^* be the subtree of T_n from u^*.

Label the sinks $A_0 = \{s_1, \ldots, s_{|A_0|}\}$ and vertices in A_1 according to which sink they are the ancestors of $A_1^i \stackrel{\text{def}}{=} \{v \in A_1 \ : \ v < s_i\}$.

We obtain a subcount of $[\vec{H}]_{T_n}$ by embedding A_2 onto the path from ρ to u^*, embedding A_0 to leaves of T^* and then for each i in turn embedding vertices in A_i on the path from u^*
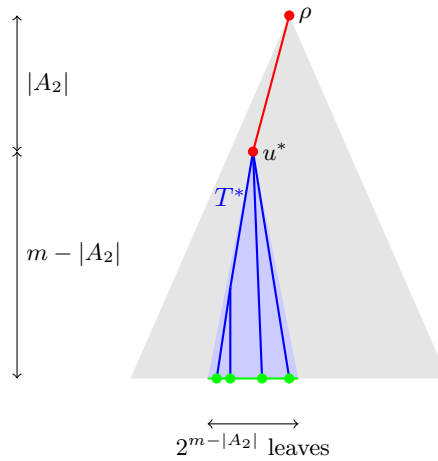
Figure 2 Schematic for the lower bound construction in Lemma 5. The colours indicate the positions in the binary tree to which the common-ancestor (red), ancestor (blue) and sink (green) vertices are embedded. Recall $A_2 = A_2(\vec{H})$ denotes the set of common-ancestor vertices of $\vec{H}$.

to the embedding of s_i. There are $m - |A_2| - 1$ vertices on the path from s_i to u^* and at most $|A_1|$ of them already have an ancestor vertex embedded onto to them (i.e. from A_1^j for some $j < i$). Thus

$$[\vec{H}]_{T_n} \geq \binom{2^{m-|A_2|}}{|A_0|} \prod_i \binom{m - |A_2| - |A_1| - 1}{|A_1^i|}$$

where the first binomial counts the number of ways to embed A_0 and the i-th binomial in the product counts the ways to embed A_1^i. Now because $\vec{H}$ is fixed $|A_2| = O(1)$ and the product over i is at least $\binom{m-|A_2|-1}{|A_1|}$ so the lower bound follows. ◄

3   Embeddings of small digraphs into the split trees

In this section we show upper and lower bounds on the number of embeddings of a fixed digraph $\vec{H}$, thought of as constant, into a random split tree with n balls. We begin by briefly listing some results on split trees from the literature that will be useful for us.

For split vector $\mathcal{V}$ define $\mu = \sum_i \mathbb{E}[V_i \ln V_i]$. The average depth of a ball is $\sim \frac{1}{\mu} \ln n$ [7][Cor 1.1]. Moreover almost all balls are very close to this depth. Define a ball v to be *good* if it has depth

$$\left| d(v) - \frac{1}{\mu} \ln n \right| \leq \ln^{0.6} n$$

and then whp $n - o(n)$ of the balls in the split tree are good [2][Thm 1.2]. That whp in a split tree all good balls have a $\Theta(n)$ depth and almost all balls are good is the only result about split trees required for the proof of the lower bound on $[\vec{H}]_{T_n}$ in Lemma 8. For the upper bound we need a bit more.

It is known that the height of a split tree with split vector $\mathcal{V}$ is whp $(c + o(1)) \ln n$ for a (known) constant c; for details see [1][Thm 2]. We write T_u to denote the subtree from bag (node) u and $|T_u|$ the number of balls in the subtree.

▶ **Lemma 6.** *Fix k. Let U be the set of bags at depth $\lfloor \alpha \ln \ln n \rfloor$ for some large enough constant $\alpha = \alpha(k)$. Then whp*

$$\sum_{u \in U} |T_u|^2 = o\left(\frac{n^2}{(\ln n)^k}\right).$$

We omit the proof of the lemma but note that it follows the same steps as Lemma 3.5 of [2].

Similarly for binary trees we show that the number of embeddings of a fixed acyclic digraph $\vec{H}$, to a good approximation, depends only on the number of 'sink' and 'ancestor' vertices in $\vec{H}$. It is a little trickier to prove the corresponding statement to the upper bound Lemma 5 in the case of split trees. However, we are rewarded by a tighter bound on the number of embeddings is determined by the numbers of 'sink' and 'ancestor' vertices up to $\ln \ln n$ factors.

▶ **Lemma 7.** *Let $\vec{H}$ be a fixed directed acyclic graph and let T_n be a split tree with split vector $\mathcal{V}$ and n balls. Then writing $|A_0| = |A_0(\vec{H})|$ for the number of sink (green) vertices, $|A_1| = |A_1(\vec{H})|$ for the number of ancestor (blue) vertices and $|A_2| = |A_2(\vec{H})|$ for the number of common-ancestor (red) vertices whp*

$$[\vec{H}]_{T_n} = O(n^{|A_0|}(\ln n)^{|A_1|}(\ln \ln n)^{|A_2|}).$$

Proof. The idea of the proof is to show that any way of embedding $A_0(\vec{H})$ into the tree can only be extended to an embedding of all the vertices in $\vec{H}$ in a limited number of ways. Note

$$[\vec{H}]_{T_n} = \sum_{\mathbf{v} = v_1, \ldots, v_{|A_0|}} f(\mathbf{v}) \tag{7}$$

where $f(\mathbf{v})$ is the number of ways to extend an embedding of $A_0(\vec{H})$ to an embedding $V(\vec{H}) \to V(T_n)$. Formally label the vertices in $A_0(\vec{H})$ by $s_1, \ldots, s_{|A_0|}$ and define

$$f(\mathbf{v}) \stackrel{\text{def}}{=} |\{\iota : \iota(s_j) = v_j \text{ for each } j = 1, \ldots, |A_0| \text{ and}$$
$$\iota : V(\vec{H}) \to V(T_n) \text{ such that if } u < v \text{ in } \vec{H} \text{ then } \iota(u) < \iota(v) \text{ in } T_n \}|.$$

We claim first that for any $\mathbf{v}$, whp $f(\mathbf{v}) = O\big((\ln n)^{|A_1|+|A_2|}\big)$ and indeed will later show a stronger bound holds for most $\mathbf{v}$.

To see this first claim recall that whp the height of a split tree on n balls is $\Theta(\ln n)$. In particular the depth of each ball v_j is $O(\ln n)$ and so v_j has $O(\ln n)$ balls as ancestors. Each vertex in $A_1(\vec{H}) \cup A_2(\vec{H})$ must be embedded to a ball which is the ancestor of some v_j (and possibly further restricted to balls which are ancestors of some set of v_j's but we will not need this). Hence there are at most $O(\ln n)$ choices of where to embed each vertex in $A_1(\vec{H}) \cup A_2(\vec{H})$ which finishes the claim.

Similarly to the proof for the case of binary trees we now exploit the fact that in split trees most pairs of balls have their least common ancestor in a bag very near the root. This will allow us to define a large set of $\mathbf{v}$ for which $f(\mathbf{v})$ is small.

Say a tuple of balls $\mathbf{v}$ is *inbred* if some pair of balls has a common ancestor at depth greater than $L \stackrel{\text{def}}{=} \lfloor \alpha \ln \ln n \rfloor$ for some α such that Lemma 6 holds with $k = |A_2|$. Denote the set of these tuples by $\mathcal{I}$. We claim that whp

$$|\mathcal{I}| \le |A_0|^2 n^2 (\ln n)^{-|A_2|}. \tag{8}$$

Before proving claim (8) let us show that it implies the theorem. If a tuple of balls is not inbred, $\mathbf{v} \notin \mathcal{I}$, then any ancestor of any pair of balls has depth at most $L = O(\ln \ln n)$. Thus whp there are at most $O(\ln \ln n)$ choices of where to embed each vertex in $A_2(\vec{H})$ when extending an embedding in which $A_0(\vec{H})$ was embedded to $\mathbf{v} \notin \mathcal{I}$. So for non inbred $\mathbf{v}$,

$$\max_{\mathbf{v} \notin \mathcal{I}} f(\mathbf{v}) = O((\ln n)^{|A_1|} (\ln \ln n)^{|A_2|}).$$

We are almost finished (modulo the claim). By (9) and recalling there are less than $n^{|A_0|}$ possible tuples of balls we get

$$[\vec{H}]_{T_n} = \sum_{\mathbf{v} \in \mathcal{I}} f(\mathbf{v}) + \sum_{\mathbf{v} \notin \mathcal{I}} f(\mathbf{v}) \leq |\mathcal{I}| O\big((\ln n)^{|A_1|+|A_2|}\big) + O(n^{|A_0|}(\ln n)^{|A_1|}(\ln \ln n)^{|A_2|}) \quad (9)$$

and so the claim $|\mathcal{I}| = O(n^{|A_0|}(\ln n)^{-|A_2|})$ does imply the theorem.

It now remains to prove the claim. Let $c(v_1, v_2)$ be the depth of the bag which is the least common ancestor of balls v_1 and v_2. To prove the claim it suffices to show

$$\sum_{v_1, v_2} \mathbf{1}[c(v_1, v_2) \geq L] \leq \frac{n^2}{(\ln n)^{|A_2|}}.$$

Trivially, if $c(v_1, v_2) \geq L$ then both v_1 and v_2 must be at depth at least L. Also notice if v_1 and v_2 have their least common ancestor at depth at least L they must have some common ancestor, u say, at depth exactly L. Let U be the set of bags at depth L. Then

$$\mathbf{1}\big[c(v_1, v_2) \geq L\big] = \mathbf{1}\big[v_1, v_2 \in T_u \text{ for some } u \in U\big]$$

and so we may apply Lemma 6 directly

$$\sum_{v_1, v_2} \mathbf{1}\big[c(v_1, v_2) \geq L\big] \ \leq \ \sum_u |T_u|^2 \ \leq \ \frac{n^2}{(\ln n)^{|A_2|}}$$

which establishes the claim. ◀

▶ **Lemma 8.** *Let $\vec{H}$ be a fixed directed acyclic graph and let T_n be a split tree with split vector $\mathcal{V} = \{V_1, \ldots, V_b\}$ and n balls. Then writing $|A_0| = |A_0(\vec{H})|$ for the number of sink (green) vertices and $|A_1| = |A_1(\vec{H})|$ for the number of ancestor (blue) vertices whp*

$$[\vec{H}]_{T_n} = \Omega(n^{|A_0|}(\ln n)^{|A_1|}).$$

Proof. *(sketch)* We describe a strategy to embed $\vec{H}$ into T_n. The details of the proof are then to show that whp this strategy can be followed to obtain a valid embedding of $\vec{H}$ and that there are sufficiently many different such embeddings to achieve the lower bound.

First embed 'common-ancestor' vertices along a path to some node u^* with $\tilde{n} = \Omega(n)$ balls. Now consider a split tree with $\tilde{n}$ balls and embed 'ancestor' and 'sink' vertices into that. Embed 'sink' vertices to 'good' balls in the tree (i.e. depth very close to the expected depth) and the 'ancestor' vertices to balls which along the path between u^* and the embedding of their descendent. See Figure 3.

We embed the common-ancestor vertices, $A_2(\vec{H})$, to the balls in the nodes on the path between a node, u^* say, at depth $|A_2| - 1$ and the root, using one ball per node. This is so far effectively the same as in the binary case. And we will later embed the 'sink' and
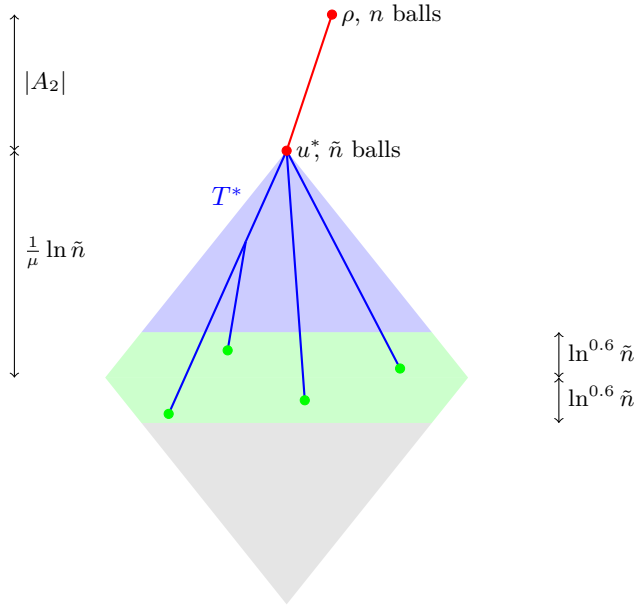
■ **Figure 3** Schematic for the construction in Lemma 8. The colours indicate the positions in the split tree to which the common-ancestor (red), ancestor (blue) and sink (green) vertices are embedded. Recall $A_2 = A_2(\vec{H})$ denotes the set of common-ancestor vertices of $\vec{H}$.

'common-ancestor' vertices to balls in the subtree T_{u^*}. We need to confirm there is some node u^* at depth $L = |A_2| - 1$ with $\tilde{n}$ balls in its subtree. Each node (bag) has capacity at most s_0 or s and at most $(b^{L+1} - 1)$ nodes, a constant number, at depth less than L, so $n - O(1)$ balls remaining. These balls are shared between b^L, a constant, number of subtrees T_u. Hence by pigeon-hole principle some vertex u^* has $\tilde{n} = \Theta(n)$ balls in its subtree.

Now work in the split tree $T_{\tilde{n}}$. Embed the 'sink' vertices to any good balls $v_1, \ldots, v_{|A_0|}$ in the split trees. There are $\Theta(\tilde{n}^{|A_0|})$ ways to embed them. Label the 'sink' vertices $s_1, \ldots, s_{|A_0|}$ and $A_1^j \subset A_1^j(\vec{H})$ to be the 'ancestor' vertices with s_j as their lone descendent. Vertices in A_1^j can be embedded to balls anywhere between v_j and u^* and so there are $\Theta((\ln \tilde{n})^{|A_1^j|})$ ways to do that for each j. All up there are $\Omega(\tilde{n}^{|A_0|}(\ln \tilde{n})^{|A_1|})$ ways to embed $A_0(\vec{H}) \cup A_1(\vec{H})$ into balls of $T_{\tilde{n}}$. But now as $\tilde{n} = \Theta(n)$ we are done. ◀

4   Star counts

After having proved the required properties of our two classes of trees, binary trees and split trees, we show these imply the desired results on cumulants of the number of appearances of a permutation in the node labellings of binary trees, respectively ball labellings in split trees.

Say a sequence of trees T_n with n nodes (resp. balls) is *explosive* if for any fixed acyclic digraph $\vec{H}$

$$\Omega(n^{|A_0|}(\ln n)^{|A_1|}) = [\vec{H}]_{T_n} = o(n^{|A_0|}(\ln n)^{|A_1|+1}).$$

Thus Section 2 was devoted to showing binary trees are explosive and Section 3 to showing split trees are explosive whp. This section proves the cumulant results using only this explosive property of the tree classes.

Now we introduce some notation in order to state Proposition 9. We use a notion of subgraph on an ordered set of vertices. For a k-tuple of vertices $V_i = (v_i^1, \ldots, v_i^k)$
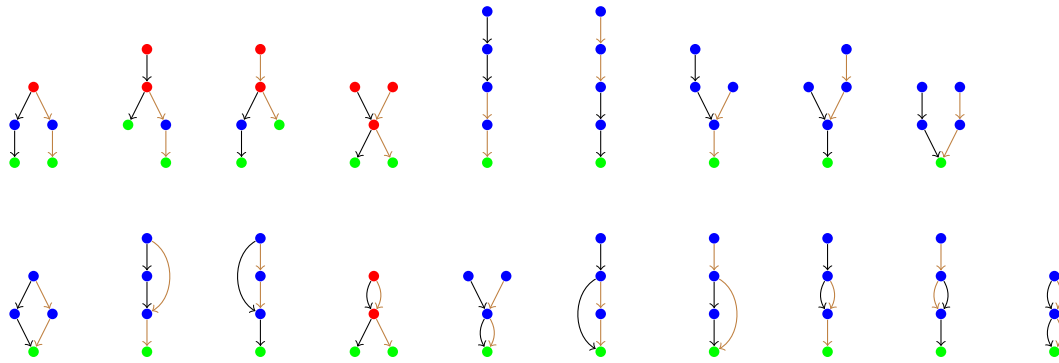
Figure 4 The set $\mathcal{G}'_{3,2}$. Labels of the first path $V_1 = (v_1^1, v_1^2, v_1^3)$ indicated by black arrows between the nodes and respectively brown arrows for labels of the second path $V_2 = (v_2^1, v_2^2, v_2^3)$. Colours of nodes indicate 'sink' (green), 'ancestor' (blue) and 'common-ancestor' (red) nodes respectively. These labelled directed acyclic graphs appear in variance calculations of $R(\sigma)$ for $|\sigma| = 3$.

we say $\vec{H}|_{V_i} = \vec{P}_k$ if the subgraph of $\vec{H}$ induced on V_i has precisely the directed edges $v_i^1 v_i^2, v_i^2 v_i^3, \ldots, v_i^{k-1} v_i^k$.

The set $\mathcal{G}_{k,r}$ is the set of acyclic digraphs which may be obtained by taking r copies of the path $\vec{P}_k$ and iteratively fusing pairs of vertices together such that each path is involved in at least one fusing operation. Likewise labelled $\vec{H}'$ in $\mathcal{G}'_{k,r}$ are those obtained by fusing together j labelled paths $\vec{P}_k$ keeping both sets of labels when a pair of vertices are fused. The set $\mathcal{G}'_{4,2}$ is illustrated in Figure 4.

Formally let $\mathcal{G}_{k,r}$ be the set of directed acyclic graphs $\vec{H}$ such that we can find (non-disjoint) vertex subsets $V_1, \ldots, V_r$ where for each i we have $\vec{H}|_{V_i} = \vec{P}_k$ and $\exists j \neq i$ with $V_i \cap V_j = \varnothing$. (The second condition is to ensure each i-th path is involved in a fusing operation.) For $\vec{H} \in \mathcal{G}_{k,r}$ write $\vec{H}'$ for $\vec{H}$ together with a labelling $V_1, \ldots, V_r$ (note some vertices have multiple labels). Likewise write $\mathcal{G}'_{k,r}$ for the labelled set of graphs.

Denote by $\vec{S}_{k,j}$ the digraph composed by taking j copies of the path $\vec{P}_k$ and fusing the j source vertices into a single vertex. Also define $\mathcal{S}^*_{k,r} = \cup_i \vec{S}_{k,r_i}$ where the disjoint union is over all S_{k,r_i} with $\sum_i r_i = r$ and $r_i \geq 2$. Observe $\mathcal{S}_{k,r} \subset \mathcal{G}_{k,r}$.

▶ **Proposition 9.** *Fix k, r and let $\vec{H} \in \mathcal{G}_{k,r}$. Suppose T_n is explosive. If $\vec{H} \notin \mathcal{S}_{k,r}$ then*

$$[\vec{H}]_{T_n} = o\Big([\vec{S}_{k,r}]_{T_n}\Big).$$

Proof. First observe that $\vec{S}_{k,r}$ has r sink vertices, $(k-2)r$ ancestor vertices and exactly one common-ancestor vertex. Thus by the explosive property of T_n

$$[\vec{S}_{k,r}]_{T_n} = \Omega(n^r (\ln n)^{(k-2)r}).$$

Fix $\vec{H} \in \mathcal{G}_{k,r} \backslash \mathcal{S}_{k,r}$ and fix a labelling $V_1, \ldots, V_r$ on $\vec{H}$. Again by the explosive property

$$[\vec{H}]_{T_n} = o(n^{|A_0(\vec{H})|} (\ln n)^{|A_1(\vec{H})|+1}). \tag{10}$$

Hence if $|A_0(\vec{H})| \leq r - 1$ then $[\vec{H}]_{T_n} = o([\vec{S}_{k,r}])$ and so we would be done. Thus we may assume that $A_0(\vec{H}) = r$ and it will suffice to show that $A_1(\vec{H}) < (k-2)r$. Consider the path labelled $V^i = (v_1^i, \ldots, v_k^i)$. We know v_k^i is a sink vertex and not fused with any other vertex otherwise we would have $A_0(\vec{H}) < r$. If vertex v_j^i is fused with another vertex, it must be a vertex on a different path to avoid a cycle, and so v_j^i and $v_{j-1}^i, \ldots, v_1^i$ would

become common-ancestors. Thus if v_j^i is fused to another vertex there are at most $(k - j - 1)$ ancestor vertices in path V_i. Hence if $A_1(\vec{H}) = (k - 2)r$ then we must have only fused the source vertices of each path but this means that $\vec{H} \in \mathcal{S}_{k,r}$ and so we are done.   ◀

By exploiting only the explosive property of binary and (whp) of split trees we prove the moments result for both classes at once. In particular observe that Theorems 2 and 4 are both implied by taking Proposition 10 along with the lemmas proving binary trees are explosive and split trees are whp explosive.

▶ **Proposition 10.** *Suppose T_n is explosive. Let $\varkappa_r = \varkappa_r(R(\sigma, T_n))$ be the r-th cumulant of $R(\sigma, T_n)$. Then for $r \geq 2$,*

$$\varkappa_r = D_{\sigma,r} \Upsilon_r^{|\sigma|}(T_n) + o(\Upsilon_r^{|\sigma|}(T_n)).$$

Proof sketch. The proof proceeds by induction on r with $r = 2$, the variance, as the base case. The variance calculation is also a simpler version of the calculations for higher r and so illustrates the key steps we use for the inductive step.

We give a rough idea of these steps. The variance (and higher centralised moments) can be written as a sum over indicator random variables for a subpermutation occuring on a set of $|\sigma|$ nodes. Almost all terms in this sum are zero or negligible. Firstly if the indicators concern disjoint sets of vertices they are independent and because we calculate centralised moments these terms drop away. This leaves only terms in the sum in which the nodes of indicator variables overlap. We group the terms by how the vertices in these sets overlap and the results about numbers of embeddings then show most groups are negligible.

For the variance only one group is non-negligible and so we will be done at this step. In the inductive step the centralised r-th moment has only one 'new' group (not occuring in smaller moment calculations) which is non-negligible as well as non-negligible groups which appeared in smaller cumulants for $j \leq r$. This occurs in such a way that we can prove this new group approximates the r-th cumulant.   ◀

References

1   Nicolas Broutin, Luc Devroye, and Erin McLeish. Weighted height of random trees. *Acta Informatica*, 45(4):237, 2008. `doi:10.1007/s00236-008-0069-0`.

2   Nicolas Broutin and Cecilia Holmgren. The total path length of split trees. *The Annals of Applied Probability*, 22(5):1745–1777, 2012. `doi:10.1214/11-aap812`.

3   Xing Shi Cai, Cecilia Holmgren, Svante Janson, Tony Johansson, and Fiona Skerman. Inversions in split trees and conditional Galton–Watson trees. *arXiv preprint arXiv:1709.00216*, 2017.

4   Luc Devroye. Universal limit laws for depths in random trees. *SIAM Journal on Computing*, 28(2):409–432, 1998. `doi:10.1137/s0097539795283954`.

5   R. Finkel and J. Bentley. Quad trees a data structure for retrieval on composite keys. *Acta informatica*, 4(1):1–9, 1974. `doi:10.1007/bf00288933`.

6   C. Hoare. Quicksort. *The Computer Journal*, 5(1):10–16, 1962. `doi:10.1145/366622.366644`.

7   C. Holmgren. Novel characteristics of split trees by use of renewal theory. *Electronic Journal of Probability*, 17, 2012. `doi:10.1214/ejp.v17-1723`.

8   R. Pyke. Spacings. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 395–449, 1965.

Analytic Combinatorics of Lattice Paths with Forbidden Patterns: Asymptotic Aspects and Borges's Theorem

Andrei Asinowski
Technische Universität Wien, Institut für Diskrete Mathematik und Geometrie,
`https://dmg.tuwien.ac.at/asinowski/`
 https://orcid.org/0000-0002-0689-0775

Axel Bacher
Université Paris 13, LIPN, UMR CNRS 7030, `http://lipn.univ-paris13.fr/~bacher`
 https://orcid.org/0000-0002-9789-7074

Cyril Banderier
Université Paris 13, LIPN, UMR CNRS 7030, `http://lipn.univ-paris13.fr/~banderier`
 https://orcid.org/0000-0003-0755-3022

Bernhard Gittenberger
Technische Universität Wien, Institut für Diskrete Mathematik und Geometrie,
`https://dmg.tuwien.ac.at/gittenberger/`
 https://orcid.org/0000-0002-2639-8227

──── **Abstract** ────

In a companion article dedicated to the enumeration aspects, we showed how to obtain closed form formulas for the generating functions of walks, bridges, meanders, and excursions avoiding any fixed word (a *pattern p*). The autocorrelation polynomial of this forbidden pattern p (as introduced by Guibas and Odlyzko in 1981, in the context of regular expressions) plays a crucial role. In this article, we get the asymptotics of these walks. We also introduce a trivariate generating function (length, final altitude, number of occurrences of p), for which we derive a closed form. We prove that the number of occurrences of p is normally distributed: This is what Flajolet and Sedgewick call an instance of Borges's theorem.

We thus extend and refine the study by Banderier and Flajolet in 2002 on lattice paths, and we unify several dozens of articles which investigated patterns like peaks, valleys, humps, etc., in Dyck and Motzkin paths. Our approach relies on methods of analytic combinatorics, and on a matricial generalization of the kernel method. The situation is much more involved than in the Banderier–Flajolet work: forbidden patterns lead to a wider zoology of asymptotic behaviours, and we classify them according to the geometry of a Newton polygon associated with these constrained walks, and we analyse what are the universal phenomena common to all these models of lattice paths avoiding a pattern.

■ **Figure 1** Some models of self-avoiding walks are encoded by partially directed lattice paths avoiding a pattern (see [2]). In this article, we analyse the asymptotics of more general walks (any jumps, any forbidden pattern).
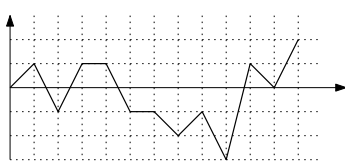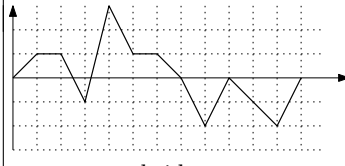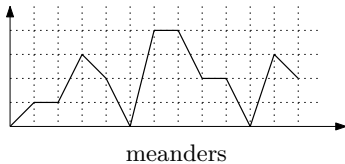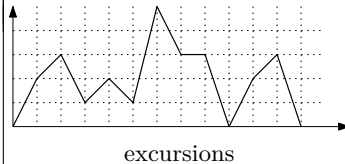
1   Introduction

Combinatorial structures having a rational or an algebraic generating function play a key role in many fields: computer science (analysis of algorithms involving trees, lists, words), computational geometry (integer points in polytopes, maps, graph decomposition), bioinformatics (RNA structure, pattern matching), number theory (integer compositions, automatic sequences and modular properties, integer solutions of varieties), probability theory (Markov chains, directed random walks), see e.g. [4, 11, 22, 32]. They are often the trace of a structure which has a recursive specification in terms of a system of tree-like structures, or of some functional equation solvable by variants of the kernel method [12].

Since the seminal article by Chomsky and Schützenberger on the link between context-free grammars and algebraic functions [15], which also holds for pushdown automata [30], many articles encoded and enumerated combinatorial structures via a formal language approach. See e.g. [20, 25, 28] for such an approach on the so-called *generalized Dyck languages*. The words generated by these languages are in bijection with directed lattice paths, and in this article, we try to understand how some of these fundamental objects can be enumerated when they have the additional constraint to avoid a given pattern. For sure, such a class of objects can be described as the intersection of a context-free language and a rational language; therefore, classical closure properties imply that they are directly generated by another (but huge and clumsy) context-free language. Unfortunately, despite the fact that the algebraic system associated with the corresponding context-free grammar is *in theory* solvable by a resultant computation or by Gröbner bases, this leads *in practice* to equations which are so big that no current computer could handle them in memory, even for generalized Dyck languages with only 20 different letters.

In this article, we generalize the asymptotics obtained by Banderier and Flajolet [5] to lattice paths avoiding a given pattern. As we shall see, the situation is much more involved, and we build on the explicit formulas that we obtained in our companion article [1]. There, we introduced a generic way to tackle the question of enumerating words avoiding a given pattern (for languages generated by pushdown automata) which bypass these intractable equations. For directed lattice paths, our method allows to handle an arbitrary number of letters (*i.e.*, allowed steps), up to alphabets of thousands of letters, computationally in a few minutes. It relies on an analytic combinatorics approach, and also on the kernel method, which we used in our investigation of enumerative and asymptotic properties of lattice paths [6–8]. This allows to unify the considerations of many articles which investigated natural patterns like peaks, valleys, humps, *etc.*, in Dyck and Motzkin words, corresponding patterns in trees, compositions..., see e.g. [9, 10, 13, 16, 17, 19, 21, 26, 27, 29].

■ **Table 1** Summary of our results from [1], which extend the Banderier–Flajolet results from [5] to lattice paths avoiding a pattern. For the four types of paths and for any set of jumps encoded by $P(u)$, we give the corresponding generating function of such lattice paths avoiding a pattern p (of length ℓ and final altitude b). The formulas involve the autocorrelation polynomial $R(t, u)$ of p, and the small roots u_i of the kernel $K(t, u) := (1 - tP(u))R(t, u) + t^\ell u^b$.

	ending anywhere	ending at 0
on $\mathbb{Z}$	 walks $W(t, u) = \dfrac{R(t, u)}{K(t, u)}$	 bridges $B(t) = -\displaystyle\sum_{i=1}^{e} \frac{u_i'}{u_i} \frac{R(t, u_i)}{K_t(t, u_i)}$
on $\mathbb{N}$	 meanders $M(t, u) = \dfrac{R(t, u)}{K(t, u)} \displaystyle\prod_{i=1}^{c}(u - u_i(t))$	 excursions $E(t) = \dfrac{(-1)^{c+1}}{t} \displaystyle\prod_{i=1}^{c} u_i(t)$

2 Definitions and notations

In their paper, Banderier and Flajolet consider the following setting. Let $\mathcal{S}$, the *set of steps* (or *jumps*), be some finite subset of $\mathbb{Z}$ that contains at least one negative and at least one positive number. A *lattice path with steps from* $\mathcal{S}$ is a finite word $w = [v_1, v_2, \ldots, v_n]$ in which all letters belong to $\mathcal{S}$, visualized as a directed polygonal line in the plane, which starts in the origin and is formed by successive appending of vectors $(1, v_1), (1, v_2), \ldots, (1, v_n)$. The letters that form the path $w = [v_1, v_2, \ldots, v_n]$ are referred to as its *steps*. The *length* of w, to be denoted by $|w|$, is the number of steps in w. The *final altitude* of w, to be denoted by $h(w)$, is the sum of all steps in w, that is $v_1 + v_2 + \ldots + v_n$.

Under this setting, it is usual to consider two restrictions: that the whole path is (weakly) above the x-axis, and that it has final altitude 0 (equivalently, terminating at the x-axis). Consequently, one considers four classes of lattice paths:

1. A *walk* is any path as described above.
2. A *bridge* is a path that terminates at the x-axis.
3. A *meander* is a path that stays (weakly) above the x-axis.
4. An *excursion* is a path that stays (weakly) above the x-axis and terminates at the x-axis.

In the generating functions, the variable t corresponds to the length of a path, and the variable u to its final altitude. $P(u)$ is the *characteristic polynomial* of the set of steps $\mathcal{S}$, defined by $P(u) = \sum_{s \in \mathcal{S}} u^s$. The smallest (negative) number in $\mathcal{S}$ is denoted by $-c$, and the largest (positive) number in $\mathcal{S}$ is denoted by d: that is[1], if one orders the terms of $P(u)$ by the powers of u, one has $P(u) = u^{-c} + u^{s_2} + \ldots + u^{s_{|\mathcal{S}|-1}} + u^d$ with $\mathcal{S} = \{-c, s_2, \ldots, s_{|\mathcal{S}|-1}, d\}$. The drift of the walk is given by the quantity $P'(1)$.

---

[1] Some weights (or probabilities, or multiplicities) could be associated with each jump, but we omit them in this article for clarity. All the proofs would be similar.

Lattice paths with forbidden patterns and the autocorrelation polynomial

We consider lattice paths with step set $\mathcal{S}$ that avoid a certain *pattern*, that is, an *a priori* fixed path $p = [a_1, a_2, \dots, a_\ell]$. To be precise, we define an *occurrence* of p in a lattice path w as a substring of w which coincides with p. If there is no occurrence of p in w, we say that w *avoids* p. For example, the path $[1, 2, 3, 1, 2]$ has two occurrences of $[1, 2]$, but it avoids $[2, 1]$.

Before we state our results, we introduce some notations.

A *presuffix* of p is a non-empty string that occurs in p both as a prefix and as a suffix. In particular, the whole word p is a (trivial) presuffix of itself. If p has one or several non-trivial presuffixes, we say that p exhibits an *autocorrelation* phenomenon. For example, for the pattern $p = [1, 1, 2, 1, 2]$ we have no autocorrelation. In contrast, the pattern $p = [1, 1, 2, 3, 1, 1, 2, 3, 1, 1]$ has three non-trivial presuffixes: $[1]$, $[1, 1]$, and $[1, 1, 2, 3, 1, 1]$, and thus in this case we have autocorrelation.

While analysing the Boyer–Moore string searching algorithm and properties of periodic words, Guibas and Odlyzko [23] introduced in 1981 what turns out to be one of the key characters of our article, the autocorrelation polynomial[2] of the pattern p: For any given word p, let $\mathcal{Q}$ be the set of its presuffixes; the *autocorrelation polynomial* of p is

$$R(t, u) = \sum_{q \in \mathcal{Q}} t^{|\bar{q}|} u^{h(\bar{q})}, \tag{1}$$

where $\bar{q}$ denotes the complement of q in p (i.e. $q\bar{q} = p$) and $h(\bar{q})$ the final altitude of a walk made of the steps of $\bar{q}$.

For example, consider the pattern $p = [1, 1, 2, 3, 1, 1, 2, 3, 1, 1]$. Its four presuffixes produce four terms of $R(t, u)$ as follows:

q	$|\bar{q}|$	$h(\bar{q})$
$[1]$	9	15
$[1, 1]$	8	14
$[1, 1, 2, 3, 1, 1]$	4	7
$[1, 1, 2, 3, 1, 1, 2, 3, 1, 1]$	0	0

Therefore, for this p we have $R(t, u) = 1 + t^4 u^7 + t^8 u^{14} + t^9 u^{15}$. Notice that if for some p no autocorrelation occurs, then we have $\mathcal{Q} = \{p\}$ and therefore $R(t, u) = 1$.

Finally, we define the *kernel* of a lattice path avoiding some pattern p as the following Laurent polynomial:

$$K(t, u) := (1 - tP(u))R(t, u) + t^{|p|} u^{h(p)}. \tag{2}$$

Also, in our case it can be shown that each root $u = u(t)$ of $K(t, u) = 0$ is either small (*i.e.*, $\lim_{t \to 0} u(t) = 0$) or large (*i.e.*, $\lim_{t \to 0} |u(t)| = +\infty$). The small roots are denoted by $u_1, \dots, u_e$. We will also refer to them as the *small branches*.

Now we can state the enumeration results. Recall that t is the variable for the length of a path, and u is the variable for its final altitude.

▶ **Theorem 1.** *Let $\mathcal{S}$ be a set of steps, and let p be a pattern with steps from $\mathcal{S}$. Denote $\ell = |p|$, $b = h(p)$. Let $R(t, u)$ be the autocorrelation polynomial of the pattern p. Let $u_1, \dots, u_e$ be the*

---

[2] A similar notion also appears in the work of Schützenberger on synchronizing words [31].
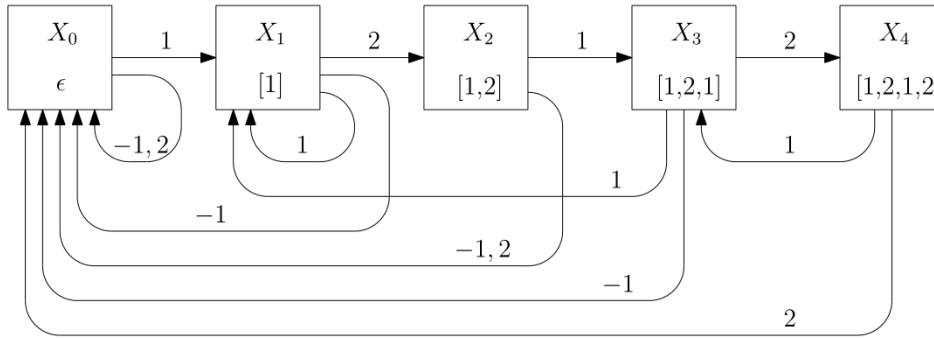
Figure 2 The automaton for the jumps $\mathcal{S} = \{-1, 1, 2\}$ and the pattern $p = [1, 2, 1, 2, -1]$. Any walk avoiding a given pattern p is associated with a similar automaton. It is in fact a *pushdown* automaton, in order to follow the positivity constraint. The matricial kernel method leads to the formulas of Theorem 1 for the corresponding generating functions, without having to solve a big algebraic system.

small roots of the kernel $K(t, u)$, as defined in (2). Then (under one additional constraint detailed in the proof), the generating functions of walks, bridges, meanders and excursions avoiding the pattern p are given by:

$$W(t) = \frac{1}{1 - tP(1) + t^\ell/R(t, 1)} \,, \tag{3}$$

$$B(t) = -\sum_{i=1}^{e} \frac{u_i'(t)}{u_i(t)} \frac{R(t, u_i)}{K_t(t, u_i)}, \tag{4}$$

$$M(t) = \frac{R(t, 1)}{K(t, 1)} \prod_{i=1}^{c} \big(1 - u_i(t)\big) \,, \tag{5}$$

$$E(t) = \begin{cases} \frac{(-1)^{c+1}}{t} \prod_{i=1}^{c} u_i(t) & \text{if } b > -c, \\ \frac{(-1)^{c+1}}{t - t^\ell} \prod_{i=1}^{c} u_i(t) & \text{if } b = -c. \end{cases} \tag{6}$$

Proof. We refer to our companion article [1] for the proofs and the complete bivariate generating functions. The kernel $K(t, u)$ is in fact the determinant of $(\mathbf{I} - tA(u))^{-1}$, where $A(u)$ is the transition matrix encoding the stack automaton associated with the constrained walk (see Figure 2 below). The formulas then follow from an extension of the kernel method to matrix equations. (In fact, we presented above the simplified formulas for M and E, when p is what we call a *pseudomeander*, *i.e.* a lattice path which does not cross the x-axis, except, possibly, at its last step. If this is not the case, then we may have more than c small roots.) ◀

▶ Remark. Notice that for these four classes of lattice paths, forbidding a pattern of length 1 or using symbolic weights for the jumps recovers the formulas from Banderier and Flajolet [5].

4   Asymptotics of excursions avoiding a given pattern

The aim of this section is to characterize the asymptotics of the number of walks with jumps $\mathcal{S}$ avoiding a given pattern p.

▶ **Lemma 2** (Location of the dominant singularity). *The dominant singularity (i.e. the nearest from zero) of $B(t)$ and $E(t)$ is ρ, the smallest real positive number where a small branch meets a large branch. (The branches refer to the solutions of $K(t, u) = 0$, as defined in* (2)).

Proof. Lattice paths avoiding a given pattern can be generated by a pushdown automaton (see Figure 2). Accordingly, they can be generated by a context-free grammar, and their generating functions therefore satisfy a "positive" system of algebraic equations (see [15]). Therefore, the asympotic number of words of length n in such languages is of the form $C\rho^{-n}n^{\alpha}$. When the system is not strongly connected, α is either an integer (if ρ is a pole), either a dyadic number (if one has an iterated square root Puiseux singularity at ρ), as proven by Banderier and Drmota in [4]. For excursions, one has a strongly connected dependency graph (see Figure 2); the dominant singularity ρ (or, possibly, the dominant singularities) thus behaves like a square root. What is more, the cycle lemma (see the discussion on this in [5]) gives a correspondence between excursions and bridges, which implies that $E(t)$ and $B(t)$ have the same radius of convergence (this still holds when there is a forbidden pattern).

Now, because of the product formula (6) for excursions, one (or several) of the small branches have to follow this square root Puiseux behaviour. By Pringsheim's theorem, this has to be at a place $0 < \rho \leq 1$; the geometry of the branches implies (see Table 2) that its location is where a large branch meets a small branch (because if the branching point comes from the intersection of small branches only, then their product will be regular). Therefore, ρ has to be the smallest real positive number where a small branch meets a large branch.   ◀

▶ **Remark.** ρ is also the radius of convergence of meanders with negative or zero drift. For meanders with positive drift, the dominant pole of $1/K(t, 1)$ will be the radius of convergence.

In order to avoid pathological cases, we now focus on *generic* walks.

▶ **Definition 3** (Generic walks). We call a constrained walk model "generic" if the following three properties hold.

- Property 1. The generating functions $B(t), M(t)$ and $E(t)$ are algebraic, not rational.
- Property 2. They have a unique domininant singularity.
- Property 3. No large negative branch (i.e. a branch such that $\lim_{z \to 0^{+}} u(z) = -\infty$) meets a small negative branch at ρ.

These three properties are very natural; we now comment more on them:

- For Property 1, it can be the case that the forbidden pattern leads to a degenerated model, in the sense that it is no more involving any stack and then we have words generated by a *regular* automaton (then, the generating functions are rational and the asymptotics are well understood). Example: $\mathcal{S} = \{-1, 1\}$ and $p = [1, -1]$ or $p = [-1, -1]$.
- For Property 2, it is proven in [3] that multiple dominant singularities appear if and only if the gcd of the pairwise differences of the jumps is not 1. In this case, the asymptotics are obtained via [8, Theorem 8.8].
- For Property 3, we conjecture that it always holds. We have a proof for many classes of walks, but some remaining cases are tricky as it is possible to exhibit cases where one small negative branch meets a large negative branch, at some $\rho' > \rho$: This is e.g. the case for $\mathcal{S} = \{-2, -1, 0, 1, 2\}$ and $p = [0, 1, -2]$. Moreover, it is also possible that two small negative branches meet at ρ: This is e.g. the case for $\mathcal{S} = \{-2, 1\}$ and $p = [1, -2, 1, -2]$.

Table 2 Plot of the real branches of the kernel equation $K(t, u) = 0$, for several pattern p. This illustrates the diversity of behaviours. In all the examples, the set of jumps is $S = \{-2, -1, 0, 1, 2\}$, and the pattern p is indicated. Note that due to a theorem of Pólya–Fatou–Carlson [14] on pure algebraic functions with integer coefficients (and therefore for generic walks), the first crossing between a small and large branch is at $0 < \rho < 1$ (*i.e.* $\rho = 1$ or any other root of $t - t^{\ell}$, cannot be the dominant singularity).

$p = [1, -2, 1, 1, -2, 1]$	$p = [0, 1, 0, 0, 1, 0]$	$p = [0, 0, 1, 2, 0, 0]$
$p = [0, 0, -1, -2, 0, 0]$	$p = [-1, -2, -1, -2, -1, -2]$	$p = [-1, -2, -1, -1, -2, -1]$
$p = [-1, -1, 0, -1, -1, 0]$	$p = [-2, -2, 0, -2, -2, 0]$	$p = [-2, -1, 1, -2, -1, 1]$
$p = [2, 2, -1, 2, 2 - 1]$	$p = [2, -1, -1, 2, -1, -1]$	$p = [2, -1, -1, 2, -1, 1]$
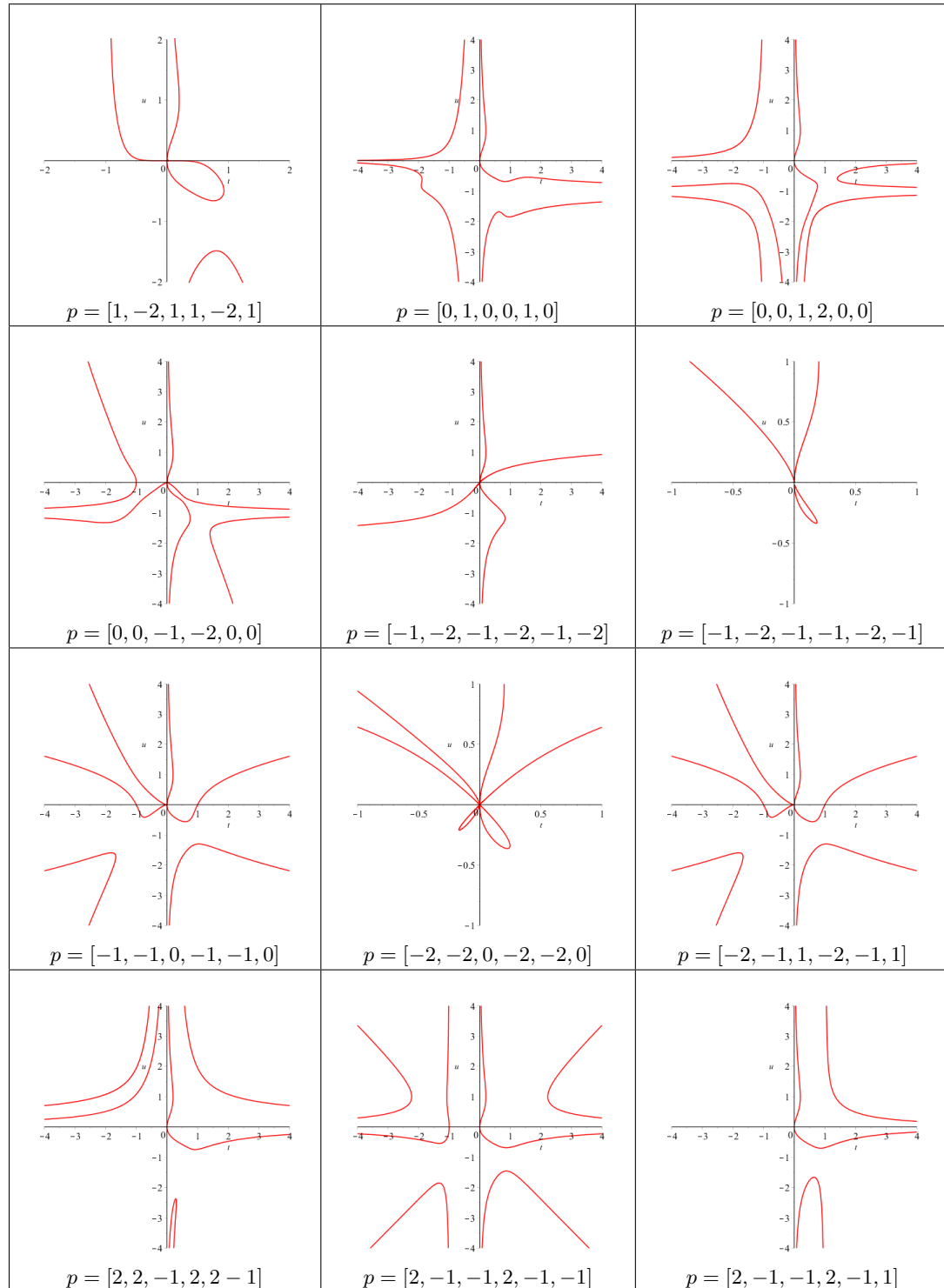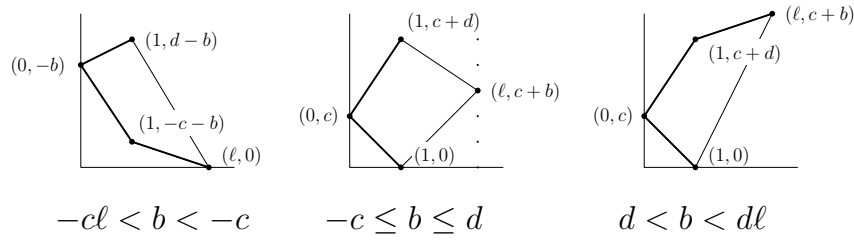
Figure 3 The three possibilities for the Newton polygon of the kernel $K(t, u)$ (the list is exhaustive if $R(t, u) = 1$). Each dot (i, j) corresponds to a monomial $t^i u^j$ of K. The slopes of the convex hull segments on the left give the Puiseux behaviour at 0 of the small and large roots u_i and v_j.

We observe that the behaviour of real branches of $K(t, u) = 0$ is much more complicated and diverse than that in the Banderier–Flajolet study. To recall, in their case there are always exactly two real *positive* branches (one small branch u_1 and one large branch v_1), and they meet at a singularity point $(t, u) = (\rho, \tau)$, where $u = \tau$ is the only positive number such that $P'(\tau) = 0$. In contrast, in our case we may have additional positive branches – even when the autocorrelation is trivial. Table 2 illustrates that we always have a small branch and one large branch whose shape in general resembles that of $u_1 \cup v_1$ from Banderier–Flajolet.

In one sense, the forbidden pattern gives a perturbation of the Banderier–Flajolet geometry of branches, and adds additional branches. A rigorous version of this intuition can be obtained by playing with a Boltzmann weight/Gibbs measure (like in statistical mechanics): moving the parameter v in a continuous way from 1 to 0 in the generating function $F(t, u, v)$ in the next section gives the explanation of these phenomena.

More information about these branches (on their Puiseux expansions) can be derived from the Newton polygon associated with the kernel (see [18] for a crisp presentation of the theory of Newton polygons for Puiseux expansion).

Equipped with all this information on the roots, and the way they cross, we can derive the following asymptotic results. Note that we use the notations $K_t(t, u)$ for $(\partial_t K)(t, u)$, and $K_{uu}(t, u)$ for $(\partial_u^2 K)(t, u)$. We start with asymptotics of walks on $\mathbb{Z}$ with a forbidden pattern.

▶ **Theorem 4** (Asymptotics of walks on $\mathbb{Z}$). *Let ρ_K be the smallest positive root of $K(t, 1)$. For any generic model, the asymptotic number of walks of length n is:*

$$W_n \sim -\rho_K K_t(\rho_K, 1) R(\rho_K, 1) \rho_K^{-n}.$$

Proof. This follows from the partial fraction decomposition of $W(t) = \frac{R(t,1)}{K(t,1)}$. ◀

▶ **Theorem 5** (Asymptotics of excursions). *Assume that we have a generic walk avoiding a pattern p which is a pseudomeander. Let $Y(t) := (-1)^{c-1} u_2(t) \cdots u_c(t)$. The number of excursions of size n satisfies*

$$E_n \sim Y(\rho) \sqrt{\frac{K_t(\rho, 1)}{2\pi \rho K_{uu}(\rho, 1)}} \cdot n^{-3/2} \rho^{-n}.$$

Proof. Since the walk is generic, $Y(t)$ is analytic for $|t| \leq \rho$. Thus the singular behaviour of $u_1(t)$ determines the singularity and the local behaviour of $E(t)$. We obtain:

$$E(t) \sim E(\rho) - Y(\rho) \sqrt{\frac{2K_t(\rho, \tau)}{\rho K_{uu}(\rho, \tau)}} \sqrt{1 - \frac{t}{\rho}}.$$ ◀

▶ **Theorem 6** (Asymptotics of bridges). *Assume that we have a generic walk avoiding a pattern p. The number of bridges of size n satisfies*

$$B_n \sim -\frac{R(\rho, \tau)}{\tau K_t(\rho, \tau)} \sqrt{\frac{K_t(\rho, 1)}{2\pi \rho K_{uu}(\rho, 1)}} \cdot n^{-1/2} \rho^{-n}.$$

Proof. We know from Lemma 2 that $B(t)$ and $E(t)$ have the same radius of convergence. Thus, the singular behaviour of $u_1(t)$ determines the singularity and the local behaviour of $B(t)$. We have therefore

$$B(t) \sim -\frac{R(t, u_1(t))}{K_t(t, u_1(t))} \frac{u_1'(t)}{u_1(t)}$$

and plugging the singular expansion of u_1 into this formula yields the result. ◀

▶ **Theorem 7** (Asymptotics of meanders). *Let ρ_K be the smallest positive root of $K(t, 1)$ (as in Theorem 4). Assume that the walk is generic and that p is a pseudomeander. Then the asymptotics of the coefficients of the meander generating function*

$$M(z) = (1 - u_1(t))Y(t)R(t, 1)/K(t, 1) \qquad with \ Y(t) := \prod_{i=2}^{c}(1 - u_i(t))$$

is given by

$$M_n \sim R(\rho, 1)Y(\rho)\sqrt{\frac{2}{\pi \rho K_t(\rho, 1)K_{uu}(\rho, 1)}} \cdot n^{-1/2}\rho^{-n} \qquad (for\ \rho_K = \rho),$$

$$M_n \sim -\frac{Y(\rho_K)R(\rho_K, 1)}{\rho_K K_t(\rho_K, 1)} \cdot \rho_K^{-n} \qquad (for\ \rho_K < \rho),$$

$$M_n \sim \frac{R(\rho, 1)Y(\rho)}{K(\rho, 1)}\sqrt{\frac{\rho K_t(\rho, 1)}{2\pi K_{uu}(\rho, 1)}} \cdot n^{-3/2}\rho^{-n} \qquad (for\ \rho_K > \rho).$$

Proof. To prove the first assertion, observe that $\rho_K = \rho$ is equivalent to $\tau = 1$. The dominant singularity of the generating function $M(t) = (1 - u_1(t))Y(t)R(t, 1)/K(t, 1)$ is at $\rho_K = \rho$ and it originates from a simple zero in the denominator $K(t, u)$ and from u_1. The singular expansion from $u_1(t)$ at ρ gives (we use $\kappa(t) := -\frac{1}{tK_t(t,1)}$):

$$M(t) \sim R(\rho, 1)Y(\rho)\kappa(\rho)\sqrt{\frac{2\rho K_t(\rho, 1)}{K_{uu}(\rho, 1)}}\left(1 - \frac{t}{\rho}\right)^{-1/2} = \frac{R(\rho, 1)Y(\rho)\sqrt{2}}{\sqrt{\rho K_t(\rho, 1)K_{uu}(\rho, 1)}}\left(1 - \frac{t}{\rho}\right)^{-1/2}.$$

In the case $\rho_K < \rho$ we have $\tau \neq 1$ and thus $K(\rho, 1) > 0$. Hence the generating function has the dominant singularity ρ_K which comes from the kernel only. This implies

$$M(t) \sim Y(\rho_K)R(\rho_K, 1)\kappa(\rho_K)\frac{1}{1 - t/\rho_K}.$$

In the last case, $\rho_K > \rho$, u_1 has a square-root type singularity before $K(t, 1)$ becomes singular. Singularity analysis thus gives the last claim of the theorem, via the following Puiseux expansion at the dominant singularity ρ

$$M(t) \sim M(\rho) + \frac{R(\rho, 1)Y(\rho)}{K(\rho, 1)}\sqrt{\frac{2\rho K_t(\rho, 1)}{K_{uu}(\rho, 1)}}\sqrt{1 - \frac{t}{\rho}}. \qquad ◀$$

Caveat: We are aware that several constants in these theorems can be further simplified, but we kept them like this in order to help the reader to follow the proofs (just sketched here, due to the page limit).

The theorems above for excursions and meanders are stated when the pattern p is a pseudomeander; there is a similar result for any pattern, but the proof goes through a wider disjunction of cases to handle, as then the closed form for the generating function is no more the same. We will handle this in the full version of this article.

These asymptotics also allow to get results on limit laws, as presented in the next section.

5   Limit law for the number of occurrences of a given pattern

Our approach also allows to count the number of occurrences of a pattern in paths. As usual, an occurrence of p in w is any substring of w that coincides with p, and when we count them we do not require that the occurrences will be disjoint. For example, the number of occurrences of 11 in 1111 is 3. We use the same notations than in Section 3. Then one has

▶ **Theorem 8** (Gaussian limit laws for occurrences). *Let X_n be the random variable which counts the number of occurrences of a pattern in a generic walk, bridge, meander, excursion model. Then X_n has a Gaussian limiting distribution with $\mathbb{E}[X_n] = \mu n + O(1)$ and $\mathrm{Var}[X_n] = \sigma^2 n + O(1)$ for some constants $\mu > 0$ and $\sigma^2 \geq 0$:*

$$\frac{1}{\sqrt{n}}\left(X_n - \mathbb{E}[X_n]\right) \to \mathcal{N}(0, \sigma^2).$$

Proof (sketch). The proof relies on the Gaussian limit laws for positive algebraic systems from [4, Theorem 9], which itself comes from following the dependency in the graph associated with the system, and applying Hwang's quasi-power theorem to each component. In this process, some positive variance conditions have to be checked on the formulas given by an equivalent of Theorem 1, with the additional variable v counting the number of occurrences of the pattern, and where the corresponding trivariate kernel is

$$K(t, u, v) := \det(I - tA) = (1 - v)((1 - tP(u))R(t, u) + t^\ell u^b) + v(1 - tP(u)). \qquad (7)$$

This comes from the associated automaton (as illustrated in Figure 4), and its adjacency matrix A. Note that for $v = 0$ we get the kernel from the avoidance case (see equation (2)), and for $v = 1$ we get $1 - tP$ (which is, as expected, the kernel from [5]).

To show the relation (7), we use a method adapted from [22, p. 60]. Let $W \equiv W(t, u, v)$ and $W_p \equiv W_p(t, u, v)$ be the generating functions of all words and words ending with p, respectively, where v counts the number of occurrences of p. We show the following two identities:

$$1 + WtP = W - W_p + v^{-1}W_p, \qquad (8)$$

$$Wt^\ell u^b = v^{-1}W_p R - (R - 1)W_p. \qquad (9)$$

This system is readily solved and gives W as a rational function with denominator the right-hand side of (7). Since it is an irreducible polynomial, with degree ℓ in t, this denominator times a polynomial factor $Q(t, u, v)$ has to be equal to $\det(I - tA)$. In fact, $Q(t, u, v) = 1$. Indeed, an inspection of the degrees of the product shows that they cannot be higher than the degrees of the determinant of $I - tA$, and multiplying the denominator by a non constant Q would contradict this. Now, setting $v = 1$ gives that $1 - tP = \det(I - tA) = (1 - tP)Q(t, u, 1)$ and thus $Q = 1$. This shows (7).
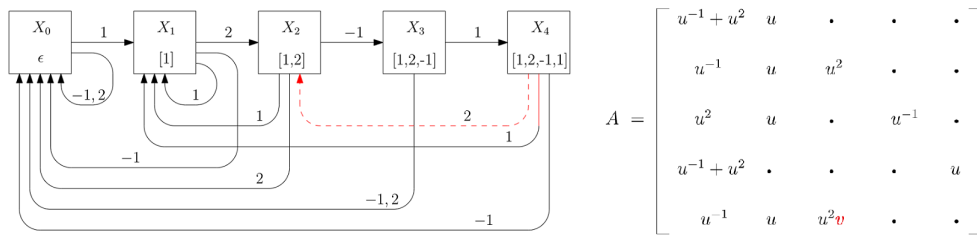
Figure 4 Pushdown automaton for the set of jumps $\mathcal{S} = \{-1, 1, 2\}$ and the pattern $p = [1, 2, -1, 1, 2]$. In dashed red we marked the arrow from the last state $(X_{\ell-1})$ labelled by the last letter of the pattern (a_ℓ). Marking this transition with v leads to formulas involving the kernel $K(t, u, v) = \det(I - tA)$ as given in Equation (7), where A is the adjacency matrix of this automaton.

To show (8), take a word and add a letter to it. If the resulting word does not end with p, it is counted by $W - W_p$; if it does, it is counted by $v^{-1}W_p$. To show (9), take a word w and add the pattern p to it. This creates a number $j \geq 1$ of new occurrences of p. The path wp can be written in j ways as $w'r$, where w' ends with a new occurence of p and r is an autocorrelation factor, or $j - 1$ ways if we impose that $r \neq \varepsilon$. It is therefore counted with a factor $v + \cdots + v^j$ by $W_p R$ and with a factor $v + \cdots + v^{j-1}$ by $(R - 1)W_p$, and the result follows. ◀

6 Conclusion

In this article, we presented a unifying way which gives the asymptotics of all families of lattice paths with a forbidden pattern, and we proved that the number of occurrences of a given pattern is normally distributed. The same approach would, for instance, allow to do the asymptotics of walks having exactly m occurrences of a given pattern, or to consider patterns which are no longer a word but a regular expression.

It is also nice that our approach gives a *method* (let us call it the *vectorial kernel method*) to solve in an efficient way the question of the enumeration and asymptotics of words generated by a pushdown automaton (or words belonging to the intersection of an algebraic language and a rational language). What is more, it is possible to use our functional equation approach to analyse the intersection of two algebraic languages. Note that testing if this intersection is empty is known to be an undecidable problem, even for deterministic context-free grammars (see e.g. [24]), so we cannot expect too much from a generic method in this case. However, we can specify a little bit more the type of system of functional equations we get: indeed this problem is related to automata with two stacks, which, in turn, are known to have the same power as a Turing machine; the evolution of these two stacks corresponds to lattice paths in the quarter plane (with steps of arbitrary length), the complexity of the problem is reflected by the fact that one can then get generating functions which are no more algebraic, D-finite, or differentially-algebraic, and we do not expect some universal nice results here, but a wider zoo of behaviours.

However, no doubt that all these cases will be new instances of what Flajolet and Sedgewick called *Borges's Theorem*: Any pattern which is not forbidden by design will appear a linear number of times in large enough structures, with Gaussian fluctuations.

For sure, it is more a metatheorem, a natural credo, so it is always worthwhile to establish this claim rigorously. Naturally, may it be with tools of probability theory or of analytic combinatorics, there is always some technical conditions to check to ensure this claim. In

this article, our closed form expressions for the generating functions were one of the keys, together with the universal behaviour of the small branches. This allowed us to prove this Gaussian behaviour for the number of occurrences of any given pattern. Year after year, this claim is established for more and more combinatorial structures (it was done for patterns in Markov chains, trees, maps, permutations, context-free grammars, and now... lattice paths!).

Let us end with the passage of Flajolet and Sedgewick [22, p. 61] which explains where Borges's Theorem comes from:

> This property is sometimes called "Borges's Theorem" as a tribute to the famous Argentinian writer Jorge Luis Borges (1899-1986) who, in his essay *The Library of Babel*, describes a library so huge as to contain:
>
> "Everything: the minutely detailed history of the future, the archangels' autobiographies, the faithful catalogues of the Library, thousands and thousands of false catalogues, the demonstration of the fallacy of those catalogues, the demonstration of the fallacy of the true catalogue, the Gnostic gospel of Basilides, the commentary on that gospel, the commentary on the commentary on that gospel, the true story of your death, the translation of every book in all languages, the interpolations of every book in all books."

References

1 Andrei Asinowski, Axel Bacher, Cyril Banderier, and Bernhard Gittenberger. Analytic combinatorics of lattice paths with forbidden patterns: enumerative aspects. In S. Klein, C. Martín-Vide, and D. Shapira, editors, *Language and Automata Theory and Applications. LATA 2018*, volume 10782 of *Lecture Notes in Computer Science*, pages 195–206. Springer, 2018. `doi:10.1007/978-3-319-77313-1_15`.

2 Axel Bacher and Mireille Bousquet-Mélou. Weakly directed self-avoiding walks. *Journal of Combinatorial Theory, Series A*, 118(8):2365–2391, 2011. `doi:10.1016/j.jcta.2011.06.001`.

3 Cyril Banderier. *Combinatoire analytique des chemins et des cartes*. Ph.D. thesis, Université Paris VI, 2001. URL: `https://lipn.univ-paris13.fr/~banderier/Papers/these.pdf`.

4 Cyril Banderier and Michael Drmota. Formulae and asymptotics for coefficients of algebraic functions. *Combinatorics, Probability and Computing*, 24(1):1–53, 2015. `doi:10.1017/S0963548314000728`.

5 Cyril Banderier and Philippe Flajolet. Basic analytic combinatorics of directed lattice paths. *Theoretical Computer Science*, 281(1-2):37–80, 2002. `doi:10.1016/S0304-3975(02)00007-5`.

6 Cyril Banderier and Bernhard Gittenberger. Analytic combinatorics of lattice paths: enumeration and asymptotics for the area. *Discrete Math. Theor. Comput. Sci. Proc.*, AG:345–355, 2006. URL: `https://hal.inria.fr/hal-01184683/document`.

7 Cyril Banderier and Pierre Nicodème. Bounded discrete walks. *Discrete Math. Theor. Comput. Sci.*, AM:35–48, 2010. URL: `https://dmtcs.episciences.org/2792/pdf`.

8 Cyril Banderier and Michael Wallner. The kernel method for lattice paths below a rational slope. In *Lattice paths combinatorics and applications*, Developments in Mathematics Series. Springer, 2018. To appear. URL: `https://arxiv.org/abs/1606.08412`.

9   Jean-Luc Baril and Armen Petrossian. Equivalence classes of Motzkin paths modulo a pattern of length at most two. *J. Integer Seq.*, 18(11):Article 15.7.1, 2015. URL: `https://cs.uwaterloo.ca/journals/JIS/VOL18/Baril/baril3.pdf`.

10   Antonio Bernini, Luca Ferrari, Renzo Pinzani, and Julian West. Pattern-avoiding Dyck paths. *Discrete Math. Theor. Comput. Sci. Proc.*, pages 683–694, 2013. URL: `https://dmtcs.episciences.org/2334/pdf`.

11   Mireille Bousquet-Mélou. Rational and algebraic series in combinatorial enumeration. In *International Congress of Mathematicians. Vol. III*, pages 789–826. EMS, 2006. URL: `https://hal.archives-ouvertes.fr/file/index/docid/277255/filename/icm.pdf`.

12   Mireille Bousquet-Mélou and Arnaud Jehanne. Polynomial equations with one catalytic variable, algebraic series and map enumeration. *Journal of Combinatorial Theory. Series B*, 96(5):623–672, 2006. `doi:10.1016/j.jctb.2005.12.003`.

13   Charlotte Brennan and Simon Mavhungu. Visits to level r by Dyck paths. *Fund. Inform.*, 117(1-4):127–145, 2012. URL: [online pdf].

14   Fritz Carlson. Über Potenzreihen mit ganzzahligen Koeffizienten. *Math. Z.*, 9(1-2):1–13, 1921. `doi:10.1007/BF01378331`.

15   Noam Chomsky and Marcel-Paul Schützenberger. The algebraic theory of context-free languages. In *Computer programming and formal systems*, pages 118–161. North-Holland, Amsterdam, 1963.

16   Nachum Dershowitz and Shmuel Zaks. More patterns in trees: up and down, young and old, odd and even. *SIAM J. Discrete Math.*, 23(1):447–465, 2008/09. `doi:10.1137/070687475`.

17   Emeric Deutsch and Louis W. Shapiro. A bijection between ordered trees and 2-Motzkin paths and its many consequences. *Discrete Math.*, 256(3):655–670, 2002. `doi:10.1016/S0012-365X(02)00341-2`.

18   Jean Dieudonné. *Calcul infinitésimal*. Hermann, Paris, 2 edition, 1980. 1st ed. in 1968: 479pp., there is also an English translation of the 1st ed. in 1971, 427pp.

19   Yun Ding and Rosena R. X. Du. Counting humps in Motzkin paths. *Discrete Appl. Math.*, 160(1-2):187–191, 2012. `doi:10.1016/j.dam.2011.08.018`.

20   Philippe Duchon. On the enumeration and generation of generalized Dyck words. *Discrete Mathematics*, 225(1-3):121–135, 2000. `doi:10.1016/S0012-365X(00)00150-3`.

21   Sen-Peng Eu, Shu-Chung Liu, and Yeong-Nan Yeh. Dyck paths with peaks avoiding or restricted to a given set. *Stud. Appl. Math.*, 111(4):453–465, 2003. `doi:10.1111/1467-9590.t01-1-00042`.

22   Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009. URL: `http://algo.inria.fr/flajolet/Publications/book.pdf`.

23   Leonidas J. Guibas and Andrew M. Odlyzko. String overlaps, pattern matching, and nontransitive games. *J. Combin. Theory Ser. A*, 30(2):183–208, 1981. `doi:10.1016/0097-3165(81)90005-4`.

24   John E. Hopcroft, Raghavan Motwani, and Jeffrey D. Ullman. *Introduction to automata theory, languages, and computation*. Pearson, 3rd edition, 2006. (First edition by Addison-Wesley in 1979). URL: `http://infolab.stanford.edu/~ullman/ialc.html`.

25   Jacques Labelle and Yeong Nan Yeh. Generalized Dyck paths. *Discrete Mathematics*, 82(1):1–6, 1990. `doi:10.1016/0012-365X(90)90039-K`.

26   Kostas Manes, Aristidis Sapounakis, Ioannis Tasoulas, and Panagiotis Tsikouras. Strings of length 3 in Grand-Dyck paths and the Chung–Feller property. *Electron. J. Combin.*, 19(2):Paper 2, 10, 2012. URL: `http://www.combinatorics.org/ojs/index.php/eljc/article/view/v19i2p2/pdf`.

27   Toufik Mansour and Mark Shattuck. Counting humps and peaks in generalized Motzkin paths. *Discrete Appl. Math.*, 161(13-14):2213–2216, 2013. `doi:10.1016/j.dam.2013.03.007`.

28   Donatella Merlini, D. G. Rogers, R. Sprugnoli, and M. Cecilia Verri. Underdiagonal lattice paths with unrestricted steps. *Discrete Appl. Math.*, 91(1-3):197–213, 1999. `doi:10.1016/S0166-218X(98)00126-7`.

29   Youngja Park and Seung Kyung Park. Enumeration of generalized lattice paths by string types, peaks, and ascents. *Discrete Math.*, 339(11):2652–2659, 2016. `doi:10.1016/j.disc.2016.04.024`.

30   Marcel-Paul Schützenberger. On context-free languages and push-down automata. *Information and Control*, 6:246–264, 1963.

31   Marcel-Paul Schützenberger. On the synchronizing properties of certain prefix codes. *Information and Control*, 7:23–36, 1964. URL: `https://www.sciencedirect.com/science/article/pii/S0019995864902323/pdf`.

32   Richard P. Stanley. *Enumerative combinatorics. Vol. 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2nd edition, 2011.

Periodic Pólya Urns and an Application to Young Tableaux

Cyril Banderier
Université Paris 13, LIPN, UMR CNRS 7030, `http://lipn.univ-paris13.fr/~banderier`
 `https://orcid.org/0000-0003-0755-3022`

Philippe Marchal
Université Paris 13, LAGA, UMR CNRS 7539, `https://math.univ-paris13.fr/~marchal`
 `https://orcid.org/0000-0001-8236-5713`

Michael Wallner
Université de Bordeaux, LaBRI, UMR CNRS 5800, `http://dmg.tuwien.ac.at/mwallner`
 `https://orcid.org/0000-0001-8581-449X`

──── **Abstract** ────

Pólya urns are urns where at each unit of time a ball is drawn and is replaced with some other balls according to its colour. We introduce a more general model: The replacement rule depends on the colour of the drawn ball *and* the value of the time $(\bmod\, p)$. We discuss some intriguing properties of the differential operators associated to the generating functions encoding the evolution of these urns. The initial *non-linear* partial differential equation indeed leads to *linear* differential equations and we prove that the moment generating functions are D-finite. For a subclass, we exhibit a closed form for the corresponding generating functions (giving the exact state of the urns at time n). When the time goes to infinity, we show that these *periodic Pólya urns* follow a rich variety of behaviours: their asymptotic fluctuations are described by a family of distributions, the generalized Gamma distributions, which can also be seen as powers of Gamma distributions. En passant, we establish some enumerative links with other combinatorial objects, and we give an application for a new result on the asymptotics of Young tableaux: This approach allows us to prove that the law of the lower right corner in a triangular Young tableau follows asymptotically a product of generalized Gamma distributions.

1 Periodic Pólya urns

Pólya urns were introduced in a simplified version by George Pólya and his PhD student Florian Eggenberger in [7, 8, 27], with applications to disease spreading and conflagrations. They constitute a powerful model, still widely used: see e.g. Rivest's recent work on auditing elections [28], or the analysis of deanonymization in Bitcoin's peer-to-peer network [9]. They are well-studied objects in combinatorial and probabilistic literature [2, 11, 22], and offer fascinatingly rich links with numerous objects like random recursive trees, m-ary search trees, branching random walks (see e.g. [3, 6, 15, 16, 30]). In this paper we introduce a variation which offers new links with another important combinatorial structure: Young tableaux. We solve the enumeration problem of this new model, derive the limit law for the evolution of the urn, and give some applications.

In the *Pólya urn model*, one starts with an urn with b_0 black balls and w_0 white balls at time 0. At every discrete time step one ball is drawn uniformly at random. After inspecting its colour it is returned to the urn. If the ball is black, a black balls and b white balls are added; if the ball is white, c black balls and d white balls are added (where $a, b, c, d \in \mathbb{N}$ are non-negative integers). This process can be described by the so-called *replacement matrix*:

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad a, b, c, d \in \mathbb{N}.$$

We call an urn and its associated replacement matrix *balanced* if $K := a + b = c + d$. In other words, in every step the same number K of balls is added to the urn. This results in a deterministic number of balls after n steps: $b_0 + w_0 + Kn$ balls.

Now, we introduce a more general model which has rich combinatorial, probabilistic, and analytic properties.

▶ **Definition 1.** A *periodic Pólya urn* of period p with replacement matrices $M_1, M_2, \ldots, M_p$ is a variant of a Pólya urn in which the replacement matrix M_k is used at steps $np + k$. Such a model is called *balanced* if each of its replacement matrices is balanced.

In this article, we illustrate the aforementioned rich properties on the following model (the results for other values of the parameters are similar to the case we now handle in detail).

▶ **Definition 2.** We call a *Young–Pólya urn* the periodic Pólya urn of period 2 with replacement matrices $M_1 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ for every odd step, and $M_2 := \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ for every even step.

Let us describe the state of the urn after n steps by pairs (number of black balls, number of white balls), starting with $b_0 = 1$ black ball and $w_0 = 1$ white ball shown in Figure 1. In the first step the matrix M_1 is used and gives the two states $(2, 1)$, and $(1, 2)$. In the second step, matrix M_2 is used, in the third step, matrix M_1 is used again, in the fourth step, matrix M_2, etc. Thus, the possible states are $(3, 2), (2, 3)$, and $(1, 4)$, at time 2, and $(4, 2), (3, 3), (2, 4)$, and $(1, 5)$, at time 3.

In fact, each of these states may be reached in different ways, and such a sequence of transitions is called a *history*. Each history comes with weight one. Implicitly, they induce a probability measure on the states at step n. So, let B_n and W_n be random variables for the number of black and white balls after n steps, respectively. As our model is balanced, $B_n + W_n$ is a deterministic process, reflecting the identity $B_n + W_n = b_0 + w_0 + n + \lfloor \frac{n}{2} \rfloor$. So, from now on, we concentrate our analysis on B_n.

M_1

$H_0 = xy$

M_2

$H_1 = x^2y + xy^2$

M_1

$H_2 = 2x^3y^2 + 2x^2y^3 + 2xy^4$

$H_3 = 6x^4y^2 + 8x^3y^3 + 8x^2y^4 + 8xy^5$
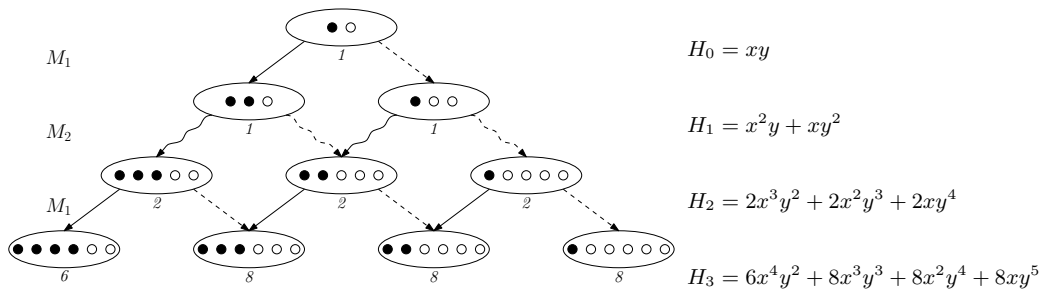
Figure 1 The evolution of a Young–Pólya urn with one initial black and one initial white ball. Black arrows mark that a black ball was drawn, dashed arrows mark that a white ball was drawn. Straight arrows indicate that the replacement matrix M_1 was used, curly arrows show that the replacement matrix M_2 was used. The number below each node is the number of possible transitions to reach such a state. In this article we give a formula for H_n (which encodes all the possible states of the urn at time n) and their asymptotic behaviour.

For the classical model of a single balanced Pólya urn, the limit law of the random variable B_n is fully known: The possible limit laws include a rich variety of distributions. To name a few, let us mention the uniform distribution [10], the normal distribution [3], and the Beta and Mittag-Leffler distributions [15]. Periodic Pólya urns (which include the classical model) lead to an even larger variety of distributions involving a product of *generalized Gamma distributions* [31].

▶ **Definition 3.** The generalized Gamma distribution $\mathrm{GenGamma}(\alpha, \beta)$ with real parameters $\alpha, \beta > 0$ is defined by the density function (having support $(0, +\infty)$)

$$f(x; \alpha, \beta) := \frac{\beta\, x^{\alpha-1} \exp(-x^\beta)}{\Gamma\left(\alpha/\beta\right)},$$

where Γ is the classical Gamma function $\Gamma(z) := \int_0^\infty t^{z-1} \exp(-t)\, dt$.

▶ Remark. Let $\boldsymbol{\Gamma}(\alpha)$ be the Gamma distribution[1] of parameter $\alpha > 0$, given by its density

$$g(x; \alpha) = \frac{x^{\alpha-1} \exp(-x)}{\Gamma(\alpha)}.$$

Then, one has $\boldsymbol{\Gamma}(\alpha) \overset{\mathcal{L}}{=} \mathrm{GenGamma}(\alpha, 1)$ and, for $r > 0$, the distribution of the r-th power of a random variable distributed according to $\boldsymbol{\Gamma}(\alpha)$ is $\boldsymbol{\Gamma}(\alpha)^r \overset{\mathcal{L}}{=} \mathrm{GenGamma}(\alpha/r, 1/r)$.

Our main results are the enumeration result from Theorem 5, the application to Young tableaux in Theorem 7, and the following result (and its generalization in Theorem 6):

▶ **Theorem 4.** *The normalized random variable $\frac{2^{2/3}}{3} \frac{B_n}{n^{2/3}}$ of the number of black balls in a Young–Pólya urn converges in law to a generalized Gamma distribution:*

$$\frac{2^{2/3}}{3} \frac{B_n}{n^{2/3}} \quad \overset{\mathcal{L}}{\longrightarrow} \quad \mathrm{GenGamma}\left(1, 3\right).$$

---

[1] Caveat: It is traditional to use the same letter for both the Γ function and the $\boldsymbol{\Gamma}$ distribution. Also, some authors add a second parameter to the distribution $\boldsymbol{\Gamma}$, which is set to 1 here.

We give a proof of this result in Section 3. Let us first mention some articles where this distribution has already appeared before:

- in Janson [17], for the analysis of the area of the supremum process of the Brownian motion,
- in Peköz, Röllin, and Ross [25], as distributions of processes on walks, trees, urns, and preferential attachments in graphs (they also consider what they call a Pólya urn with immigration, which is a special case of a periodic Pólya urn),
- in Khodabin and Ahmadabadi [19] following a tradition to generalize special functions by adding parameters in order to capture several probability distributions, such as e.g. the normal, Rayleigh, and half-normal distribution, as well as the MeijerG function (see also the addendum of [17], mentioning a dozen of other generalizations of special functions).

In the next section we translate the evolution process into the language of generating functions by encoding the dynamics of this process into partial differential equations.

2   A functional equation for periodic Pólya urns

Let $h_{n,k,\ell}$ be the number of histories of a periodic Pólya urn after n steps with k black balls and ℓ white balls, with an initial state of b_0 black balls and w_0 white balls, and with replacement matrices M_1 for the odd steps and M_2 for the even steps. We define the polynomials

$$H_n(x,y) := \sum_{k,\ell \geq 0} h_{n,k,\ell} x^k y^\ell.$$

Note that these are indeed polynomials as there are just a finite number of histories after n steps. We collect all these histories in the trivariate exponential generating function

$$H(x,y,z) := \sum_{n \geq 0} H_n(x,y) \frac{z^n}{n!}.$$

In particular, we get for the first 3 terms of $H(x,y,z)$ the expansion (compare Figure 1)

$$H(x,y,z) = xy + \left(xy^2 + x^2y\right) z + \left(2xy^4 + 2x^2y^3 + 2x^3y^2\right) \frac{z^2}{2} + \ldots$$

Observe that the polynomials $H_n(x,y)$ are homogeneous, as we have a balanced urn model.

Now it is our goal to derive a partial differential equation describing the evolution of the periodic Pólya urn model. For a comprehensive introduction to the method we refer to [10].

In order to capture the periodic behaviour we split the generating function $H(x,y,z)$ into odd and even steps. We define

$$H_e(x,y,z) := \sum_{n \geq 0} H_{2n}(x,y) \frac{z^{2n}}{(2n)!} \qquad \text{and} \qquad H_o(x,y,z) := \sum_{n \geq 0} H_{2n+1}(x,y) \frac{z^{2n+1}}{(2n+1)!},$$

such that $H(x,y,z) = H_e(x,y,z) + H_o(x,y,z)$. Next, we associate to the replacement matrices M_1 and M_2 from Definition 2 the differential operators $\mathcal{D}_1$ and $\mathcal{D}_2$, respectively. We get

$$\mathcal{D}_1 := x^2 \partial_x + y^2 \partial_y \qquad \text{and} \qquad \mathcal{D}_2 := x^2 y \partial x + y^3 \partial_y,$$

where ∂_x and ∂_y are defined as the partial derivatives $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$, respectively. These model the evolution of the urn. For example, in the term $x^2 \partial_x$, the derivative ∂_x represents drawing

a black ball and the multiplication by x^2 returning the black ball and an additional black ball into the urn. The other terms have analogous interpretations.

With these operators we are able to link odd and even steps with the following system

$$\partial_z H_o(x, y, z) = \mathcal{D}_1 H_e(x, y, z) \qquad \text{and} \qquad \partial_z H_e(x, y, z) = \mathcal{D}_2 H_o(x, y, z). \tag{1}$$

Note that the derivative ∂_z models the evolution in time. This system of partial differential equations naturally corresponds to recurrences at the level of coefficients $h_{n,k,\ell}$, and *vice versa*. This philosophy is well explained in the *symbolic method* part of [12] and see also [10].

As a next step we want to eliminate the y variable in these equations. This is possible as the number of balls in each round and the number of black and white balls are connected due to the fact that we are dealing with balanced urns. First, as observed previously, one has

$$\text{number of balls after } n \text{ steps} = b_0 + w_0 + n + \left\lfloor \frac{n}{2} \right\rfloor. \tag{2}$$

Therefore, for any $x^k y^\ell z^n$ appearing in $H(x, y, z)$ with $b_0 = w_0 = 1$ we have

$$k + \ell = 2 + \frac{3n}{2} \quad \text{(if } n \text{ is even)} \qquad \text{and} \qquad k + \ell = 2 + \frac{3n}{2} - \frac{1}{2} \text{ (if } n \text{ is odd)}.$$

This translates directly into

$$\begin{cases} x\partial_x H_e(x, y, z) + y\partial_y H_e(x, y, z) = 2H_e(x, y, z) + \dfrac{3}{2}z\partial_z H_e(x, y, z), \\ x\partial_x H_o(x, y, z) + y\partial_y H_o(x, y, z) = \dfrac{3}{2}H_o(x, y, z) + \dfrac{3}{2}z\partial_z H_o(x, y, z). \end{cases} \tag{3}$$

Finally, combining (1) and (3), we eliminate $\partial_y H_e$ and $\partial_y H_o$. After that it is legitimate to insert $y = 1$ as there appears no differentiation with respect to y anymore. As the urns are balanced, the exponents of y and x in each monomial are bound (see Equation (2)), so we are losing no information on the trivariate generating functions by setting $y = 1$. Hence, from now on we use the notation $H(x, z)$, $H_e(x, z)$, and $H_o(x, z)$ instead of $H(x, 1, z)$, $H_e(x, 1, z)$, and $H_o(x, 1, z)$, respectively. All of this leads to our first main enumeration theorem:

▶ **Theorem 5** (Linear differential equations and hypergeometric expressions for histories)**.** *The generating functions describing the 2-periodic Young–Pólya urn at even and odd time satisfy the following system of differential equations:*

$$\begin{cases} \partial_z H_e(x, z) = x(x - 1)\partial_x H_o(x, z) + \dfrac{3}{2}z\partial_z H_o(x, z) + \dfrac{3}{2}H_o(x, z), \\ \partial_z H_o(x, z) = x(x - 1)\partial_x H_e(x, z) + \dfrac{3}{2}z\partial_z H_e(x, z) + 2H_e(x, z). \end{cases} \tag{4}$$

Moreover, all these functions satisfy linear *differential equations (they are D-finite, see e.g. [12, Appendix B.4] for more on this notion), which in return implies that $H = H_e + H_o$ satisfies the equation $L.H(x, z) = 0$, where L is a differential operator of order 3 in ∂_z, and then one has the hypergeometric closed forms for $h_n := [z^n]H(1, z)$:*

$$h_n = \begin{cases} 3^n \dfrac{\Gamma\left(\frac{n}{2}+1\right)\Gamma\left(\frac{n}{2}+\frac{2}{3}\right)}{\Gamma(2/3)} & \text{if } n \text{ is even,} \\ 3^n \dfrac{\Gamma\left(\frac{n}{2}+1/2\right)\Gamma\left(\frac{n}{2}+7/6\right)}{\Gamma(2/3)} & \text{if } n \text{ is odd.} \end{cases} \tag{5}$$

Alternatively, this sequence satisfies $h(n + 2) = \frac{2}{3}h(n + 1) + \frac{1}{4}(9n^2 + 21n + 12)h(n)$. This sequence is not found in the OEIS[2], we added it there, it is now `A293653`, and it starts like this: $1, 2, 6, 30, 180, 1440, 12960, 142560, 1710720, 23950080, 359251200, \ldots$

---

[2] On-Line Encyclopedia of Integer Sequences, https://oeis.org.

In the next section we will use Equations (4) to iteratively derive the moments of the distribution of black balls after n steps.

3   Moments of periodic Pólya urns

In this section, we give a proof via the method of moments of Theorem 4 stated in the introduction. Let $m_r(n)$ be the r-th factorial moment of the distribution of black balls after n steps, i.e.

$$m_r(n) := \mathbb{E}\left(B_n(B_n - 1)\cdots(B_n - r + 1)\right).$$

Expressing them in terms of the generating function $H(x, z)$, it holds that

$$m_r(n) = \frac{[z^n]\left.\frac{\partial^r}{\partial x^r}H(x, z)\right|_{x=1}}{[z^n]H(1, z)}.$$

Splitting them into odd and even moments, we have access to these quantities via the partial differential equation (4). As a first step we compute $h_n := [z^n]H(1, z)$, the total number of histories after n steps. We substitute $x = 1$, which makes the equation independent of the derivative with respect to x. Then, the idea is to transform (4) into two independent differential equations for $H_e(1, z)$ and $H_o(1, z)$. This is achieved by differentiating the equations with respect to z and substituting the other one to eliminate $H_e(1, z)$ or $H_o(1, z)$, respectively. This decouples the system, but increases the degree of differentiation by 1. We get

$$\left(9z^2 - 4\right)\partial_z^2 H_e(1, z) + 39z\partial_z H_e(1, z) + 24H_e(1, z) = 0,$$
$$\left(9z^2 - 4\right)\partial_z^2 H_o(1, z) + 39z\partial_z H_o(1, z) + 21H_o(1, z) = 0.$$

In this case it is easy to extract the underlying recurrence relations and solve them explicitly. This also leads to the closed forms (5) for h_n, from which it is easy to compute the asymptotic number of histories for $n \to \infty$. Interestingly, the first two terms in the asymptotic expansion are the same for odd and even number of steps, only the third ones differ. We get

$$h_n = n!\frac{\sqrt{\pi}}{2^{1/6}\Gamma\left(\frac{2}{3}\right)}\left(\frac{3}{2}\right)^n n^{1/6}\left(1 + O\left(\frac{1}{n}\right)\right).$$

As a next step we compute the mean. Therefore, we differentiate (4) once with respect to x, substitute $x = 1$, decouple the system, derive the recurrence relations of the coefficients, and solve them. Note again that the factor $(x-1)$ prevents higher derivatives from appearing and is therefore crucial for this method. After normalization by h_n we get

$$m_1(n) = \begin{cases} \frac{3^{3/2}\Gamma\left(\frac{2}{3}\right)^2}{2\pi}\frac{\Gamma\left(\frac{n}{2} + \frac{4}{3}\right)}{\Gamma\left(\frac{n}{2} + \frac{2}{3}\right)} & \text{if } n \text{ is even,} \\ \frac{3^{3/2}\Gamma\left(\frac{2}{3}\right)^2}{4\pi}\frac{(n+1)\Gamma\left(\frac{n}{2} + \frac{5}{6}\right)}{\Gamma\left(\frac{n}{2} + \frac{7}{6}\right)} & \text{if } n \text{ is odd.} \end{cases}$$

For the asymptotic mean we discover again the same phenomenon that the first two terms in the asymptotic expansion are equal for odd and even n.

Differentiating (4) to higher orders allows to derive higher moments in a mechanical way (this however requires further details, which will be included in the expanded version of this article). In general we get the closed form for the r-th factorial moment

$$m_r(n) = \frac{3^r}{2^{2r/3}}\frac{\Gamma\left(\frac{r}{3} + \frac{1}{3}\right)}{\Gamma\left(\frac{1}{3}\right)}n^{2r/3}\left(1 + O\left(\frac{1}{n}\right)\right). \tag{6}$$

Therefore we see that the moments $\mathbb{E}\left(B_n^{*r}\right)$ of the rescaled random variable $B_n^* := \frac{2^{2/3}}{3}\frac{B_n}{n^{2/3}}$ converge for n to infinity to the limit

$$m_r := \frac{\Gamma\left(\frac{r}{3}+\frac{1}{3}\right)}{\Gamma\left(\frac{1}{3}\right)}. \tag{7}$$

Note that one has $m_r^{-1/(2r)} = \left(\frac{3e}{r}\right)^{1/6}(1+o(1))$ for large r, so the following sum diverges:

$$\sum_{r>0} m_r^{-1/(2r)} = +\infty. \tag{8}$$

Therefore, a result by Carleman (see [5, pp. 189-220] or [33, p. 330])[3] implies that there exists a unique distribution (let us call it $\mathcal{D}$) with such moments m_r.

Furthermore, by the asymptotic result from Equation (6) there exist an $n_0 > 0$ and constants a_r and b_r independent of n such that $a_r < m_r(n) < b_r$, for all $n \geq n_0$. Thus, by the limit theorem of Fréchet and Shohat [13][4] there exists a limit distribution (which therefore has to be $\mathcal{D}$) to which a *subsequence* of our rescaled random variables B_n^* converge to. And as we know via Carleman's criterion above that $\mathcal{D}$ is uniquely determined by its moments, it is in fact the *full sequence* of B_n^* which converges to $\mathcal{D}$.

Now it is easy to check that if $X \sim \mathrm{GenGamma}(d,p)$ is a generalized Gamma distributed random variable (as defined in Definition 3), then it is a distribution determined by its moments, which are given by $\mathbb{E}(X^r) = \Gamma\left(\frac{d+r}{p}\right)/\Gamma\left(\frac{d}{p}\right)$.

In conclusion, the structure of m_r in Formula (7) implies that the normalized random variable B_n^* of the number of black balls in a Young–Pólya urn converges to $\mathrm{GenGamma}(1,3)$. This completes the proof of Theorem 4. ◀

The same approach allows us to study the distribution of black balls for the urn with replacement matrices $M_1 = M_2 = \cdots = M_{p-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $M_p = \begin{pmatrix} 1 & \ell \\ 0 & 1+\ell \end{pmatrix}$. We call this model the *Young–Pólya urn of period p and parameter ℓ*.

▶ **Theorem 6.** *The renormalized distribution of black balls in the Young–Pólya urn of period p and parameter ℓ is asymptotically a distribution, which we call $\mathrm{ProdGenGamma}(p,\ell,b_0,w_0)$, defined as the following product of independent distributions:*

$$\frac{p^\delta}{p+\ell}\frac{B_n}{n^\delta} \xrightarrow{\mathcal{L}} \mathrm{Beta}(b_0,w_0)\prod_{i=0}^{\ell-1}\mathrm{GenGamma}(b_0+w_0+p+i,p+\ell) \tag{9}$$

with $\delta = p/(p+\ell)$, and where $\mathrm{Beta}(b_0,w_0)$ is as usual the law with support $[0,1]$ and density $\frac{\Gamma(b_0+w_0)}{\Gamma(b_0)\Gamma(w_0)}x^{b_0-1}(1-x)^{w_0-1}$.

Sketch. This follows from the following r-th (factorial) moment computation:

$$\mathbb{E}\left(B_n^r\right) = \frac{(p+\ell)^r}{p^{\delta r}}\frac{\Gamma(b_0+r)\Gamma(b_0+w_0)}{\Gamma(b_0)\Gamma(b_0+w_0+r)}\prod_{i=0}^{\ell-1}\frac{\Gamma\left(\frac{b_0+w_0+p+r+i}{p+\ell}\right)}{\Gamma\left(\frac{b_0+w_0+p+i}{p+\ell}\right)}n^{\delta r}\left(1+O\left(\frac{1}{n}\right)\right),$$

---

[3] Note that there is no typo in Formula 8: if the support of the density is $[0,+\infty[$ the moments in the sum have index r and exponent $-1/(2r)$, while they have index $2r$ and exponent $-1/(2r)$ if the support is $]-\infty,+\infty[$.

[4] As a funny coincidence, Fréchet and Shohat mention in [13] that the generalized Gamma distribution with parameter $p \geq 1/2$ is uniquely characterized by its moments.

which in turn characterizes the ProdGenGamma distribution. Indeed, if for some independent random variables X, Y, Z, one has $\mathbb{E}(X^r) = \mathbb{E}(Y^r)\mathbb{E}(Z^r)$ (and if Y and Z are determined by their moments), then $X \stackrel{\mathcal{L}}{=} YZ$.   ◀

This is consistent with our results on the Young–Pólya urn introduced in Section 1. Indeed, there one has $w_0 = b_0 = 1, p = 2, \ell = 1$, and therefore the renormalized distribution of black balls $\frac{p^\delta}{p+\ell} B_n/n^\delta$ is asymptotically $\mathrm{Unif}(0,1) \cdot \mathrm{GenGamma}(4,3) = \mathrm{GenGamma}(1,3)$.

We will now see what are the implications of this result on an apparently unrelated topic: Young tableaux.

4   Urns, trees, and Young tableaux

As predicted by Anatoly Vershik in [32], the 21st century should see a lot of challenges and advances on the links of probability theory with (algebraic) combinatorics. A key role is played here by Young tableaux[5], because of their ubiquity in representation theory. Many results on their asymptotic shape have been collected, but very few results are known on their asymptotic content when the shape is fixed (see e.g. the works by Pittel and Romik, Angel *et al.*, Marchal [1, 24, 26, 29], who have studied the distribution of the values of the cells in random rectangular or staircase Young tableaux, while the case of Young tableaux with a more general shape seems to be very intricate). It is therefore pleasant that our work on periodic Pólya urns allows us to get advances on the case of a triangular shape, with any slope.

For any fixed integers $n, \ell, p \geq 1$, we introduce the quantity $N := p\ell n(n+1)/2$. We define a *triangular Young tableau* of slope $-\ell/p$ and of size N as a classical Young tableau with N cells with length $n\ell$ and height np such that the first p rows (from the bottom) have length $n\ell$, the next p lines have length $(n-1)\ell$ and so on (see Figure 2). We now study what is the typical value of its lower right corner (with the French convention for drawing Young tableaux, see [21] but take however care that on page 2 therein, Macdonald advises readers preferring the French convention to "read this book upside down in a mirror"!).

It could be expected (e.g. via the Greene–Nijenhuis–Wilf hook walk algorithm for generating Young tableaux, see [14]) that the entries near the hypotenuse should be $N - o(N)$. Can we expect a more precise description of these $o(N)$ fluctuations? Our result on periodic urns enables us to exhibit the right critical exponent, and the limit law in the corner:

▶ **Theorem 7.** *Choose a uniform random triangular Young tableau $\mathcal{Y}$ of slope $-\ell/p$ and size $N = p\ell n(n+1)/2$ and put $\delta = p/(p+\ell)$. Let X_n be the entry of the lower right. Then $(N - X_n)/n^{1+\delta}$ converges in law to the same limiting distribution as the number of black balls in the periodic Young–Pólya urn with initial conditions $w_0 = \ell$, $b_0 = p$ and with replacement matrices $M_1 = \cdots = M_{p-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $M_p = \begin{pmatrix} 1 & \ell \\ 0 & 1+\ell \end{pmatrix}$, i.e. we have the convergence in law, as n goes to infinity:*

$$\frac{p^\delta}{p+\ell} \frac{N - X_n}{n^{1+\delta}} \stackrel{\mathcal{L}}{\longrightarrow} \mathrm{ProdGenGamma}(p, \ell, b_0, w_0).$$

(Recall that $\mathrm{ProdGenGamma}$ *is defined by Formula 9.)*

---

[5] A Young tableau of size n is an array with columns of (weakly) decreasing height, in which each cell is labelled, and where the labels run from 1 to n and are strictly increasing along rows from left to right and columns from bottom to top, see Figure 2. We refer to [21] for a thorough discussion on these objects.

▶ **Remark.** The simplest case $\ell = p = 1$ relates to the Young–Pólya urn model which we analysed in the previous sections.

Sketch of proof. We first establish a link between Young tableaux and linear extensions of trees. Then we will be able to conclude via a link between these trees and periodic Pólya urns. Let us start with Figure 2, which describes the main characters of this proof.

The bottom part of Figure 2 presents two trees (the "big" tree $\mathcal{T}$, which contains the "small" tree $\mathcal{S}$). More precisely, we define the rooted planar tree $\mathcal{S}$ as follows

- The left-most branch of $\mathcal{S}$ has $n\ell + 1$ vertices, which we call $v_1, v_2, \ldots, v_{n\ell+1}$, where v_1 is the root and $v_{n\ell+1}$ is the left-most leaf of the tree.
- For $2 \leq k \leq n - 1$, the vertex $v_{k\ell}$ has $p + 1$ children.
- The vertex $v_{n\ell}$ has $p - 1$ children.
- All other vertices v_j (for $j < n\ell, j \neq k\ell$) have exactly one child.

Now, define $\mathcal{T}$ as the "big" tree obtained from the "small" tree $\mathcal{S}$ by adding a vertex v_0 as the father of v_1 and adding $N + 1 - n(p + \ell)$ children to v_0 (see Figure 2). Remark that the number of vertices of $\mathcal{T}$ is equal to $1 +$ the number of cells of $\mathcal{Y}$. Moreover, the hook length of each cell in the first row (from the bottom) of $\mathcal{Y}$ is equal to the hook length of the corresponding vertex in the left-most branch of $\mathcal{S}$.

Let us now introduce a linear extension $E_\mathcal{T}$ of $\mathcal{T}$, i.e. a bijection from the set of vertices of $\mathcal{T}$ to $\{0, 1, \ldots, N\}$ such that $E_\mathcal{T}(u) < E_\mathcal{T}(u')$ whenever u is an ancestor of u'. A key result, which will be proved in the expanded version of this abstract, is the following: if $E_\mathcal{T}$ is a uniformly random linear extension of $\mathcal{T}$, then X_n (the entry of the lower right corner in a uniformly random Young tableau with shape $\mathcal{Y}$) has the same law as $E_\mathcal{T}(v_{n\ell})$:

$$X_n \overset{\mathcal{L}}{=} E_\mathcal{T}(v_{n\ell}). \tag{10}$$

What is more, recall that $\mathcal{T}$ was obtained from $\mathcal{S}$ by adding a root and some children to this root. Therefore, one can obtain a linear extension of the "big" tree $\mathcal{T}$ from a linear extension of the "small" tree $\mathcal{S}$ by a simple insertion procedure. This allows us to construct a uniformly random linear extension $E_\mathcal{T}$ of $\mathcal{T}$ and a uniformly random linear extension $E_\mathcal{S}$ of $\mathcal{S}$ such that

$$\left| \frac{2(p+\ell)}{n\ell p} (N - E_\mathcal{T}(v_{n\ell})) - (n\ell + p - E_\mathcal{S}(v_{n\ell})) \right| \to 0 \qquad \text{(in probability).}$$

So, to summarize, we have now

$$E_\mathcal{T}(v_{n\ell}) \overset{\mathcal{L}}{=} E_\mathcal{S}(v_{n\ell}) + \text{deterministic quantity} + \text{smaller order error terms.} \tag{11}$$

The last step (which we just state here, see our forthcoming long version for its full proof) is that

$$E_\mathcal{S}(v_{n\ell}) \overset{\mathcal{L}}{=} \text{distribution of periodic Pólya urn} + \text{deterministic quantity.} \tag{12}$$

Indeed, more precisely $N - E_\mathcal{S}(v_{n\ell})$ has the same law as the number of black balls in a periodic urn after $(n - 1)p$ steps (an urn with period p, with adding parameter ℓ, and with initial conditions $w_0 = \ell$ and $b_0 = p$). Thus, our results on periodic urns from Section 3 and the conjunction of Equations (10), (11), and (12) gives the convergence in law for X_n which we wanted to prove. ◀

Figure 2 In this section, we see that there is a relation between Young tableaux with a given periodic shape, some trees, and the periodic Young–Pólya urns. The lower right corner of these Young tableaux is thus following the same generalized Gamma distribution we proved for urns.

5   Conclusion and further work

In this article, we introduced Pólya urns with periodic replacements, and showed that they can be exactly solved with generating function techniques, and that the initial non-linear equation encoding their dynamics leads to linear (D-finite) moment generating functions, which we identify as a product of generalized Gamma distributions. Note that [20, 23] involve the asymptotics of a related process (by grouping p units of time at once of our periodic Pólya urns). This related process is therefore "smoothing" the irregularities created by our periodic model, and allows us to connect with the usual famous key quantities for urns, such as the quotient of eigenvalues of the substitution matrix, etc. Our approach has the advantage to describe each unit of time (and not just what happens after "averaging" p units of time at once), giving more asymptotic terms, and also exact enumeration.

In the full version of this work we will consider arbitrary periodic balanced urn models, and their relationship with Young tableaux. It remains a challenge to understand the asymptotic landscape of Young tableaux, even if it could be globally expected that they behave like a Gaussian free field, like for many other random surfaces [18]. As a first step, understanding the fluctuations and the universality of the critical exponent at the corner could help to get a more global picture. Note that our results on the lower right corner directly imply similar results on the upper right corner: just use our formulae by exchanging ℓ and p, i.e. for a slope corresponding

to the complementary angle to 90°. Thus the critical exponent for the upper right corner is $2 - \delta$. In fact, it is a nice surprise that there is even more structure: there is a *duality* between the limit laws X and X' of these two corners and we get the factorization as independent random variables (up to renormalization and slight modifications of the boundary conditions) $XX' \stackrel{\mathcal{L}}{=} \mathbf{\Gamma}(b_0)$. Similar factorizations of the exponential law, which is a particular case of the Gamma distribution, have appeared recently in relation with functionals of Lévy processes, following [4].

─── **References** ───

1   Omer Angel, Alexander E. Holroyd, Dan Romik, and Bálint Virág. Random sorting networks. *Adv. Math.*, 215(2):839–868, 2007. `doi:10.1016/j.aim.2007.05.019`.

2   Krishna B. Athreya and Peter E. Ney. *Branching processes*. Springer, 1972.

3   Amitava Bagchi and Asim K. Pal. Asymptotic normality in the generalized Pólya-Eggenberger urn model, with an application to computer data structures. *SIAM J. Algebraic Discrete Methods*, 6(3):394–405, 1985. `doi:10.1137/0606041`.

4   Jean Bertoin and Marc Yor. On subordinators, self-similar Markov processes and some factorizations of the exponential variable. *Electron. Comm. Probab.*, 6:95–106, 2001. URL: `https://projecteuclid.org/euclid.ecp/1461097555`.

5   Torsten Carleman. *Sur les équations intégrales singulières à noyau réel et symétrique.* Uppsala Universitets Årsskrift, 1923.

6   Brigitte Chauvin, Cécile Mailler, and Nicolas Pouyanne. Smoothing equations for large Pólya urns. *J. Theoret. Probab.*, 28(3):923–957, 2015. `doi:10.1007/s10959-013-0530-z`.

7   Florian Eggenberger and George Pólya. Über die Statistik verketteter Vorgänge. *Z. Angew. Math. Mech.*, 3:279–290, 1923. `doi:10.1002/zamm.19230030407`.

8   Florian Eggenberger and George Pólya. Sur l'interprétation de certaines courbes de fréquence. *C. R. Acad. Sc.*, 187:870–872, 1928. URL: `http://visualiseur.bnf.fr/CadresFenetre?O=NUMM-3140&I=2&M=tdm`.

9   Giulia Fanti and Pramod Viswanath. Deanonymization in the Bitcoin P2P Network. *Proceedings of the 31st Conference on Neural Information Processing Systems*, 2017. URL: `https://papers.nips.cc/paper/6735-deanonymization-in-the-bitcoin-p2p-network.pdf`.

10   Philippe Flajolet, Philippe Dumas, and Vincent Puyhaubert. Some exactly solvable models of urn process theory. *Discrete Math. Theor. Comput. Sci. Proc.*, AG:59–118, 2006. Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities. URL: `https://hal.inria.fr/hal-01184710/document`.

11   Philippe Flajolet, Joaquim Gabarró, and Helmut Pekari. Analytic urns. *Ann. Probab.*, 33(3):1200–1233, 2005. `doi:10.1214/009117905000000026`.

12   Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009. URL: `http://algo.inria.fr/flajolet/Publications/book.pdf`.

13   Maurice Fréchet and James Shohat. A proof of the generalized second-limit theorem in the theory of probability. *Trans. Amer. Math. Soc.*, 33(2):533–543, 1931. `doi:10.2307/1989421`.

14   Curtis Greene, Albert Nijenhuis, and Herbert S. Wilf. Another probabilistic method in the theory of Young tableaux. *J. Comb. Theory, Ser. A*, 37:127–135, 1984. `doi:10.1016/0097-3165(84)90065-7`.

15   Svante Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.*, 110(2):177–245, 2004. `doi:10.1016/j.spa.2003.12.002`.

16   Svante Janson. Asymptotic degree distribution in random recursive trees. *Random Structures Algorithms*, 26(1-2):69–83, 2005. `doi:10.1002/rsa.20046`.

17   Svante Janson. Moments of Gamma type and the Brownian supremum process area. *Probab. Surv.*, 7:1–52, 2010. With addendum on pages 207–208. `doi:10.1214/10-PS160`.

18   Richard Kenyon. Dominos and the Gaussian free field. *Ann. Probab.*, 29(3):1128–1137, 2001. `doi:10.1214/aop/1015345599`.

19   Morteza Khodabin and Alireza Ahmadabadi. Some properties of generalized gamma distribution. *Mathematical Sciences Quarterly Journal*, 4(1):9–27, 2010. URL: `http://mathscience.kiau.ac.ir/Content/Vol4No1/2.pdf`.

20   Markus Kuba and Henning Sulzbach. On martingale tail sums in affine two-color urn models with multiple drawings. *J. Appl. Probab.*, 54(1):96–117, 2017. `doi:10.1017/jpr.2016.89`.

21   Ian G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Classic Texts in the Physical Sciences. The Clarendon Press, Oxford University Press, second edition, 2015.

22   Hosam M. Mahmoud. *Pólya urn models*. CRC Press, 2009.

23   Cécile Mailler. Describing the asymptotic behaviour of multicolour Pólya urns via smoothing systems analysis. *arXiv*, 2014. URL: `https://arxiv.org/abs/1407.2879`.

24   Philippe Marchal. Rectangular Young tableaux and the Jacobi ensemble. *Discrete Math. Theor. Comput. Sci. Proceedings*, BC:839–850, 2016. URL: `http://www.lix.polytechnique.fr/~pilaud/FPSAC16/final_3`.

25   Erol A. Peköz, Adrian Röllin, and Nathan Ross. Generalized gamma approximation with rates for urns, walks and trees. *Ann. Probab.*, 44(3):1776–1816, 2016. `doi:10.1214/15-AOP1010`.

26   Boris Pittel and Dan Romik. Limit shapes for random square Young tableaux. *Adv. Appl. Math.*, 38(2):164–209, 2007. `doi:10.1016/j.aam.2005.12.005`.

27   George Pólya. Sur quelques points de la théorie des probabilités. *Annales de l'Institut Henri Poincaré*, 1:117–161, 1930. URL: `http://www.numdam.org/article/AIHP_1930__1_2_117_0.pdf`.

28   Ronald L. Rivest. Tabulation audits: Explained and extended. *arXiv*, 2018. URL: `https://arxiv.org/abs/1801.00528`.

29   Dan Romik. *The surprising mathematics of longest increasing subsequences.* Institute of
         Mathematical Statistics Textbooks. Cambridge University Press, New York, 2015. URL:
         `https://www.math.ucdavis.edu/~romik/download-book.php`.

30   Robert T. Smythe and Hosam M. Mahmoud. A survey of recursive trees. *Theory Probab.
         Math. Statist.*, 51:1–29, 1994.

31   Edney W. Stacy. A generalization of the gamma distribution. *Ann. Math. Statist.*, 33:1187–
         1192, 1962. `doi:10.1214/aoms/1177704481`.

32   Anatoly M. Vershik. Randomization of algebra and algebraization of probability. In *Math-
         ematics unlimited—2001 and beyond.*, pages 1157–1166. Springer, 2001.

33   Hubert Stanley Wall. *Analytic Theory of Continued Fractions.* D. Van Nostrand Company,
         Inc., New York, 1948. (See also the reprint by Chelsea in 1967.).

Diagonal Asymptotics for Symmetric Rational Functions via ACSV

Yuliy Baryshnikov[1]

University of Illinois, Department of Mathematics
273 Altgeld Hall 1409 W. Green Street (MC-382), Urbana, IL 61801, USA
ymb@illinois.edu

Stephen Melczer[2]

University of Pennsylvania, Department of Mathematics
209 South 33rd Street, Philadelphia, PA 19104, USA
smelczer@sas.upenn.edu
 https://orcid.org/0000-0002-0995-3444

Robin Pemantle[3]

University of Pennsylvania, Department of Mathematics
209 South 33rd Street, Philadelphia, PA 19104, USA
pemantle@math.upenn.edu

Armin Straub[4]

University of South Alabama, Department of Mathematics and Statistics
411 University Blvd N, MSPB 325, Mobile, AL 36688, USA
straub@southalabama.edu
 https://orcid.org/0000-0001-6802-6053

Abstract

We consider asymptotics of power series coefficients of rational functions of the form $1/Q$ where Q is a symmetric multilinear polynomial. We review a number of such cases from the literature, chiefly concerned either with positivity of coefficients or diagonal asymptotics. We then analyze coefficient asymptotics using ACSV (Analytic Combinatorics in Several Variables) methods. While ACSV sometimes requires considerable overhead and geometric computation, in the case of symmetric multilinear rational functions there are some reductions that streamline the analysis. Our results include diagonal asymptotics across entire classes of functions, for example the general 3-variable case and the Gillis–Reznick–Zeilberger (GRZ) case, where the denominator in terms of elementary symmetric functions is $1 - e_1 + ce_d$ in any number d of variables. The ACSV analysis also explains a discontinuous drop in exponential growth rate for the GRZ class at the parameter value $c = (d - 1)^{d-1}$, previously observed for $d = 4$ only by separately computing diagonal recurrences for critical and noncritical values of c.

1 Introduction

We study the power series coefficients of rational functions of the form $F(x_1, \ldots, x_d) = 1/Q(x_1, \ldots, x_d)$ where Q is a symmetric multilinear function with $Q(\mathbf{0}) \neq 0$. Let

$$F(\mathbf{x}) = \frac{1}{Q(\mathbf{x})} = \sum_{\mathbf{r} \in \mathbb{Z}^d} a_{\mathbf{r}} \mathbf{x}^{\mathbf{r}},$$

converging in some polydisk $\mathcal{D} \subset \mathbb{C}^d$. Often one focuses on the diagonal coefficients $\delta_n := a_{n,\ldots,n}$, whose univariate generating function $\mathsf{diag}_F(z) := \sum_n \delta_n z^n$ satisfies a linear differential equation with polynomial coefficients, but may be transcendental. A number of questions are natural, including nonnegativity (are all coefficients nonnegative), eventual nonnegativity (all but finitely many coefficients nonnegative), diagonal extraction (computing diag_F from Q), diagonal asymptotics, multivariate asymptotics and phase transitions in the asymptotics of $\{a_{\mathbf{r}}\}$.

The positivity (nonnegativity) question is the most classical, dating back at least to Szegő's work in [26]. The techniques, some of which are indicated in the next section, used in the literature are diverse and include integral methods and special functions, positivity preserving operators, combinatorial identities, computer algebra such as cylindrical algebraic decomposition, or determinantal methods. Contrasting to these methods are analytic combinatorial several-variable methods (ACSV) as developed in [20]. These are typically asymptotic, rather than exact, and therefore less useful for proving classical positivity statements, though they can be used to disprove them. Their chief advantages are their broad applicability and, increasingly, the level to which they have been automated. Our aim in this paper is to apply ACSV methods to a number of previously studied families of rational coefficient sequences, thereby extending what is known as well as illuminating the relative advantages of each method.

1.1 Previously studied instances

Let $\mathcal{M}_d$ denote the class of symmetric functions of d variables that are multilinear (degree 1 in each variable). This class of generating functions $F(\mathbf{x}) := 1/Q(\mathbf{x})$ where $Q \in \mathcal{M}_d$ includes a great number of previously studied cases, some of which we now review. Here and in the following, we use d for the number of variables and boldface $\mathbf{x}, \mathbf{y}, \mathbf{z}$, etc., for vectors of length d of integer, real or complex numbers. When d is small we use x, y, z, w for x_1, x_2, x_3, x_4. Let $e_k = e_{k,d}$ denote the kth elementary symmetric function of d variables, the sum of all distinct k element products from the set of d variables. An equivalent description of the class $\mathcal{M}_d$ is that it contains all linear combinations of $\{e_{k,d} : 0 \leq k \leq d\}$.

The Askey-Gasper rational function is

$$A(x, y, z) := \frac{1}{1 - x - y - z + 4xyz}, \tag{1}$$

which, in the previous notation, is $A(\mathbf{x}) = F(\mathbf{x})$ when $d = 3$ and $Q = 1 - e_1 + 4e_3$. Gillis, Reznick and Zeilberger [11] deduce positivity of A from positivity of a 4-variate extension due to Koornwinder [15], for which they give a short elementary proof using a positivity preserving operation. Gillis, Reznick and Zeilberger also provide an elementary proof of the

stronger result by Askey and Gasper [3] that A^β is positive for $\beta \geq (\sqrt{17} - 3)/2 \approx 0.56$, by deriving a recurrence relation for the coefficients that makes positivity apparent.

Specific functions in $\mathcal{M}_4$ that have shown up in the literature include the Szegő rational function

$$S(x, y, z, w) := \frac{1}{e_3(1 - x, 1 - y, 1 - z, 1 - w)} \tag{2}$$

as well as the Lewy-Askey function

$$L(x, y, z, w) := \frac{1}{e_2(1 - x, 1 - y, 1 - z, 1 - w)}, \tag{3}$$

which is a rescaled version of $1/Q(\mathbf{x})$ with $d = 4$ and $Q = 1 - e_1 + \frac{2}{3}e_2$. Szegő [26] proved that (2) is positive. In fact, he showed that $e_{d-1,d}^{-\beta}(1-\mathbf{x})$ is nonnegative if $\beta \geq 1/2$. His proof relates the power series coefficients to integrals of products of Bessel functions and, among other ingredients, employs the Gegenbauer–Sonine addition theorem. Scott and Sokal [22] establish a vast and powerful generalization of this result by showing that, if T_G is the spanning-tree polynomial of a connected series-parallel graph, then $T_G^{-\beta}(1 - \mathbf{x})$ is nonnegative if $\beta \geq 1/2$. In the simplest non-trivial case, if G is a d-cycle, then $T_G = e_{d-1,d}$, thus recovering Szegő's result. Relaxing the condition on β, Scott and Sokal further extend their results to spanning-tree polynomials of general connected graphs. They do so by realizing that Kirchhoff's matrix-tree theorem implies that these polynomials can be expressed as determinants, and by proving that determinants of this kind are nonnegative. As another consequence of this determinantal nonnegativity, Scott and Sokal conclude that (3) is nonnegative, thus answering a question originating with Lewy [2] (with positivity replaced by nonnegativity). Kauers and Zeilberger [14] show that positivity of the Lewy-Askey rational function (3) would follow from positivity of the four variable function

$$K(x, y, z, w) := \frac{1}{1 - e_1 + 2e_3 + 4e_4}. \tag{4}$$

However, the conjectured positivity (or even nonnegativity) of (4) remains open.

As noted above, $e_{d-1,d}^{-\beta}(1 - \mathbf{x})$ is nonnegative if $\beta \geq 1/2$. The asymptotics of $e_{k,d}^{-\beta}(1 - \mathbf{x})$ are computed in [5] for $(k, d) = (2, 3)$. In the cone $2(rs + rt + st) > r^2 + s^2 + t^2$, the coefficient $a_{r,s,t}$ is asymptotically positive when $\beta > 1/2 = (d - k)/2$ and not when $\beta < 1/2$. A conjecture of Scott and Sokal that remains open in both directions is that, for general k and d, the condition $\beta \geq (d - k)/2$ is necessary and sufficient for nonnegativity of the coefficients of $e_{k,d}^{-\beta}(1 - \mathbf{x})$.

Gillis, Reznick and Zeilberger [11] consider the family

$$F_{c,d}(x_1, \ldots, x_d) := \frac{1}{1 - e_1 + c\,e_d} \tag{5}$$

of rational functions, where c is a real parameter. When $c < 0$, the coefficients are trivially positive, therefore it is usual to assume $c > 0$. Gillis, Reznick and Zeilberger show that $F_{c,3}$ has nonnegative coefficients if $c \leq 4$ (and this condition is shown to be necessary in [23]), but they conjecture that the threshold for $d \geq 4$ has a different form, namely that $F_{c,d}$ has nonnegative coefficients if and only if $c \leq d!$. It is claimed in [11], but the proof is omitted due to its length, that nonnegativity of $F_{d!,d}$ is implied by nonnegativity of the diagonal power series coefficients. In the cases $d = 4, 5, 6$, Kauers [13] proved nonnegativity of these diagonal coefficients by applying cylindrical algebraic decomposition (CAD) to the respective recurrences. On the other hand, it is suggested in [25] that the diagonal coefficients are eventually positive if $c < (d - 1)^{d-1}$.

1.2   Previous questions and results on diagonals

The diagonal generating function diag_F and the sequence $\delta_n := a_{n,\ldots,n}$ it generates have received special attention. One reason is that the question of multivariate asymptotics in the diagonal direction is simply stated, whereas the question of asymptotics in all possible directions requires discussion of different possible phase regimes, a notion of uniformity over directions, degeneracies when the coordinates are not of comparable magnitudes, and so forth. Another reason is that there are effective methods for determining diag_F from Q, transferring the problem to the familiar univariate realm.

We briefly recall the theory of diagonal extraction. A d-variate power series F is said to be D-finite if the formal derivatives $\{\partial_{\mathbf{r}}F : \mathbf{r} \in (\mathbb{Z}^+)^d\}$ form a finite dimensional vector space over $\mathbb{C}[\mathbf{x}]$. In one variable, this is equivalent to F satisfying a linear differential equation with polynomial coefficients,

$$\sum_{i=0}^{k} q_i(z)\frac{d^i}{dz^i}F = 0, \quad q_i \in \mathbb{C}[z].$$

▶ **Proposition 1** (D-finite closure under diagonals [17]). *Let $F(\boldsymbol{x})$ be a D-finite power series. Then $\mathsf{diag}(z) := \sum_n \delta_n z^n$ is D-finite, where $\delta_n := a_{n,\ldots,n}$.*

When F is a rational function and $d = 2$, it was known that diag is algebraic (and thus D-finite) at least by the late 1960's [10, 12], and in special cases by Pólya in the 1920's [21]. In the rational function $F(x, y) = P(x, y)/Q(x, y)$ one substitutes $y = 1/x$ and computes a residue integral to extract the constant coefficient. The basis for Lipshitz' proof was the realization that the complex integration can be viewed as purely formal. With the advent of computer algebra this formal D-module computation was automated, with an early package in Macaulay and more widely used modern implementations in Magma, Mathematica and Maple. Due to advances in software and processor speed, these computations are often completable on functions arising in applications. Christol [8] was the first to show that diagonals of *rational* functions are D-finite.

The following relationship between D-finiteness of a univariate function and the existence of a polynomial recursion satisfied by its coefficient sequence is the result of translating a formal differential equation into a relation among the coefficients.

▶ **Proposition 2.** *The series $f(z) = \sum_{n \geq 0} a_n z^n$ is D-finite if and only if it is polynomially recursive, meaning that there is a $k > 0$ and there are polynomials $p_0, \ldots, p_k$, not all zero, such that for all but finitely many n,*

$$\sum_{i=0}^{k} p_i(n)f(n + i) = 0 \,.$$

Let f be a D-finite power series in one variable. If f has positive finite radius of convergence and integer coefficients, then it is a so-called *G-function* and has well behaved asymptotics according to following result.

▶ **Proposition 3** (Asymptotics of G-Function Coefficients). *Suppose f is D-finite with finite radius of convergence and integer coefficients annihilated by a minimal order linear differential operator $\mathcal{L}$ with polynomial coefficients. Then $\mathcal{L}$ has only regular singular points in the Frobenius sense. Consequently, the coefficients $\{a_n\}$ are given asymptotically by a formula*

$$a_n \sim \sum_{\alpha} C_\alpha n^{\beta_\alpha} \rho_\alpha^{-n}(\log n)^{k_\alpha} \tag{6}$$

where the sum is over quadruples $(C_\alpha, b_\alpha, \rho_\alpha, k_\alpha)$ as α ranges over a finite set A with the following properties. The base ρ_α is an algebraic number, a root of the leading polynomial coefficient of $\mathcal{L}$. The β_α are rational and for each value of ρ_α can be determined as roots of an explicit polynomial constructed from ρ_α and $\mathcal{L}$. The log powers k_α are nonnegative integers, zero unless for fixed ρ_α there exist two values of β_α differing by an integer (including multiplicities in the construction of β_α). The C_α are not in general closed form analytic expressions, but may be determined rigorously to any desired accuracy.

Proof. The discussion in [18, page 37] gives references to several published results that together establish this proposition; see also Flajolet and Sedgewick [9, Section VII. 9]. Determination of all rational and algebraic numbers other than C_α is known to be effective. ◄

Because there are computational methods for the study of diagonals, it is of interest to reduce positivity questions to those involving only diagonals. For the Gillis-Reznick-Zeilberger class $F_{c,d}$, such a result is conjectured.

▶ **Conjecture 4** ([11]). *For $d \geq 4$, the following three statements are equivalent.*
   (i) $c \leq d!$
  (ii) *The diagonal coefficients of $F_{c,d}$ are nonnegative*
 (iii) *All coefficients of $F_{c,d}$ are nonnegative*

To be precise, $(iii) \Rightarrow (ii) \Rightarrow (i)$ is trivial (look at δ_1); nonnegativity of all coefficients of $F_{c,d}$ holds for some interval $c \in [0, c_{\max}]$, therefore the conjecture comes down to nonnegativity of $F_{d!,d}$. A proof for $(ii) \Rightarrow (iii)$ in the case $c = d!$ is claimed in [11] but omitted from the paper due to length. This question is generalized in [25] to all of $\mathcal{M}_d$.

▶ **Question 5** ([25, Question 1.1 and following]). *For $Q \in \mathcal{M}_d$ and $F = 1/Q$, under what conditions does nonnegativity of the coefficients of diag_F imply nonnegativity of all coefficients of F?*

More specifically, with nonnegativity in place of positivity, the authors of that paper wonder whether positivity of F is equivalent to positivity of diag_F together with positivity of $F(x_1, \ldots, x_{d-1}, 0)$. They prove that this is true for $d = 2$ and, with additional evidence, conjecture this to be true for $d = 3$ as well. Combined with [23, Conjecture 1] and [25, Conjecture 3.3], we obtain the following explicit predictions on the diagonal coefficients.

▶ **Conjecture 6.** *Let $F = 1/Q$ where $Q = 1 - e_1 + ae_2 + be_3$, which is, up to rescaling, the general element of $\mathcal{M}_3$. Then diag_F is nonnegative if and only if*

$$b \leq \begin{cases} 6(1-a) & a \leq a_0 \\ 2 - 3a + 2(1-a)^{3/2} & a_0 \leq a \leq 1 \\ -a^3 & a \geq 1, \end{cases} \tag{7}$$

where $a_0 \approx -1.81$ is characterized by $6(1 - a_0) = 2 - 3a_0 + 2(1 - a_0)^{3/2}$.

1.3 Present results

In the present work we use ACSV to answer asymptotic versions of these questions. Aside from computing special cases, the main new results are (1) simplification for diagonals with symmetric denominators via the Grace-Walsh-Szegő Theorem (Lemma 15 below); (2) an easy further simplification for the Gillis-Reznick-Zeilberger class (Lemma 18 below); and

(3) a topological computation to explain the drop in magnitude of coefficients at critical parameter values (Theorem 22 below).

The first special case we look at is the diagonal of the general element of $\mathcal{M}_3$, corresponding to Conjecture 6.

▶ **Theorem 7.** *Let $Q = 1 - e_1 + ae_2 + be_3$, let $F = 1/Q = \sum_{\boldsymbol{r}} a_{\boldsymbol{r}} \boldsymbol{z}^{\boldsymbol{r}}$ and let $\delta_n = a_{n,\dots,n}$ be the diagonal coefficients of F. Then δ_n is eventually positive when*

$$b < \begin{cases} -9a & a \leq -3 \\ 2 - 3a + 2(1-a)^{3/2} & -3 \leq a \leq 1 \\ -a^3 & a \geq 1 \end{cases} \tag{8}$$

while, when the inequality is reversed, δ_n attains an infinite number of positive and negative values.

Theorem 7 is obtained by examining asymptotic regimes, captured in the following result.

▶ **Theorem 8.** *Let Q, F, and δ_n be as in Theorem 7. Assuming that b is not equal to the piecewise function in Equation (8),*

$$\delta_n = \sum_{x \in E} \left(\frac{x^{-3n}}{n} \cdot \left| \frac{1 - 2ax - bx^2}{1 - ax} \right| \cdot \frac{1}{2\sqrt{3}(1 - 2x + ax^2)} \right) \left(1 + O\left(\frac{1}{n}\right) \right), \tag{9}$$

where E consists of the minimal modulus roots of the polynomial $Q(x,x,x) = 1 - 3x + 3ax^2 + bx^3$.

The situation for eventual positivity on the diagonal when equality holds in Equation (8) is more delicate. When $a < -3$ it follows from seeing that there are two diagonal minimal points, (r, r, r) and $(-r, -r, -r)$, with a greater constant at the positive point. When $-3 < a < 1$, it follows from a dominant positive real cone point. When $a = -3$ a quadratically degenerate smooth point at $(-1/3, -1/3, -1/3)$ may be shown via rigorous numerical diagonal extraction to dominate the cone point at $(1/3, 1/3, 1/3)$, leading to alternation. When $a = 1$, $a_{\boldsymbol{r}} \equiv 1$. Finally, when $a > 1$, there are three smooth points on the unit circle, with nonnegativity conjectured because the positive real point is degenerate and should dominate.

Our second set of results concern the diagonal of the general element of the GRZ rational function $F_{c,d}$. Let

$$c_* = c_*(d) := (d-1)^{d-1}. \tag{10}$$

The following corresponds to Conjecture 4.

▶ **Theorem 9.** *Let $d \geq 4$. Then the diagonal coefficients of $F_{c,d}$ are eventually positive when $c < c_*$ and contain an infinite number of positive and negative values when $c > c_*$. When $c < c_*$, there is a conical neighborhood $\mathcal{N}$ of the diagonal such that $a_{\boldsymbol{r}} > 0$ for all but finitely many $\boldsymbol{r} \in \mathcal{N}$.*

Again, the result is obtained through an explicit asymptotic analysis.

▶ **Theorem 10.** *Let δ_n be the diagonal coefficients of $F_{c,d}$. Then when $c \neq c_*$,*

$$\delta_n = \sum_{x \in E} \left(\frac{x^{-dn}}{n^{(d-1)/2}} \cdot \left(\frac{2\pi(1 - (d-1)r)}{r^{(d-1)/2}} \right)^{(d-1)/2} \cdot \frac{1}{d^{1/2}(1 - (d-1)r)} \right) \left(1 + O\left(\frac{1}{n}\right) \right),$$

where E consists of the minimal modulus roots of the polynomial $1/F_{c,d}(x, \dots, x) = 1 - dx + cx^d$.

These theorems are proven in Section 4, using ACSV smooth point methods summarized in Section 2, however the case $c = c_*$ for the GRZ rational function requires the more delicate results of Section 5.

1.4 Exponential drop and further results

In the GRZ family, for even values of $d \geq 4$ the exponential growth rate of the coefficients drops at the special value $c = (d-1)^{d-1}$. This special value, and the corresponding drop in exponential growth, may be identified for each fixed d from the differential equation annihilating the diagonal. For example, when $d = 4$ an annihilating differential equation for the diagonal of $F_{c,4}$ is computed by D-module integration in the Mathematica package of Koutschan [16] producing the annihilating operator $\mathcal{L}$, of order 3 and maximum coefficient degree 8, such that $\mathcal{L}\mathrm{diag}_{F_{c,4}} = 0$:

$$
\begin{aligned}
\mathcal{L} = {} & z^2(c^4 z^4 + 4c^3 z^3 + 6c^2 z^2 + 4cz - 256z + 1)(3cz - 1)^2 \partial_z^3 \\
& + 3z(3cz - 1)(6c^5 z^5 + 15c^4 z^4 + 8c^3 z^3 - 6c^2 z^2 - 384cz^2 - 6cz + 384z - 1)\partial_z^2 \\
& + (cz + 1)(63c^5 z^5 - 3c^4 z^4 - 66c^3 z^3 + 18c^2 z^2 + 720cz^2 + 19cz - 816z + 1)\partial_z \\
& + 9c^6 z^5 - 3c^5 z^4 - 6c^4 z^3 + 18c^3 z^2 - 360c^2 z^2 + 13c^2 z - 384cz + c - 24.
\end{aligned}
\tag{11}
$$

When $c = 27$, all coefficients in (11) acquire enough zeros at $z = 1/81$ that the quantity $(81z - 1)^4$ may be factored out of the entire operator, leaving the following operator of order 3 and maximum degree 4:

$$
\begin{aligned}
\mathcal{L}_{27} := {} & z^2(81z^2 + 14z + 1)\,\partial_z^3 + 3z(162z^2 + 21z + 1)\,\partial_z^2 \\
& + (21z + 1)(27z + 1)\partial_z + 3(27z + 1).
\end{aligned}
\tag{12}
$$

Asymptotics for δ_n may be extracted via the methodology described in Proposition 3. In the special case $d = 4, c = 27$, the recursion may be found on the OEIS (entry A125143) and identifies $\{\delta_n\}$ as the *Almkvist–Zudilin numbers*[5] from [1, sequence (4.12)(δ)]. The known asymptotic formula implies that $|\delta_n|^{1/n} \to 9$. However, as $c \neq 27$ approaches 27 from either side, we have

$$
\lim_{c \to 27} \lim_{n \to \infty} |\delta_n|^{1/n} = 81;
$$

in other words, the growth rate at $c = 27$ drops suddenly from 81 to 9. The occurrence of a phase change at $(d-1)^{d-1}$ for all d and drop in exponential rate for even $d \geq 4$ had not previously been proved. The special role of the case $c = (d-1)^{d-1}$ was observed in [25, Example 4.4] and claimed to agree with intuition from hypergeometric functions. We verify this, first by identifying the singularity from an ACSV point of view and then by checking that this singularity indeed produces the observed dimension drop.

▶ **Theorem 11** (exponential growth approaching criticality). *For all $d \geq 2$,*

$$
\lim_{c \to c_*} \limsup_{n \to \infty} |\delta_n|^{1/(dn)} = d - 1 \,.
$$

▶ **Theorem 12** (dimension drop at criticality). *When $c = c_*$ and $d \geq 4$ is even,*

$$
\limsup_{n \to \infty} |\delta_n|^{1/(dn)} < d - 1 \,.
$$

Theorem 12 is proved in Section 5.

---

[5] That these are the diagonals of the rational function $F_{27,4}$ was observed in [24], where it is further conjectured that the coefficients of $F_{27,4}$ satisfy very strong congruences.

2   ACSV

In this section we describe the basic setup for ACSV and state some existing results. Definitions for the topological and geometric quantities used below can be found in Pemantle and Wilson [20]. Throughout this section let $F(\mathbf{z}) = P(\mathbf{z})/Q(\mathbf{z}) = \sum_{\mathbf{r}} a_{\mathbf{r}}\mathbf{z}^{\mathbf{r}}$ denote a rational series in d variables, with P and Q co-prime polynomials. Assume that F has a (finite) positive radius of convergence; that is, $Q(\mathbf{0}) \neq 0$ and P/Q is not a polynomial. Let $\mathcal{V} := \{\mathbf{z} \in \mathbb{C}^d : Q(\mathbf{z}) = 0\}$ denote the singular variety for F and let $\mathcal{M} = (\mathbb{C}^*)^d \setminus \mathcal{V}$ where $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. Coefficients $a_{\mathbf{r}}$ are extracted via the multivariate Cauchy formula

$$a_{\mathbf{r}} = \frac{1}{(2\pi i)^d} \int_{\mathbf{T}} \mathbf{z}^{-\mathbf{r}} F(\mathbf{z}) \frac{d\mathbf{z}}{\mathbf{z}}, \tag{13}$$

where $d\mathbf{z}/\mathbf{z}$ denotes the holomorphic logarithmic volume form $(dz_1/z_1) \wedge \cdots \wedge (dz_d/z_d)$ and $\mathbf{T}$ denotes a small torus (a product of sufficiently small circles about the origin in each coordinate, so that the product of the corresponding disks is disjoint from $\mathcal{V}$). The fundamental insight of ACSV is that the integral depends only on the homology class of $\mathbf{T}$ in $H_d(\mathcal{M})$. Therefore, one tries to replace $\mathbf{T}$ by some homologous chain $\mathcal{C}$ over which the integral is easier, typically via some combination of residue reductions and saddle point estimates.

A *direction* of asymptotics is an element $\hat{\mathbf{r}} \in (\mathbb{RP}^d)^+$; that is, a projective vector in the positive orthant. If $\mathbf{r} \in (\mathbb{R}^d)^+$ we write $\hat{\mathbf{r}}$ to denote the representative $\mathbf{r}/|\mathbf{r}|$ of the projective equivalence class containing $\mathbf{r}$, where $|\mathbf{r}| = |\mathbf{r}|_1 := r_1 + \cdots + r_d$. Given a Whitney stratification of $\mathcal{V}$ into smooth manifolds, the *critical set* $\mathsf{crit}(\hat{\mathbf{r}})$ for a direction $\hat{\mathbf{r}}$ is the set of $\mathbf{z} \in \mathcal{V}$ such that $\hat{\mathbf{r}}$ is orthogonal to the tangent space of the stratum of $\mathbf{z}$ in $\mathcal{V}$. If $\mathbf{z}$ is a smooth point of $\mathcal{V}$ and Q is square-free, this means $\hat{\mathbf{r}}$ should be parallel to the logarithmic gradient $(z_1 \partial Q/\partial z_1, \ldots, z_d \partial Q/\partial z_d)$. A *minimal* point for direction $\hat{\mathbf{r}}$ is a point $\mathbf{z} \in \mathsf{crit}(\hat{\mathbf{r}})$ such that the open polydisk $\mathcal{D}(\mathbf{z}) := \{\mathbf{w} : |w_j| < |z_j| \forall 1 \leq j \leq d\}$ does not intersect $\mathcal{V}$. The minimal point $\mathbf{z}$ is called *strictly minimal* if the closed polydisk $\overline{\mathcal{D}(\mathbf{z})}$ intersects $\mathcal{V}$ only at $\mathbf{z}$.

For any $\boldsymbol{\beta} \in \mathbb{R}^d$, let $\mathbf{T}(\boldsymbol{\beta}) = \{\mathbf{w} : |w_j| = \exp(\beta_j) \forall 1 \leq j \leq d\}$ denote the torus of points with log modulus vector $\boldsymbol{\beta}$. The *amoeba* of $Q(\mathbf{z})$ is the image of $\mathcal{V}$ under the map $\mathrm{Relog}(\mathbf{z}) = (\log|z_1|, \ldots, \log|z_d|)$, while the *height* of a point $\mathbf{z}$ is $h_{\hat{\mathbf{r}}}(\mathbf{z}) = -\hat{\mathbf{r}} \cdot \mathrm{Relog}(\mathbf{z})$. Except in Section 5, all ACSV computations are based on the following result.

▶ **Theorem 13** (smooth point formula). *Fix* $F = P/Q = \sum_{\boldsymbol{r}} a_{\boldsymbol{r}} \boldsymbol{z}^{\boldsymbol{r}}$ *and vector* $\boldsymbol{r} \in (\mathbb{R}^d)^+$ *in direction* $\hat{\boldsymbol{r}}$. *Assume there exists* $\boldsymbol{\beta} \in \mathbb{R}^d$ *such that the following two hypotheses hold.*

1 **Finite critical points on the torus.** *The set* $E := \boldsymbol{T}(\boldsymbol{\beta}) \cap \mathsf{crit}(\hat{\boldsymbol{r}})$ *is finite, nonempty and contains only minimal smooth points.*
2 **Quadratic nondegeneracy.** *At each* $\boldsymbol{z} \in E$ *fix* $k = k(\boldsymbol{z})$ *such that* $\partial Q/\partial z_k(\boldsymbol{z}) \neq 0$ *and let* $z_k = g(z_1, \ldots, \hat{z}_k, \ldots, z_d)$ *be a smooth local parametrization of* z_k *on* $\mathcal{V}$ *as a function of* $\{z_j : j \neq k\}$. *We assume that the Hessian determinant* $\mathcal{H}_{k(\boldsymbol{z})}$ *of second partial derivatives of* $g\left(w_1 e^{i\theta_1}, \ldots, w_d e^{i\theta_d}\right)$ *with respect to the* θ_j *at the origin is non-zero for each* $\boldsymbol{z} \in E$.

Then there exists a closed neighborhood $\mathcal{N}$ *of* $\hat{\boldsymbol{r}}$ *in* $(\mathbb{R}^d)^+$ *on which all the above hypotheses hold and, for any* $\boldsymbol{r}$ *with* $\hat{\boldsymbol{r}}$ *in this neighborhood,*

$$a_{\boldsymbol{r}} = (2\pi)^{(1-d)/2} \sum_{\boldsymbol{z} \in E} \det \mathcal{H}_{k(\boldsymbol{z})}^{-1/2} \frac{P(\boldsymbol{z})}{z_k (\partial Q/\partial z_k)(\boldsymbol{z})} r_k^{(1-d)/2} \boldsymbol{z}^{-\boldsymbol{r}} + O\left(r_k^{-d/2} \boldsymbol{z}^{-\boldsymbol{r}}\right). \tag{14}$$

▶ Remark. A number of other formulae for $a_{\mathbf{r}}$ are equivalent to this one and hold under the same hypotheses. An explicit formula for $\mathcal{H}_k$ in terms of partial derivatives of Q is given

in [18, Theorem 54]. The following coordinate-free formula for the constants involved in terms of the complexified Gaussian curvature $\mathcal{K}$ at a smooth point $\mathbf{z} \in \mathcal{V}$ is given in [20, (9.5.2)] as

$$a_{\mathbf{r}} = (2\pi)^{(1-d)/2} \left[\sum_{\mathbf{z} \in E} \mathcal{K}(\mathbf{z})^{-1/2} \, |\nabla_{\log} Q(\mathbf{z})|^{-1} P(\mathbf{z}) \, |\mathbf{r}|^{(1-d)/2} \, \mathbf{z}^{-\mathbf{r}} \right] + O\left(|\mathbf{r}|^{-d/2} |\mathbf{z}|^{-\mathbf{r}} \right) \quad (15)$$

Proof. Assume first that $\log |\mathbf{w}|$ is the unique minimizer of $\mathbf{r} \cdot \mathbf{x}$ on the boundary of the log domain of convergence (this being a component of the complement of the amoeba). Under no assumptions on E or $\mathcal{K}$, Theorem 9.3.2 of [20] writes the multivariate Cauchy integral 13 as the integral of a residue form ω over an intersection cycle, $\mathcal{C}$. Taking into account that E is finite, and assuming an extra hypothesis that $\mathbf{r}$ is a *proper direction* (see [5, Definition 2.3]), Theorem 9.4.2 of [20] identifies $\mathcal{C}$ as a sum of quasi-local cycles near the points of E. For each such $\mathbf{z}$, if $\partial Q / \partial z_k$ and $\det \mathcal{H}_k$ do not vanish, Theorem 9.2.7 of [20] identifies the integral as the corresponding summand in (14). Nonvanishing of $\mathcal{H}_k$ is equivalent to nonvanishing of $\mathcal{K}$, leading to the coordinate-free formula (15), which may be found in [20, Theorem 9.3.7]. This proves the theorem under an extra hypothesis on the amoeba boundary.

To remove the properness hypothesis, consider the intersection cycle $\mathcal{C}$ obtained from expanding the torus $\mathbf{T}(\boldsymbol{\beta} - \epsilon\mathbf{r})$ inside the domain of convergence of F to a torus $\mathbf{T}(\boldsymbol{\beta} + \epsilon\mathbf{r})$. The construction in [20, Section A4] gives a compact $(d-1)$-chain representing a relative cycle in $H_{d-1}(\mathcal{V}^{c+\varepsilon}, \mathcal{V}^{c-\varepsilon})$; that is, a chain of maximum height $c + \varepsilon$ with maximum boundary height $c - \varepsilon$. Applying the downward gradient flow of $h_{\hat{\mathbf{r}}}$ on $\mathcal{V}$ for arbitrarily small time, we arrive again at a chain satisfying the conclusions of [20, Theorem 9.4.2]. Because the deformed chain has nonvanishing boundary, one must add a term for the chain swept out by the deformation applied to this boundary, but the elements of this chain have height at most $c - \varepsilon$ so the resulting integral will be within the error term above. ◀

▶ **Corollary 14.** *Assume the hypotheses of Theorem 13, and fix a vector $\mathbf{v}$ in direction $\hat{\mathbf{r}}$.*

  (i) *If $E = \{\mathbf{z}\}$ for some $\mathbf{z}$ in the positive real orthant in $\mathbb{C}^d$ and the leading constant of Equation (14) is positive, then there exists a neighbourhood of $\hat{\mathbf{r}}$ such that all but finitely many coefficients $\{a_{\mathbf{r}} : \hat{\mathbf{r}} \in \mathcal{N}\}$ are positive.*

 (ii) *If $E = \{\mathbf{z}\}$ for some $\mathbf{z}$ such that $\mathbf{z}^{\mathbf{v}} := \prod_{j=1}^{d} z_j^{v_j}$ is positive real and the leading constant of Equation (14) is positive, then all but finitely many coefficients $a_{n\mathbf{v}}$ are positive.*

(iii) *If E does not contain a point $\mathbf{z}$ with $\mathbf{z}^{\mathbf{v}}$ positive real and the sum in Equation (14) is not identically zero, then infinitely many coefficients $a_{n\mathbf{v}}$ are positive and infinitely many $a_{n\mathbf{v}}$ are negative.*

▶ **Remark.** When E contains a point in the positive real orthant but it is not a singleton, the corollary does not provide information as to eventual positivity.

Proof. Conclusions (i) and (ii) follow immediately from (14) because the sum is a single positive term.

For conclusion (iii), grouping the elements of E by conjugate pairs we note that up to scaling by $\mathbf{z}^{n\mathbf{v}} n^{d/2}$ the asymptotic leading term of $a_{n\mathbf{v}}$ has the form

$$l_n = \sum_{i=1}^{|E|} a_i \cos(2\pi\theta_i n + \beta_i),$$

where each θ_i, a_i, β_i is real, and $\theta_i \in (0, 1)$. If r_n is any sequence satisfying a linear recurrence relation with constant coefficients, and $r_n = O(1/n)$, then Bell and Gerhold [6, Section 3]

show that $l_n > r_n$ infinitely often. Since the modulus of the error term in Equation (14) can be bounded by a linear recurrence sequence with growth $O(1/n)$, we see that $a_{n\mathbf{v}}$ is positive infinitely often. Repeating the argument with $-l_n$ shows that $a_{n\mathbf{v}}$ is negative infinitely often.   ◀

Any computer algebra system can compute the set of smooth critical points in $\mathsf{crit}(\hat{\mathbf{r}})$ by solving the $d-1$ equations $(\nabla_{\log}Q)(\mathbf{z}) \parallel \hat{\mathbf{r}}$ together with the equation $Q(\mathbf{z}) = 0$, where $\nabla_{\log}Q = (z_1\partial Q/\partial z_1, \ldots, z_d\partial Q/\partial z_d)$. Identifying which points in crit are minimal is more difficult, although still effective [19]. For our cases, we can use results about symmetric functions to help with the computations. For any polynomial Q in d variables, let δ^Q denote the codiagonal: the univariate polynomial defined by $\delta^Q(x) = Q(x, \ldots, x)$.

▶ **Lemma 15** (polynomials in $\mathcal{M}_d$ have diagonal minimal points). *Let $F = 1/Q$ with $Q \in \mathcal{M}_d$. Let x be a zero of δ^Q of minimal modulus. Then $\boldsymbol{x} := (x, \ldots, x)$ is a minimal point for F in* $\mathsf{crit}(1, \ldots, 1)$.

This follows directly from the classical Grace-Walsh-Szegő Theorem, a modern proof of which is contained in the following.

Proof. Let $\alpha_1, \ldots, \alpha_k$ be the roots of δ^Q, where $k \leq d$ is the common degree of Q and δ^Q and $|\alpha_1|$ is minimal among $\{|\alpha_j| : j \leq k\}$. For any $\varepsilon > 0$, the polynomial

$$M(\mathbf{x}) := \prod_{j=1}^{k}(x_j - \alpha_j)$$

has no zeros in the polydisk $\mathcal{D}$ centered at the origin whose radii are $\alpha_1 - \varepsilon$. The symmetrization of M (see [7]) is defined to be the multilinear symmetric function m such that $m(x, \ldots, x) = M(x, \ldots, x)$. In our case $M(x, \ldots, x) = \delta^Q(x)$, and it immediately follows that $m = Q$. By the Borcea-Brändén symmetrization lemma (see [7, Theorem 2.1]), the polynomial Q has no zeros in the polydisk $\mathcal{D}$. We conclude that the zero $\mathbf{x}$ of Q is a minimal point of F.   ◀

3   Symmetric multilinear functions of three variables

In this section we determine the diagonal asymptotics for general $Q = 1 - e_1 + ae_2 + be_3 \in \mathcal{M}_3$. Taking the coefficient of e_1 to be 1 loses no generality because of the rescaling $x_j \to \lambda x_j$ which preserves $\mathcal{M}_d$ and affects coefficient asymptotics in a trivial way. In order to use Theorem 13, we begin by identifying minimal points. Lemma 15 dictates that our search should be on the diagonal.

To that end, let $\delta^Q(x) = Q(x, x, x) = 1 - 3x + 3ax^2 + bx^3$. The discriminant of δ^Q is a positive real multiple of $p(a, b) := 4a^3 - 3a^2 + 6ab + b^2 - 4b = (a - 1 + 3(b-1))^2 - 4(b-1)^3$, and the zero set of δ^Q is obtained from that of the cubic $4b^3 = -a^2$ by centering at $(1, -1)$ and shearing via $(a, b) \mapsto (a + 3b, b)$. The discriminant $p(a, b)$ vanishes along the red curve (solid and dashed) in Figure 1. Let $r_1(a)$ and $r_2(a)$ denote respectively the upper and lower branches of the solution to $p(a, b) = 0$.

▶ **Lemma 16.** *Let p be a minimal modulus root of δ^Q. Then any critical point of $1/Q$ on the torus $T(p, p, p)$ has the form (q, q, q) where $\delta^Q(q) = 0$.*

Proof. Gröbner basis computations show nondiagonal critical points to be permutations of $\left(\frac{1}{a}, \frac{1}{a}, \frac{a(1-a)}{a^2+b}\right)$, occurring when $b = a^2(a-2)$. When $a \leq 1$, the only time the positive root
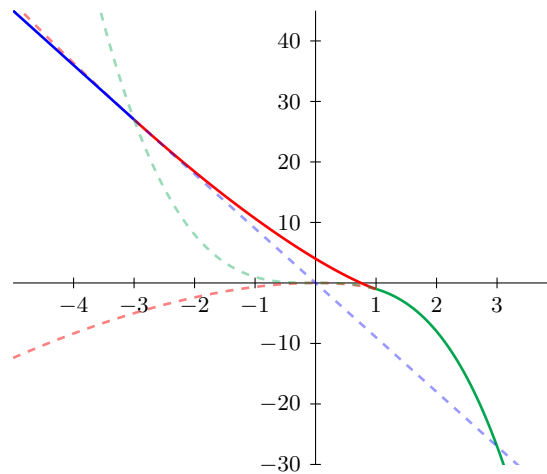
■ **Figure 1** The three regimes defined by Proposition 17, made up of the curves $b = -9a$, $p(a,b) = 0$, and $b = -a^3$. Dashed lines represent the curves where they do not determine positivity of coefficients; note smoothness in the transitions between regimes.

of $\delta^Q(x)$ has modulus $1/|a|$ is the trivial case $(a,b) = (1,-1)$. When $b = a^2(a-2)$ and $a > 1$, the modulus of the product of the roots of $\delta^Q(x)$ equals $\frac{1}{a^2(a-2)}$ and the minimal roots of $\delta^Q(x)$ are a pair of complex conjugates. If this pair has modulus $1/a$, then the real root of $\delta^Q(x)$ is $\pm\frac{1}{a^4(a-2)}$, but $\delta^Q\left(\pm\frac{1}{a^4(a-2)}\right) \neq 0$ for $a > 1$. ◄

Determining asymptotics is thus a matter of determining the minimal modulus roots of $\delta^Q(x)$. The following may be proved by comparing moduli of roots, separating cases according to the sign of $p(a,b)$.

▶ **Proposition 17.** *The function δ^Q has a minimal positive real zero if and only if*

$$b \leq \begin{cases} -9a & a \leq -3 \\ r_1(a) & -3 \leq a \leq 1 \\ -a^3 & a \geq 1 \end{cases}$$

This corresponds to the set of points lying on and below the solid curve in Figure 1.

Proof of Theorems 7 and 8: Suppose b is greater than the piecewise expression in the proposition; then δ^Q has no minimal positive zero, so the product of the three coordinates of the minimal points determined above do not lie in the positive orthant. By part (iii) of Corollary 14, the diagonal coefficients are not eventually positive. Asymptotics of δ_n are determined by Theorem 13, and when b is less than the piecewise expression it can be verified that the dominant term is positive. ◄

4 The Gillis-Reznick-Zeilberger classes

Throughout this section, let $F = F_{c,d} = 1/Q_{c,d} = 1/(1 - e_1 + ce_d)$ and recall that $c_* = (d-1)^{d-1}$. Lemma 15 implies that for $Q \in \mathcal{M}_d$, in the diagonal direction, one may find diagonal minimal points. For $F_{c,d}$, things are even simpler: all critical points for diagonal asymptotics are diagonal points.

▶ **Lemma 18.** *Let $F_{c,d} = 1/Q_{c,d}$. If $\mathbf{z} \in \mathsf{crit}(1,\ldots,1)$ then $z_i = z_j$ for all $1 \leq i, j \leq d$.*

Proof. From $Q = Q_{c,d} = 1 - e_1 + ce_d$ we see that $(\nabla_{\log}Q)_j = -z_j - ce_d$ and hence that $(\nabla_{\log}Q)_i = (\nabla_{\log}Q)_j$ if and only if $z_i = z_j$. ◀

▶ **Proposition 19** (Smoothness of $F_{c,d}$ for $c \neq c_*$). *Let $F_{c,d} = 1/Q_{c,d}$. If $c \neq c_*$ then $\mathcal{V}$ is smooth. If $c = c_*$ then $\mathcal{V}$ fails to be smooth at the single point $\mathbf{z}_* = (1/(d-1), \ldots, 1/(d-1))$. When $c = c_*$, the singularity at $\mathbf{z}_*$ has tangent cone e_2.*

Proof. Checking smoothness of $\mathcal{V}$ we observe that for d fixed and c and $x_1, \ldots, x_d$ variable, vanishing of the gradient of $Q_{c,d}$ with respect to the x variables implies $x_j = ce_d$ for all j. This common value, x, cannot be zero, hence $x_j \equiv x$ and $c = x^{1-d}$. Vanishing of $Q_{c,d}$ then implies vanishing of $1 - dx + x$, hence $x = 1/(d-1)$ and $c = c_*$. This proves the first two statements. Setting $c = c_*$ and $x_j = 1/(d-1) + y_j$ centers $Q_{c_*,d}$ at the singularity and produces a leading term of $(d-1)e_2(\mathbf{y})$, proving the third statement. ◀

4.1   Proof of Theorems 9 and 10 in the case $c < c_*$

When $c \leq 0$, the denominator of $F_{c,d}$ is one minus the sum of positive monomials, which leaves no doubt as to positivity. Assume, therefore, that $0 < c < c_*$. Apply Lemma 15 to see that if x is a minimum modulus zero of $\delta^Q := Q_{c,d}(x, \ldots, x)$ then $(x, \ldots, x)$ is a minimal point for $F_{c,d}$ in the diagonal direction. Apply Lemma 18 to conclude that the set E in Theorem 13 of minimal critical points on $\mathbf{T}(|x|, \ldots, |x|)$ consists only of points $(y, \ldots, y)$ such that y is a root of δ^Q. By part (i) of Corollary 14, it suffices to check that $\delta^Q = 1 - dx + cx^d$ has a unique minimal modulus root ρ and that $\rho \in \mathbb{R}^+$. Thus, the conclusion follows from the following proposition.

▶ **Proposition 20.** *For $c \in (0, c_*)$, the polynomial $\delta^Q = 1 - dx + cx^d$ has a root $\rho \in \left[\frac{1}{d}, \frac{1}{d-1}\right]$ which is the unique root of δ^Q of modulus less than $1/(d-1)$.*

Proof. Checking signs we find that $\delta^Q(1/d) = cd^{-d} > 0$ while $\delta^Q(1/(d-1)) = -(d-1)^{-1} + c(d-1)^{-d} < -(d-1)^{-1} + c_*(d-1)^{-d} = 0$, therefore there is at least one root, call it ρ, of δ^Q in the interval $[1/d, 1/(d-1)]$. On the other hand, when $|z| = 1/(d-1)$, we see that $|dz| \geq |1+cz^d|$ and therefore, by applying Rouché's theorem to the functions $-dz$ and $1+cz^d$, we see that δ^Q has as many zeros on $|z| < 1/(d-1)$ as does $-dz$: precisely one root, ρ. ◀

4.2   Proof of Theorems 9 and 10 in the case $c > c_*$

Again, by Lemmas 15 and 18, we may apply part (iii) of Corollary 14 to the set E of points $(y, \ldots, y)$ for all minimal modulus roots y of δ^Q. The result then reduces to the following proposition.

▶ **Proposition 21.** *For $c > c_*$, the set of minimal modulus roots of the polynomial $\delta^Q = 1 - dx + cx^d$ contains no point whose d^{th} power is real and positive.*

Proof. First, if z^d is real then the imaginary part of $\delta^Q(z)$ is equal to the imaginary part of $-dz$, hence any root z of δ^Q with z^d real is itself real.

Next we check that δ^Q has no positive real roots. Differentiating $\delta^Q(x)$ with respect to x gives the increasing function $d(-1 + cx^{d-1})$ with a unique zero at $c^{-1/(d-1)}$. This gives the location of the minimum of δ^Q on $\mathbb{R}^+$, where the function value is $1 - dc^{-1/(d-1)} + c^{1-d/(d-1)} = 1 - (d-1)/c^{1/(d-1)}$ which is positive because $c > (d-1)^{d-1}$.

If d is even, δ^Q clearly has no negative real roots, hence no real roots at all, finishing the proof in this case. If d is odd δ^Q will have a negative real root u, however because d is odd, the product of the coordinates of $(u, \ldots, u)$ is $u^d < 0$. ◀

We conjecture that the roots of minimal modulus when $c > c_*$ are always a complex conjugate pair, however this determination does not affect our positivity results.

4.3 Proof of Theorem 11

When $c < c_*$ we have seen that there is a single real minimal point $(\rho_c, \ldots, \rho_c)$ in the diagonal direction and that $\rho_c \uparrow 1/(d-1)$ as $c \uparrow c_*^-$. The limit from below in Theorem 11 then follows directly from Theorem 10.

For the limit from above, it suffices to show that in the diagonal direction, for c sufficiently close to c_* and greater, E consists of a single diagonal complex conjugate pair $(\zeta_c, \ldots, \zeta_c)$ and $(\overline{\zeta_c}, \ldots, \overline{\zeta_c})$, and that $\overline{\zeta_c} \to 1/(d-1)$ as $c_* \downarrow c$. First, we check that at $c = c_*$ the unique minimum modulus root of δ^Q is the doubled root at $1/(d-1)$. For $c = c_*$, the first and third terms of $\delta^Q = 1 - dz + c_* z^d$ have modulus 1 and $1/(d-1)$ when $|z| = 1/(d-1)$, respectively, summing to the modulus of the middle term; therefore if $\delta^Q(z) = 0$ and $|z| = 1/(d-1)$ then the third term is positive real. But then the second term must be positive real too, hence the unique solution of modulus at most $1/(d-1)$ is $z = 1/(d-1)$. A quick computation shows the multiplicity to be precisely 2. We know that for $c > c_*$ there are no real roots. Therefore, as c increases from c_*, the minimum modulus doubled root splits into two conjugate roots, which, in a neighborhood of c_*, are still the only minimum modulus roots.

5 Lacuna computations

Theorem 22 is the subject of forthcoming work [4]. Theorem 12 follows immediately, with the specifications: $d \geq 4$ and even, $c = c_*$, $k = 1$, $P = 1$, $Q = Q_{c,d}$, $\mathbf{z}_* = (1/d, \ldots, 1/d)$, $\hat{\mathbf{r}} = (1, \ldots, 1)$, B is the component of the complement of the amoeba of Q containing $(a, \ldots, a)$ for $a < -\log d$, $\mathbf{x}_* = (-\log d, \ldots, -\log d)$, $\mathbf{y}_* = \mathbf{0}$ and $\mathcal{N}$ taken to be the diagonal. Proposition 19 guarantees the correct shape for the tangent cone to Q at $\mathbf{z}_*$.

▶ **Theorem 22.** *Suppose $F = P/Q^k$ with P a holomorphic function and Q a real Laurent polynomial. Fix $\hat{\mathbf{r}} \in \mathbb{RP}^d$, let B be a component of the complement of the amoeba of Q, let $\sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{z}^{\mathbf{r}}$ be the Laurent expansion for F convergent for $\mathbf{z} = \exp(\mathbf{x} + i\mathbf{y})$ and $\mathbf{x} \in B$. Let $\mathbf{x}_* \in \partial B$ be a maximizing point for $\mathbf{r} \cdot \mathbf{x}$ on ∂B. Assume that $\mathcal{V}$ has a unique singularity $\mathbf{z}_* = \exp(\mathbf{x}_* + i\mathbf{y}_*)$, and that the tangent cone of Q at $\mathbf{z}$ transforms by a real linear map to $z_d^2 - \sum_{j=1}^{d-1} z_j^2$. Let $\mathcal{N}$ be any closed cone such that $\mathbf{x}_*$ maximizes $\mathbf{r} \cdot \mathbf{x}$ for all $\mathbf{r} \in \mathcal{N}$.*

If $d > 2k$ is even then there is an $\varepsilon > 0$ and a chain Γ contained in the set $\mathcal{V}_\varepsilon := \{ \mathbf{z} \in \mathcal{V} : |\mathbf{z}^{-\mathbf{r}}| \leq \exp(-\mathbf{r} \cdot \mathbf{x}_* - \varepsilon |\mathbf{r}|) \}$ such that

$$a_{\mathbf{r}} = \int_\Gamma \mathbf{z}^{-\mathbf{r}} \frac{P}{Q^k} \frac{d\mathbf{z}}{\mathbf{z}} \,. \tag{16}$$

In other words, the chain of integration can be slipped below the height of the singular point.

Sketch of proof: Expand the torus $\mathbf{T}$ of integration to $\mathbf{z}_*$ and just beyond. The integral (13) turns into a residue integral over an intersection cycle swept out by the expanding torus; see, e.g. [20, Appendix A.4]. For small perturbations Q_ε of Q, the residue cycle is the union of a sphere surrounding $\mathbf{z}_*$ and a hyperboloid intersecting the sphere. As $Q_\varepsilon \to Q$, this cycle may be deformed so that the sphere shrinks to a point while the hyperboloid's neck also constricts to a point. The hyperboloid may then be folded back on itself so that in a neighborhood of $\mathbf{z}_*$, the chain vanishes, leaving a chain Γ supported below the height of $\mathbf{z}_*$. ◀

───── **References** ─────

1 G. Almkvist, D. van Straten, and W. Zudilin. Generalizations of Clausen's formula and algebraic transformations of Calabi–Yau differential equations. *Proc. Edin. Math. Soc.*, 54:273–295, 2011.

2 R. Askey and G. Gasper. Certain rational functions whose power series have positive coefficients. *Amer. Math. Monthly*, 79:327–341, 1972.

3 R. Askey and G. Gasper. Convolution structures for Laguerre polynomials. *J. D'Analyse Math.*, 31:48–68, 1977.

4 Y. Baryshnikov, S. Melczer, and R. Pemantle. Asymptotics of multivariate sequences in the presence of a lacuna. In preparation, 2018.

5 Y. Baryshnikov and R. Pemantle. Asymptotics of multivariate sequences, part III: quadratic points. *Adv. Math.*, 228:3127–3206, 2011.

6 J. P. Bell and S. Gerhold. On the positivity set of a linear recurrence sequence. *Israel J. Math.*, 157:333–345, 2007.

7 J. Borcea and P. Brändén. The Lee-Yang and Pólya-Schur programs, II: Theory of stable polynomials and applications. *Comm. Pure Appl. Math.*, 62:1595–1631, 2009.

8 G. Christol. Diagonales de fractions rationnelles et equations différentielles. In *Study group on ultrametric analysis, 10th year: 1982/83, No. 2*, pages Exp. No. 18, 10. Inst. Henri Poincaré, Paris, 1984.

9 P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009. URL: `http://algo.inria.fr/flajolet/Publications/books.html`.

10 H. Furstenberg. Algebraic functions over finite fields. *J. Algebra*, 7:271–277, 1967.

11 J. Gillis, B. Reznick, and D. Zeilberger. On elementary methods in positivity theory. *SIAM J. Math. Anal.*, 14:396–398, 1983.

12 M. Hautus and D. Klarner. The diagonal of a double power series. *Duke Math. J.*, 23:613–628, 1971.

13 M. Kauers. Computer algebra and power series with positive coefficients. In *Proc. FPSAC 2007*, 2007.

14 M. Kauers and D. Zeilberger. Experiments with a positivity-preserving operator. *Exper. Math.*, 17:341–345, 2008.

15 T. Koornwinder. Positivity proofs for linearization and connection coefficients of orthogonal polynomials satisfying an addition formula. *J. London Math. Soc. (2)*, 18(1):101–114, 1978.

16 C. Koutschan. HolonomicFunctions (User's Guide). Technical report, no. 10-01 in RISC Report Series, University of Linz, Austria, January 2010.

17 L. Lipshitz. The diagonal of a D-finite power series is D-finite. *J. Algebra*, 113(2):373–378, 1988.

18 S. Melczer. *Analytic combinatorics in several variables: effective asymptotics and lattice path enumeration.* PhD thesis, University of Waterloo, 2017. URL: `https://arxiv.org/abs/1709.05051`.

19 S. Melczer and B. Salvy. Symbolic-numeric tools for analytic combinatorics in several variables. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '16, pages 333–340, New York, NY, USA, 2016. ACM.

20 R. Pemantle and M. Wilson. *Analytic Combinatorics in Several Variables*, volume 340 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, New York, 2013.

21 G. Pólya. Sur les séries entières, dont la somme est une fonction algébrique. *L'Enseignement Mathématique*, 22:38–47, 1921.

22 A. Scott and A. Sokal. Complete monotonicity for inverse powers of some combinatorially defined poynomials. *Acta Math.*, 213:323–392, 2013.

23 A. Straub. Positivity of Szegö's rational function. *Adv. Appl. Math.*, 41(2):255–264, 2008.

24   A. Straub. Multivariate Apéry numbers and supercongruences of rational functions. *Algebra Number Theory*, 8:1985–2008, 2014.

25   A. Straub and W. Zudilin. Positivity of rational functions and their diagonals. *J. Approx. Theory*, 195:57–69, 2015.

26   G. Szegö. Über gewisse Potenzreihen mit lauter positiven Koeffizienten. *Math. Zeit.*, 37:674–688, 1933.

A   Appendix A: Maple Code

Maple worksheets going through the calculations discussed above can be found at `https://github.com/smelczer/SymmetricRationalFunctionsAofA` ; we include the main component of those worksheets, code giving dominant smooth asymptotics, here for archival purposes.

```
smoothASM := proc(G, H, vars, pt)
   local N, i, j, M, HES, C, U, lambda, sbs:
   N := nops(vars) :

   # Get the Hessian determinant of the phase implicitly
   for i from 1 to N do for j from 1 to N do
      U[i, j] := vars[i] · vars[j] · diff(Q, vars[i], vars[j]) :
   od: od:
   lambda := x · diff(Q, x) :
   for i from 1 to N − 1 do for j from 1 to N − 1 do
      if i <> j then M[i, j] := 1 + 1/lambda · (U[i, j] − U[i, N] − U[j, N] + U[N, N]) :
      else M[i, j] := 2 + 1/lambda · (U[i, i] − 2 · U[i, N] + U[N, N]) :
      fi:
   od: od:
   HES := LinearAlgebra[Determinant](Matrix([seq([seq(M[i, j], i = 1..N − 1)], j = 1..N − 1)])) :

   C := simplify(−G/vars[−1]/diff(H, vars[−1]) · HES^(−1/2) · (2 · Pi)^((1 − N)/2));
   sbs := seq(vars[j] = pt[j], j = 1..N) :
   return eval(1/mul(j, j = pt))^n · n^((1 − N)/2) · eval(subs(sbs, C)) :
end:
```

Asymptotic Distribution of Parameters in Random Maps

Olivier Bodini
UMR CNRS 7030, LIPN, Université Paris 13, Av. Jean Baptiste Clément, Villetaneuse, France
olivier.bodini@lipn.fr

Julien Courtiel
UMR 6072, AMACC, Université de Caen Normandie, Bd du Maréchal Juin, Caen, France
julien.courtiel@unicaen.fr

Sergey Dovgal
UMR CNRS 7030, LIPN, Université Paris 13, Av. Jean Baptiste Clément, Villetaneuse, France
IRIF, Université Paris 7, Rue Thomas Mann Paris, France
Moscow Institute of Physics and Technology, Institutskiy per. 9, Dolgoprudny, Russia
dovgal@lipn.fr

Hsien-Kuei Hwang
Institute of Statistical Science, Academia Sinica, 128, Section 2, Academia Rd, Nangang
District, Taipei City, Taiwan.
hkhwang@stat.sinica.edu.tw

── Abstract ──
We consider random rooted maps without regard to their genus, with fixed large number of edges, and address the problem of limiting distributions for six different parameters: vertices, leaves, loops, root edges, root isthmus, and root vertex degree. Each of these leads to a different limiting distribution, varying from (discrete) geometric and Poisson distributions to different continuous ones: Beta, normal, uniform, and an unusual distribution whose moments are characterised by a recursive triangular array.

1 Introduction

Rooted maps form a ubiquitous family of combinatorial objects, of considerable importance in combinatorics, in theoretical physics, and in image processing. They describe the possible ways to embed graphs into compact oriented surfaces [17].

The present paper focuses on asymptotic enumeration of basic parameters in rooted maps with no restriction on genus. From a generating function point of view, if the genus of the maps is not fixed, then the generating function of rooted maps is non-analytic (namely, convergent only at zero) and often satisfies a Riccati differential equation, in contrast to *planar maps* for which analytic (convergent) generating functions abound. The divergent Riccati equations appear frequently in enumerative combinatorics. For example, at least 39 entries in Sloane's OEIS [20] were found containing sequences whose generating functions
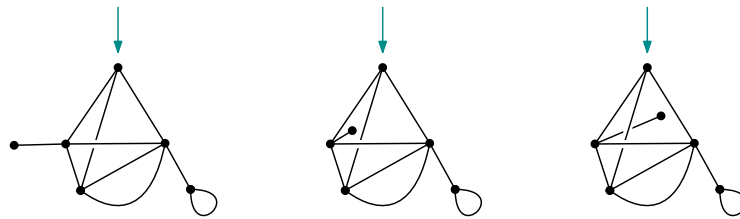
■ **Figure 1** Three rooted maps. Each root is marked by an arrow. The two last maps are equal.

satisfy Riccati equations, including some entries related to the families of indecomposable combinatorial objects, moments of probability distributions, chord diagrams [9, 10, 14], Feynman diagrams [11], etc. Some of these are closely connected to maps. Indeed, it is known that rooted maps with no genus restriction also encode different combinatorial families such as chord diagrams and Feynman diagrams on the one hand, and different fragments of lambda calculus [5, 21] on the other hand. Thus most asymptotic information obtained on maps can often be transferred to the aforementioned objects and lead to a better understanding of them in the corresponding domains.

While the asymptotics and stochastics on planar maps have been extensively studied (see for example [2, 4, 3, 12, 18]), those on rooted maps with no genus restriction have received comparatively much less attention in the literature. Of closest connection to our study here is the paper by Arquès and Béraud [1], which contains several characterisations of the number of rooted maps and their generating functions. In particular, they give an explicit formula for the number of maps, expressed as an infinite sum, from which the asymptotic number of maps with n edges can be deduced (which is $(2n + 1)!!$). Recently, Carrance [7] obtained the distribution of genus in bipartite random maps. To our knowledge, no other asymptotic distribution properties of map statistics have been properly examined so far. Along a different direction, Flajolet and Noy [14] investigated basic statistics on chord diagrams, and Courtiel and Yeats [9] studied the distribution of *terminal chords*.

From an asymptotic point of view, for planar enumeration, as Bender and Richmond put it in [3]: "*The two most successful techniques for obtaining asymptotic information from functional equations of the sort arising in planar enumeration are Lagrange inversion and the use of contour integration.*" An equally useful analytic technique is the saddle-point method as large powers of generating functions are ubiquitous in map asymptotics; see [2, 13] for more detailed information. In contrast, for divergent series, Odlyzko writes in his survey [19]: "*There are few methods for dealing with asymptotics of formal power series, at least when compared to the wealth of techniques available for studying analytic generating functions.*" We show however that a few simple linearizing techniques are very helpful in deriving the diverse limit laws mentioned in the Abstract; the approaches we use may also be of potential application to other closely related problems.

For a rigorous definition of a rooted combinatorial map we refer, for example, to [17, 1]. For our purposes in this extended abstract we use a less formal but more intuitive definition.

▶ **Definition 1** (Maps). A *map* is a connected multigraph endowed with a cyclic ordering of consecutive half-edges incident to each vertex. Multiple edges and loops are allowed. Around each vertex, each pair of adjacent half-edges is said to form a *corner*. If there is only one half-edge, there is only one corner. A *rooted map* is a map with a distinguished corner.

Figure 1 shows some examples of rooted maps. Observe that the first two maps are different since the cyclic ordering is not the same: in the first map, the pendant edge follows counterclockwise the edge after the root (the corner pointed to by an arrow), while in the
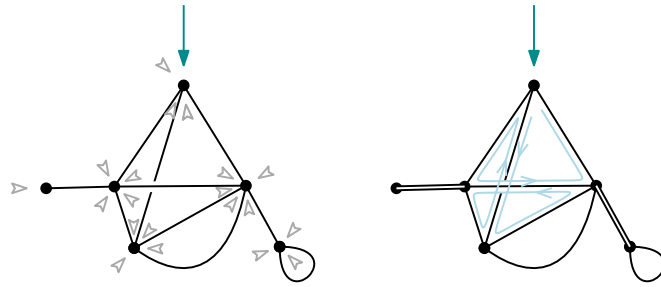
Figure 2 *Left:* The small triangles point at every corner of the map. *Right:* The light-blue line marks the contour of one face of the map. The double-lined edges are the isthmi of the map. The only loop of the map is adjacent to the rightmost isthmus, and the vertex incident to this loop has degree 3.

Table 1 The six map statistics and their limit laws studied in this extended abstract.

Statistics	Differential equation	Mean $\sim$	Limit law
leaves	$L = v + (2 - u)zL + zL^2 + 2z^2\partial_z L + z(1 - v)\partial_v L$	1	Poisson(1)
root isthmic parts	$C = 1 + zC + vzC\|_{v=1}C + 2z^2\partial_z C$	2	Geometric($\frac{1}{2}$)
vertices	$X = v + zX + zX^2 + 2z^2\partial_z X$	$\log n$	$\mathcal{N}(\log n, \log n)$
loops	$Y = v + vzY + vzY\|_{v=1}Y + 2vz^2\partial_z Y + v^2z(vw - 1)\partial_v Y$	$\frac{1}{2}n$	A new law*
root edges	$E = 1 + vzE + vzE\|_{v=1}E + 2vz^2\partial_z E$	$\frac{2}{3}n$	Beta($1, \frac{1}{2}$)
root degree	$D = 1 + v^2zD + vzD\|_{v=1}D + 2vz^2\partial_z D - v^2(1 - v)z\partial_v D$	n	Uniform[0, 2]

second map it precedes in counterclockwise order. In contrast, the last two maps are equal: although the leaves are at different positions, one can find an isomorphism between the two maps preserving the vertices, the root and the cyclic orderings around each vertex. The corners of the leftmost map are displayed in Figure 2 (left), showing all the possible rootings of this map.

▶ **Definition 2** (Map features). A *face* can be obtained by starting at some corner, moving along an incident half-edge, then switching to the next clockwise half-edge and repeating the procedure until the starting corner is met. A *loop* is an edge that connects the same vertex. An *isthmus* is an edge such that the deletion of this edge increases the number of connected components of the underlying graph. The *degree* of a vertex is the number of half-edges incident to this vertex.

These definitions are illustrated in Figure 2 (right).

Arquès and Béraud [1] prove that the generating function of maps $M(z) := \sum_{n\geqslant 0} m_n z^n$, where m_n enumerates the number of maps with n edges, satisfies

$$2z^2 M'(z) = (1 - z)M(z) - 1 - zM(z)^2, \tag{1}$$

a typical Riccati equation whose first few Taylor coefficients read $M(z) = 1 + 2z + 20z^2 + 444z^3 + 16944z^4 + \cdots$.

We address in this paper the analysis of the extended equations of (1) for bivariate (and in one case, trivariate) generating functions $M(z, v) := \sum_{n,k\geqslant 0} m_{n,k} z^n v^k$, where $m_{n,k}$ stands
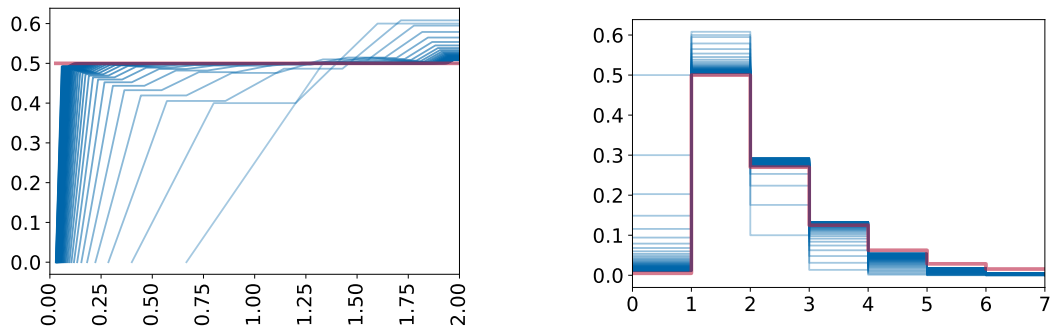
Figure 3 *Left:* Root vertex degree. *Right:* Number of root isthmic parts.
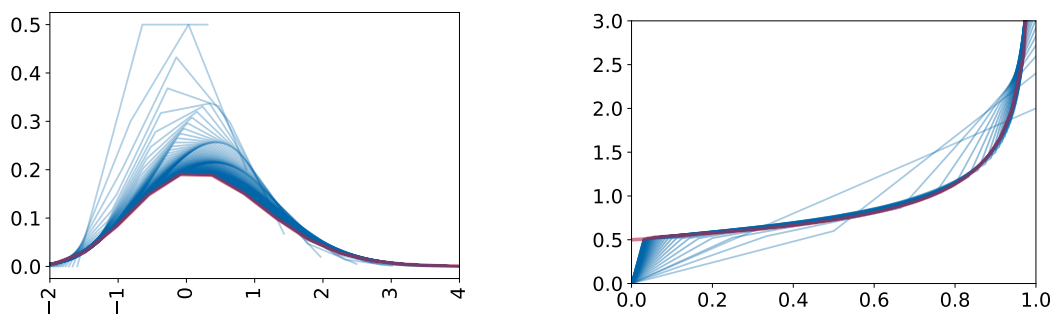
Figure 4 *Left:* Number of vertices. *Right:* Number of root edges.

for the number of maps with n edges and the value of the shape parameter equal to k. We obtain limit laws for the distributions of six different parameters (see Figures 3 to 5).

We collect the statistics and their limit laws studied here in Table 1 for comparison. We see that some of the limit laws are discrete (Poisson and Geometric), one of them (the number of vertices) is Gaussian with a logarithmic mean, which is denoted by $\mathcal{N}(\log n, \log n)$, and the others are continuous. For the number of root edges, root degree and loops, the corresponding limit laws are normalized by n, the total number of edges. The distribution of the number of loops follows a rather unusual limit law (see Figure 5) in the sense that we can only characterise the limit law by its moment sequence, η_l, which satisfies $\eta_l = \eta_{0,l}$ with $\eta_{k,l}$ computable only through a recurrence involving $\eta_{k-1,l}$ and $\eta_{k+1,l-1}$. The corresponding probability density function of this law remains unknown and does not have an explicit expression at this stage (see Figure 5). Finally, by the bijection from [10] and a known property of chord diagrams in [14], it is possible to deduce the limit laws for the number of leaves.

One technique we use several times in our proofs consists in linearising the differential equations satisfied by the generating functions, by choosing a suitable transformation, inspired from the resolution of Riccati equations. Once the dominant term is identified, the analysis for the limit law becomes more or less straightforward. When such a technique fails, we rely then on the method of moments, which establishes weak convergence by computing all higher derivatives of $M(z, v)$ at $v = 1$ and by examining asymptotically the ratios $[z^n]\partial_v^k M(z, v)|_{v=1}/[z^n]M(z, 1)$ (which correspond to factorial moments of random variable). Such a procedure also linearises to some extent the more complicated bivariate nature of the differential equations and facilitates the resolution complexity of the asymptotic problem.
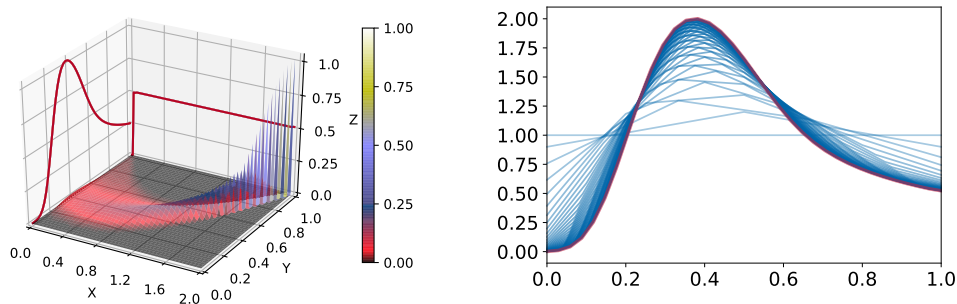
Figure 5 *Left:* Joint distribution of root vertex degree and the number of loops. *Right:* Number of loops.
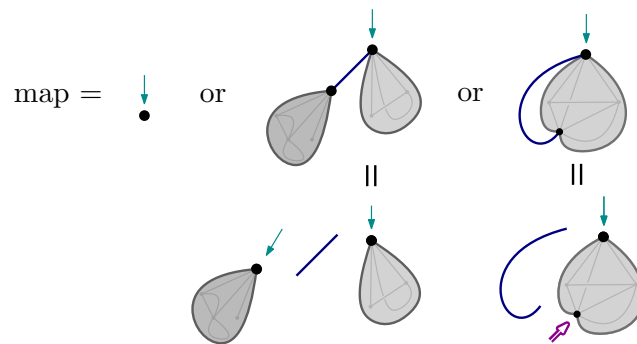
Figure 6 A symbolic construction of rooted maps.

Structure of the Paper.   In Section 2 we derive the nonlinear differential equations satisfied by the generating functions of the map statistics. Then in Section 3 we sketch the proofs for the limit laws of five statistics based on generating functions. The Poisson law for the number of leaves (together with the root face degree and the number of trivial loops) will be proved by a direct combinatorial approach in the last section.

2   Differential equations for maps

In this section, we derive the differential equations satisfied by the bivariate or trivariate generating functions with the additional variable(s) marking the shape statistics.

Univariate generating function of maps.   Since the Riccati equation (1) lies at the basis of all other extended equations in Table 1, we give a quick proof of it via the recurrence satisfied by m_n, the number of maps with n edges (see Figure 6):

$$m_n = \mathbf{1}_{[n=0]} + \sum_{0 \leqslant k < n} m_k m_{n-1-k} + (2n-1)m_{n-1}, \tag{2}$$

which then implies the Riccati equation (1).

First, $m_0 = 1$ because there is only one map with 0 edges. Then a map with n edges can be formed either by connecting the roots of two maps (with k and $n - k - 1$ edges, respectively) with an isthmus, or by adding an edge to a map with $n - 1$ edges, connecting

the root and a corner. The number of possible ways to insert an edge in this way is equal to $2n - 1$, because there are $2n - 2$ corners in a map of size $n - 1$, and there are two possible ways to insert a new edge at the root corner (either before, or after the root). This proves (2).

Vertices.   Consider now the bivariate generating function $X(z, v) = \sum_{n,k \geqslant 0} x_{n,k} z^n v^k$, where $x_{n,k}$ is equal to the number of rooted maps with n edges and k vertices. Arquès and Béraud [1] showed that

$$X(z, v) = v + zX(z, v) + zX(z, v)^2 + 2z^2 \partial_z X(z, v). \tag{3}$$

This recurrence can be obtained from (2) by noticing that no new vertex is created when we connect two maps with an isthmus, nor when we add a new root edge to a map. Note that $X(z, v)$ satisfies another functional equation (see [1])

$$X(z, v) = v + zX(z, v)X(z, v + 1),$$

which seems less useful from an asymptotic point of view.

Root isthmic parts.   We count here the *root isthmic parts*, which are the number of isthmic constructions used at the root vertex. Note that an isthmic part may not be a bridge because the additional edge constructor may induce additional connections. Then the bivariate generating function $C(z, v) = \sum_{n,k \geqslant 0} c_{n,k} z^n v^k$, where $c_{n,k}$ enumerates the number of maps with n edges and k root isthmic parts, satisfies

$$C(z, v) = 1 + zC(z, v) + vzC(z, v)C(z, 1) + 2z^2 \partial_z C(z, v). \tag{4}$$

In Figure 6, the number of root isthmic parts only changes whenever two maps are connected by an isthmus. This yields $vzC(z, v)C(z, 1)$ instead of zC^2.

Root edges.   Similarly, consider $E(z, v) = \sum_{n,k \geqslant 0} e_{n,k} z^n v^k$, where $e_{n,k}$ counts the number of rooted maps with n edges and k root edges. Then $E(z, v)$ satisfies

$$E = 1 + vzE + vzE|_{v=1}E + 2vz^2 \partial_z E. \tag{5}$$

This again results from the recurrence (2) and from Figure 6: the non-root edges come from the bottom map in the isthmic construction, yielding the term $vzE(z, v)E(z, 1)$.

Root Degree.   Consider the degree of the root vertex. Note that this may be different from the number of root edges because for the root degree, each loop edge is counted twice, therefore the degree of the root vertex varies from 0 to $2n$. By duality, the distribution of the root face degree is the same as the distribution of the root vertex degree.

Let $D(z, v) = \sum_{n,k \geqslant 0} d_{n,k} z^n v^k$ denote the bivariate generating function for maps with variable v marking root degree. Then

$$D = 1 + v^2 zD + vzD|_{v=1}D + 2vz^2 \partial_z D - v^2(1 - v)z\partial_v D. \tag{6}$$

In this case, the original construction in Figure 6 is insufficient, and we need to consider further cases in Figure 7. When an additional edge becomes a loop, it increases the degree of the root vertex by 2; otherwise, the root degree is increased merely by 1. Note that the equation (6) is now a *bona fide* partial differential equation, making the analysis more difficult.
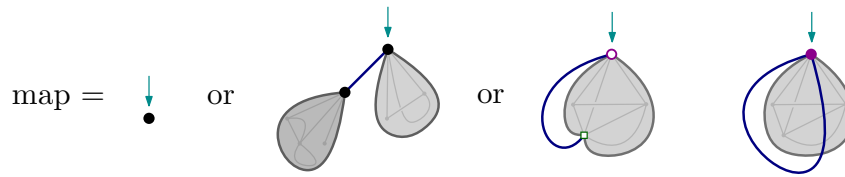
Figure 7 Symbolic method to count root degree and loops in rooted maps.

Leaves. The differential equation for the bivariate generating function of maps with variable v marking leaves (see Table 1) can be obtained in a similar way by considering different cases in the new edge constructor. The number of special leaf corners is equal to the number of leaves.

Loops. Finally, we look at the number of loops whose enumeration necessitates the consideration of the joint distribution of the number of loops and the number of root edges, namely, we consider the trivariate generating function $Y(z,v,w) = \sum_{n,k,m} y_{n,k,m} z^n v^k w^m$, where $y_{n,k,m}$ denotes the number of rooted maps with n edges, root degree equal to k, and m loops. Then $Y(z,v,w)$ satisfies a partial differential equation

$$Y = 1 + zvY + zvY|_{v=1}Y + 2z^2 v \partial_z Y + zv^2(vw-1)\partial_v Y. \tag{7}$$

As in the symbolic construction of Figure 7, a new edge becomes a loop only if it is attached to one of the corners incident to the root vertex. The differential equation (7) is then a modification of (6) with an additional variable marking the number of loops.

Note that Equation (7) is *catalytic* with respect to the variable v, i.e. putting $v = 1$ introduces a new unknown object $\partial_v Y|_{v=1}$ to the differential equation. One of the strategies for dealing with catalytic equations was developed by Bousquet-Mélou and Jehanne [6], generalising the so-called kernel method and quadratic method. However, their method does not work in our case because our equation is differentially algebraic.

3 Limit laws

This section describes the techniques we employ to establish the limit laws.

From now on, by a random map (with n edges) we assume that all rooted map with n edges are equally likely. For notational convention, we use $X' = \partial_z X$ to denote derivative with respect to z. Due to space limit, we give only the sketches of the proofs.

3.1 Transformation into a linear differential equation

For most of the equations in the previous section, it turns out that a transformation similar to that used for Riccati equations largely simplifies the resolution and leads to solvable recurrences, which are then suitable for our asymptotic purposes. We begin by solving the standard Riccati equation (1) and see how a similar idea extends to other differential equations.

▶ **Proposition 3.** *The number m_n of maps with n edges satisfies*

$$\frac{m_n}{\phi_n} = 2n - 1 + O(n^{-1}), \qquad where \quad \phi_n = \frac{(2n)!}{2^n n!} = (2n-1)!!. \tag{8}$$

Proof. We solve the Riccati equation (1) by considering the transformation

$$M(z) = 1 + \frac{2z\phi'(z)}{\phi(z)}, \tag{9}$$

for some function $\phi(z)$ with $\phi(0) = 1$. Substituting this form into the equation (9), we get the second-order differential equation $2z^2\phi'' + (5z - 1)\phi' + \phi = 0$. From this equation, the coefficients $\phi_n := [z^n]\phi(z)$ satisfy the recurrence $\phi_{n+1} = (2n+1)\phi_n$, which implies the double factorial form of ϕ_n by $\phi_0 = 1$.

Moreover, by extracting the coefficient of z^n in (9), we obtain a relation between the coefficients m_k and ϕ_ℓ. By the inequality $m_n \geqslant (2n - 1)m_{n-1}$ (see (2)), we then deduce the asymptotic relation (8). ◀

▶ **Theorem 4.** *Let X_n denote the number of vertices in a random rooted map with n edges. Then X_n follows a central limit theorem with logarithmic mean and logarithmic variance:*

$$\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} \xrightarrow{d} \mathcal{N}(0,1), \quad \mathbb{E}(X_n) \sim \log n, \quad \mathbb{V}(X_n) \sim \log n. \tag{10}$$

Proof. Similar to (9), we define a bivariate generating function $S(z,v) = \sum_{n\geqslant 0} s_n(v)z^n$ such that

$$X(z,v) = v + \frac{2zS'}{S}, \quad S(0) = 1.$$

Substituting this $X(z,v)$ into (3) leads to a linear differential equation from which one can extract the recurrence

$$s_n(v) = \frac{(2n + v - 2)(2n + v - 1)}{2n} s_{n-1}(v).$$

We then get an explicit expression for $s_n(v)$, from which we deduce, by singularity analysis, that

$$\mathbb{E}(v^{X_n}) = \frac{2^{v-1}}{\Gamma(v)} n^{v-1}(1 + O(n^{-1})),$$

and conclude by applying the Quasi-Powers Theorem [13, 15]. ◀

A finer Poisson$(\log n + c)$ approximation, for a suitably chosen c, is also possible, which results in a better convergence rate $O(\log n)^{-1}$ instead of $(\log n)^{-\frac{1}{2}}$; see [16] for details.

▶ **Theorem 5.** *Let C_n denote the number of root isthmic parts in a random rooted map with n edges. Then,*

$$C_n \xrightarrow{d} \mathrm{Geometric}\left(\tfrac{1}{2}\right).$$

Proof. Since $C(z,1) = M(z)$, we use again the substitution (9) and apply it to (4):

$$2z^2(\phi C' + v\phi'C) = (1 - (1 + v)z)\phi C - \phi.$$

The trick here is to multiply both sides by $\phi(z)^{v-1}$ and set $Q(z,v) = \phi(z)^v C(z,v)$. We then obtain

$$2z^2 Q' = (1 - (1 + v)z)Q - \phi^v.$$

Using the recurrence for the normalised coefficients $\widehat{q}_n(v) := q_n(v)/\phi_n$ and dominant-term approximations, we find that the n-th coefficient of Q is proportional to

$$\widehat{q}_n(v) = \frac{v}{2n} \sum_{1 \leqslant k \leqslant n} \left(\frac{n}{k}\right)^{v/2} + O(n^{-1/2}) = \frac{v}{2-v} + O(n^{-\frac{1}{2}}).$$

This corresponds to a (shifted by 1) geometric distribution with parameter $\frac{1}{2}$. By the definition $Q(z, v) = \phi(z)^v C(z, v)$, we deduce that the limiting distribution of C_n is also geometric with parameter $\frac{1}{2}$. ◀

▶ **Theorem 6.** *Let E_n denote the number of edges incident to the root vertex in a random rooted map with n edges. Then E_n follows asymptotically a Beta distribution:*

$$\frac{E_n}{n} \xrightarrow{d} \mathrm{Beta}\left(1, \tfrac{1}{2}\right), \tag{11}$$

with the density function $\frac{1}{2}(1-t)^{-\frac{1}{2}}$ for $t \in [0, 1)$.

Proof. We use again the substitution $E(z, 1) = M(z) = 1 + 2z\frac{\phi'}{\phi}$ in (5), giving

$$2vz^2(\phi E' + \phi' E) = (1 - 2vz)\phi E - \phi.$$

With $Q(z, v) = \phi(z)E(z, v)$, we then obtain

$$2vz^2 Q' = (1 - 2vz)Q - \phi. \tag{12}$$

This linear differential equation translates into a recurrence for the coefficients $q_n(v)$ of $Q(z, v)$, which yields the closed-form expression

$$q_n(v) = 2^n n! \sum_{0 \leqslant j \leqslant n} \binom{2j}{j} 4^{-j} v^{n-j}. \tag{13}$$

Returning to $E(z, v)$, we see that its coefficients behave asymptotically like $q_n(v)$. This implies the Beta limit law (11) for the random variable E_n/n since $\binom{2j}{j} 4^{-j} \sim (\pi j)^{-1/2}$ for large j. ◀

▶ **Theorem 7.** *Let D_n denote the degree of the root vertex in a random rooted map with n edges. Then, D_n, divided by the number of edges, converges in law to the uniform distribution on $[0, 2]$:*

$$\frac{D_n}{n} \xrightarrow{d} \mathrm{Uniform}\,[0, 2]\,. \tag{14}$$

Proof. The substitutions

$$D(z, 1) = M(z) = 1 + \frac{2z\phi'}{\phi}, \quad \text{and} \quad D(z, v) = \frac{Q(z, v)}{\phi(z)}$$

lead to a partial differential equation, which in turn yields the recurrence for the coefficients $q_n(v) := [z^n]Q(z, v)$:

$$q_n(v) = v(2n - 1 + v)q_{n-1} - v^2(1 - v)q'_{n-1}(v) + \phi_n.$$

We then get the exact solution $q_n(v) = \phi_n(1 + v + \cdots + v^{2n})$. Accordingly, $d_n(v) := [z^n]D(z, v) \sim q_n(v)$. This implies the uniform limit law (14). ◀

A more intuitive interpretation of this uniform limit law is given in the next section.

3.2   Approximation and method of moments

Unlike all previous proofs, we use the method of moments to establish the limiting distribution of the number of loops. The situation is complicated by the presence of the term involving $\partial_v Y$ in (7), which introduces higher order derivatives with respect to v at $v = 1$ when computing the asymptotic of the moments.

▶ **Theorem 8.** *Let Y_n denote the total number of loops in a random rooted map with n edges. Then*

$$\frac{Y_n}{n} \xrightarrow{d} \mathcal{L}, \tag{15}$$

where $\mathcal{L}$ is a continuous law with a computable density on $[0, 1]$.

Proof. First, we show by induction that there exist constants $\eta_{k,\ell}$, such that as $n \to \infty$,

$$[z^n]\partial_v^k \partial_w^\ell Y(z, v, w)\big|_{v=w=1} \sim \eta_{k,\ell} n^{k+\ell+1} \phi_n, \quad k, \ell \geqslant 0. \tag{16}$$

For $k = \ell = 0$ the statement clearly holds. Let $y_n^{(k,\ell)} := [z^n]\partial_v^k \partial_w^\ell Y(z, v, w)\big|_{v=w=1}$ for larger $k, \ell \geqslant 0$. By translating (7) into the corresponding recurrence for the coefficients and by collecting the dominant terms (using the induction hypothesis (16)), we deduce that

$$y_n^{(k,\ell)} \sim (2n+k)y_{n-1}^{(k,\ell)} + \ell y_{n-1}^{(k+1,\ell-1)} + (2kn - 2k)y_{n-1}^{(k-1,\ell)} + \mathbf{1}_{[k=0]}y_{n-1}^{(k,\ell)}.$$

Accordingly, we are led to the recurrence

$$\eta_{k,\ell} = \frac{1}{k + 2\ell + \mathbf{1}_{[k>0]}}(2k\eta_{k-1,\ell} + \ell\eta_{k+1,\ell-1}),$$

for $k + \ell > 0$ (provided that we interpret $\eta_{k,\ell} = 0$ when any index becomes negative). In particular, when $\ell = 0$, we obtain the moments of the random variable E_n, the number of root edges: $\eta_{k,0} = \frac{2^{k+1}}{k+1}$, which coincides with the moments of the uniform random variable Uniform$[0, 2]$. Finally, it is not complicated to check that the numbers $\eta_{0,\ell}$ satisfy the condition of Hausdorff moment problem, i.e. $\eta_{0,\ell}$ uniquely determine the limiting random variable defined on $[0, 1]$. ◀

4   Combinatorics of map statistics

We examine briefly the combinatorial aspect of the map statistics, relying our arguments on the close connection between maps and chord diagrams (see [8]).

Recall that a *chord diagram* [14] with n chords is a set of vertices labelled with the numbers $\{1, 2, \ldots, 2n\}$ equipped with a perfect matching. A chord diagram is *indecomposable* if it cannot be expressed as a concatenation of two smaller diagrams.

Why the root degree follows a uniform law?   We begin with Cori's bijection [8] between rooted maps and indecomposable diagrams. In this bijection, each chord connecting labels i and j corresponds to matching of the half-edges with labels i and j. The set of half-edges incident to each vertex of the resulting map corresponds to the set of nodes to the right of the starting points of the so-called *outer chords*, i.e. chords that do not lie under any other chord.

▶ **Proposition 9.** *There exists a bijection between rooted maps of root degree d with n edges, and indecomposable diagrams with $n + 1$ chords such that the vertex $d - 2$ is matched with vertex 1.*
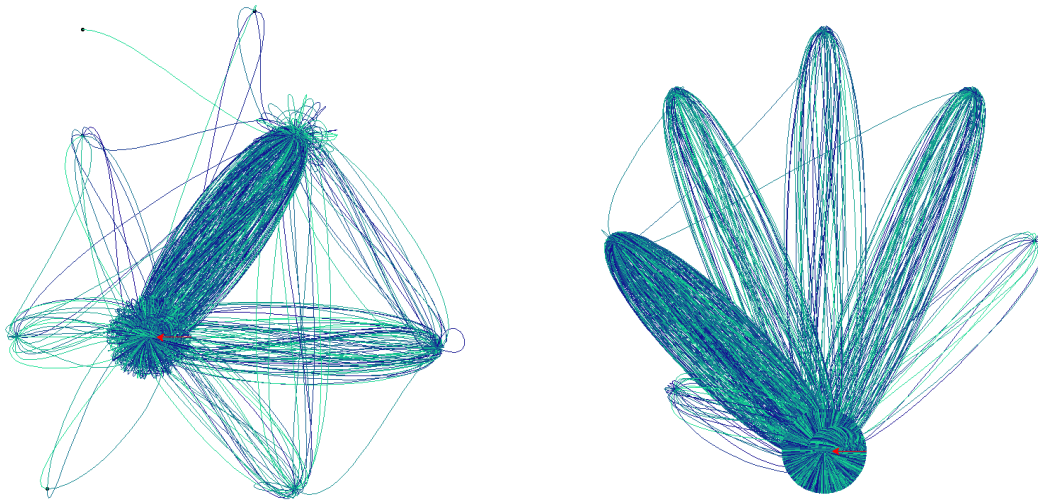
■ **Figure 8** Random rooted maps, respectively with 1000 and 20000 edges.

Once this proposition is available, it leads to a simpler and more intuitive proof of Theorem 7 as follows. In a (not necessarily indecomposable) diagram, the label of the vertex matched with 1 follows exactly a uniform law on $\{2, \ldots, 2n\}$. But a diagram is almost surely an indecomposable diagram (because its cardinality is asymptotically the same); thus the label of the vertex matched with 1 divided by $2n$ obeys asymptotically a uniform law on $[0, 1]$ (or Uniform$[0, 2]$ if divided by n as in Theorem 7).

Uniform random generation.   Cori's bijection is also useful for generating random rooted maps. Uniformly sampling a random diagram can be achieved by adding the chords sequentially one after another. If this procedure results in a decomposable diagram, it is rejected (which occurs with asymptotic probability 0). A successful sampled diagram is then transformed into a map using Cori's bijection [8]. Figure 8 shows two instances of random maps thus generated.

The number of leaves.   Another bijection in [10] is useful in proving the Poisson limit law of the number of leaves. This bijection sends leaves of a map into the *isolated* chords (namely, edges connecting vertices k and $k + 1$) of an indecomposable chord diagram. According to [14, Theorem 2], the number of isolated edges in a random chord diagram has a Poisson distribution with parameter 1. We can then deduce the following theorem.

▶ **Theorem 10.** *The number of leaves in a random map with n edges follows asymptotically a Poisson law with parameter* 1.

Two dual parameters.   We briefly remark that two other parameters, namely *root face degree* and the number of *trivial loops* do not seem easily dealt with by the method of generating functions because marking them requires additional nested information such as the degrees of all the faces. However, such parameters can be easily marked in their corresponding dual maps. Their limit distributions are uniform and Poisson, respectively.

References

1   Didier Arquès and Jean-François Béraud. Rooted maps on orientable surfaces, Riccati's equation and continued fractions. *Discrete mathematics*, 215(1-3):1–12, 2000.

2   Cyril Banderier, Philippe Flajolet, Gilles Schaeffer, and Michele Soria. Random maps, coalescing saddles, singularity analysis, and Airy phenomena. *Random Structures & Algorithms*, 19(3-4):194–246, 2001.

3   Edward A Bender and L.Bruce Richmond. A survey of the asymptotic behaviour of maps. *Journal of Combinatorial Theory, Series B*, 40(3):297–329, 1986. `doi:10.1016/0095-8956(86)90086-9`.

4   Olivier Bernardi and Mireille Bousquet-Mélou. Counting coloured planar maps: differential equations. *Communications in Mathematical Physics*, 354(1):31–84, 2017.

5   Olivier Bodini, Danielle Gardy, and Alice Jacquot. Asymptotics and random sampling for BCI and BCK lambda terms. *Theoretical Computer Science*, 502:227–238, 2013.

6   Mireille Bousquet-Mélou and Arnaud Jehanne. Polynomial equations with one catalytic variable, algebraic series and map enumeration. *Journal of Combinatorial Theory, Series B*, 96(5):623–672, 2006.

7   Ariane Carrance. Uniform random colored complexes. *arXiv preprint arXiv:1705.11103*, 2017.

8   Robert Cori. Indecomposable permutations, hypermaps and labeled Dyck paths. *Journal of Combinatorial Theory, Series A*, 116(8):1326–1343, 2009.

9   Julien Courtiel and Karen Yeats. Terminal chords in connected chord diagrams. *Annales de l'Institut Henri Poincaré D*, 4(4):417–452, 2017.

10   Julien Courtiel, Karen Yeats, and Noam Zeilberger. Connected chord diagrams and bridgeless maps. *arXiv preprint arXiv:1611.04611*, 2016.

11   Predrag Cvitanović, B. Lautrup, and Robert B. Pearson. Number and weights of Feynman diagrams. *Phys. Rev. D*, 18:1939–1949, Sep 1978. `doi:10.1103/PhysRevD.18.1939`.

12   Michael Drmota and Konstantinos Panagiotou. A central limit theorem for the number of degree-k vertices in random maps. *Algorithmica*, 66(4):741–761, 2013.

13   P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Camb. Univ. Press, 2009.

14   Philippe Flajolet and Marc Noy. Analytic combinatorics of chord diagrams. In *Formal Power Series and Algebraic Combinatorics*, pages 191–201. Springer, 2000.

15   Hsien-Kuei Hwang. On convergence rates in the central limit theorems for combinatorial structures. *European Journal of Combinatorics*, 19(3):329–343, 1998.

16   Hsien-Kuei Hwang. Asymptotics of poisson approximation to random discrete distributions: an analytic approach. *Advances in Applied Probability*, 31(2):448–491, 1999.

17   Sergei K. Lando and Alexander K. Zvonkin. *Graphs on Surfaces and Their Applications*. Number 141 in Encyclopaedia of Mathematical Sciences. Springer-Verlag, 2004.

18   Valery A Liskovets. A pattern of asymptotic vertex valency distributions in planar maps. *Journal of Combinatorial Theory, Series B*, 75(1):116–133, 1999.

19   Andrew M Odlyzko. Asymptotic enumeration methods. *Handbook of combinatorics*, 2(1063-1229):1229, 1995.

20   N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences. URL: `http://oeis.org`.

21   Noam Zeilberger and Alain Giorgetti. A correspondence between rooted planar maps and normal planar lambda terms. *Logical Methods in Computer Science*, 11(3:22), 2015.

Beyond Series-Parallel Concurrent Systems: The Case of Arch Processes

Olivier Bodini
Laboratoire d'Informatique de Paris-Nord, CNRS UMR 7030 - Institut Galilée - Université Paris-Nord, 99, avenue Jean-Baptiste Clément, 93430 Villetaneuse, France.
Olivier.Bodini@lipn.univ-paris13.fr

Matthieu Dien
Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan.
dien@stat.sinica.edu.tw

Antoine Genitrini
Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6 -LIP6- UMR 7606, F-75005 Paris, France.
Antoine.Genitrini@lip6.fr

Alfredo Viola
Universidad de la República, Montevideo, Uruguay.
viola@fing.edu.uy

Abstract

In this paper we focus on concurrent processes built on *synchronization* by means of *futures*. This concept is an abstraction for processes based on a main execution thread but allowing to delay some computations. The structure of a general concurrent process is a directed acyclic graph (DAG). Since the quantitative study of increasingly labeled DAG (directly related to processes) seems out of reach (this is a ♯P-complete problem), we restrict ourselves to the study of *arch processes*, a simplistic model of processes with futures. They are based on two parameters related to their sizes and their numbers of arches. The increasingly labeled structures seems not to be specifiable in the classical sense of Analytic Combinatorics, but we manage to derive a recurrence equation for the enumeration.

For this model we first exhibit an *exact* and an *asymptotic* formula for the number of runs of a given process. The second main contribution is composed of a *uniform random sampler* algorithm and an *unranking* one that allow efficient generation and exhaustive enumeration of the runs of a given arch process.

1 Introduction

Concurrent processes are logic units running independently in the same environment which share resources (processing time, file inputs or outputs, etc). To guarantee the good behavior of a concurrent program (a set of processes), a mechanism of *synchronization* has to be set up. For example, synchronization can be used to avoid a process to write in a file currently read by another process.

To deal with concurrent programs, researchers in the concurrency theory community formalize the programs in an abstract language called *process algebra*. Different formalisms exhibit different properties of concurrent programs. We are mainly interested in the so-called *Calculus of Communicating Systems* introduced in [18], because of its popularity in concurrency theory and the simplicity to reason about.

In this context, one of the main goals in concurrency is to check the good behavior of such programs. A very popular method to do the verification is the *model checking*: several logical properties (the specification of the program behavior) are checked for all the possible runs of the program (see [3] as a reference book).

A common problem in such a method is the *combinatorial explosion* phenomenon: the huge number of runs to check. To deal with that explosion, a statistical method has been introduced: the *Monte-Carlo model checking* (see [13]). Here, the idea is to check the specification only for few runs randomly sampled. Thus, the result of the method is not anymore a proof of good behavior but a statistical certificate.

In this paper we investigate this phenomenon from a combinatorial point of view. We consider concurrent programs as sets of *atomic* actions (executed computations) constrained by a partial order relation: some actions have to finish before others start their computations. Thus in this setting, a concurrent program can be modeled as a partial order (*a.k.a.* poset). Then, its runs (possible execution flows) are its linear extensions (i.e. the total orders compatible with the partial order) and the combinatorial explosion phenomenon (for a given family of posets) is the fast growth of the number linear extensions as its number of atomic actions increases.

The problem of counting the number of linear extensions of a poset is known to be ♯P-complete [10]. As a consequence, an analytic approach to study this counting problem for general posets seems out of reach. We thus limit the difficulty by considering restricted classes of partial orders. In previous works we dealt with tree-like processes [9], tree-like processes with non-deterministic choice [8] and Series-Parallel processes [5, 7]. In these papers, like in the present one, we take the point of view to model a partial order by its covering directed acyclic graph – DAG – (*a.k.a.* Hasse diagram). Then a linear extension becomes an increasing labeling of the covering DAG. Previously this consideration let us to use symbolic method to specify our models and so to use tools of Analytic Combinatorics (see [11]).

In the present work we focus on processes built on *synchronization* by means of *futures* or *promises* (see [4]). This concept is an abstraction for processes based on a main execution thread but allowing to delay some computations. These computations are run asynchronously and are represented as an object that can be queried in two ways: `finish?` to know if the computation has terminated and `get` to retrieve the result of the computation (and properly proceed the synchronization). This quite old principle aroses recently in many programming languages, especially in the very popular Javascript language (see [1]).

To emphasize these paradigm we consider *arch processes*: a simplistic model of processes with futures. An arch process is composed of a main *trunk* from which start several *arches*
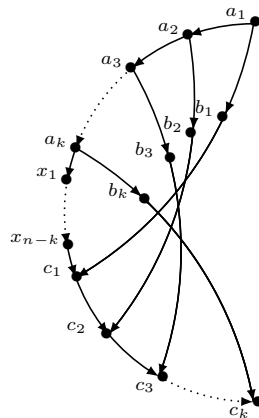
(modeling futures). The general shape of such a process is given in the Figure 1. Arch processes are based on two parameters related to their sizes and their numbers of arches. A combinatorial and recursive specification (as in [11]) for these increasing labeled structures seems out of the reach at the moment. As a consequence we present here a different approach to specify the problem.

For this model we exhibit *exact* and *asymptotic* formula for the number of increasing labelings.As a second main contribution is the design of two algorithms. The first one is an uniform random sampler for runs of a given arch process and the second one is an *unranking* algorithm which allows to obtain an exhaustive builder of runs. The design of these algorithms is motivated by the possible applications to (statistical) model checking.

The paper is organized as follows. The next section is devoted to the formal description of (n,k)-arch processes and gives the solution of the recurrence equation driving their numbers of runs. In Section 3 we prove the algebraicity of the bivariate generating function, we give a closed form formula for it and the asymptotic behaviors of the diagonal coefficients of the functions. Section 4 carefully describes both algorithms.

2   The arch processes and their runs

A concurrent program is seen as a *partially ordered set (poset)* of atomic actions where the order relation define the precedence constraints over the executions of the actions.
A run of a concurrent program is a *linear extension* of the corresponding poset: i.e. a total order compatible with the partial order relation.

Note that many other models of concurrent program exist but we have chosen to use this one because it is well-suited to study the combinatorial explosion phenomenon.
We introduce now the model of arch processes, a family of restricted concurrent program encoding synchronization by means of futures.

▶ **Definition 1.** Let n and k be two positive integers with $k \leq n+1$. The (n,k)-*arch process*, denoted by $A_{n,k}$, is built in the following way:
■ the trunk of the process: a sequence of $(n+k)$ actions $a_1, \ldots, a_k, x_1, \ldots, x_{n-k}, c_1, \ldots, c_k$
   and represented in Figure 1 on a semicircle;
■ the k arches that correspond to the triplets, for all $i \in \{1, \ldots, k\}$, $a_i \to b_i \to c_i$.
Thus k is the number of arches in the process, and n is the length (along the trunk) between both extremities of each arch a_i and c_i (for all i). There are two extreme cases: when $k = n$,
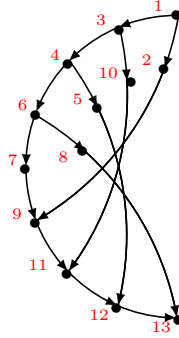
■ **Figure 2** A run of the $(5, 4)$-arch process.

it corresponds to the arch processes that do not contain any node x_i in the trunk, and the case $k = n + 1$ that corresponds to the case where both the nodes a_k and c_1 are merged into a single node (and thus there is no node x_i).

In Figure 1 representing the (n, k)-arch process, the precedence constraints are encoded with the directed edges such that $a \to b$ means that the action a precedes b. We remark that the (n, k)-arch process contains exactly $(n + 2k)$ actions. Due to the intertwining of the arches, we immediately observe when k is larger than 1 then the arch processes are not Series-Parallel processes. Hence the results we exhibited in our papers [6, 7] cannot be applied in this context.

▶ **Definition 2.** An *increasing labeling* for a concurrent process containing ℓ actions is a bijection between the integers $\{1, \ldots, \ell\}$ and the actions of the process, satisfying the following constraint: if an action a precedes an action b then the label associated to a is smaller than the one related to b.

In Figure 2 we have represented an increasing labeling of the $(5, 4)$-arch process $A_{5,4}$ corresponding to the run $\langle a_1, b_1, a_2, a_3, b_3, a_4, x_1, b_4, c_1, b_2, c_2, c_3, c_4 \rangle$. As one can see, every directed path (induced by the precedence relation) is increasingly labeled. Our quantitative goal is to calculate the number of runs for a given arch process.

▶ **Proposition 3.** *The number of runs of a concurrent process is the number of increasing labelings of the actions of the process.*

Thus, each increasing labeling is in bijection with a single linear extension.

While there is the classical hook-length formula for tree-processes [15, 9] and its generalization for Series-Parallel processes [6], to the best of our knowledge, no closed form formula is known for more general classes of processes. In the rest of the paper, for a given process A, we denote by $\sigma(A)$ its number of runs.

First, let us easily exhibit a lower bound and an upper bound (in the case $k < n + 1$) in order to obtain a first idea for the growth of the numbers of runs for the arch processes. We remark that a similar approach could be used for the case when $k = n+1$. We first enumerate the runs where all the b_i nodes are preceded by a_k, and all of them precede the node c_1. This imposes new precedence constraints for the process, and thus its number of runs is a lower bound for the total number of runs. In this case the b_i's permute without any constraint, i.e. $k!$ possibilities and then each permutation of the b_i's shuffles with the sequence $x_1, \ldots, x_{n-k}$. Thus we get the following lower bound for the number of runs of $A_{n,k}$:

$$\sigma(A_{n,k}) \geq k! \binom{k + n - k}{k} = \frac{n!}{(n - k)!}.$$
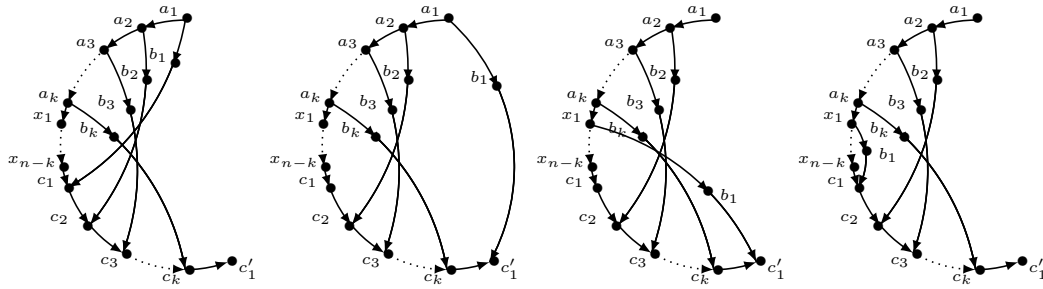
Figure 3 From left to right, the processes denoted $D_{n,k}$, $\overline{D}_{n,k}$, $\overline{D}^1_{n,k}$ and $\overline{D}^2_{n,k}$.

We now focus on an upper bound for the number of runs of $A_{n,k}$. Here again we suppose that all the permutations of the b_i's are possible, but we allow each b_i to appear everywhere between a_1 and c_k. This constraint is satisfied by all the runs, but some possibilities are not valid runs: thus we are computing an upper bound. Once the permutation of the b_i's is calculated, we shuffle it into the trunk (containing $n + k$ nodes):

$$\sigma(A_{n,k}) \leq k!\binom{k+n+k-1}{n+k-1} = \frac{(n+2k-1)!}{(n+k-1)!}.$$

A refinement of these ideas for the bounds computation allows to exhibit a recurrence formula for the value $\sigma(A_{n,k})$.

▶ **Theorem 4.** *Let n and k be two integers such that $0 \leq k \leq n + 1$. The number $\sigma(A_{n,k})$ of runs of the process $A_{n,k}$ is equal to $t_{n,k}$ that satisfies:*

$$t_{n,k} = \frac{n+2k-1}{2}t_{n,k-1} + \frac{n-k}{2}t_{n+1,k-1} \quad and \quad t_{n,0} = 1. \tag{1}$$

In order to provide the proof, we first introduce the four processes in Figure 3. Notice that they are not arch processes. From left to right, the first process, denoted by $D_{n,k}$, is almost the process $A_{n,k}$. In fact, the single difference is that $D_{n,k}$ contains exactly one more action, denoted by c'_1, that is preceded by all the other actions. The second process $\overline{D}_{n,k}$ is related to $D_{n,k}$ in the following way: the precedence relation starting at b_1 is replaced, instead of having $b_1 \rightarrow c_1$, it is $b_1 \rightarrow c'_1$. Finally, for the two last processes $\overline{D}^1_{n,k}$ and $\overline{D}^2_{n,k}$, it is also the relations $a_1 \rightarrow b_1 \rightarrow c_1$ which are modified.

Proof. The extreme case $A_{n,0}$ corresponds to a process without any arch: just a trunk. Obviously it admits a single increasing labeling: it has a single run.

Suppose first that $k < n+1$. The number $\sigma(A_{n,k})$ is equal to the number of runs $\sigma(D_{n,k})$ because for all runs, the integer associated to c'_1 is inevitably the largest one: $2k + n + 1$. Then, using a inclusion/exclusion principle, we obtain the following formula for the number $\sigma(D_{n,k})$:

$$\sigma(D_{n,k}) = \sigma(\overline{D}_{n,k}) - \left(\sigma(\overline{D}^1_{n,k}) - \sigma(\overline{D}^2_{n,k})\right). \tag{2}$$

In fact we are focusing on the action preceded by b_1. In $D_{n,k}$ it corresponds to c_1. By modifying it to c'_1 in $\overline{D}_{n,k}$ we allow runs where b_1 appears after c_1, thus that are not valid for $D_{n,k}$. We remove this number of non-valid runs with $\sigma(\overline{D}^1_{n,k}) - \sigma(\overline{D}^2_{n,k})$, by playing with both actions x_1 and c_1. To compute $\sigma(\overline{D}_{n,k})$, first omit the action b_1 (and its incoming

and outgoing edges) ; the remaining process is a $(n, k-1)$-arch process, up to renaming, with added top and bottom actions (a_1 and c'_1) which do not modify the number of runs of $A_{n,k-1}$. It remains to insert b_1 in this "almost" $A_{n,k-1}$, somewhere between a_1 and c'_1: there are $(2 \cdot (k-1) + n - 1) + 2 = 2k + n - 1$ possibilities. The term $(2 \cdot (k-1) + n - 1)$ are the cases where b_1 is put between a_2 and c_k and the term 2 corresponds to the cases where b_1 is either before a_2 or after c_k. The process $\overline{D}^1_{n,k}$ is similar to the arch process $A_{n,k}$, there is only an action a_1 that precedes it, so $\sigma(\overline{D}^1_{n,k}) = t_{n,k}$. Lastly, for the process $\sigma(\overline{D}^2_{n,k})$, forgetting b_1 we recognize $A_{n+1,k-1}$ up to renaming, so b_1 can be inserted between x_1 and c_1: there are $n - k$ possibilities. Finally we obtain the following equation

$$\sigma(A_{n,k}) = (n + 2k - 1) \cdot \sigma(A_{n,k-1}) - \sigma(A_{n,k}) + (n - k) \cdot \sigma(A_{n+1,k-1}).$$

Suppose now that $k = n + 1$. Here there is no action x_i and both the nodes a_k and c_1 are merged into a single node. We can adapt equation (2) and obtain the same recurrence, but via a small difference in the computation: $\sigma(A_{k-1,k}) = 3k \cdot \sigma(A_{k-1,k-1}) - \sigma(A_{k-1,k}) - \sigma(A_{k,k-1})$. But since $k = n + 1$, this recurrence is equal to equation (1) too. ◀

When $k \geq n + 1$, one can think to the arch process $A_{n,k}$ as an arch process where the last $(k - n)$ actions a_{n-i} are merged with the first $(k - n)$ actions c_i. But the recursive formula (1) does not apply to such models: once $k > n + 1$ the recurrence loses its combinatorial meaning.

The next result exhibits a closed form formula for the number of runs of the arch processes.

▶ **Theorem 5.** *Let n and k be integers such that $0 < k \leq n + 1$. The number* [1] *of runs of the (n, k)-arch process is*

$$\sigma(A_{n,k}) = \frac{(2k + n - 1)!!}{2^{k-1}} \sum_{s=0}^{k-1} \frac{(n + s) \, \boldsymbol{par}(n, s)}{(n + s + 1)!!}$$

$$\sum_{1 \leq i_1 < \cdots < i_s \leq k} \prod_{j=1}^{s} (i_j + j + n - k - 1) \frac{\Gamma\left(\frac{2(k - i_j) + n + j + 2}{2}\right)}{\Gamma\left(\frac{2(k - i_j) + n + j + 3}{2}\right)},$$

$$\text{where } \boldsymbol{par}(n, s) = \begin{cases} (2^{s/2})^{-1} & \text{if } s \text{ is even} \\ \sqrt{\pi}(2^{(s+1)/2})^{-1} & \text{if } s \text{ is odd and } n \text{ is even} \\ (2^{(s-1)/2}\sqrt{\pi})^{-1} & \text{if } s \text{ is odd and } n \text{ is odd.} \end{cases}$$

Let us recall the double factorial notation: for $n \in \mathbb{N}$, $n!! = n \cdot (n-2)!!$ with $0!! = 1!! = 1$. We remark that the ratio of the two Γ-functions is related to the central binomial coefficient. The asymptotic behavior of the sequence does not seem immediate to obtain using this formula.

key-ideas. The formula for $\sigma(A_{n,k})$ is obtained by resolving the recurrence stated in equation (1). First remark that the calculation of $\sigma(A_{n,k})$ requires the values of $\sigma(A_{i,j})$ in the triangle such that $n \leq i \leq n + k$ and $0 \leq j \leq k - (i - n)$. The formula is computed by unrolling k times the recurrence. In particular, the index s in the formula corresponds to the number of times we have used the second term of equation (1), to reach the final term $\sigma(A_{n+s,0})$. The i_j values indicate in which iteration the second terms of equation (1) have been chosen. They describe the path from (n, k) to $(n + s, 0)$. The brute formula obtained in this way is composed of a product of truncated double factorials that can be written as ratios

---

[1] In Theorem 5 we use the convention that the sum over the sequence of i_j's is equal to 1 when $s = 0$.

of double factorial numbers. Finally, by coupling the adequate numerators and denominators in the product we exhibit several Wallis's ratios [2] that are easily simplified by using the Γ-function: $\dfrac{(2n-1)!!}{(2n)!!} = \dfrac{1}{\sqrt{\pi}}\dfrac{\Gamma\left(n+\frac{1}{2}\right)}{\Gamma\left(n+1\right)}.$ ◀

By using this closed form formula, or the bivariate recurrence (cf. equation (1)), we easily compute the first diagonals of the recurrence. The values of a given diagonal correspond to the class of arch processes with the same number of actions x_i in the trunk.

$$(\sigma(A_{k-1,k}))_{k\in\mathbb{N}\setminus\{0,1\}} = (1, 12, 170, 2940, 60760, 1466640, 40566680, 1266064800, \dots)$$
$$(\sigma(A_{k,k}))_{k\in\mathbb{N}^*} = (1, 5, 44, 550, 8890, 176120, 4130000, 111856360, \dots)$$
$$(\sigma(A_{k+1,k}))_{k\in\mathbb{N}^*} = (2, 11, 100, 1270, 20720, 413000, 9726640, 264279400, \dots)$$
$$(\sigma(A_{k+2,k}))_{k\in\mathbb{N}^*} = (3, 19, 186, 2474, 41670, 850240, 20386800, 561863960, \dots)$$

We remark that the first terms of the sequence $(\sigma(A_{k+1,k}))_{k\in\mathbb{N}^*}$ coincide with the first terms of the sequence `A220433` (shifted by 2) in OEIS [2] . This sequence is related to a specific Alia algebra and is exhibited in the paper of Khoroshkin and Piontkovski [14]. In their paper, the exponential univariate generating function naturally appears as an algebraic function. This motivates us to study in detail the bivariate generating function for $(t_{n,k})$ and in particular its diagonals.

3 Algebraic generating functions

Let us associate to the bivariate sequence $(t_{n,k})_{n,k}$ the generating function, denoted by $A(z,u)$, exponential in u and ordinary in z:

$$A(z,u) = \sum_{n\geq 0, k\geq 0} \frac{t_{n,k}}{k!}z^n u^k.$$

Recall this series enumerates the increasing labelings of the arch processes, when $k \leq n+1$, but has no combinatorial meaning beyond this bound.

▶ **Proposition 6.** *The bivariate generating function $A(z,u)$ is holonomic and satisfies the following differential equation.*

$$(2zu - 2z - u)\frac{\partial}{\partial u}A(z,u) + (z-2)A(z,u) + z(z+1)\frac{\partial}{\partial z}A(z,u) + C(u) = 0.$$

where $C(u)$ is an algebraic function determined by the initial conditions of the equation.

The differential equation can be exhibited since the recursive behavior of $(t_{n,k})$ is not disturbed beyond the bound $k > n+1$.

key-ideas. The differential equation is directly obtained from the recurrence equation (1). The function $C(u)$ encodes the initial conditions of the equation. The differential equation satisfied by $A(z,u)$ ensures its holonomicity (cf. [21, 11]). ◀

---

[2] OEIS corresponds to the On-line Encyclopedia of Integer Sequences: `http://oeis.org/`.

It is important to remark that $C(u)$ is holonomic. In fact we have $C(u) = u\frac{\partial}{\partial u}A(0, u) + 2A(0, u)$ and consequently $C(u)$ is holonomic as a specialization of a holonomic bivariate generating function. A direct computation for $C(u)$ exhibits the following differential equation

$$4\left(24u^2 + 3u + 1\right)C(u) - 4u\left(84u^2 - 3u + 1\right)\frac{\mathrm{d}}{\mathrm{d}u}C(u)$$

$$- 2u^2\left(216u^2 - 151u + 13\right)\frac{\mathrm{d}^2}{\mathrm{d}u^2}C(u)$$

$$- 2u^2\left(58u^3 - 75u^2 + 33u - 2\right)\frac{\mathrm{d}^3}{\mathrm{d}u^3}C(u)$$

$$- u^3\left(8u^3 - 15u^2 + 12u - 4\right)\frac{\mathrm{d}^4}{\mathrm{d}u^4}C(u) - 8\left(3u + 1\right) = 0.$$

Note that we prove also that $C(u)$ is solution of an algebraic equation. This fact is really not obvious from a combinatorial point of view. But it is deduced through the fact that the function $A(0, u)$ is algebraic:

$$(8u^3 - 15u^2 + 12u - 4)A(0, u)^3 + (12u^2 - 12u + 6)A(0, u) - 2u^3 = 0. \tag{3}$$

The equation is obtained by a *guess and prove approach*. Once it has been guessed it remains to prove it by using the holonomic equation proven in Proposition 6. Thus we get

$$32\left(9u^2 - 12u + 8\right)(u - 1)^3$$

$$+ 48\left(36u^6 - 120u^5 + 202u^4 - 199u^3 + 123u^2 - 44u + 8\right)(u - 1)^2 C(u)$$

$$+ \left(8u^3 - 15u^2 + 12\,u - 4\right)^3 C(u)^3 = 0.$$

▶ **Theorem 7.** *The function $A(z, u)$ is an algebraic function in (z and u) whose annihilating polynomial has degree* 3:

$$2 + 6\left(12zu^3 - 18zu^2 - 2u^2 + 13zu + 2u - 3z - 1\right)A(z, u)$$

$$+ 6z^2\left(8u^3 - 15u^2 + 12u - 4\right)A(z, u)^2$$

$$+ \left(8u^3 - 15u^2 + 12u - 4\right)\left(z^3 + 6zu + 3z^2 - 3z - 1\right)A(z, u)^3 = 0.$$

Note that the choice to use a doubly exponential generating function (in u and z) for $(t_{n,k})$ would have made sense and would be holonomic too (closure property of Borel transform). But it would not be algebraic because of the inappropriate asymptotic expansion (cf. Theorem 9).

Proof. The fact that the initial conditions and a diagonal of $A(z, u)$ are algebraic suggests that it could also be algebraic as a function of z and u. Applying a bivariate guessing procedure, we observe that the bivariate function $H(z, u) = (u + 1)(z^3 + 3z^2 + 6zu - 3z - 1)A(z, u)$ is such that $[z^n]H(z, u) = 0$ for $n > 2$. Furthermore $[z^j]H(z, u)$ is algebraic for $j = \{0, 1, 2\}$. So, let us calculate these z-extractions. First recall that $[z^0]A(z, u)$ satisfies the algebraic equation (3). In the same vein, $[z^1]A(z, u)$ satisfies the algebraic equation

$$\left(8u^3 - 15u^2 + 12u - 4\right)f(u)^3 + 3\left(8u^3 - 15u^2 + 12u - 4\right)f(u)^2$$

$$+ 3\left(8u^3 - 15u^2 + 10u - 2\right)f(u) + 8u^3 - 15u^2 + 6u = 0,$$

and finally $[z^2]A(z, u)$ satisfies the algebraic equation

$$\left(8u^3 - 15u^2 + 12u - 4\right)f(u)^3 + \left(-24u^3 + 45u^2 - 36u + 12\right)f(u)^2$$

$$+ \left(-72u^3 + 135u^2 - 84u + 18\right)f(u) - 40u^3 + 75u^2 - 36u = 0.$$

Thus we obtain

$$[z^0]H(z,u) = -(1+u)A(0,u)$$

$$[z^1]H(z,u) = -1 + (u+1)\left((6u-3)A(0,u) - [z^1]A(z,u)\right)$$

$$[z^2]H(z,u) = (u+1)\left((6u-3)[z^1]A(z,u) - [z^2]A(z,u) + 3A(0,u) + (6u-4)\right).$$

Finally we get $A(z,u) = \dfrac{[z^{\leq 2}]H(z,u)}{(u+1)\left(z^3 + 3z^2 + 6uz - 3z - 1\right)}$. By using the elimination theory, we get a closed form algebraic equation for $A(z,u)$ of degree 27, that obviously cannot fit in the conference paper format. Nevertheless, this equation is not minimal. Simplifying it, we get a minimal polynomial of degree 3 which annihilates $A(z,u)$:

$$\left(8u^3 - 15u^2 + 12u - 4\right)\left(z^3 + 3z^2 + 6zu - 3z - 1\right)A(z,u)^3$$

$$+ 6z^2\left(8u^3 - 15u^2 + 12u - 4\right)A(z,u)^2$$

$$+ 6\left(12zu^3 - 18zu^2 - 2u^2 + 13zu + 2u - 3z - 1\right)A(z,u) + 2 = 0.$$

A direct proof by recurrence confirms the validity of this equation.   ◄

We remark in the previous section that the diagonals of the function $A(z,u)$ are of particular interest because they define subclasses of arch processes with a fixed number of x_i actions covered by all the arches. In order to extract the generating functions of this subclass, we could use the Cauchy formula to compute $[u^0]A(z/u,u)$ and so on; we would keep the holonomicity property of the sequences but not their algebraicity. So, we prefer to define the generating function $B(z,u) = A(z/u,u)$. A similar proof as for the case $A(z,u)$ can be done to prove the algebraicity of $B(z,u)$. In particular, it exhibits the following algebraic equation satisfied by $B(z,u)$

$$\left(9u^2 + 12u - 4\right)\left(z^3 + 3z^2 + 6u - 3z - 1\right)B(z,u)^3 + 6z^2\left(9u^2 + 12u - 4\right)B(z,u)^2$$

$$+ 6\left(18u^2z - 18u^2 + 6uz + 9u - 3z - 1\right)B(z,u) + 2\left(6u - 1\right)^2 = 0$$

In particular, $B(0,u)$ is associated to the sequence $(t_{k,k})_k$, $[z^1]B(z,u)$ corresponds to the sequence $(t_{k-1,k})_k$ and so on. By specializing $z = 0$ in the latter algebraic equation then by resolving it through the Viète-Descartes approach for the resolution of cubic equation (detailed in the paper [19]), we obtain the following closed form formula corresponding to the branch that is analytic in 0:

$$B(0,u) = \sqrt{2}\sqrt{\frac{1-3u}{1-3u-\frac{9}{4}u^2}}\cos\left(\frac{1}{3}\arccos\left(\frac{6u-1}{\sqrt{2}(1-3u)}\sqrt{\frac{1-3u-\frac{9}{4}u^2}{1-3u}}\right)\right).$$

Even if the way we represented $B(0,u)$ could suggest a singularity when the argument of the arccos function is equal to 1, the function admits an analytic continuation up to its dominant singularity ρ : the solution of $1 - 3u - \frac{9}{4}u^2 = 0$, thus corresponding to $\rho = \frac{2}{3}\left(\sqrt{2} - 1\right)$. Furthermore, by studying the global generating function $B(z,u)$, we obtain its singular expansion.

▶ **Lemma 8.** *Near the singularity when u tends to ρ, the function $B(z,u)$ satisfies*

$$B(z,u) \underset{u\to\rho}{=} a(z) + \frac{b(z)}{\sqrt{\rho - u}} + o\left((\rho - u)^{-1/2}\right),$$

with $a(z)$ and $b(z)$ two functions independent from u.

By using this result we deduce the asymptotic behaviors of the diagonal coefficients of $A(z, u)$.

▶ **Theorem 9.** *Let i be a given integer greater than -1, and k tend to infinity:*

$$t_{k+i,k} \underset{k \to \infty}{\sim} \gamma_i \frac{\rho^{-k}}{\sqrt{k}} \, k! \quad with \; \gamma_0 = \frac{1}{2} \sqrt{\frac{3}{\sqrt{2}\pi} \left(\sqrt{2} - 1 \right)} \quad and \quad \gamma_i = \left(\frac{1}{\sqrt{2} - 1} \right)^i \gamma_0.$$

This theorem is a direct consequence of Lemma 8. The $(\gamma_i)_i$ can be deduced by asymptotic matching (using an *Ansatz*).

Finally, by computing $[z^1]B(z, u)$ with the algebraic equation it satisfies, we prove that its second derivative is solution of the algebraic equation exhibited in OEIS `A220433`.

4   Uniform random generation of runs

We now introduce an algorithm to uniformly sample runs of a given arch process $A_{n,k}$. Our approach is based on the recursive equations (1) and (2) for the sequence $(t_{n,k})$. Here we deal with the cases $k \le n$ and avoid the limit case $k = n + 1$. Although the latter limit case satisfies this equation too, its proof is based on another combinatorial approach, and so the construction of a run cannot be directly deduced form the combinatorial approach proposed for the cases $k \le n$. Of course, a simple adaptation of the algorithm presented below would allow to sample in $A_{k-1,k}$, but the lack of space prevent us to present it here.

Our algorithm is a *recursive generation algorithm*. But since the objects are not specified in a classical Analytic Combinatorics way, we cannot use the results of [12]. As usual for recursive generation, the first step consists in the computation and the memorization of the value $t_{n,k}$ and all the intermediate values $(t_{i,j})$ needed for the calculation of $t_{n,k}$.

▶ **Proposition 10.** *In order to compute the value $t_{n,k}$, it is sufficient to calculate the values in the bi-dimensional set $\{t_{i,j} \mid n \le i \le n + k$ and $0 \le j \le k - (i - n)\}$. This computation is done with $\mathcal{O}\left(k^2\right)$ arithmetic operations.*

Recall that the coefficient computations are done only once for a given pair (n, k), and then many runs can be drawn uniformly for $A_{n,k}$ by using the recursive generation algorithm.

Let us present the way we exploit the recurrence equation (2) to design the sampling method. The main problem that we encounter is the presence of a minus sign in the recurrence equation. Let us rewrite it in a slightly different way: $\sigma(D_{n,k}) + \sigma(\overline{D}^1_{n,k}) = \sigma(\overline{D}_{n,k}) + \sigma(\overline{D}^2_{n,k})$.

Recall that the structures under consideration are depicted in Figure 3. We introduce the classes of increasingly labeled structures from $D_{n,k}, \overline{D}^1_{n,k}, \overline{D}_{n,k}$ and $\overline{D}^2_{n,k}$, respectively denoted by $I_{n,k}, \overline{I}^1_{n,k}, \overline{I}_{n,k}$ and $\overline{I}^2_{n,k}$. Remark that the number of runs of $A_{n,k}$ is equal to $|I_{n,k}|$, where the function $|\cdot|$ corresponds to the cardinality of the considered class. Obviously the equation on the cardinalities can be written directly on the classes $I_{n,k} \cup \overline{I}^1_{n,k} = \overline{I}_{n,k} \cup \overline{I}^2_{n,k}$ (since their intersections are empty: $I_{n,k}$ and $\overline{I}^1_{n,k}$ are distinct even if they are isomorphic). Thus, we consider the problem of sampling the class $I_{n,k} \cup \overline{I}^1_{n,k}$ where we bijectively replace the runs belonging to $\overline{I}^1_{n,k}$ by runs of $I_{n,k}$ (which can be performed recursively during the sampling procedure). The Algorithm SAMPLING(n, k) is based on the correspondence depicted in the Figure 3 and its adaptation presented above on the classes $I_{n,k} \cup \overline{I}^1_{n,k}$. In each case the algorithm completes a recursively drawn run and applies some renaming on the actions of that run. Then, it inserts the action b_1 according to the cases $\overline{I}_{n,k} \backslash \overline{I}^1_{n,k}, \overline{I}^1_{n,k}$ or $\overline{I}^2_{n,k}$. In the specific case $\overline{I}^1_{n,k}$, instead of b_1, it is the action b_k that is inserted and the renaming occurs in a similar fashion to obtain a run of $I_{n,k}$ from the one of $\overline{I}^1_{n,k}$.

Algorithm 1 Uniform random sample for $I_{n,k}$.

 1: **function** SAMPLING(n, k)
 2:    **if** $k = 0$ **then**
 3:       **return** $\langle x_1, x_2, \dots, x_n \rangle$
 4:    $r := \text{RAND\_INT}(0, 2 \cdot t_{n,k} - 1)$ ▷ a uniform integer between 0 and $2 \cdot t_{n,k} - 1$ in r
 5:    **if** $r < |\bar{I}_{n,k}|$ **then**   ▷ generation in $\bar{I}_{n,k}$
 6:       $U := \text{SAMPLING}(n, k - 1)$
 7:       $p_b := 1 + r // t_{n,k-1}$ ▷ The position of the new b to insert
 8:       **if** $p_b > p_{x_1}$ **then**   ▷ generation in $\bar{I}_{n,k}^1$
 9:          Rename x_1 by a_k ; and each x_i with $i > 1$ by x_{i-1}
10:          Insert b_k at position p_b ; and c_k at the end of U
11:       **else**   ▷ generation in $\bar{I}_{n,k} \backslash \bar{I}_{n,k}^1$
12:          In U, rename each a_i (resp. c_i and b_i) by a_{i+1} (resp. c_{i+1} and b_{i+1})
13:          Rename x_{n-k+1} by c_1
14:          Insert b_1 at position p_b ; and a_1 at the head of U
15:    **else**   ▷ generation in $\bar{I}_{n,k}^2$
16:       $U := \text{SAMPLING}(n + 1, k - 1)$
17:       $p_b := 2 + (r - (n + 2k - 1) \cdot t_{n,k-1}) // t_{n+1,k-1}$
18:       Rename x_{p_b} by b_1 and x_{n-k+2} by c_1 ; and each x_i with $i > p_b$ by x_{i-1}
19:       Insert a_1 at the head of U
20:    **return** U

Line 4 and 17 : the binary operator $//$ denotes the Euclidean division.
The position of an action in a run is its arrival number (from 1 to the number of actions).

▶ **Theorem 11.** *The Algorithm* SAMPLING(n, k) *builds uniformly at random a run of $A_{n,k}$ in k recursive calls, once the coefficients computations and memorizations have been done.*

Since each object of $I_{n,k}$ is sampled in two distinct ways, the uniform sampling in $I_{n,k} \cup \bar{I}_{n,k}^1$ induces the uniform sampling of $I_{n,k}$.

Focus on the run of $A_{5,4}$ depicted in Figure 2: $\langle a_1, b_1, a_2, a_3, b_3, a_4, x_1, b_4, c_1, b_2, c_2, c_3, c_4 \rangle$. It is either obtained from a (renamed) run of $\bar{I}_{5,4}^1$: $\langle a_1, b_1, a_2, a_3, b_3, x_1, x_2, c_1, b_2, c_2, c_3 \rangle$ with $p_b = 8$ (Line 8 of the algorithm). Or it is built from $\langle a_1, a_2, b_2, a_3, x_1, b_3, x_2, b_1, c_1, c_2, c_3 \rangle$ of $\bar{I}_{5,4} \backslash \bar{I}_{5,4}^1$, with $p_b = 1$ (Line 11). But it cannot be built from a run of $\bar{I}_{5,4}^2$.

In Figure 4, we have uniformly sampled 1000 runs for $A_{1000,1000}$ and we have represented in blue points every pair (k, n) corresponding to an increasing sub-structure from $A_{n,k}$ that has been built during the algorithm (k for abscissa and n for ordinate). Only around $4.78 \cdot 10^4$ sub-structures have been built among the $50 \cdot 10^4$ inside the red lines which are calculated for the value $t_{1000,1000}$. At the beginning $n \approx k$ and the **if** branch on Line 5 is preferred (instead of the **else** one on Line 15) because the number of x_i actions is too small. After some recursive calls, the number of x_i actions has increased and then both branches of the algorithm are taken with probabilities of the same order. Recall that the constants γ_i (cf. Theorem 9) are evolving with an exponential growth. Finally, we observe that only a small number of diagonals are necessary for the samplings. Since the diagonals (t_{n_i,k_i}) for increasing sequences $(n_i)_i$ and $(k_i)_i$ follow P-recurrences (cf. [16]), a lazy calculation of the terms of the necessary diagonals that envelop the blue points would allow to minimize the pre-computations of Proposition 10.

We close this section with the presentation of an *unranking* algorithm for the construction
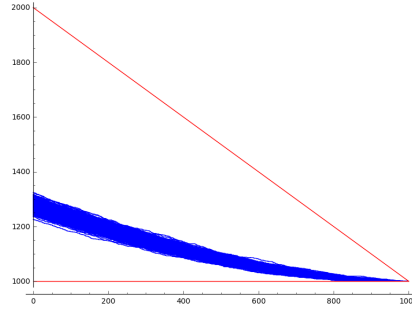
■ **Figure 4** The terms $t_{i,j}$ needed for the sampling of 1000 runs of $A_{1000,1000}$.

of the runs of a given arch process $A_{n,k}$. This type of algorithm has been developed during the 70's by Nijenhuis and Wilf [20] and introduced in the context of Analytic Combinatorics by Martínez and Molinero [17]. Our algorithm is based on a bijection between the set of integers $\{0, \ldots, t_{n,k} - 1\}$ and the set of runs of $A_{n,k}$. Here again we restrict ourselves to the values $k \leq n$. As usual for unranking algorithms, the first step consists in the computation and the memorization of the values of a sequence. But compared to the uniform random sampling, here we need more information than the one given by the sequence $(t_{n,k})$.

To be able to reconstruct the run associated to a given rank, we need to know the position of the action x_1 in the recusively drawn run in order to decide if the action b_1 appears before or after it. First suppose $k < n$ and let $t_{n,k,\ell}$ be the number of runs in $\mathcal{A}_{n,k}$ whose action x_1 appears at position ℓ. Let us denote by $I_{n,k,\ell}$ the associated combinatorial class. We obtain directly a constructive recurrence for the sequence.

$$t_{n,k,\ell} = (\ell - 2)\, t_{n,k-1,\ell-2} + (n - k)\, t_{n+1,k-1,\ell-1} \quad \text{and} \quad t_{n,0,1} = 1;\ t_{n,0,\ell>1} = 0.$$

▶ **Proposition 12.** *The computation of $t_{n,k,\ell}$ is done with $\mathcal{O}\left(k^2\right)$ arithmetic operations.*

The UNRANKING algorithm computes a run given its rank in the following total order:

$$\alpha \preceq_{n,k} \beta \quad \text{iff.} \quad \begin{cases} \alpha \in I_{n,k,i_0} \text{ and } \beta \in I_{n,k,i_1} & \wedge & i_0 < i_1, \\[1mm] \text{or} \quad \alpha, \beta \in I_{n,k,i} & \wedge & \begin{array}{l}\alpha \text{ is built recursively from } I_{n,k-1,i-2} \text{ and} \\ \beta \text{ is built recursively from } I_{n+1,k-1,i-1}\end{array} \\[2mm] \text{or} \quad \alpha, \beta \in I_{n,k,i} & \wedge & \begin{array}{l}\alpha, \beta \in I_{n,k-1,i-2} \text{ (resp. } I_{n+1,k-1,i-1}) \text{ and} \\ \alpha_0, \beta_0 \text{ inducing } \alpha, \beta \text{ satisfy } \alpha_0 \preceq_{n,k-1} \beta_0.\end{array} \end{cases}$$

The run example of Figure 2 has rank 479 among the 1270 runs of $A_{5,4}$. Note that in the case $k = n$ (at the end there is no x_1) the algorithm is easily extended by considering the position of b_1 as the one of x_1.

▶ **Theorem 13.** *The Algorithm* UNRANKING(n, k, r) *builds the r-th run of $A_{n,k}$ in k recursive calls, once the coefficient memorizations $t_{n,k,\ell}$, for all ℓ such that $k + 1 \leq \ell \leq 2k + 1$ (and the necessary n and k), have been done.*

Note that the implementation of both algorithms can be much more efficient than the pseudocode exhibited above. Actually, only the absolute positions of the b_i actions are important in a run, because all other actions have their positions determined by the positions of the b_i actions. However, such implementations are much more cryptic to read, and so we preferred to present here easy-to-read algorithms.

Algorithm 2 Unranking for $I_{n,k}$.

1: **function** UNRANKING(n, k, r)
2:     $\ell := k + 1$
3:     **while** $r \geq 0$ **do**
4:         $r := r - t_{n,k,l}$
5:         $\ell := \ell + 1$
6:     **return** CONS(n, k, ℓ, r)

7: **function** CONS(n, k, ℓ, r)
8:     **if** $k = 0$ **then**
9:         **return** $\langle x_1, x_2, \ldots, x_n \rangle$
10:     **if** $r < (\ell - 2) \cdot t_{n,k-1,\ell-2}$ **then**   ▷ generation in $I_{n,k-1,\ell-2}$
11:         $rr := r \% t_{n,k-1,\ell-2}$
12:         $U := $ CONS$(n, k - 1, \ell - 2, rr)$
13:         $p_b := 1 + r // t_{n,k-1,\ell-2}$ ▷ The position of the new b to insert
14:         In U, rename each a_i (resp. c_i and b_i) by a_{i+1} (resp. c_{i+1} and b_{i+1})
15:         Rename x_{n-k+1} by c_1
16:         Insert b_1 at position p_b ; and a_1 at the head of U
17:     **else**   ▷ generation in $I_{n+1,k-1,\ell-1}$
18:         $r' := r - (\ell - 2) \cdot t_{n,k-1,\ell-2}$
19:         $rr := r' \% t_{n+1,k-1,\ell-1}$
20:         $U := $ CONS$(n + 1, k - 1, \ell - 1, rr)$
21:         $p_b := 2 + r' // t_{n+1,k-1,\ell-1}$
22:         Rename x_{p_b} by b_1 and x_{n-k+2} by c_1 ; and each x_i with $i > p_b$ by x_{i-1}
23:         Insert a_1 at the head of U
24:     **return** U

Line 11 and 19 : the binary operator $\%$ denotes the Euclidean division remainder.

References

1 The ecmascript 2015 language specification. Technical report, ECMA-262 6th Edition, 2015. URL: http://www.ecma-international.org/ecma-262/6.0/index.html.

2 M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York, ninth dover printing, tenth gpo printing edition, 1964.

3 C. Baier and J.-P. Katoen. *Principles of Model Checking (Representation and Mind Series).* The MIT Press, 2008.

4 H. C. Baker, Jr. and C. Hewitt. The incremental garbage collection of processes. In *Proceedings of the 1977 Symposium on Artificial Intelligence and Programming Languages*, pages 55–59. ACM, 1977.

5 O. Bodini, M. Dien, X. Fontaine, A. Genitrini, and H.-K. Hwang. Increasing diamonds. In *Latin American Symposium on Theoretical Informatics*, pages 207–219. Springer, Berlin, Heidelberg, 2016.

6 O. Bodini, M. Dien, A. Genitrini, and F. Peschanski. Entropic uniform sampling of linear extensions in series-parallel posets. In *12th International Computer Science Symposium in Russia (CSR)*, pages 71–84, 2017.

7 O. Bodini, M. Dien, A. Genitrini, and F. Peschanski. The Ordered and Colored Products in Analytic Combinatorics: Application to the Quantitative Study of Synchronizations in

Concurrent Processes. In *14th SIAM Meeting on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 16–30, 2017.

8   O. Bodini, A. Genitrini, and F. Peschanski. The combinatorics of non-determinism. In *FSTTCS'13*, volume 24 of *LIPIcs*, pages 425–436. Schloss Dagstuhl, 2013.

9   O. Bodini, A. Genitrini, and F. Peschanski. A Quantitative Study of Pure Parallel Processes. *Electronic Journal of Combinatorics*, 23(1):P1.11, 39 pages, (electronic), 2016.

10   G. Brightwell and P. Winkler. Counting linear extensions is ♯P-Complete. In *STOC*, pages 175–181, 1991.

11   P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009. URL: `http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=9780521898065`.

12   P. Flajolet, P. Zimmermann, and B. Van Cutsem. A calculus for the random generation of labelled combinatorial structures. *Theoretical Computer Science*, 132(1-2):1–35, 1994.

13   R. Grosu and S. A. Smolka. Monte carlo model checking. In *TACAS'05*, volume 3440 of *LNCS*, pages 271–286. Springer, 2005.

14   A. Khoroshkin and D. Piontkovski. On generating series of finitely presented operads. *Journal of Algebra*, 426:377–429, 2015.

15   D. E. Knuth. *The art of computer programming, volume 3: (2nd ed.) sorting and searching*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1998.

16   L. Lipshitz. The diagonal of a d-finite power series is d-finite. *Journal of Algebra*, 113(2):373–378, 1988.

17   C. Martínez and X. Molinero. Generic algorithms for the generation of combinatorial objects. In *MFCS'03*, pages 572–581. Springer Berlin Heidelberg, 2003.

18   R. Milner. *A Calculus of Communicating Systems*. Springer Verlag, 1980.

19   R.W.D. Nickalls. Viète, descartes and the cubic equation. *The Mathematical Gazette*, 90(518):203–208, 2006.

20   A. Nijenhuis and H.S. Wilf. *Combinatorial algorithms*. Computer science and applied mathematics. Academic Press, New York, NY, 1975.

21   R.P. Stanley. *Enumerative Combinatorics:*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2001.

Inversions in Split Trees and Conditional Galton–Watson Trees

Xing Shi Cai

Department of Mathematics, Uppsala University, Uppsala, Sweden
xingshi.cai@math.uu.se
https://orcid.org/0000-0002-3768-8078

Cecilia Holmgren

Department of Mathematics, Uppsala University, Uppsala, Sweden
cecilia.holmgren@math.uu.se
https://orcid.org/0000-0003-0717-4671

Svante Janson

Department of Mathematics, Uppsala University, Uppsala, Sweden
svante.janson@math.uu.se
https://orcid.org/0000-0002-9680-2790

Tony Johansson

Department of Mathematics, Uppsala University, Uppsala, Sweden
tony.johansson@math.uu.se

Fiona Skerman

Department of Mathematics, Uppsala University, Uppsala, Sweden
fiona.skerman@math.uu.se
https://orcid.org/0000-0003-4141-7059

— Abstract

We study $I(T)$, the number of inversions in a tree T with its vertices labeled uniformly at random. We first show that the cumulants of $I(T)$ have explicit formulas. Then we consider X_n, the normalized version of $I(T_n)$, for a sequence of trees T_n. For fixed T_n's, we prove a sufficient condition for X_n to converge in distribution. For T_n being split trees [6], we show that X_n converges to the unique solution of a distributional equation. Finally, when T_n's are conditional Galton–Watson trees, we show that X_n converges to a random variable defined in terms of Brownian excursions. Our results generalize and extend previous work by Panholzer and Seitz [20].

1 Introduction

1.1 Inversions in a fixed tree

Let $\sigma_1, \ldots, \sigma_n$ be a permutation of $\{1, \ldots, n\}$. If $i < j$ and $\sigma_i > \sigma_j$, then the pair (σ_i, σ_j) is called an inversion. The concept of inversions was introduced by Cramer [5] (1750) due to its connection with solving linear equations. More recently, the study of inversions has been motivated by its applications in the analysis of sorting algorithms ([15, Section 5.1]). Many authors, e.g., Feller [7, pp. 256], have shown that the number of inversions in uniform random permutations has a central limit theorem.

The concept of inversions can be generalized as follows. Consider an unlabeled rooted tree T on node set V. Let ρ denote the root. Write $u < v$ if u is a *proper ancestor* of v, i.e., the unique path from ρ to v passes through u and $u \neq v$. Write $u \leq v$ if u is an ancestor of v, i.e., either $u < v$ or $u = v$. Given a bijection $\lambda : V \to \{1, \ldots, |V|\}$ (a *node labeling*), define the number of *inversions*

$$I(T, \lambda) \stackrel{\text{def}}{=} \sum_{u < v} \mathbf{1}_{\lambda(u) > \lambda(v)}.$$

Note that if T is a path, then $I(T, \lambda)$ is nothing but the number of inversions in a permutation. Our main object of study is the random variable $I(T)$, defined by $I(T) = I(T, \lambda)$ where λ is chosen uniformly at random from the set of bijections from V to $\{1, \ldots, |V|\}$.

The enumeration of trees with a fixed number of inversions has been studied by Mallows and Riordan [16] and Gessel et al. [9] using the so called *inversions polynomial*. While analyzing linear probing hashing, Flajolet et al. [8] noticed that the numbers of inversions in Cayley trees with uniform random labeling converges to an Airy distribution. Panholzer and Seitz [20] showed that this is true for conditional Galton–Watson trees, which encompasses the case of Cayley trees.

For a node v, let z_v denote the size of the subtree rooted at v. The following representation of $I(T)$ is the basis of most of our results:

▶ **Lemma 1.** *Let T be a fixed tree. Then*

$$I(T) \stackrel{\text{d}}{=} \sum_{v \in V} Z_v,$$

where $\{Z_v\}_{v \in V}$ are independent random variables, and $Z_v \sim \text{Unif}\{0, 1, \ldots, z_v - 1\}$.

We will generally be concerned with the centralized number of inversions, i.e., $I(T) - \mathbb{E}[I(T)]$. For any $u < v$ we have $\mathbb{P}\{\lambda(u) > \lambda(v)\} = 1/2$. Let $h(v)$ denote the *depth* of v, i.e., the distance from v to the root ρ. It immediately follows that,

$$\mathbb{E}[I(T)] = \sum_{u < v} \mathbb{E}\left[\mathbf{1}_{\lambda(u) > \lambda(v)}\right] = \frac{1}{2}\Upsilon(T), \tag{1.1}$$

where $\Upsilon(T) \stackrel{\text{def}}{=} \sum_v h(v)$ is called the *total path length* (or *internal path length*) of T.

Let $\varkappa_k = \varkappa_k(X)$ denote the k-th cumulant of a random variable X (provided it exists); thus $\varkappa_1(X) = \mathbb{E}[X]$ and $\varkappa_2(X) = \text{Var}(X)$. We now define $\Upsilon_k(T)$, *the k-total common ancestors* of T, which allows us to generalize (1.1) to higher cumulants of $I(T)$. For k nodes $v_1, \ldots, v_k$ (not necessarily distinct), let $c(v_1, \ldots, v_k)$ be the number of ancestors that they share, i.e.,

$$c(v_1, \ldots, v_k) \stackrel{\text{def}}{=} |\{u \in V : u \leq v_1, u \leq v_2, \ldots, u \leq v_k\}|.$$

We define

$$\Upsilon_k(T) \stackrel{\text{def}}{=} \sum_{v_1,\ldots,v_k} c(v_1,\ldots,v_k),$$

where the sum is over all ordered k-tuples of nodes in the tree. For a single node v, $h(v) = c(v) - 1$, since v itself is counted in $c(v)$. So $\Upsilon(T) = \Upsilon_1(T) - |V|$; i.e., we recover the usual notion of total path length. Using Lemma 1, it is easy to show the following:

▶ **Theorem 2.** *Let T be a fixed tree. Let $\varkappa_k(I(T))$ be the k-th cumulant of $I(T)$. Then*

$$\mathbb{E}\left[I(T)\right] = \varkappa_1(I(T)) = \frac{1}{2}\Upsilon(T) = \frac{1}{2}(\Upsilon_1(T) - |V|),$$

$$\mathrm{Var}\left(I(T)\right) = \varkappa_2(I(T)) = \frac{1}{12}(\Upsilon_2(T) - |V|),$$

and, more generally, for $k \geq 1$,

$$\varkappa_{2k+1}(I(T)) = 0, \qquad \varkappa_{2k}(I(T)) = \frac{B_{2k}}{2k}(\Upsilon_{2k}(T) - |V|),$$

where B_k denotes the k-th Bernoulli number. Moreover, $I(T)$ has the moment generating function

$$\mathbb{E}\left[e^{tI(T)}\right] = \prod_{v\in V} \frac{e^{z_v t} - 1}{z_v(e^t - 1)},$$

and for the centralized variable we have the estimate

$$\mathbb{E}\left[e^{t(I(T)-\mathbb{E}[I(T)])}\right] \leq \exp\left(\tfrac{1}{8}t^2 \sum_{v\in T}(z_v - 1)^2\right) \leq \exp\left(\tfrac{1}{8}t^2 \sum_{v\in T} z_v^2\right) = \exp\left(\tfrac{1}{8}t^2 \Upsilon_2(T)\right), \qquad t \in \mathbb{R}.$$

1.2   Inversions in sequences of trees

The total path length $\Upsilon(T)$ has been studied for random trees like split trees [3] and conditional Galton–Watson trees [1, Corollary 9]. This leads us to focus on the deviation

$$X_n = \frac{I(T_n) - \mathbb{E}\left[I(T_n)\right]}{s(n)},$$

under some appropriate scaling $s(n)$, for a sequence of (random or fixed) trees T_n.

Fixed trees

The following theorem follows easily from Theorem 2:

▶ **Theorem 3.** *Let T_n be a sequence of fixed trees on n nodes. Let*

$$X_n = \frac{I(T_n) - \mathbb{E}\left[I(T_n)\right]}{\sqrt{\Upsilon_2(T_n)}}.$$

Assume that for all $k \geq 1$,

$$\frac{\Upsilon_{2k}(T_n)}{\Upsilon_2(T_n)^k} \to \zeta_{2k},$$

for some sequence (ζ_{2k}). Then there exists a unique distribution X with

$$\varkappa_{2k-1}(X) = 0, \qquad \varkappa_{2k}(X) = \frac{B_{2k}}{2k}\zeta_{2k}, \qquad k \geq 1,$$

such that $X_n \stackrel{\mathrm{d}}{\longrightarrow} X$ and, moreover, $\mathbb{E}\left[e^{tX_n}\right] \to \mathbb{E}\left[e^{tX}\right] < \infty$ for every $t \in \mathbb{R}$.

▶ **Example 4.** When P_n is a path of n nodes, we have for fixed $k \geq 1$

$$\Upsilon_k(P_n) \sim \frac{1}{k+1} n^{k+1}.$$

Thus $\Upsilon_{2k}(P_n)/\Upsilon_2(P_n)^k \to \varkappa_{2k} = 0$ for $k \geq 2$. So by Theorem 3, X_n converges to a normal distribution, and we recover the central limit law for inversions in permutations. Also, the vertices have subtree sizes $1, \ldots, n$ and so we also recover from Theorem 2 the moment generating function $\prod_{j=1}^{n}(e^{jt} - 1)/(j(e^t - 1))$ [22, 17].

Other examples where Theorem 2 can be easily applied include complete b-ary trees and stars (a star is a tree containing only a root and leaves).

Random trees

We move on to random trees. We consider generating a random tree T_n and, conditioning on T_n, labeling its nodes uniformly at random. The relation (1.1) is maintained for random trees:

$$\mathbb{E}\left[I(T_n)\right] = \mathbb{E}\left[\mathbb{E}\left[I(T_n) \mid T_n\right]\right] = \frac{1}{2}\mathbb{E}\left[\Upsilon(T_n)\right].$$

The deviation of $I(T_n)$ from its mean can be taken to mean two different things. Consider for some scaling function $s(n)$,

$$X_n = \frac{I(T_n) - \mathbb{E}\left[I(T_n)\right]}{s(n)}, \qquad Y_n = \frac{I(T_n) - \mathbb{E}\left[I(T_n) \mid T_n\right]}{s(n)} = \frac{I(T_n) - \frac{1}{2}\Upsilon(T_n)}{s(n)}.$$

Then X_n and Y_n each measure the deviation of $I(T_n)$, unconditionally and conditionally. They are related by the identity

$$X_n = Y_n + W_n/2, \tag{1.2}$$

where

$$W_n = \frac{\Upsilon(T_n) - \mathbb{E}\left[\Upsilon(T_n)\right]}{s(n)}.$$

In the case of fixed trees $W_n = 0$ and $X_n = Y_n$, but for random trees we consider the sequences separately.

Split trees

The first class of random trees which we study are split trees. They were introduced by Devroye [6] to encompass many families of trees that are frequently used in algorithm analysis, e.g., binary search trees, m-ary search trees, digital search trees, etc.

A split tree can be constructed as follows. Consider a rooted infinite b-ary tree where each node is a bucket of finite capacity s. We place n balls at the root, and the balls individually trickle down the tree in a random fashion until no bucket is above capacity. Each node draws a *split vector* $\mathcal{V} = (V_1, \ldots, V_b)$ from a common distribution, where V_i describes the probability that a ball passing through the node continues to the ith child. The trickle-down procedure is defined precisely in Section 2. Any node u such that the subtree rooted at u contains no balls is then removed, and we consider the resulting tree T_n.

In the context of split trees we differentiate between $I(T_n)$ (the number of inversions on *nodes*), and $\hat{I}(T_n)$ (the number of inversions on *balls*). In the former case, the nodes (buckets)

are given labels, while in the latter the individual balls are given labels. For balls β_1, β_2, write $\beta_1 < \beta_2$ if the node containing β_1 is a proper ancestor of the node containing β_2; if β_1, β_2 are contained in the same node we do not compare their labels. Define

$$\hat{I}(T_n) = \sum_{\beta_1 < \beta_2} \mathbf{1}_{\lambda(\beta_1) > \lambda(\beta_2)}.$$

Similarly define $\hat{\Upsilon}(T_n)$ as the total path length on balls, i.e., the sum of the depth of all balls. And let

$$\hat{X}_n = \frac{\hat{I}(T_n) - \mathbb{E}\left[\hat{I}(T_n)\right]}{n}, \quad \hat{Y}_n = \frac{\hat{I}(T_n) - s_0 \hat{\Upsilon}(T_n)/2}{n}, \quad \hat{W}_n = \frac{\hat{\Upsilon}(T_n) - \mathbb{E}\left[\hat{\Upsilon}(T_n)\right]}{n}. \quad (1.3)$$

Here s_0 is a fixed integer denoting the number of balls in any internal node, and we have $\hat{X}_n = \hat{Y}_n + s_0 \hat{W}_n/2$ (justified in Section 2). The following theorem gives the limiting distributions of the random vector $(\hat{X}_n, \hat{Y}_n, \hat{W}_n)$. In a longer version of this paper [4], we also have a similar result for (X_n, Y_n, W_n) under stronger assumptions. Note that the concepts are identical for any class of split trees where each node holds exactly one ball, such as binary search trees and digital search trees.

Let d_2 denote the Mallows metric, also called the minimal ℓ_2 metric (defined in Section 2). Let $\mathcal{M}_{0,2}^d$ be the set of probability measures on $\mathbb{R}^d$ with zero mean and finite second moment.

▶ **Theorem 5.** *Let T_n be a split tree and let $\mathcal{V} = (V_1, \ldots, V_b)$ be a split vector. Define*

$$\mu = -\sum_{i=1}^{b} \mathbb{E}\left[V_i \ln V_i\right], \qquad \text{and} \qquad D(\mathcal{V}) = \frac{1}{\mu} \sum_{i=1}^{b} V_i \ln V_i.$$

Assume that $\mathbb{P}\{\exists i : V_i = 1\} < 1$ and $s_0 > 0$. Let $(\hat{X}, \hat{Y}, \hat{W})$ be the unique solution in $\mathcal{M}_{0,2}^3$ for the system of fixed-point equations

$$\begin{bmatrix} \hat{X} \\ \hat{Y} \\ \hat{W} \end{bmatrix} \stackrel{d}{=} \begin{bmatrix} \sum_{i=1}^{b} V_i \hat{X}^{(i)} + \sum_{j=1}^{s_0} U_j + \frac{s_0}{2} D(\mathcal{V}) \\ \sum_{i=1}^{b} V_i \hat{Y}^{(i)} + \sum_{j=1}^{s_0} (U_j - 1/2) \\ \sum_{i=1}^{b} V_i \hat{W}^{(i)} + 1 + D(\mathcal{V}) \end{bmatrix}. \qquad (1.4)$$

Here $(V_1, \ldots, V_b)$, $U_1, \ldots, U_{s_0}$, $(\hat{X}^{(1)}, \hat{Y}^{(1)}, \hat{W}^{(1)}), \ldots, (\hat{X}^{(b)}, \hat{Y}^{(b)}, \hat{W}^{(b)})$ are independent, with $U_j \sim \mathrm{Unif}[0,1]$ for $j = 1, \ldots, s_0$, and $\left(\hat{X}_n^{(i)}, \hat{Y}_n^{(i)}, \hat{W}_n^{(i)}\right) \sim (\hat{X}, \hat{Y}, \hat{W})$ for $i = 1, \ldots, b$. Then the sequence $(\hat{X}_n, \hat{Y}_n, \hat{W}_n)$ defined in (1.3) converges to $(\hat{X}, \hat{Y}, \hat{W})$ in d_2 and in moment generating function within a neighborhood of the origin.

The proof of Theorem 5 uses the contraction method, introduced by Rösler [21] for finding the total path length of binary search trees. The technique has been applied to d-dimensional quad trees by Neininger and Rüschendorf [19] and to split trees in general by Broutin and Holmgren [3].

Conditional Galton–Watson trees

A conditional Galton–Watson tree T_n is a Galton–Watson tree conditioned on having n nodes, which we define in details in Section 3. It generalizes many uniform random tree models, e.g., Cayley trees, Catalan trees, binary trees, b-ary trees, and Motzkin trees. For a comprehensive survey, see Janson [12].

Aldous [1] showed that many asymptotic properties of conditional Galton–Watson trees, such as the height and the total path length, can be derived from properties of Brownian excursions. Our analysis of inversions follows a similar route. In particular, we relate $I(T_n)$ to the *Brownian snake* studied by e.g., Janson and Marckert [14].

In the context of Galton–Watson trees, Aldous [1, Corollary 9] showed that $n^{-3/2}\Upsilon(T_n)$ converges to an Airy distribution. We will see that the standard deviation of $I(T_n) - \frac{1}{2}\Upsilon(T_n)$ is of order $n^{5/4} \ll n^{3/2}$, which by the decomposition (1.2) implies that $n^{-3/2}I(T_n)$ converges to the same Airy distribution, recovering one of the main results of Panholzer and Seitz [20, Theorem 5.3]. Our contribution for conditional Galton–Watson trees is a detailed analysis of Y_n under the scaling function $s(n) = n^{5/4}$.

Let $e(s), s \in [0,1]$ be the random path of a standard Brownian excursion, and define $C(s,t) \stackrel{\text{def}}{=} C(t,s) \stackrel{\text{def}}{=} 2\min_{s \le u \le t} e(u)$ for $0 \le s \le t \le 1$.

We define a random variable, see [11],

$$\eta \stackrel{\text{def}}{=} \int_{[0,1]^2} C(s,t)\mathrm{d}s\,\mathrm{d}t = 4\int_{0 \le s \le t \le 1} \min_{s \le u \le t} e(u). \tag{1.5}$$

▶ **Theorem 6.** *Suppose T_n is a conditional Galton–Watson tree with offspring distribution ξ such that $\mathbb{E}[\xi] = 1$, $\mathrm{Var}(\xi) = \sigma^2 \in (0,\infty)$, and $\mathbb{E}[e^{\alpha\xi}] < \infty$ for some $\alpha > 0$, and define*

$$Y_n = \frac{I(T_n) - \frac{1}{2}\Upsilon(T_n)}{n^{5/4}}.$$

Then we have

$$Y_n \stackrel{\mathrm{d}}{\longrightarrow} Y \stackrel{\text{def}}{=} \frac{1}{\sqrt{12\sigma}}\sqrt{\eta}\,\mathcal{N}, \tag{1.6}$$

where $\mathcal{N}$ is a standard normal random variable, independent from the random variable η defined in (1.5). Moreover, $\mathbb{E}[e^{tY_n}] \to \mathbb{E}[e^{tY}] < \infty$ for all fixed $t \in \mathbb{R}$.

In the rest of the paper, we outline the proofs of our main results, Theorem 5 and 6. The proofs of Theorem 2 and 3 are omitted. The details of the proofs can be found in the longer version of this paper [4].

2   A sequence of split trees

In this section we outline how one can apply the *contraction method* to prove Theorem 5.

We will now define split trees introduced by Devroye [6]. The random split tree T_n has parameters $b, s, s_0, s_1, \mathcal{V}$ and n. The integers b, s, s_0, s_1 are required to satisfy the inequalities

$$2 \le b, \quad 0 < s, \quad 0 \le s_0 \le s, \quad 0 \le bs_1 \le s + 1 - s_0. \tag{2.1}$$

and $\mathcal{V} = (V_1, \ldots, V_b)$ is a random non-negative vector with $\sum_{i=1}^{b} V_i = 1$. Consider an infinite b-ary tree $\mathcal{U}$. The split tree T_n is constructed by distributing n balls (pieces of information) among nodes of $\mathcal{U}$. For a node u, let n_u be the number of balls stored in the subtree rooted at u. Once n_u are all decided, we take T_n to be the largest subtree of $\mathcal{U}$ such that $n_u > 0$

for all $u \in T_n$. Let the split vector $\mathcal{V} \in [0,1]^b$ be as before. Let $\mathcal{V}_u = (V_{u,1}, \ldots, V_{u,b})$ be the independent copy of $\mathcal{V}$ assigned to u. Let $u_1, \ldots, u_b$ be the child nodes of u. Conditioning on n_u and $\mathcal{V}_u$, if $n_u \leq s$, then $n_{u_i} = 0$ for all i; if $n_u > s$, then

$$(n_{u_1}, \ldots, n_{u_b}) \sim \mathrm{Mult}(n - s_0 - bs_1, V_{u,1}, \ldots, V_{u,b}) + (s_1, s_1, \ldots, s_1),$$

where Mult denotes multinomial distribution, and b, s, s_0, s_1 are integers satisfying (2.1). Note that $\sum_{i=1}^b n_{u_i} \leq n$ (hence the "splitting"). Naturally for the root ρ, $n_\rho = n$. Thus the distribution of $(n_u, \mathcal{V}_u)_{u \in V(\mathcal{U})}$ is completely defined.

Once all n balls have been placed in $\mathcal{U}$, we obtain T_n by deleting all nodes u such that the subtree rooted at u contains no balls. Note that an internal node of T_n contains exactly s_0 balls, while a leaf contains a random amount in $\{1, \ldots, s\}$. We assume, as previous authors, that $\mathbb{P}\{\exists i : V_i = 1\} < 1$. We can assume that $\mathcal{V}$ has a permutation invariant distribution without loss of generality, since a uniform random permutation of subtree order does not change the number of inversions.

2.1 Outline

Recall that in (1.3), we define $\hat{X}_n, \hat{Y}_n$ and $\hat{W}_n$. Let $\overline{n} = (n_1, \ldots, n_b)$ denote the vector of the (random) number of balls in each of the b subtrees of the root. Broutin and Holmgren [3] showed that, conditioning on $\overline{n}$,

$$\hat{W}_n \overset{\mathrm{d}}{=} \sum_{i=1}^b \frac{n_i}{n} \hat{W}_{n_i} + \frac{n - s_0}{n} + \hat{D}_n(\overline{n}), \quad \hat{D}_n(\overline{n}) \overset{\mathrm{def}}{=} -\frac{\mathbb{E}\left[\hat{\Upsilon}(T_n)\right]}{n} + \sum_{i=1}^b \frac{\mathbb{E}\left[\hat{\Upsilon}(T_{n_i})\right]}{n}.$$

We derive similar recursions for $\hat{X}_n$ and $\hat{Y}_n$. Conditioning on $\overline{n}$, $\hat{I}(T_n)$ satisfies the recursion

$$\hat{I}(T_n) \overset{\mathrm{d}}{=} \hat{Z}_\rho + \sum_{i=1}^b \hat{I}(T_{n_i}),$$

where $\hat{Z}_\rho$ denotes the number of inversions involving balls contained in the root ρ. Therefore, still conditioning on $\overline{n}$, we have

$$\hat{X}_n \overset{\mathrm{d}}{=} \sum_{i=1}^b \frac{n_i}{n} \hat{X}_{n_i} + \frac{\hat{Z}_\rho}{n} - \frac{\mathbb{E}\left[\hat{I}(T_n)\right]}{n} + \sum_{i=1}^b \frac{\mathbb{E}\left[\hat{I}(T_{n_i})\right]}{n}$$

$$= \sum_{i=1}^b \frac{n_i}{n} \hat{X}_{n_i} + \frac{\hat{Z}_\rho}{n} - \frac{s_0}{2} \frac{\mathbb{E}\left[\hat{\Upsilon}(T_n)\right]}{n} + \frac{s_0}{2} \sum_{i=1}^b \frac{\mathbb{E}\left[\hat{\Upsilon}(T_{n_i})\right]}{n}$$

$$= \sum_{i=1}^b \frac{n_i}{n} \hat{X}_{n_i} + \frac{\hat{Z}_\rho}{n} + \frac{s_0}{2} \hat{D}_n(\overline{n}),$$

where we use that

$$\mathbb{E}\left[\hat{I}(T_n) \mid T_n\right] = \frac{s_0}{2} \hat{\Upsilon}(T_n). \tag{2.2}$$

It follows also from (2.2) that $\hat{X}_n = \hat{Y}_n + \frac{s_0}{2} \hat{W}_n$ and

$$\hat{Y}_n \overset{\mathrm{d}}{=} \sum_{i=1}^b \frac{n_i}{n} \hat{Y}_{n_i} + \frac{\hat{Z}_\rho}{n} - \frac{s_0}{2} \frac{n - s_0}{n}.$$

It is not difficult to see that

$$\frac{\hat{Z}_\rho}{n} \xrightarrow{L^2} U_1 + \cdots + U_{s_0},$$

where $U_1, \ldots, U_{s_0}$ are independent and uniformly distributed in $[0, 1]$. Broutin and Holmgren [3] have shown that $\hat{D}_n(\overline{n}) \xrightarrow{\text{a.s.}} D(\mathcal{V})$, where

$$\mu = -\sum_{i=1}^{b} \mathbb{E}\left[V_i \ln V_i\right], \qquad \text{and} \qquad D(\mathcal{V}) = \frac{1}{\mu} \sum_{i=1}^{b} V_i \ln V_i.$$

Together with $(n_1/n, \ldots, n_b/n) \xrightarrow{\text{a.s.}} (V_1, \ldots, V_b)$ (by the law of large number), we arrive at the fixed-point equations (1.4) presented in Theorem 5.

For a random vector $X \in \mathbb{R}^d$, let $\|X\|$ be the Euclidean norm of X. Let $\|X\|_2 \overset{\text{def}}{=} \sqrt{\mathbb{E}\left[\|X\|^2\right]}$. Recall that $\mathcal{M}_{0,2}^d$ denotes the set of probability measures on $\mathbb{R}^d$ with zero mean and finite second moment. The Mallows metric on $\mathcal{M}_{0,2}^d$ is defined by

$$d_2(\nu, \lambda) = \inf\left\{\|X - Y\|_2 : X \sim \lambda, Y \sim \nu\right\}.$$

Using the contraction method, Broutin and Holmgren [3] proved that $\hat{W}_n \xrightarrow{d_2} \hat{W}$, the unique solution of the last equation of (1.4) in $\mathcal{M}_{0,2}^1$.

We can apply the same contraction method to show that the vector $(\hat{X}_n, \hat{Y}_n, \hat{W}_n) \xrightarrow{d_2} (\hat{X}, \hat{Y}, \hat{W})$, the unique solution of (1.4) in $\mathcal{M}_{0,2}^3$. Assume that the independent vectors $\left(\hat{X}^{(i)}, \hat{Y}^{(i)}, \hat{W}^{(i)}\right)$, $i = 1, \ldots, b$ share some common distribution $\mu \in \mathcal{M}_{0,2}^3$. Let $F(\mu) \in \mathcal{M}_{0,2}^3$ be the distribution of the random vector given by the right hand side of (1.4). Using a coupling argument, we can show that for all $\nu, \lambda \in \mathcal{M}_{0,2}^3$,

$$d_2(F(\nu), F(\lambda)) < c d_2(\nu, \lambda),$$

where $c \in (0, 1)$ is a constant. Thus F is a contraction and by Banach's fixed point theorem, (1.4) must have a unique solution $(\hat{X}, \hat{Y}, \hat{W}) \in \mathcal{M}_{0,2}^3$. Finally, we can use a similar coupling argument to show that $(\hat{X}_n, \hat{Y}_n, \hat{W}_n) \xrightarrow{d_2} (\hat{X}, \hat{Y}, \hat{W})$.

Note that in [4], instead of carrying out the above argument in details, we actually used a result by Neininger [18] which gives us a shortcut.

3   A sequence of conditional Galton–Watson trees

Let ξ be a random variable with $\mathbb{E}[\xi] = 1$, $\text{Var } \xi = \sigma^2 < \infty$, and $\mathbb{E}\left[e^{\alpha\xi}\right] < \infty$ for some $\alpha > 0$, (The last condition is used in the proof below, but is presumably not necessary.) Let G^ξ be a (possibly infinite) Galton–Watson tree with offspring distribution ξ. The *conditional Galton–Watson tree* T_n^ξ on n nodes is given by

$$\mathbb{P}\left\{T_n^\xi = T\right\} = \mathbb{P}\left\{G^\xi = T \mid G^\xi \text{ has } n \text{ nodes}\right\}$$

for any rooted tree T on n nodes. The assumption $\mathbb{E}[\xi] = 1$ is justified by noting that if ζ is such that $\mathbb{P}\{\xi = i\} = c\theta^i \mathbb{P}\{\zeta = i\}$ for all $i \geq 0$ then T_n^ξ and T_n^ζ are identically distributed; hence it is typically possible to replace an offspring distribution ζ by an equivalent one with mean 1, see [12, Sec. 4].

We fix some ξ and drop it from the notation, writing $T_n = T_n^\xi$.

In a fixed tree T with root ρ and n total nodes, for each node $v \neq \rho$ let $Q_v \sim$ Unif$(-1/2, 1/2)$, all independent, and let $Q_\rho = 0$. For each node v define

$$\Phi_v \overset{\text{def}}{=} \sum_{u \leq v} Q_u, \quad \text{and let} \quad J(T) \overset{\text{def}}{=} \sum_{v \in T} \Phi_v.$$

In other words, Φ_u is the sum of Q_v for all v on the path from the root to u. For each $v \neq \rho$ also define $Z_v = \lfloor (Q_v + 1/2)z_v \rfloor$, where z_v denotes the size of the subtree rooted at v. Then Z_v is uniform in $\{0, 1, \ldots, z_v - 1\}$, and by Lemma 1, the quantity

$$I^*(T) \overset{\text{def}}{=} \sum_{v \neq \rho} \big(Z_v - \mathbb{E}[Z_v]\big)$$

is equal in distribution to the centralized number of inversions in the tree T, ignoring inversions involving ρ. The main part (1.6) of Theorem 6 will follow from arguing that for a conditional Galton–Watson tree T_n,

$$\frac{J(T_n)}{n^{5/4}} \overset{\text{d}}{\longrightarrow} Y \overset{\text{def}}{=} \frac{1}{\sqrt{12}\sigma}\sqrt{\eta}\mathcal{N}. \tag{3.1}$$

Indeed, under the coupling of Q_v and Z_v above,

$$J(T_n) = \sum_v \Phi_v = \sum_v \sum_{u:u \leq v} Q_u = \sum_u Q_u \sum_{v:u \leq v} 1 = \sum_u Q_u z_u$$

$$\leq \sum_{u \neq \rho} \left(Z_u - \frac{z_u}{2} + 1\right) < n + I^*(T_n),$$

and similarly $J(T_n) > I^*(T_n) - n$. As ρ contributes at most n inversions to $I(T_n)$, it follows from the triangle inequality that $|J(T_n) - (I(T_n) - \Upsilon(T_n)/2)| \leq 2n = o(n^{5/4})$. Thus (3.1), once proved, will imply that

$$Y_n \overset{\text{def}}{=} \frac{I(T_n) - \Upsilon(T_n)/2}{n^{5/4}} = o(1) + \frac{J(T_n)}{n^{5/4}} \overset{\text{d}}{\longrightarrow} Y.$$

The quantity $J(T_n)$ and the limiting distribution (3.1) have been considered by several authors. In the interest of keeping this section self-contained, we will now outline the proof of (3.1) which relies on the concept of a *discrete snake*, a random curve which under proper rescaling converges to a *Brownian snake*, a curve related to a standard Brownian excursion. This convergence was shown by Gittenberger [10], and later in more generality by Janson and Marckert [14], whose notation we use.

Define $f : \{0, \ldots, 2(n-1)\} \to V$ by saying that $f(i)$ is the location of a depth-first search (under some fixed ordering of nodes) at stage i, with $f(0) = f(2(n-1)) = \rho$. Also define $V_n(i) = d(\rho, f(i))$ where d denotes distance. The process $V_n(i)$ is called the depth-first walk, the Harris walk or the tour of T_n. For non-integer values t, $V_n(t)$ is given by linearly interpolating adjacent values. See Figure 1.

Finally, define $R_n(i) \overset{\text{def}}{=} \Phi_{f(i)}$ to be the value at the vertex visited after i steps. For non-integer values t, $R_n(t)$ is defined by linearly interpolating the integer values. Also define $\widetilde{R}_n(t)$ by $\widetilde{R}_n(t) \overset{\text{def}}{=} R_n(t)$ when $t \in \{0, 1, \ldots, 2n\}$, and

$$\widetilde{R}_n(t) \overset{\text{def}}{=} \begin{cases} R_n(\lfloor t \rfloor), & \text{if} \quad V_n(\lfloor t \rfloor) > V_n(\lceil t \rceil), \\ R_n(\lceil t \rceil), & \text{if} \quad V_n(\lfloor t \rfloor) < V_n(\lceil t \rceil). \end{cases}$$
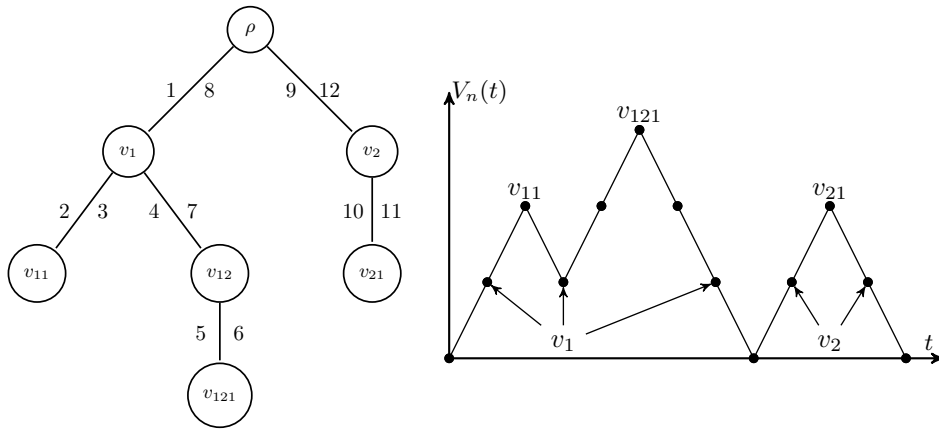
Figure 1 The depth-first walk $V_n(t)$ of a fixed tree.

In other words, $\widetilde{R}_n(t)$ takes the value of node $f(\lfloor t \rfloor)$ or $f(\lceil t \rceil)$, whichever is further from the root. We can recover $J(T_n)$ from $\widetilde{R}_n(t)$ via

$$2J(T_n) = \int_0^{2(n-1)} \widetilde{R}_n(t)dt.$$

Indeed, for each non-root node v there are precisely two unit intervals during which $\widetilde{R}_n(t)$ draws its value from v, namely the two unit intervals during which the parent edge of v is being traversed. Now, since $Q_v \sim \mathrm{Unif}(-1/2, 1/2)$ we have $|R_n(i) - R_n(i-1)| \leq 1/2$ for all $i > 0$ and

$$\frac{J(T_n)}{n^{5/4}} = \frac{1}{2n^{5/4}} \int_0^{2(n-1)} \widetilde{R}_n(t)\mathrm{d}t = \frac{1}{2n^{5/4}} \int_0^{2(n-1)} R_n(t)dt + O(n^{-1/4}) = \int_0^1 r_n(s)ds + o(1),$$

where $r_n(s) \stackrel{\text{def}}{=} n^{-1/4}R_n(2(n-1)s)$. Also normalize $v_n(s) \stackrel{\text{def}}{=} n^{-1/2}V_n(2(n-1)s)$. Theorem 2 of [14] (see also [10]) states that $(r_n, v_n) \stackrel{\mathrm{d}}{\longrightarrow} (r, v)$ in $C[0,1] \times C[0,1]$, with r, v to be defined shortly.

Before defining r and v, we will briefly motivate what they ought to be. Firstly, as the offspring distribution ξ of T_n satisfies $\mathbb{E}[\xi] = 1$, we expect the tour V_n to be roughly a random walk with zero-mean increments, conditioned to be non-negative and return to the origin at time $2(n-1)$, and the limiting law v ought to be a Brownian excursion (up to a constant scale factor). Secondly, consider a node u and the path $\rho = u_0, u_1 \ldots, u_d = u$, where d is the depth of u. We can define a random walk $\Phi_u(t)$ for $t = 0, \ldots, d$ by $\Phi_u(0) = 0$ and $\Phi_u(t) = \sum_{i=1}^t Q_{u_i}$ for $t > 0$, noting that $\Phi_u = \Phi_u(d)$. Under rescaling, the random walk $\Phi_u(t)$ will behave like Brownian motion. For any two nodes u_1, u_2 with last common ancestor at depth m, the processes Φ_{u_1}, Φ_{u_2} agree for $t = 0, \ldots, m$, while any subsequent increments are independent. Hence $\mathrm{Cov}(\Phi_{u_1}, \Phi_{u_2}) = cm$ for some constant $c > 0$. Now, for any $i, j \in \{0, \ldots, 2(n-1)\}$, the nodes $f(i), f(j)$ at depths $V_n(i), V_n(j)$ have last common ancestor $f(k)$, where k is such that $V_n(k)$ is minimal in the range $i \leq k \leq j$. Hence $r(s)$ should be normally distributed with variance given by $v(s)$, and the covariance of $r(s), r(t)$ proportional to $\min_{s \leq u \leq t} v(u)$.

We now define r, v precisely. If $\mathrm{Var}\, \xi = \sigma^2$, then $v(s) \stackrel{\text{def}}{=} 2\sigma^{-1}e(s)$, where $e(s)$ is a standard Brownian excursion, as shown by Aldous [1, 2]. Conditioning on v, we define r as a centered

Gaussian process on $[0, 1]$ with

$$\text{Cov}(r(s), r(t) \mid v) = \frac{1}{12} \min_{s \le u \le t} v(u) = \frac{1}{12\sigma} C(s, t), \qquad s \le t.$$

The constant $1/12$ appears as the variance of the random increments Q_v. Again, Theorem 2 of [14] states that $(r_n, v_n) \overset{d}{\longrightarrow} (r, v)$ in $C[0, 1]^2$. We conclude that

$$\lim_{n \to \infty} \frac{J(T_n)}{n^{5/4}} = \int_0^1 r_n(t)dt + o(1) \overset{d}{\longrightarrow} \int_0^1 r(t)dt \overset{\text{def}}{=} Y.$$

This integral is the object of study in [13], wherein it is shown that

$$Y \overset{\text{def}}{=} \int_0^1 r(t)dt \overset{d}{=} \frac{1}{\sqrt{12\sigma}} \sqrt{\eta}\, \mathcal{N},$$

where $\mathcal{N}$ is a standard normal variable, η is given by

$$\eta = \int_{[0,1]^2} C(s, t)\mathrm{d}s\,\mathrm{d}t,$$

and $\eta, \mathcal{N}$ are independent. The odd moments of Y are zero, as this is the case for $\mathcal{N}$, and by [13, Theorem 1.1], for $k \ge 0$

$$\mathbb{E}\left[Y^{2k}\right] = \frac{1}{(12\sigma)^k} \frac{(2k!)\sqrt{\pi}}{2^{(9k-4)/2}\Gamma((5k-1)/2)} a_k,$$

where $a_1 = 1$ and for $k \ge 2$,

$$a_k = 2(5k - 4)(5k - 6)a_{k-1} + \sum_{i=1}^{k-1} a_i a_{k-i}.$$

In particular ([13, Theorem 1.2]),

$$\mathbb{E}\left[Y^{2k}\right] \sim \frac{1}{(12\sigma)^k} \frac{2\pi^{3/2}\beta}{5} (2k)^{1/2}(10e^3)^{-2k/4}(2k)^{\frac{3}{4}\cdot 2k},$$

as $k \to \infty$, where $\beta = 0.981038\ldots$. Further analysis of the moments of η and Y, including the moment generating function and tail estimates, can be found in [13].

The last bit of Theorem 6 which remains to be proved is that $\mathbb{E}\left[e^{tY_n}\right] \to \mathbb{E}\left[e^{tY}\right]$ for all fixed $t \in \mathbb{R}$. Since we have already shown $Y_n \overset{d}{\longrightarrow} Y$, we can apply the Vitali convergence theorem once we have shown that the sequence e^{tY_n} is uniformly integrable. See Section 5.1 of [4] for details.

---- **References** ----

1   David Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge Univ. Press, Cambridge, 1991.

2   David Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.

3   Nicolas Broutin and Cecilia Holmgren. The total path length of split trees. *Ann. Appl. Probab.*, 22(5):1745–1777, 2012.

4   Xing Shi Cai, C. Holmgren, S. Janson, T. Johansson, and F. Skerman. Inversions in split trees and conditional Galton–Watson trees. *ArXiv e-prints*, 2017. `arXiv:1709.00216`.

5 G. Cramer. *Introduction à l'analyse des lignes courbes algébriques*. Cramer, 1750.

6 Luc Devroye. Universal limit laws for depths in random trees. *SIAM J. Comput.*, 28(2):409–432, 1999.

7 William Feller. *An introduction to probability theory and its applications. Vol. I.* John Wiley & Sons, Inc., New York-London-Sydney, 3rd edition, 1968.

8 P. Flajolet, P. Poblete, and A. Viola. On the analysis of linear probing hashing. *Algorithmica*, 22(4):490–515, 1998.

9 Ira M. Gessel, Bruce E. Sagan, and Yeong Nan Yeh. Enumeration of trees by inversions. *J. Graph Theory*, 19(4):435–459, 1995.

10 Bernhard Gittenberger. A note on: "State spaces of the snake and its tour—convergence of the discrete snake". *J. Theoret. Probab.*, 16(4):1063–1067 (2004), 2003.

11 Svante Janson. The Wiener index of simply generated random trees. *Random Structures Algorithms*, 22(4):337–358, 2003.

12 Svante Janson. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation: extended abstract. In *23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12)*, Discrete Math. Theor. Comput. Sci. Proc., AQ, pages 479–490. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2012.

13 Svante Janson and Philippe Chassaing. The center of mass of the ISE and the Wiener index of trees. *Electron. Comm. Probab.*, 9:178–187, 2004.

14 Svante Janson and Jean-François Marckert. Convergence of discrete snakes. *J. Theoret. Probab.*, 18(3):615–647, 2005.

15 Donald E. Knuth. *The art of computer programming. Vol. 3.* Addison-Wesley, Reading, MA, 1998.

16 C. L. Mallows and John Riordan. The inversion enumerator for labeled trees. *Bull. Amer. Math. Soc.*, 74:92–94, 1968.

17 Barbara H. Margolius. Permutations with inversions. *J. Integer Seq.*, 4(2):Article 01.2.4, 13, 2001.

18 Ralph Neininger. On a multivariate contraction method for random recursive structures with applications to Quicksort. *Random Structures Algorithms*, 19(3-4):498–524, 2001. Analysis of algorithms (Krynica Morska, 2000).

19 Ralph Neininger and Ludger Rüschendorf. On the internal path length of d-dimensional quad trees. *Random Structures Algorithms*, 15(1):25–41, 1999.

20 Alois Panholzer and Georg Seitz. Limiting distributions for the number of inversions in labelled tree families. *Ann. Comb.*, 16(4):847–870, 2012.

21 Uwe Rösler. A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl.*, 25(1):85–100, 1991.

22 Vladimir N. Sachkov. *Probabilistic methods in combinatorial analysis*, volume 56 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1997. Translated from the Russian, Revised by the author.

The Cover Time of a Biased Random Walk on a Random Cubic Graph

Colin Cooper[1]

Department of Informatics, King's College, University of London, London WC2R 2LS, UK
colin.cooper@kcl.ac.uk

Alan Frieze[2]

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA 15213, USA
alan@random.math.cmu.edu

Tony Johansson[3]

Department of Mathematics, Uppsala University, 751 05 Uppsala, Sweden
tony.johansson@math.uu.se

──── Abstract ────

We study a random walk that prefers to use unvisited edges in the context of random cubic graphs, i.e., graphs chosen uniformly at random from the set of 3-regular graphs. We establish asymptotically correct estimates for the vertex and edge cover times, these being $n \log n$ and $\frac{3}{2} n \log n$ respectively.

1 Introduction

Our aim in this paper is to analyse a variation on the simple random walk that may tend to speed up the cover time of a connected graph. A simple random walk on a graph is a walk which repeatedly moves from its currently occupied vertex v to one of its neighbours, chosen uniformly at random. The vertex cover time $T_{\text{COV}}^V(G)$ of a simple random walk on a graph G is the expected number of steps needed to visit each vertex of G, defined as the maximum over all starting vertices. Feige [9, 10] showed that for any graph G on n vertices,

$$(1 - o(1))n \log n \leq T_{\text{COV}}^V(G) \leq (1 + o(1))\frac{4}{27}n^3.$$

When G is chosen uniformly at random from the set of d-regular graphs, Cooper and Frieze [6] showed that w.h.p.[4] G is such that $T_{\text{COV}}^V(G)$ is asymptotically equal to $\frac{d(d-1)}{2(d-2)} n \log n$.

In recent years, variations of the simple random walk have been introduced with the aim of achieving faster cover times. In this paper we do this by choosing to walk along unvisited edges whenever possible. This variation is just one of several possible approaches which include non-backtracking walks, see Alon, Benjamini, Lubetzky and Sodin [3], or walks that

---

are biased toward low degree vertices, see Cooper, Frieze and Petti [8], or any number of other ideas.

The papers [4], [13] describe the random walk model considered here, which uses unvisited edges when available at the currently occupied vertex. If there are unvisited edges incident with the current vertex, the walk picks one u.a.r.[5] and makes a transition along this edge. If there are no unvisited edges incident with the current vertex, the walk moves to a random neighbour. In [4] this walk was called an *unvisited edge process* (or edge-process), and in [13], a *greedy random walk*. We use the name *biased random walk* for the same process. For random d-regular graphs where $d = 2k$ (d even), it was shown in [4] that the biased random walk has vertex cover time $\Theta(n)$, which is best possible up to a constant. The paper also gives an upper bound of $O(n\omega)$ for the edge cover time. The ω factor comes from the fact that cycles of length at most ω exist w.h.p. In [7], the constant for the vertex cover time was shown to be $d/2$.

▶ **Theorem 1.** *Let $d \geq 4$ be even and suppose G is chosen u.a.r. from the set of d-regular graphs. W.h.p., G is such that the vertex cover time of the biased random walk is[6]*
$$T_{cov}^V(G) \approx dn/2.$$

This is faster than any of the other random graph models mentioned here by a factor of $\log n$, and the biased random walk generally performs well on even-degree graphs. Orenshtein and Shinkar [13, Lemma 2.9] showed that in an even-degree graph, the biased random walk has cover time at most that of the simple random walk plus the number of edges in the graph. Briefly, this is because there are at most two vertices incident to an odd number of unvisited edges at any time. In the random setting this means that the most likely scenario is that traversing an unvisited (random) edge will bring the walk to a vertex incident to at least one more unvisited edge, and the walk will find a large number of unvisited edges in succession. This is no longer true in odd-degree graphs. The paper [4] included experimental data for the performance of red-blue walks on odd degree regular graphs. Namely, for $d = 3$ the vertex cover time is $\Theta(n \log n)$ and decreases rapidly with increasing d.

Random walks have applications in networks where each vertex only has local information, e.g. each vertex knows only of its immediate neighbours. For example, random walks provide efficient routing algorithms in Wireless Sensor Networks [15]. The vertex cover time measures the expected number of steps needed to spread information to each vertex of the network. A drawback of biased random walks in general applications is that it requires $O(|E|)$ additional memory usage, but in networks with independently acting agents, the additional memory for each agent is $O(\Delta)$ where Δ denotes the maximum degree of the network.

1.1   Our results

Let $G = (V, E)$ be a connected cubic (i.e. 3-regular) (multi)graph on an even number n of vertices. Consider the following random walk process, called a *biased random walk*. Initially color all edges red, and pick a starting vertex v_0. At any time, if the walk occupies a vertex incident to at least one red edge, then the walk traverses one of those red edges chosen uniformly at random, and re-colors it blue. If no such edge is available, the walk traverses a blue edge chosen uniformly at random. For $s \in \{1, \ldots, n\}$ let $C_V(s)$ denote the number of steps taken by the walk until it has visited s vertices, and similarly let $C_E(t)$ denote the number of steps taken to visit $t \in \{1, \ldots, 3n/2\}$ edges.

---

[5] We use u.a.r. for *uniformly at random*.
[6] We say that $a_n \approx b_n$ if $\lim a_n/b_n = 1$.

We will let G be a random graph, and we use $\mathbb{E}_G(X)$ to denote the expectation of X with the underlying graph G fixed. Note that a cubic graph on n vertices contains exactly $3n/2$ edges.

▶ **Theorem 2.** *Let s, t be such that $n - n \log^{-1} n \le s \le n$ and $(1 - \log^{-2} n)\frac{3n}{2} \le t \le 3n/2$. Let $\varepsilon > 0$ also be fixed. Suppose G is chosen uniformly at random from the set of cubic graphs on n vertices. Then w.h.p., G is connected and*

$$\mathbb{E}_G(C_V(s)) = (1 \pm \varepsilon)n \log\left(\frac{n}{n-s+1}\right) + o(n \log n), \tag{1}$$

$$\mathbb{E}_G(C_E(t)) = \left(\frac{3}{2} \pm \varepsilon\right) n \log\left(\frac{3n}{3n-2t+1}\right) + o(n \log n). \tag{2}$$

Here $a = b \pm c$ is taken to mean $a \in [b-c, b+c]$. Note in particular that this shows that the expected vertex and edge cover times are asymptotically $n \log n$ and $\frac{3}{2}n \log n$ w.h.p., respectively. The same statement is true with the word "graphs" replaced by "configuration multigraphs" (defined in Section 3). Thus, taking $s = n$ and $t = 3n/2$ we have the following corollary.

▶ **Corollary 3.** *Suppose G is chosen uniformly at random from the set of cubic graphs on n vertices. W.h.p., G is such that the vertex cover time $T_{cov}^V(G)$ of G is asymptotically equal to $n \log n$ and the edge cover time $T_{cov}^E(G)$ is asymptotically equal to $\frac{3}{2}n \log n$.*

Cooper and Frieze [6] showed that w.h.p. the vertex cover time for a simple random walk on a random d-regular graph on n vertices is asymptotically equal to $\frac{d-1}{d-2}n \log n$. The argument there also shows that the edge cover time of a random d-regular graph on n vertices is asymptotically equal to $\frac{d(d-1)}{2(d-2)}n \log n$. For $d = 3$ these values are $2n \log n$ and $3n \log n$ respectively and are to be compared with $n \log n$ and $\frac{3}{2}n \log n$. For a non-backtracking random walk, Cooper and Frieze [7] show that the vertex and edge cover times are asymptotically $n \log n$ and $\frac{3}{2}n \log n$ respectively. Interestingly, these values coincide with the results in Corollary 3.

1.2 Outlook

Our proof relies on the fact that the set of vertices incident to exactly one unvisited edge coincides with the set of vertices visited exactly once by the biased random walk, modulo the head and tail of the walk. This is no longer true when $d \ge 5$, and additional analysis would be required to extend the method to larger degrees. We expect the walk to behave similarly for higher degrees and conjecture that Corollary 3 generalizes to $T_{\text{cov}}^V(G_d) \approx \frac{1}{d-2}n \log n$ and $T_{\text{cov}}^E(G_d) \approx \frac{d}{2(d-2)}n \log n$ for the random d-regular graph G_d, for any odd $d \ge 3$.

For fixed graphs, the behaviour of the greedy random walk is not well understood. See [13] for a list of open problems, including questions regarding transience and recurrence on infinite lattices.

2 Outline proof of Theorem 2

We will choose the multigraph G according to the configuration model. Each vertex v of G is associated with a set $\mathcal{P}(v)$ of 3 configuration points. We set $\mathcal{P} = \cup_v \mathcal{P}(v)$ and generate G by choosing a pairing μ of $\mathcal{P}$ uniformly at random. The pairing μ is exposed along with the biased random walk. See Section 3 for more details on the configuration model.

Starting at a uniformly random configuration point $x_1 \in \mathcal{P}$, we define $W_0 = (x_1)$. Given a walk $W_k = (x_1, x_2, \ldots, x_{2k+1})$, the walk proceeds as follows. Set $x_{2k+2} = \mu(x_{2k+1})$, thus exposing the value of $\mu(x_{2k+1})$ if not previously exposed. If x_{2k+2} belongs to a vertex v which is incident to some red edge (other than (x_{2k+1}, x_{2k+2}) which is now recoloured blue), the walk chooses one of the red edges uniformly at random, setting x_{2k+3} to be the corresponding configuration point. Otherwise, x_{2k+3} is chosen uniformly at random from $\mathcal{P}(v)$. Set $W_{k+1} = (x_1, \ldots, x_{2k+3})$. We will refer to x_1 and x_{2k+1} (and the vertices to which they belong) as the *tail* and *head* of W_k, respectively. We will also refer to $\{x_1, x_2, \ldots, x_{2k+1}\}$ as the points of $\mathcal{P}$ that have been *visited*.

Define partial edge and vertex cover times

$$C_E(t) = \min\{k : W_k \text{ spans } t \text{ edges}\}, \tag{3}$$

$$C_V(t) = \min\{k : W_k \text{ spans } t \text{ vertices}\}. \tag{4}$$

We will mainly be concerned with the partial edge cover time, and write $C(t) = C_E(t)$ from this point on.

For $t \in \{1, 2, \ldots, \frac{3n}{2}\}$ we define a subsequence of walks by

$$W(t) = W_{C(t)-1} = (x_1, x_2, \ldots, x_{2k+1}) \tag{5}$$

where k is the smallest integer such that $|\{x_1, x_2, \ldots, x_{2k+1}\}| = 2t - 1$. In other words, $W(t)$ denotes the walk up to the point when $2t - 1$ of the members of $\mathcal{P}$ have been visited. Thus throughout the paper:

- Time t is measured by the number of edges t that have been visited at least once.
- The parameter $\delta = \delta(t)$ is given by the equation

$$t = (1 - \delta)\frac{3n}{2}. \tag{6}$$

  $\delta(t)$ is important as a measure of how close we are to the edge cover time.
- The walk length k is measured by the number of steps taken so far. Equation (5) relates t and k.

A cubic graph G chosen u.a.r. is connected w.h.p. (this follows from Lemma 8 (i) below) and we will implicitly condition on this in what follows. The bulk of the paper will be spent proving the following lemma.

▶ **Lemma 4.** *For any fixed $\varepsilon > 0$ and $(1 - \log^{-2} n)\frac{3n}{2} \leq t \leq \frac{3n}{2}$,*

$$\mathbb{E}\left(C(t)\right) = \left(\frac{3}{2} \pm \varepsilon\right) n \log\left(\frac{3n}{3n - 2t + 1}\right) + o(n \log n) \tag{7}$$

for n large enough. Furthermore, for $n - \frac{n}{\log n} \leq s \leq n$,

$$(1 - \varepsilon) n \log\left(\frac{n}{n - s + 1}\right) \leq \mathbb{E}\left(C_V(s)\right) \leq (1 + \varepsilon) n \log\left(\frac{n}{n - s + 1}\right). \tag{8}$$

Expectations in Lemma 4 are taken over the full probability space. In particular, if $\mathcal{G}$ denotes the set of graphs,

$$\frac{3}{2} n \log\left(\frac{3n}{3n - 2t + 1}\right) \approx \mathbb{E}\left(C(t)\right) = \frac{1}{|\mathcal{G}|} \sum_{G \in \mathcal{G}} \mathbb{E}_G\left(C(t)\right).$$

We can strengthen Lemma 4 to stating that almost every G satisfies $\mathbb{E}_G(C(t)) \approx \mathbb{E}(C(t))$, and similarly for $C_V(s)$ (proof omitted in this extended abstract). Theorem 2 will then follow.

An essential part of the proof of Lemma 4 is a set of recurrences for the random variables $X_i(t)$, where $X_i(t)$ is the number of vertices incident with $i = 0, 1, 2, 3$ untraversed edges at *time* t, $t = 1, 2, ..., 3n/2$ (note that the graph contains exactly $3n/2$ edges). Ignoring in this extended abstract the set $X_2(t)$, which can only contain the tail vertex, the recurrences are

$$\mathbb{E}(X_3(t+1) \mid W(t)) = X_3(t) - \frac{3X_3(t)}{3n - 2t + 1}, \tag{9}$$

$$\mathbb{E}(X_1(t+1) \mid W(t)) = X_1(t) - \frac{2X_1(t)}{3n - 2t + 1} + \frac{3X_3(t)}{3n - 2t + 1}, \tag{10}$$

and we have $X_0(t) = n - X_1(t) - X_2(t) - X_3(t)$. These recurrences suggest that at time $t = (1 - \delta)\frac{3n}{2}$ with $\delta = o(1)$ we have $X_1(t) \approx 3n\delta$ and $X_3(t) \approx n\delta^{3/2}$, and this is proven in the full paper version.

We will argue that for most of the process, it takes approximately $3n/(3n - 2t + 1)$ steps of the walk to increase time by one. As the process finishes at time $3n/2$ we see that the edge cover time should be approximately

$$\sum_{t=1}^{3n/2} \frac{3n}{3n - 2t + 1} \approx \frac{3}{2} n \log n,$$

as claimed in Corollary 3.

Given that $X_3(t) \approx n\delta^{3/2}$, we would expect $X_3(t)$ to be zero when δ is smaller than $n^{-2/3}$ or equivalently, when $3n/2 - t$ is less than $n^{1/3}$. Thus we would expect that vertex cover time to be

$$\sum_{t=1}^{3n/2 - n^{1/3}} \frac{3n}{3n - 2t + 1} \approx n \log n,$$

as claimed in Corollary 3. In this extended abstract we omit further details in calculating the vertex cover time.

We separate the proof of Lemma 4 into phases. Define

$$\delta_0 = \frac{1}{\log \log n}, \ \delta_1 = \log^{-1/2} n, \ \delta_2 = \log^{-2} n, \ \delta_3 = n^{-2/3} \log^4 n \text{ and } \delta_4 = n^{-1} \log^{11} n$$

and set

$$t_i = (1 - \delta_i)\frac{3n}{2} \text{ for } i = 0, 1, 2, 3, 4. \tag{11}$$

The first phase, in which the first t_1 edges are discovered, will not contribute significantly to the cover time.

▶ **Lemma 5.** *Let $\delta_1 = \log^{-1/2} n$ and $t_1 = (1 - \delta_1)\frac{3n}{2}$. Then*

$$\mathbb{E}(C(t_1)) = o(n \log n).$$

Between times t_1 and t_4 we bound the time taken between discovering new edges. The proof, in Section 6, will be split into the ranges $t_1 \le t \le t_3$ and $t_3 \le t \le t_4$.

▶ **Lemma 6.** *Let $\varepsilon > 0$. For $t_1 \le t \le t_4$ and n large enough,*

$$\mathbb{E}\left(C(t+1) - C(t)\right) = (3 \pm \varepsilon)\frac{n}{3n - 2t} + O(\log n).$$

Note that because $\frac{3n}{2} - t_1 = O(\delta_1 n)$, the $O(\log n)$ term only contributes an amount $O(n\delta_1 \log n) = o(n \log n)$ to the the edge cover time.

Finally, the following lemma shows that the final $\log^{11} n$ edges can be found in time $o(n \log n)$.

▶ **Lemma 7.** *For $t > t_4$ and n large enough,*

$$\mathbb{E}\left(C(t) - C(t_4)\right) = o(n \log n).$$

We note now that Lemma 4 follows from Lemmas 5, 6 and 7.

3   Structural properties of random cubic graphs

The random cubic graph is chosen according to the *configuration model*, introduced by Bollobás [5]. Each vertex $v \in [n]$ is associated with a set $\mathcal{P}(v)$ of 3 *configuration points*, and we let $\mathcal{P} = \cup_v \mathcal{P}(v)$. We choose u.a.r. a perfect matching μ of the points in $\mathcal{P}$. Each μ induces a multigraph G on $[n]$ in which u is adjacent to v if and only if $\mu(x) \in \mathcal{P}(v)$ for some $x \in \mathcal{P}(u)$, allowing parallel edges and self-loops. Here we collect some properties of random cubic graphs, chosen according to the configuration model. Any simple cubic graph is equally likely to be chosen under this model.

▶ **Lemma 8.** *Let G denote the random cubic graph on vertex set $[n]$, chosen according to the configuration model. Let ω tend to infinity arbitrarily slowly with n. Its value will always be small enough so that where necessary, it is dominated by other quantities that also go to infinity with n. Then w.h.p.,*
  (i) *In absolute value, the second largest eigenvalue of the transition matrix for a simple random walk on G is at most 0.99.*
  (ii) *G contains at most $\omega 3^\omega$ cycles of length at most ω,*
  (iii) *The probability that G is simple is $\Omega(1)$.*

Friedman [11] showed that for any $\varepsilon > 0$, the second largest eigenvalue of the transition matrix is at most $2\sqrt{2}/3 + \varepsilon$ w.h.p., which gives (i). Property (ii) follows from the Markov inequality, given that the expected number of cycles of length $k \le \omega$ can be bounded by $O(3^k)$. For the proof of (iii) see Frieze and Karoński [12], Theorem 10.3. Note that (iii) implies that any property which holds w.h.p. for a configuration multigraph chosen u.a.r., also holds w.h.p. for a simple cubic graph chosen u.a.r.

Let $G(t)$ denote the random graph formed by the edges visited by $W(t)$. Let $X_i(t)$ denote the set of vertices incident to i red edges in $G(t)$ for $i = 0, 1, 2, 3$. Let $\overline{X}(t) = X_1(t) \cup X_2(t) \cup X_3(t)$. Let $G^*(t)$ denote the graph obtained from $G(t)$ by contracting the set $\overline{X}(t)$ into a single vertex, retaining all edges. Define $\lambda^*(t)$ to be the second largest eigenvalue of the transition matrix for a simple random walk on $G^*(t)$.

We note that if Γ is a graph obtained from G by contracting a set of vertices, retaining all edges, then $\lambda(\Gamma) \le \lambda(G)$, see [2, Corollary 3.27]. This implies that $\lambda^*(t) = \lambda(G^*(t)) \le \lambda(G) \le 0.99$ for all t. Initially, for small t, we find that w.h.p. $G^*(t)$ consists of a single vertex. In this case there is no second eigenvalue and we take $\lambda^*(t) = 0$. This is in line with the fact that a random walk on a one vertex graph is always in the steady state, as the only possible probability measure on a singleton is the trivial measure.

4 Hitting times for simple random walks

We are interested in calculating $\mathbb{E}\left(C(t+1) - C(t)\right)$, i.e. the expected time taken between discovering the tth and the $(t+1)$th edge. Between the two discoveries, the biased random walk can be coupled to a simple random walk on the graph induced by $W(t)$ which ends as soon as it hits a vertex of $\overline{X}$. We will be able to calculate the hitting time as a consequence of $\overline{X}$ having a special structure as in the following definition.

▶ **Definition 9.** Let $G = (V, E)$ be a cubic graph. A set $S \subseteq V$ is a *root set of order ℓ* if (i) $|S| \geq \ell^5$, (ii) the number of edges with both endpoints in S is between $|S|/2$ and $(1/2 + \ell^{-3})|S|$, and (iii) there are at most $|S|/\ell^3$ paths of length at most ℓ between vertices of S that contain no edges between a pair of vertices in S.

Root sets of large order may be thought of as sets that contain an almost-perfect matching, and most of whose vertices are otherwise separated by a large distance. We can calculate the expected hitting time for such sets.

▶ **Lemma 10.** *Let ω tend to infinity arbitrarily slowly with n. Suppose G is a cubic graph on n vertices with positive eigenvalue gap, containing at most $\omega 3^\omega$ cycles of length at most ω. If S is a root set of order ω, then the expected hitting time of S for a simple random walk starting at a uniformly chosen vertex is*

$$\mathbb{E}\left(H(S)\right) \approx \frac{3n}{|S|}.$$

5 The structure of $\overline{X}$

Eventually the biased random walk will spend the majority of its time at vertices in X_0, i.e. vertices with no red incident edges. To bound the cover time, we will bound the time taken to hit $\overline{X} = X_1 \cup X_2 \cup X_3$, which may be thought of as the boundary of X_0.

Let $W_k, k \geq 0$ denote the biased random walk after $2k + 1$ walk steps have been taken. Say that a fixed finite walk W is *feasible* if $\Pr\{W_k = W\} > 0$ for some $k \geq 0$, and fix a feasible walk W. Let t be the time associated with W as indicated in (5). Let Y denote the subset of vertices in $X_1(t)$ that were visited and left exactly once by W. Note that $|Y \triangle X_1| \leq 1$, as the tail v_0 and head v_k of the walk are the only vertices which may be in X_1 after being visited twice and then only when $v_0 = v_k$. Indeed, the first time a vertex v is visited, a feasible walk must enter and exit v via distinct edges. Color all vertices of Y green. We can write $Y = X_1(t) \setminus \{v_0\}$.

Given a feasible walk W, define a *green bridge* to be a part of the walk starting and ending in $V \setminus Y$, with any internal vertices being in Y. Note also that it is not necessary for a green bridge to contain any vertices of Y. Form the *contracted walk* $\langle W \rangle$ by replacing any green bridge by a single green edge between the two endpoints of the bridge, with the walk orientation intact. Let $[W]$ denote the pair of (contracted walk, set), $[W] = (\langle W \rangle, Y)$, noting that $\langle W \rangle$ contains no vertex of Y.

We define an equivalence relation on the set of feasible walks by saying that $W \sim W'$ if and only if $[W] = [W']$. See Figure 1. Thus the only way that W, W' differ is as to where the vertices in Y are placed on the green bridges.

▶ **Lemma 11.** *Let $k > 0$ and suppose W is such that $Pr\{W_k = W\} > 0$. If $[W] = (\langle W \rangle, Y)$ and $\langle W \rangle$ contains ϕ green edges, then*

$$Pr\{W_k = W \mid [W_k] = [W]\} = \frac{1}{|[W]|} = \frac{1}{(\phi + |Y| - 1)_{|Y|}},$$

where $(a)_b = a(a-1)\cdots(a-b+1)$.
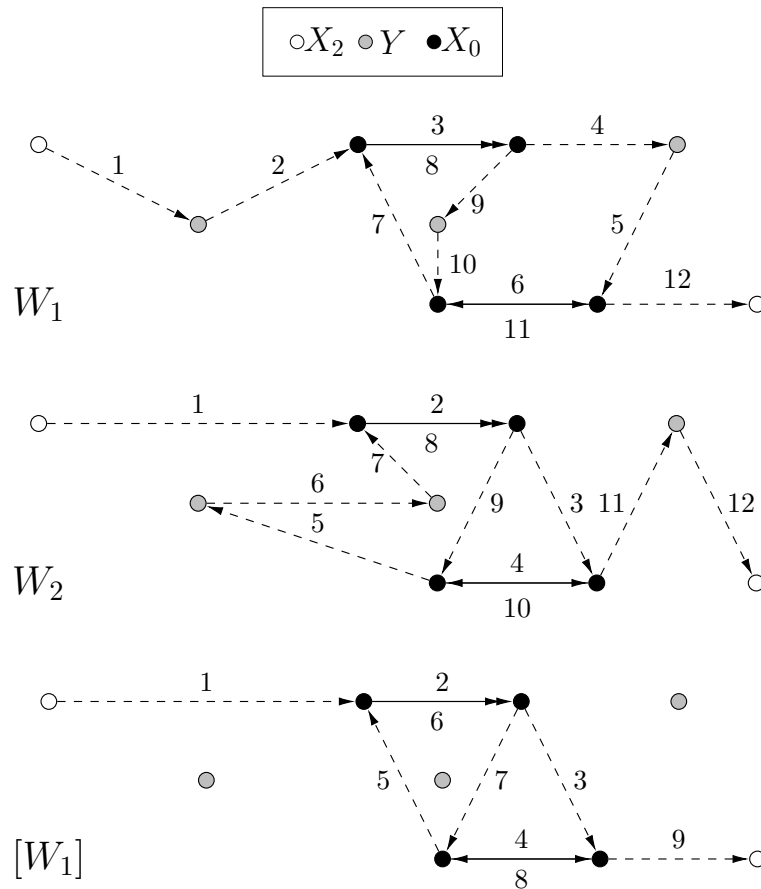
Figure 1 Two equivalent walks, and a visual representation of their equivalence class. Numbers represent order of traversal. Unvisited edges and vertices are not displayed, and edges visited exactly once are dashed. Lemma 11 shows that the walks are equiprobable.

We can now view the biased random walk as a walk on the equivalence class $[W(t)]$. Any time a green edge in $[W(t)]$ is visited, the probability that the edge corresponds to a nontrivial path in a randomly chosen $W(t) \in [W(t)]$ is about $X_1(t)/\Phi(t)$, where $\Phi(t)$ denotes the number of green edges in $W(t)$. This provides a precise recursion for $\mathbb{E}(\Phi(t))$ similar to those for $X_1(t), X_3(t)$, which we use to prove the following. Recall $\delta_0 = 1/\log\log n$. W.h.p.,

$$|X_1(t)| \sim 3n\delta \quad \text{when } \delta \leq \delta_1, \tag{12}$$

$$|X_3(t)| \sim n\delta^{3/2} \quad \text{when } \delta \leq \delta_1, \tag{13}$$

$$\Phi(t) \geq n(\delta_0\delta)^{1/2} \quad \text{when } \delta_3 \leq \delta \leq \delta_1. \tag{14}$$

Suppose $\delta_3 \leq \delta \leq \delta_1$. As $X_1(t) = o(\Phi(t))$, when $W(t) \in [W(t)]$ is chosen uniformly at random, the vertices of $X_1(t)$ are sprinkled into the much larger set of green edges, and are expected to be spread far apart. This will imply that $X_1(t)$ is a root set of order ω, and as $X_1(t)$ makes up almost all of $\overline{X}(t)$ by (12), the latter is also a root set of order ω. When $\delta \leq n^{-2/3}$, the same technique can be applied with a little more work.

6 Calculating the cover time

6.1 Early stages

With $t_1 = (1 - \log^{-1/2} n)\frac{3n}{2}$, we show that $\mathbb{E}(C(t_1)) = o(n \log n)$. Suppose $W(t) = (x_1, x_2, \ldots, x_{2k-1})$ for some t and $k \geq 1$. If $x_{2k-1} \in \mathcal{P}(\overline{X}(t))$ then $x_{2k} = \mu(x_{2k-1})$ is uniformly random inside $\mathcal{P}(\overline{X}(t))$, and since $C(t+1) = C(t) + 1$ in the event of $x_{2k} \in \mathcal{P}(X_2 \cup X_3)$, we have

$$\mathbb{E}(C(t+1) - C(t)) \leq 1 + \mathbb{E}(C(t+1) - C(t) \mid x_{2k} \in \mathcal{P}(X_1)) \Pr\{x_{2k} \in \mathcal{P}(X_1)\}, \qquad (15)$$

We use the following theorem of Ajtai, Komlós and Szemerédi [1] to bound the expected change when $x_{2k} \in \mathcal{P}(X_1)$.

▶ **Theorem 12.** *Let $G = (V, E)$ be a d-regular graph on n vertices, and suppose that each of the eigenvalues of the adjacency matrix with the exception of the first eigenvalue are at most λ_G (in absolute value). Let Z be a set of cn vertices of G. Then for every ℓ, the number of walks of length ℓ in G which avoid Z does not exceed $(1 - c)n((1 - c)d + c\lambda_G)^\ell$.*

The set Z of Theorem 12 is fixed. In our case the exit vertex u of the red walk is chosen randomly from $X_1(t)$. This follows from the way the red walk constructs the graph in the configuration model. The subsequent walk now begins at vertex u and continues until it hits a vertex of $Y_u = X_1(t) \setminus \{u\}$ (or more precisely $Y_u \cup X_2(t)$). Because the exit vertex u is random, the set $B_u = Y_u \cup X_2(t) \cup X_3(t)$ differs for each possible exit vertex $u \in X_1(t)$. To apply Theorem 12, we split $X_1(t)$ into two disjoint sets A, A' of (almost) equal size. For $u \in A$, instead of considering the number of steps needed to hit B_u, we can upper bound this by the number of steps needed to hit $B' = A' \cup X_2 \cup X_3$.

Let $Z(\ell)$ be a simple random walk of length ℓ starting from a uniformly chosen vertex of A. Thus $Z(\ell)$ could be any of $|A|3^\ell$ uniformly chosen random walks. Let $c = |B'|/n$. The probability p_ℓ that a randomly chosen walk of length ℓ starting from A has avoided B' is at most

$$p_\ell \leq \frac{1}{(|X_1(t)|/2)3^\ell}(1 - c)n(3(1 - c) + c\lambda_G)^\ell \leq \frac{2(1 - c)n}{|X_1(t)|}((1 - c) + c\lambda)^\ell,$$

where $\lambda \leq .99$ (see Lemma 8) is the absolute value of the second largest eigenvalue of the transition matrix of Z. Thus

$$\mathbb{E}_A(H(B')) \leq \sum_{\ell \geq 1} p_\ell \leq \frac{2(1 - c)n}{|X_1(t)|} \frac{1}{c(1 - \lambda)}. \qquad (16)$$

As $|B'| = |X_1|/2 + |X_3|$, we have

$$\mathbb{E}(C(t+1) - C(t) \mid x_{2k} \in \mathcal{P}(X_1(t))) = O\left(\frac{(n - |X_3|)n}{|X_1|(|X_1| + |X_3|)}\right). \qquad (17)$$

Using (12), (13), and other bounds for $|X_1(t)|, |X_3(t)|$,

$$\mathbb{E}(C(t_1)) = \sum_{t=1}^{t_1} \mathbb{E}(C(t) - C(t-1)) = o(n \log n).$$

Details are omitted in this extended abstract.

6.2   Later Stages

We will now use Lemmas 10 and 11, together with Definition 9 and equations (12) – (14). For $t = (1 - \delta)\frac{3n}{2}$ with $\delta \leq \delta_1 = \log^{-1/2} n$ we set $\omega = \omega(t) = \log(-\log \delta)$ and define the events (with $\overline{X}(t) = X_1(t) \cup X_2(t) \cup X_3(t)$)

$$\mathcal{A}(t) = \{|X_1(t) - 3n\delta| = O(\omega^{-1}\delta n)\}, \tag{18}$$

$$\mathcal{B}(t) = \{\overline{X}(t) \text{ is a root set of order } \omega\}. \tag{19}$$

and set $\mathcal{E}(t) = \mathcal{A}(t) \cap \mathcal{B}(t)$. As a consequence of Lemma 10, equation (16) and the fact that $\mathbb{E}\left(\overline{X}(t)\right) = 3n - 2t + o(3n - 2t)$, we have

$$\mathbb{E}\left(C(t+1) - C(t)\right) = (3 \pm \varepsilon)\frac{n}{3n - 2t}\Pr\left\{\mathcal{E}(t)\right\} + O\left(\frac{n}{3n - 2t}\right)\Pr\left\{\overline{\mathcal{E}(t)}\right\} + O(\log n). \tag{20}$$

Here the $O(\log n)$ and ε terms account for the number of steps needed to take for the random walk Markov chain to mix to within variation distance ε of the stationary distribution π, at which time we apply Lemma 10. Here we rely on $\lambda^*(t) \leq 0.99$. In the event of $\overline{\mathcal{E}(t)}$ we use the fact that $\overline{X}(t) = \Omega(3n - 2t)$, which follows from (13) and the well-known hitting time bound $\frac{1}{1-\lambda}\frac{n}{\overline{X}(t)}$ (see e.g. Jerrum and Sinclair [14]) to conclude that the hitting time is $O(n/(3n - 2t))$.

The bound (12) for $|X_1(t)|$ implies that $\mathcal{A}(t)$ occurs w.h.p. for any fixed $t \geq t_1$ and we will prove that $\mathcal{B}(t)$ also occurs w.h.p. Lemma 6 will follow. The relatively simple proof of Lemma 7 is sketched at the end.

▶ **Lemma 13.** *Fix t and let $\delta = (3n - 2t)/3n$. If $\delta_1 = \log^{-1/2} n \geq \delta \geq \delta_4 = n^{-1} \log^{11} n$ then,*

$$Pr\{\mathcal{E}(t)\} = 1 - o(1).$$

Proof. Fix some t, δ with $t_1 \leq t \leq t_3$. Expose $[W(t)]$. As in (12) and (14), w.h.p.,

$$\Phi(t) \geq (\delta_0\delta)^{1/2}n, \tag{21}$$

$$|X_1(t)| = 3\delta n + O(\omega^{-1}\delta n). \tag{22}$$

As already remarked, this shows that $\Pr\left\{\mathcal{A}(t)\right\} = 1 - o(1)$. By (13), w.h.p. $X_3(t) \approx n\delta^{3/2} = o(X_1(t))$. We can now show that $\overline{X}(t) = X_1(t) \cup X_2(t) \cup X_3(t)$ is a root set of order ω w.h.p. Here ω is chosen to satisfy (25) below.

Let E_t denote the set of t edges discovered by the walk, and E_t^c the set of (random) edges yet to be discovered. The number of edges inside $\overline{X}(t)$ is given by

$$e(\overline{X}(t)) = |E_t^c| + |E(X_1 \cup X_2) \cap E_t| \tag{23}$$

where $|E_t^c| = (X_1 + 2X_2 + 3X_3)/2$, so

$$|E_t^c| = \frac{|X_1|}{2} + O(\delta_1^{1/2}) = \frac{|X_1|}{2} + O(\omega^{-3})$$

for $\omega^3 = o(\delta_0^{-1/2})$.

We bound the number of paths of length at most ω between vertices of X_1 on edges of E_t, showing that the number is $O(|X_1|/\omega^3)$. Note that such paths include $E(X_1) \cap E_t$, so that the bound implies $|E(X_1) \cap E_t| = O(|X_1|/\omega^3)$.

Let $u, v \in X_1$. Suppose u is placed on some green edge f_1. There are at most 3^ω green edges at distance at most ω from f_1, so as v is placed in a random green edge,

$$\Pr\{d(u,v) \leq \omega\} = O\left(\frac{3^\omega}{\Phi}\right) = O\left(\frac{3^\omega}{n(\delta_0\delta)^{1/2}}\right).$$

So the expected number of pairs $u, v \in X_1$ at distance at most ω is bounded by

$$\sum_{u,v \in X_1} \Pr\{d(u,v) \leq \omega\} = O\left(\frac{|X_1|^2 3^\omega}{n(\delta_0\delta)^{1/2}}\right) = O(n\delta_0^{-1/2}\delta^{3/2}3^\omega) = o(|X_1|/\omega^3), \quad (24)$$

if we choose

$$\omega^3 3^\omega = o\left((\delta_0/\delta)^{1/2}\right). \quad (25)$$

w.h.p. the number of paths is $O(|X_1|/\omega^3)$ by the Markov inequality. This shows that $\overline{X}(t)$ is a root set of order ω w.h.p.

We show in the full paper version that w.h.p., $\mathcal{E}(t_3)$ holds with enough room to spare so that $\mathcal{E}(t)$ must hold for $t_3 \leq t \leq t_4$. ◄

For $t \geq t_4$, we use the bound

$$\mathbb{E}\left(C(t+1) - C(t)\right) \leq \frac{1}{1-\lambda}\frac{n}{|\overline{X}(t)|},$$

see e.g. Jerrum and Sinclair [14], to conclude that $\mathbb{E}\left(C(3n/2) - C(t_4)\right) = o(n\log n)$.

──── **References** ────

1   Miklós Ajtai, János Komlós, and Endre Szemerédi. Deterministic simulation in LOG-SPACE. In Alfred V. Aho, editor, *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 132–140. ACM, 1987. `doi:10.1145/28395.28410`.

2   David Aldous and James Allen Fill. Reversible markov chains and random walks on graphs, 2002. Unfinished monograph, recompiled 2014, available at `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

3   Noga Alon, Itai Benjamini, Eyal Lubetzky, and Sasha Sodin. Non-backtracking random walks mix faster. *Communications in Contemporary Mathematics*, 09(04):585–603, 2007. `doi:10.1142/S0219199707002551`.

4   Petra Berenbrink, Colin Cooper, and Tom Friedetzky. Random walks which prefer unvisited edges: Exploring high girth even degree expanders in linear time. *Random Struct. Algorithms*, 46(1):36–54, 2015. `doi:10.1002/rsa.20504`.

5   Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Comb.*, 1(4):311–316, 1980. `doi:10.1016/S0195-6698(80)80030-8`.

6   Colin Cooper and Alan M. Frieze. The cover time of random regular graphs. *SIAM J. Discrete Math.*, 18(4):728–740, 2005. `doi:10.1137/S0895480103428478`.

7   Colin Cooper and Alan M. Frieze. Vacant sets and vacant nets: Component structures induced by a random walk. *SIAM J. Discrete Math.*, 30(1):166–205, 2016. `doi:10.1137/14097937X`.

8   Colin Cooper, Alan M. Frieze, and Samantha Petti. The cover time of a biased random walk on $G_{n,p}$. In Markus E. Nebel and Stephan G. Wagner, editors, *Proceedings of the Fifteenth Workshop on Analytic Algorithmics and Combinatorics, ANALCO 2018, New Orleans, LA, USA, January 8-9, 2018.*, pages 158–167. SIAM, 2018. `doi:10.1137/1.9781611975062.14`.

9   Uriel Feige. A tight lower bound on the cover time for random walks on graphs. *Random Struct. Algorithms*, 6(4):433–438, 1995. `doi:10.1002/rsa.3240060406`.

10   Uriel Feige. A tight upper bound on the cover time for random walks on graphs. *Random Struct. Algorithms*, 6(1):51–54, 1995. `doi:10.1002/rsa.3240060106`.

11   Joel Friedman. A proof of Alon's second eigenvalue conjecture. In *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 720–724, New York, NY, USA, 2003. ACM. `doi:10.1145/780542.780646`.

12   Alan M. Frieze and Michal Karonski. *Introduction to Random Graphs*. Cambridge University Press, Cambridge, UK, 2015. `doi:10.1017/CBO9781316339831.011`.

13   Tal Orenshtein and Igor Shinkar. Greedy random walk. *Combinatorics, Probability & Computing*, 23(2):269–289, 2014. `doi:10.1017/S0963548313000552`.

14   Alistair Sinclair and Mark Jerrum. Approximate counting, uniform generation and rapidly mixing markov chains. *Inf. Comput.*, 82(1):93–133, 1989. `doi:10.1016/0890-5401(89) 90067-9`.

15   Hui Tian, Hong Shen, and Teruo Matsuzawa. RandomWalk routing for wireless sensor networks. In *Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT 2005), 5-8 December 2005, Dalian, China*, pages 196–200. IEEE Computer Society, 2005. URL: `http://ieeexplore.ieee.org/xpl/ mostRecentIssue.jsp?punumber=10544`, `doi:10.1109/PDCAT.2005.193`.

The Genus of the Erdős-Rényi Random Graph and the Fragile Genus Property

Chris Dowden[1]

Institute of Discrete Mathematics, Graz University of Technology, 8010 Graz, Austria
dowden@math.tugraz.at

Mihyun Kang[2]

Institute of Discrete Mathematics, Graz University of Technology, 8010 Graz, Austria
kang@math.tugraz.at

Michael Krivelevich[3]

School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University,
Tel Aviv 6997801, Israel
krivelev@post.tau.ac.il

── Abstract ──

We investigate the genus $g(n, m)$ of the Erdős-Rényi random graph $G(n, m)$, providing a thorough description of how this relates to the function $m = m(n)$, and finding that there is different behaviour depending on which 'region' m falls into.

Existing results are known for when m is at most $\frac{n}{2} + O(n^{2/3})$ and when m is at least $\omega\left(n^{1+\frac{1}{j}}\right)$ for $j \in \mathbb{N}$, and so we focus on intermediate cases.

In particular, we show that $g(n, m) = (1 + o(1))\frac{m}{2}$ whp (with high probability) when $n \ll m = n^{1+o(1)}$; that $g(n, m) = (1 + o(1))\mu(\lambda)m$ whp for a given function $\mu(\lambda)$ when $m \sim \lambda n$ for $\lambda > \frac{1}{2}$; and that $g(n, m) = (1 + o(1))\frac{8s^3}{3n^2}$ whp when $m = \frac{n}{2} + s$ for $n^{2/3} \ll s \ll n$.

We then also show that the genus of fixed graphs can increase dramatically if a small number of random edges are added. Given any connected graph with bounded maximum degree, we find that the addition of ϵn edges will whp result in a graph with genus $\Omega(n)$, even when ϵ is an arbitrarily small constant! We thus call this the 'fragile genus' property.

1 Introduction

1.1 Background and motivation

The Erdős-Rényi random graph $G(n, m)$ (taken uniformly at random from the set of all labelled graphs with vertex set $[n] = \{1, 2, \ldots, n\}$ and exactly m edges) and the binomial

---

random graph $G_{n,p}$ (the graph on $[n]$ where every edge occurs independently at random with probability p) have been a source of fascination for many decades, producing numerous exciting results (see, for example, [3], [5], and [6]).

In this work, we are interested in the genus of a graph. A graph is said to have genus g if this is the minimum number of handles that must be attached to a sphere in order to be able to embed the graph without any crossing edges. Hence, the simplest case when $g = 0$ corresponds to planar graphs.

The genus is one of the most fundamental properties of a graph, and plays an important role in a number of applications and algorithms (e.g. colouring problems and the manufacture of electrical circuits). It is naturally intriguing to consider the genus of a random graph, and such matters are also related to random graphs on surfaces (see, for example, Question 8.13 of [7] and Section 9 of [4]).

The genus of the binomial random graph $G_{n,p}$ was first studied in [1], and it was shown that this is $(1 + o(1))\frac{pn^2}{12}$ with high probability (whp for short, meaning with probability tending to 1 as $n \to \infty$ – see Definition 7) if $p^2(1 - p^2) \geq \frac{8(\log n)^4}{n}$. A particularly notable consequence of this result (by taking $p = \frac{1}{2}$) is that the classical uniform random graph $G(n)$ (taken uniformly at random from the set of all labelled graphs on $[n]$) must then have genus $(1 + o(1))\frac{n^2}{24}$ whp.

As noted in [1], results for the genus of $G_{n,p}$ can be transferred into analogous results for the genus $g(n, m)$ of $G(n, m)$. Taking into account the work in both [1] and [16] (the latter of which deals with a substantially wider range for p), these show that $g(n, m) = (1 + o(1))\frac{m}{6}$ whp when $m = \Theta(n^2)$ and that $g(n, m) = (1 + o(1))\frac{jm}{2(j+2)}$ whp when $n^{1+\frac{1}{j+1}} \ll m \ll n^{1+\frac{1}{j}}$ for $j \in \mathbb{N}$.

Separately, important work has also been carried out to determine the probability that $G(n, m)$ is planar (i.e. has zero genus) when m is comparatively small. In particular, it is now well-known that $G(n, m)$ is planar whp when $m < \frac{n}{2} - \omega\left(n^{2/3}\right)$ (see [14]) and that $\liminf \mathbb{P}[G(n, m) \text{ is planar}] > 0$ when $m = \frac{n}{2} + O\left(n^{2/3}\right)$ (see [14], and see [15] for exact limiting probabilities).

It is our aim here to now bridge the gap between the $m \gg n^{1+\frac{1}{j+1}}$ and $m = \frac{n}{2} + O\left(n^{2/3}\right)$ results. We provide a thorough description of this intermediate region, finding that there is different behaviour depending on whether (i) $n \ll m = n^{1+o(1)}$, (ii) $m \sim \lambda n$ for $\lambda > \frac{1}{2}$, or (iii) $m = \frac{n}{2} + s$ for $s > 0$ satisfying $n^{2/3} \ll s \ll n$.

We then turn our attention to a related problem, concerning the genus of a graph that is partially random. Here, we take an arbitrary connected graph H with bounded maximum degree, and examine the supergraph G formed by adding some random edges to H (this type of model is sometimes called 'smoothed analysis' or a 'randomly perturbed' graph, see e.g. [2], [8], [9], and [11]).

Rather surprisingly, we find that G will whp have high genus, even if H has low genus and the number of random edges added is relatively small. We thus call this the 'fragile genus' property.

1.2   Main results and techniques

The main contributions of this paper are two-fold. Firstly, we obtain a complete picture of $g(n, m)$ for all values of m by producing precise results for the previously uncharted regions. Secondly, we then initiate the study of how the genus of a fixed graph is affected when random edges are added, discovering the fragile genus property.

We now present our main results in detail. In the first of these, we consider $g(n, m)$ for the region when $n \ll m = n^{1+o(1)}$ (e.g. this would be the case for a function such as $m = n \ln n$). Note that this is not an area that is covered by existing work. However, we obtain the following tight bounds:

▶ **Theorem 1.** *Let $m = m(n)$ satisfy $n \ll m = n^{1+o(1)}$. Then with high probability*

$$(1 - o(1))\frac{m}{2} \leq g(n, m) \leq \frac{m}{2}.$$

Perhaps the most obvious gap in previous knowledge concerns the case when m is linear in n, but above the threshold for planarity. We show that the genus behaves smoothly in this region:

▶ **Theorem 2.** *Let $m(n) \sim \lambda n$ for some fixed $\lambda > \frac{1}{2}$. Then with high probability*

$$g(n, m) = (1 + o(1))\mu(\lambda)m,$$

where

$$\mu(\lambda) = \frac{1}{4\lambda^2} \sum_{r=1}^{\infty} \frac{r^{r-2}}{r!} (2\lambda e^{-2\lambda})^r + \frac{1}{2}\left(1 - \frac{1}{\lambda}\right)$$

is a strictly positive, monotonically increasing, continuous function satisfying $\mu(\lambda) \to 0$ as $\lambda \to \frac{1}{2}$ and $\mu(\lambda) \to \frac{1}{2}$ as $\lambda \to \infty$.

One of the most fascinating areas of study in random graphs has been the behaviour of $G(n, m)$ when m is close to $\frac{n}{2}$, as many important features have been found to emerge around this key point. Here, we examine in detail the slightly supercritical regime when $m = \frac{n}{2} + s$ for $s > 0$ satisfying $n^{2/3} \ll s \ll n$ (i.e. precisely the region between the planarity threshold and the linear case dealt with in Theorem 2), showing exactly how the genus grows:

▶ **Theorem 3.** *Let $m(n) = \frac{n}{2} + s(n)$, where $s = s(n)$ satisfies $s > 0$ for all n and $n^{2/3} \ll s \ll n$. Then with high probability*

$$g(n, m) = (1 + o(1))\frac{8s^3}{3n^2}.$$

All these results are summarised in Table 1, which gives an exciting picture of how the genus $g = g(n, m)$ behaves as m grows. In particular, it is intriguing to see that the ratio of g to m increases from 0 to $\frac{1}{2}$ until m becomes superlinear in n, after which it then decreases from $\frac{1}{2}$ to $\frac{1}{6}$.

Our proofs typically utilise Euler's formula. Given a graph G, this states that the genus $g(G)$ satisfies

$$g(G) = \frac{1}{2}(e(G) - |G| - f(G) + \kappa(G) + 1), \tag{1}$$

where $e(G)$ is the number of edges of G, $|G|$ is the number of vertices of G, $f(G)$ is the number of faces of G when embedded on a surface of minimal genus (i.e. a sphere to which $g(G)$ handles have been attached), and $\kappa(G)$ is the number of components of G.

Consequently, our results often involve first establishing new bounds for the number of faces of $G(n, m)$ (for instance, via the number of short cycles).

For the proof of Theorem 3, in order to attain the required level of accuracy, we note that we actually find it better to deal directly with the 2-core of $G(n, m)$, rather than with the entire graph.

■ **Table 1** A summary of $g := g(n, m)$.

$m = \Theta(n^2)$	$g = (1 + o(1))\frac{m}{6}$ whp	See [16]
$n^{1+\frac{1}{j+1}} \ll m \ll n^{1+\frac{1}{j}}$	$g = (1 + o(1))\frac{jm}{2(j+2)}$ whp	See [16]
$m = \Theta\left(n^{1+\frac{1}{j}}\right)$	$(1 + o(1))\frac{(j-1)m}{2(j+1)}$ $\leq g \leq (1 + o(1))\frac{jm}{2(j+2)}$ whp	See [16]
$n \ll m = n^{1+o(1)}$	$(1 - o(1))\frac{m}{2} \leq g \leq \frac{m}{2}$ whp	Theorem 1
$m \sim \lambda n, \lambda > \frac{1}{2}$	$g = (1 + o(1))\mu(\lambda)m$ whp, where $\mu(\lambda) \to 0$ as $\lambda \to \frac{1}{2}$ and $\mu(\lambda) \to \frac{1}{2}$ as $\lambda \to \infty$	Theorem 2
$m = \frac{n}{2} + s,$ $s > 0$ and $n^{2/3} \ll s \ll n$	$g = (1 + o(1))\frac{8s^3}{3n^2}$ whp	Theorem 3
$m - \frac{n}{2} \sim cn^{2/3}$	$\lim_{n\to\infty} \mathbb{P}(g = 0) = r(c) \in (0, 1),$ where $r(c) \to 1$ as $c \to -\infty$ and $r(c) \to 0$ as $c \to \infty$	See [14]
$m < \frac{n}{2} - \omega\left(n^{2/3}\right)$	$g = 0$ whp	See [14]

We now turn our attention to our final main result, which concerns the fragile genus property. Here, we take an arbitrary connected graph H with bounded maximum degree, and a random graph R on the same vertex set, and we consider the genus $g(G)$ of the graph $G = H \cup R$. We make an interesting discovery, finding that $g(G)$ will whp be rather large, even if H and R are both planar:

▶ **Theorem 4.** *Let Δ be a fixed constant, and let $H = H(n, \Delta)$ be a connected graph with n vertices and maximum degree at most Δ. Let $k = k(n) \to \infty$ as $n \to \infty$, and let $R = R(n, k)$ be a random graph on $V(H)$ consisting of exactly k edges chosen uniformly at random from $\binom{V(H)}{2}$. Let $G = G(n, \Delta, k) = H \cup R$. Then with high probability*

$$g(G) = \Theta\left(\max\{g(H), k\}\right).$$

The proof of Theorem 4 exploits a result from [10] for decomposing the base graph H into various pieces. We construct a particular minor of G where each of these pieces is condensed into a vertex (note that the genus of G is at least the genus of any of its minors), and we find that we can obtain our result by applying Theorem 2 to this minor.

The remainder of the paper is structured as follows: in Section 2, we state the relevant terminology, notation, and key facts; in Section 3, we focus on $g(n, m)$, proving Theorem 1 and providing sketch-proofs of Theorem 2 and Theorem 3; in Section 4, we present the fragile genus property, giving a sketch-proof of Theorem 4; and finally, in Section 5, we discuss the contiguity of $G(n)$ and $G(n, m)$ with random graph models of given genus.
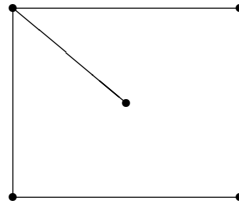
Figure 1 An embedding with faces of length six and four.

2   Preliminaries and notation

In this section, we provide details of the notation and definitions that will be used throughout, and then also present two results concerning the equivalence of $G_{n,p}$ and $G(n,m)$.

Let us first note that we shall always take n and $m = m(n)$ to be integers satisfying $n > 0$ and $m \geq 0$, even if this is not always explicitly stated.

We start with the definitions of the standard random graph models:

▶ **Definition 5.** We shall let $G(n,m)$ denote a graph taken uniformly at random from the set of all labelled graphs on the vertex set $[n] := \{1, 2, \ldots, n\}$ with exactly m edges.

We shall let $G_{n,p}$ denote a graph on $[n]$ where every edge occurs independently at random with probability p, and we shall use $G(n)$ to denote $G_{n,\frac{1}{2}}$ (i.e. a graph taken uniformly at random from the set of all labelled graphs on $[n]$).

Next, we state the notation to be used for various key characteristics:

▶ **Definition 6.** Given a graph G, we shall use $g(G)$ to denote the genus of G, $\kappa(G)$ to denote the number of components of G, and $f(G)$ to denote the number of faces of G when embedded on a surface of genus $g(G)$.

We also define $g(n,m) := g(G(n,m))$, $\kappa(n,m) := \kappa(G(n,m))$, and $f(n,m) := f(G(n,m))$.

Given a particular embedding of a graph, we shall use the *length* of a face to mean the number of edges with a side in the face, counting an edge twice if both sides are in the face (for example, the embedding shown in Figure 1 has one face of length 6 and one face of length 4).

We now also provide details of our order notation:

▶ **Definition 7.** Given non-negative functions $a(n)$ and $b(n)$, we shall use the following notation:

- $a(n) = \Omega(b(n))$ means there exists a constant $c > 0$ such that $a(n) \geq cb(n)$ for all large n;
- $a(n) = O(b(n))$ means there exists a constant C such that $a(n) \leq Cb(n)$ for all large n;
- $a(n) = \Theta(b(n))$ means $a(n) = \Omega(b(n))$ and $a(n) = O(b(n))$;
- $a(n) = \omega(b(n))$ or $a(n) \gg b(n)$ means that, given any constant K, we have $\frac{a(n)}{b(n)} > K$ for all large n (i.e. $\frac{a(n)}{b(n)} \to \infty$ as $n \to \infty$);
- $a(n) = o(b(n))$ or $a(n) \ll b(n)$ means that, given any constant $\epsilon > 0$, we have $\frac{a(n)}{b(n)} < \epsilon$ for all large n (i.e. $\frac{a(n)}{b(n)} \to 0$ as $n \to \infty$);
- $a(n) \sim b(n)$ means $a(n) = (1 + o(1))b(n)$.

We shall say that a random event X_n happens *with high probability (whp)* if $\mathbb{P}(X_n) \to 1$ as $n \to \infty$. Given a non-negative random variable $a(n)$ and a non-negative function $b(n)$, we adapt the above deterministic definitions by replacing 'for all large n' with 'whp'. For

example, $a(n) = o(b(n))$ whp or $a(n) \ll b(n)$ whp means that, given any constant $\epsilon > 0$, we have $\frac{a(n)}{b(n)} < \epsilon$ whp.

We shall always take all asymptotics to be as $n \to \infty$, even if this is not always explicitly stated.

We shall shortly present the two aforementioned equivalence results for $G_{n,p}$ and $G(n, m)$. However, we first require the following definition:

▶ **Definition 8.** We say that a property is *monotone increasing* if whenever an edge is added to a graph with the property, then the resulting graph also has the property. Similarly, we say that a property is *monotone decreasing* if whenever an edge is deleted from a graph with the property, then the resulting graph also has the property. We say that a property is *monotone* if it is either monotone increasing or monotone decreasing.

We may now state the main equivalence result:

▶ **Theorem 9** (see, for example, Proposition 1.15 of [6]). *Given $m = m(n)$, let $p = p(n) = \frac{m}{\binom{n}{2}}$. Then if a monotone property holds whp for $G_{n,p}$, it also holds whp for $G(n, m)$.*

Note that, for any function $x = x(n)$, the property that $g(G) \leq x$ is monotone, as is the property that $g(G) \geq x$.

Unfortunately, the same cannot be said if we replace $g(G)$ with $f(G)$, the number of faces of G when embedded on a surface of minimal genus. For instance, let C_4^+ denote the graph formed be adding one edge to C_4, let K_5^- denote the graph formed be removing one edge from K_5, and note that we have $f(C_4) = 2$, $f\left(C_4^+\right) = 3$, $f\left(K_5^-\right) = 6$, and $f(K_5) = 5$ (observe that the first three graphs are planar, while K_5 has genus one). Hence, adding an edge can actually increase or decrease (or have no impact on) $f(G)$.

However, the function $g(G) - e(G)$ is certainly monotone decreasing (one way to see this is to note that Euler's formula gives $f(G) - e(G) = \kappa(G) + 1 - |G| - 2g(G)$, and $\kappa(G)$ and $g(G)$ are clearly monotone decreasing and monotone increasing, respectively). Using this, we may in fact still apply Theorem 9 to derive a useful equivalence result for the number of faces:

▶ **Corollary 10.** *Let $m = m(n) \to \infty$ as $n \to \infty$, let $p = p(n) = \frac{m}{\binom{n}{2}}$, and suppose $x = x(n)$ is a function such that $f\left(G_{n,p}\right) \leq x$ whp. Then*

$$f(n, m) \leq x + o(m) \ \text{whp}.$$

Proof. We are required to show that, given any constant $\epsilon > 0$, we have $f(n, m) < x + \epsilon m$ whp.

Note that $e\left(G_{n,p}\right)$ has variance $\binom{n}{2}p(1 - p) \leq m$, and hence has standard deviation at most $m^{1/2}$, which is $o(m)$ since $m \to \infty$. Thus, since $e\left(G_{n,p}\right)$ has expectation exactly m, it follows that, given any constant $\epsilon > 0$, we have $e\left(G_{n,p}\right) > (1 - \epsilon)m$ whp. Therefore, since $f\left(G_{n,p}\right) \leq x$ whp, we then have $f\left(G_{n,p}\right) - e\left(G_{n,p}\right) < x - (1 - \epsilon)m$ whp.

Now recall our observation that $f(G) - e(G)$ is a monotone decreasing function, from which it follows that the property that a graph satisfies $f(G) - e(G) < x - (1 - \epsilon)m$ is monotone increasing. Hence, we may apply Theorem 9, thus obtaining $f(n, m) - m < x - (1 - \epsilon)m$ whp, i.e. $f(n, m) < x + \epsilon m$ whp, as desired.   ◀

3   The genus of G(n,m)

In this section, we provide an illustration of some of the techniques employed in the exploration of $g(n, m)$ by proving Theorem 1, which will involve first obtaining bounds on $f(n, m)$ in Lemma 11 and Corollary 12. We then also give sketch-proofs of Theorem 2 and Theorem 3.

▶ **Lemma 11.** *Let* $m = m(n)$ *satisfy both* $m \to \infty$ *as* $n \to \infty$ *and* $m \ll n^{1+\frac{1}{j}}$ *for some fixed* $j \in \mathbb{N}$. *Then*

$$f(n, m) \leq (1 + o(1)) \frac{2}{j + 2} m \ whp.$$

Proof. We will use the $G_{n,p}$ model with $p = \frac{m}{\binom{n}{2}}$, and show that the number of faces is at most $(1 + o(1)) \frac{2}{j+2} \binom{n}{2} p$ whp (we will then be done, by an application of Corollary 10). Thus, we are required to show that, given any constant $\epsilon > 0$, the number of faces is at most $(1 + \epsilon) \frac{2}{j+2} \binom{n}{2} p$ whp.

We will follow a similar argument to that used in the proof of Theorem 1.2 of [16], which involves showing that whp $G_{n,p}$ will have few short cycles, and hence few small faces, and hence few faces in total.

Note that the expected number of cycles in $G = G_{n,p}$ of length at most $j + 1$ is

$$
\begin{aligned}
\sum_{i=1}^{j+1} \binom{n}{i} \frac{i!}{2i} p^i \ &\leq \ \sum_{i=1}^{j+1} \frac{n^i p^i}{2i} \ \leq \ \sum_{i=1}^{j+1} (np)^i \\
&\leq \ (j + 1) \max \left\{ np, (np)^{j+1} \right\} \ \text{(since either } np \leq 1 \text{ or } np \geq 1\text{)} \\
&= \ O\left(\max \left\{ np, (np)^{j+1} \right\} \right) \\
&= \ O\left((np) \max \left\{ 1, (np)^j \right\} \right) \\
&= \ o(n^2 p) \quad \left(\text{since } 1 \ll n \text{ and } np = \frac{nm}{\binom{n}{2}} \ll \frac{n^{2+\frac{1}{j}}}{n^2} = n^{\frac{1}{j}} \right).
\end{aligned}
$$

Thus, by Markov's inequality, we can say that whp G has no more than $\frac{1}{2(j+2)} \epsilon \binom{n}{2} p$ cycles of length at most $j + 1$.

Let us now consider an embedding of G. Note that the statement of this lemma is certainly true if G is acyclic (since then there is only one face), so we may assume that G is not acyclic, in which case every face of the embedding must contain a cycle.

Let f' denote the number of faces in this embedding with length at most $j + 1$. Then every such face must contain a cycle of length at most $j + 1$, and every such cycle can only be included in at most two faces. Hence, whp we have

$$f' \leq \frac{1}{j + 2} \epsilon \binom{n}{2} p. \tag{2}$$

Now let f denote the total number of faces in this embedding, and observe that

$$2e(G) \ \geq \ 3f' + (j + 2)(f - f') \ = \ (j + 2)f - (j - 1)f'.$$

Thus, we have

$$
\begin{aligned}
f & \le & \frac{2}{j+2}e(G) + \frac{j-1}{j+2}f' \le \frac{2}{j+2}e(G) + f' \\
& \le & \frac{2}{j+2}e(G) + \frac{1}{j+2}\epsilon\binom{n}{2}p \quad \text{whp by (2)} \\
& \le & \frac{2}{j+2}\left(1 + \frac{\epsilon}{2}\right)\binom{n}{2}p + \frac{2}{j+2}\frac{\epsilon}{2}\binom{n}{2}p \quad \text{whp} \\
& = & (1+\epsilon)\frac{2}{j+2}\binom{n}{2}p,
\end{aligned}
$$

and so we are done.   ◀

We now obtain our aforementioned corollary, which gives a useful bound on the number of faces $f(n, m)$ when $m \le n^{1+o(1)}$:

▶ **Corollary 12.** *Let $m = m(n)$ satisfy both $m \to \infty$ as $n \to \infty$ and $m \le n^{1+o(1)}$. Then*

$$f(n, m) = o(m) \text{ whp.}$$

Proof. We are required to show that, given any constant $\epsilon > 0$, we have $f(n, m) < \epsilon m$ whp.

We may simply choose a value $j \in \mathbb{N}$ such that $j > \frac{2(1+\epsilon)}{\epsilon} - 2$, in which case $\frac{2}{j+2} < \frac{\epsilon}{1+\epsilon}$. Then, by Lemma 11, we have

$$
\begin{aligned}
f(n, m) & < & (1+\epsilon)\frac{2}{j+2}m \text{ whp} \\
& < & (1+\epsilon)\frac{\epsilon}{1+\epsilon}m \\
& = & \epsilon m,
\end{aligned}
$$

and so we are done.   ◀

We may now easily derive Theorem 1:

Proof of Theorem 1. The upper bound holds for all m – we simply use Euler's formula

$$g(n, m) = \frac{1}{2}(m - n - f(n, m) + \kappa(n, m) + 1)$$

from (1), and observe that $n \ge \kappa(n, m)$ and $f(n, m) \ge 1$.

The lower bound also follows from Euler's formula, using $n = o(m)$ and $f(n, m) = o(m)$ whp by Corollary 12.   ◀

Theorem 2 is obtained similarly:

Sketch of Proof of Theorem 2. The proof again utilises Euler's formula and Corollary 12. For Theorem 1, the role of the number of components $\kappa(n, m)$ was insignificant, since we had $m \gg n \ge \kappa(n, m)$. However, since we now have $m = O(n)$, this time we find that we do require accurate information on $\kappa(n, m)$, and we extract this from Theorem 6.12 of [3].   ◀

The proof of Theorem 3 is more intricate:
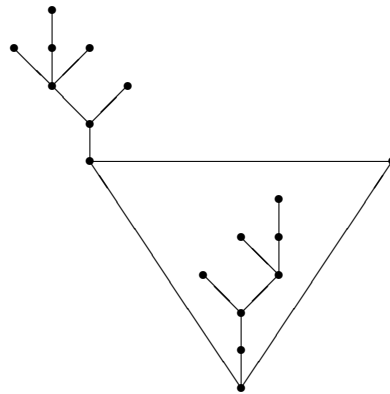
Figure 2 An embedding with two large faces.

Sketch of Proof of Theorem 3. Recall that the proof of Lemma 11 involved bounding the number of small faces via calculations on the number of short cycles. Clearly, this may greatly over-estimate the number of small faces if there are actually many large faces that consist of a short cycle with large trees rooted on the cycle (see Figure 2).

Consequently, in order to achieve the desired level of accuracy for Theorem 3, much of the proof will this time involve working directly with the 2-core of the giant component instead of with the entire graph (note that whp this determines the overall genus, and also alleviates any need to compute the number of components).

Thus, we shall proceed towards an application of Euler's formula for the 2-core of the giant component, which will involve us first needing to establish bounds for the number of faces in the 2-core of the giant.

Our strategy here is to begin by examining the number of 'short' cycles in $G(n, m)$ that also satisfy some additional properties; then to use this to bound the number of such cycles in the giant component; then (by careful consideration of the additional properties) to bound the total number of short cycles in the giant component (and hence in the 2-core of the giant); then to use this to bound the number of short faces in the 2-core of the giant; and then, finally, to separately bound the number of large faces in the 2-core of the giant (via the number of edges).

We start by utilising work from [13] on cycles in $G(n, m)$. For our region when $m = \frac{n}{2} + s$ for $s > 0$ satisfying $n^{2/3} \ll s \ll n$, results are given here concerning the number of cycles in $G(n, m)$ that both (a) have length at most $\frac{in}{s}$, for fixed $i \in \mathbb{N}$, and (b) satisfy certain technical properties involving the neighbouring vertices and the trees rooted on the cycles.

In particular, it is shown that the number of these cycles tends in probability to a random variable that has a Poisson distribution with mean $\lambda(i)$, for a given monotonically increasing function λ, and hence that the number of such cycles is concentrated around $\lambda(i)$.

We extend this latter result to cover the case when $i(n)$ is a function of n, as long as $i(n)$ grows sufficiently slowly.

Next, we construct a specific function $i(n)$ with $i(n) \to \infty$ for which $\lambda(i(n))$ also grows very slowly. For this function, we manage to show that the number of cycles in $G(n, m)$ of length at most $\frac{i(n)n}{s}$ with the aforementioned technical properties is still only $o\left(\frac{s^3}{n^2}\right)$ whp. The key factor here is that $s^3 \gg n^2$.

We then move to the giant component of $G(n, m)$. By utilising further useful results from [13], we are able to show that whp all cycles of length at most $\frac{i(n)n}{s}$ in the giant component will actually satisfy all of the various technical properties.

Thus, we find that the total number of cycles of length at most $\frac{i(n)n}{s}$ in the giant component of $G(n, m)$ (or, equivalently, in the 2-core of the giant component) must be $o\left(\frac{s^3}{n^2}\right)$ whp. Hence, the number of faces of length at most $\frac{i(n)n}{s}$ in the 2-core of the giant must be $o\left(\frac{s^3}{n^2}\right)$ whp too.

We then consider the number of faces of length at *least* $\frac{i(n)n}{s}$ in the 2-core of the giant. By results from [12] and [13], the number of edges in the 2-core of the giant is known to be $\Theta\left(\frac{s^2}{n}\right)$ whp, and so the number of such faces can only be $O\left(\frac{s^3}{i(n)n^2}\right)$ whp. Crucially, our earlier work to ensure that $i(n) \to \infty$ then implies that this is $o\left(\frac{s^3}{n^2}\right)$.

Hence, putting everything together, we find that the *total* number of faces in the 2-core of the giant is also $o\left(\frac{s^3}{n^2}\right)$ whp.

We then finish with an appropriate application of Euler's formula, using existing results from [12] and [13] on the number of vertices and edges in the 2-core of the giant. Note that our bound for the number of faces is sufficiently precise to achieve the desired level of accuracy. ◀

4 The fragile genus property

In this section, we provide a sketch-proof of Theorem 4, which shows that the genus of any given connected graph with bounded degree can increase dramatically if a small number of random edges are added.

Sketch of Proof of Theorem 4. Note that adding an edge can only increase the genus by at most one, so we certainly have $g(G) \leq g(H) + k \leq 2\max\{g(H), k\}$. Also, we clearly have $g(G) \geq g(H)$. Hence, it just remains to show that $g(G) \geq \Omega(k)$ whp.

The result for $\liminf_{n\to\infty} \frac{k}{n} > \frac{1}{2}$ can be obtained simply by applying our results on the genus of $G(n, m)$ to R (with $m = k$), and so we may assume that $k \leq n$, say.

Our proof involves contracting carefully chosen identically-sized pieces of the graph (we use 'piece' to mean a connected subgraph) into 'super-vertices' – note that this cannot increase the genus. We then show that the uniform random graph induced by these super-vertices and the random edges will whp be sufficiently dense for us to be able to apply Theorem 2.

We start by splitting the base graph H into $t = \Theta(k)$ connected pieces $V_1, V_2, \ldots, V_t$ (hence our earlier assumption that $k = O(n)$), plus a few ($o(n)$) vertices that do not belong to any of these pieces.

Using a decomposition result from [10], we may select the pieces in such a way that we have

$$\frac{5n\Delta^2}{k} \leq |V_i| \leq \frac{10n\Delta^3}{k} \tag{3}$$

for all i, and so

$$(1 - o(1))\frac{k}{10\Delta^3} \leq t \leq \frac{k}{5\Delta^2}. \tag{4}$$

Let us note that the value of t has been carefully arranged here. It will be crucial for our later calculations that we have $t = \Omega(k)$, but also that $\frac{t}{k}$ is not too large.

As mentioned, we shall wish to condense pieces of our graph into super-vertices. However, rather than contracting the entire pieces $V_1, V_2, \ldots, V_t$, for each i we instead select a connected subpiece $U_i \subset V_i$ with size $|U_i| = s := \min_i |V_i|$. This will be important to ensure uniformity, so that we can later apply Theorem 2.
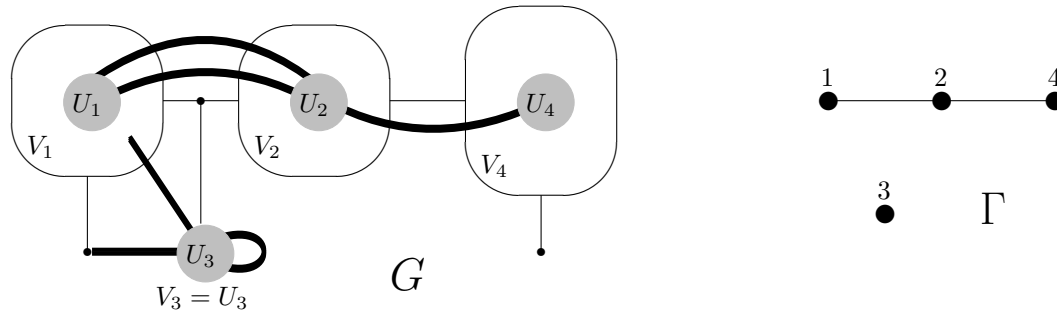
Figure 3 A graph G and the corresponding graph Γ.

We then contract each of these subpieces $U_1, U_2, \ldots, U_t$, and ignore any vertices not in these subpieces. Formally, this means that we define an auxiliary random graph Γ with vertex set $[t]$, where two vertices $i, j \in [t]$ are connected by an edge if and only if there is an edge of R going between U_i and U_j (see Figure 3, where thick lines denote the edges of R – note in particular that this example has no edge in Γ between vertex 1 and vertex 3, as there is no edge in G between U_1 and U_3).

Observe that Γ is a minor of G, and hence that $g(G) \geq g(\Gamma)$. Thus, since we deliberately chose t to be large enough that $t = \Omega(k)$ (recall (4)), it will suffice to show that $g(\Gamma) = \Omega(t)$ whp.

Note that the number of edges in Γ is not equal to k, since we only include edges between our chosen subpieces, and we only include at most one edge for each such pair U_i, U_j.

In order to obtain a bound for $e(\Gamma)$, we consider the edges of R one-by-one (in a random order). Note that $e(\Gamma)$ is then precisely equal to the number of edges of R which satisfy the two properties that

- (a) the edge lies between a vertex of U_i and a vertex of U_j for $i \neq j$;
- and (b) no previous edges of R lie between these same two sets U_i and U_j.

Observe that the probability that an edge of R satisfies both (a) and (b) is always at least

$$\frac{\left(\binom{t}{2} - (k-1)\right) s^2}{\binom{n}{2}},$$

since there are at least $\binom{t}{2} - (k-1)$ ways to choose a pair U_i, U_j which do not already have an edge of R between them, and then s ways to choose a vertex from U_i, and s ways to choose a vertex from U_j.

By (3) and (4), this probability is at least $(1 - o(1))\frac{1}{4\Delta^2}$, and so we can certainly say that whp at least $\frac{k}{5\Delta^2}$ of the k edges of R will satisfy conditions (a) and (b). Thus, $e(\Gamma) \geq \frac{k}{5\Delta^2}$ whp.

Crucially, the fact that we chose t so that $t \leq \frac{k}{5\Delta^2}$ (recall (4)) consequently means that we have $e(\Gamma) \geq t$ whp.

We then let Γ^* be the random graph formed by considering just t (randomly chosen) edges of Γ. Since each set U_i had exactly the same number of vertices, this graph Γ^* is in fact a *uniform* random graph with t vertices and t edges. Thus, by Theorem 2, we have $g(\Gamma^*) = \Theta(t)$ whp, and so $g(\Gamma) = \Omega(t)$ whp, as required.   ◀

Note that Theorem 4 implies the remarkable fact that whp $G = H \cup R$ will have $\Omega(n)$ genus even if H is a planar graph and $k = \epsilon n$ for some very small (but positive) ϵ! We thus call this the 'fragile genus' property.

Let us conclude this section by also remarking that for $\limsup_{n\to\infty} \frac{k}{n} < \frac{1}{2}$, the restriction on the maximum degree in Theorem 4 is essential, since otherwise we could take H to be a star (note that whp the random graph R would consist only of trees and unicyclic components, and would consequently be outerplanar, and so the overall graph G would then have genus zero).

5   Discussion

As mentioned in the introduction, one of our motivations for studying the genus of $G(n,m)$ comes from recent work on random graphs on surfaces. In particular, one may define $S_g(n)$ to be a graph taken uniformly at random from the set of all labelled graphs on $[n]$ with genus at most g, and $S_g(n,m)$ to be a graph taken uniformly at random from the set of all labelled graphs on $[n]$ with exactly m edges and with genus at most g. It is then natural to ask when these graphs will be contiguous with $G(n)$ and $G(n,m)$.

It immediately follows from the work in [1] that $G(n)$ and $S_g(n)$ are certainly contiguous for any $g(n)$ satisfying $g(n) \geq (1+\epsilon)\frac{n^2}{24}$ for any $\epsilon > 0$ (and also for some $g(n)$ satisfying $g(n) = (1+o(1))\frac{n^2}{24}$), since $G(n)$ will have genus at most $(1+o(1))\frac{n^2}{24}$ whp.

Conversely, $G(n)$ and $S_g(n)$ are certainly not contiguous for any $g(n)$ satisfying $g(n) \leq (1-\epsilon)\frac{n^2}{24}$ for any $\epsilon > 0$ (and also not for some $g(n)$ satisfying $g(n) = (1+o(1))\frac{n^2}{24}$), since there is then a discrepancy with respect to the property of having genus greater than g (note $\mathbb{P}[S_g(n)$ has genus $> g] = 0$, by definition, but $\mathbb{P}[G(n)$ has genus $> g] \to 1$ as $n \to \infty$).

By the same arguments, similarly precise results for the contiguity of $G(n,m)$ and $S_g(n,m)$ for the various different regions of m can now also be obtained.

References

1   A. Archdeacon and D. Grable. The genus of a random graph. *Discrete Math.*, 142:21–37, 1995.

2   T. Bohman, A. Frieze, and R. Martin. How many random edges make a dense graph hamiltonian? *Random Struct. Algor.*, 22:33–42, 2003.

3   B. Bollobás. *Random Graphs*. Cambridge University Press, Cambridge, 2001.

4   C. Dowden, M. Kang, and P. Sprüssel. The evolution of random graphs on surfaces. *SIAM J. Discrete Math.*, 32:695–727, 2018.

5   A. Frieze and M. Karoński. *Introduction to Random Graphs*. Cambridge University Press, Cambridge, 2015.

6   S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. Wiley, New York, 2000.

7   M. Kang, M. Moßhammer, and P. Sprüssel. Phase transitions in graphs on orientable surfaces. submitted, arXiv:1708.07671.

8   M. Krivelevich, M. Kwan, and B. Sudakov. Cycles and matchings in randomly perturbed digraphs and hypergraphs. *Combin. Probab. Comput.*, 25:909–927, 2016.

9   M. Krivelevich, M. Kwan, and B. Sudakov. Bounded-degree spanning trees in randomly perturbed graphs. *SIAM J. Discrete Math.*, 31:155–171, 2017.

10   M. Krivelevich and A. Nachmias. Colouring complete bipartite graphs from random lists. *Random Struct. Algor.*, 29:436–449, 2006.

11   M. Krivelevich, D. Reichman, and W. Samotij. Smoothed analysis on connected graphs. *SIAM J. Discrete Math.*, 29:1654–1669, 2015.

12   T. Łuczak. Component behaviour near the critical point of the random graph process. *Random Struct. Algor.*, 1:287–310, 1990.

13   T. Łuczak. Cycles in a random graph near the critical point. *Random Struct. Algor.*, 2:421–439, 1991.

14   T. Łuczak, B. Pittel, and J. C. Wierman. The structure of a random graph near the point
          of the phase transition. *Trans. Amer. Math. Soc.*, 341:721–748, 1994.

15   M. Noy, V. Ravelomanana, and J. Rué. The probability of planarity of a random graph
          near the critical point. *Proc. Amer. Math. Soc.*, 143:925–936, 2015.

16   V. Rödl and R. Thomas. On the genus of a random graph. *Random Struct. Algor.*, 6:1–12,
          1995.

Maximal Independent Sets and Maximal Matchings in Series-Parallel and Related Graph Classes

Michael Drmota[1]

Institute of Discrete Mathematics and Geometry, Technisches Universität Wien, Austria.
michael.drmota@tuwien.ac.at

Lander Ramos

Departament de Matemàtiques, Universitat Politècnica de Catalunya, Barcelona, Spain.
lander.ramos@upc.edu

Clément Requilé[2]

Institute for Algebra, Johannes Kepler Universität Linz, Austria.
clement.requile@jku.at

Juanjo Rué[3]

Departament de Matemàtiques, Universitat Politècnica de Catalunya and Barcelona Graduate
School of Mathematics, Spain.
juan.jose.rue@upc.edu

── Abstract ──

We provide combinatorial decompositions as well as asymptotic tight estimates for two maximal parameters: the number and average size of maximal independent sets and maximal matchings in series-parallel graphs (and related graph classes) with n vertices. In particular, our results extend previous results of Meir and Moon for trees [Meir, Moon: On maximal independent sets of nodes in trees, Journal of Graph Theory 1988]. We also show that these two parameters converge to a central limit law.

29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018).
Editors: James Allen Fill and Mark Daniel Ward; Article No. 18; pp. 18:1–18:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1   Introduction

In this extended abstract we consider labelled, loopless and simple graphs only. For a graph $G = (V(G), E(G))$, a subset J of $V(G)$ is said to be *independent* if, for any pair of vertices x and y contained in J, the edge $\{x, y\}$ does not belong to $E(G)$. An independent set J of a graph G is said to be *maximal* if any other vertex of G that is not contained in J is adjacent to at least one vertex of J. A subset N of the edge set $E(G)$ is called a *matching* if every vertex x of G is incident to at most one edge of N. A matching N is called *maximal* if it cannot be extended to a bigger matching by adding an edge from $E(G) \setminus N$.

The purpose of this paper is to enumerate maximal independent sets and maximal matchings (by means of symbolic methods) and to study their size distribution (using complex analytic tools) in certain classes of graphs including trees, cactus graphs, outerplanar graphs and series-parallel graphs. For simplicity we will only consider vertex labelled graphs, thus making the combinatorial analysis as well as the analytic one considerably simpler. However, in principle it is also possible to consider unlabelled graphs. We use the concept of generating function in order to follow the classical connectivity-decomposition scheme, first starting with the rooted *blocks*, i.e. maximal 2-connected components, then going to the level of rooted connected graphs and finally to general (not necessarily connected and unrooted) graphs.

Let $\mathcal{G}$ denote a proper class of vertex labelled graphs, which means that the vertices of a graph with n vertices are labelled with the labels $\{1, 2, \ldots, n\}$. We denote by $\mathcal{G}_n$ the set of graphs in $\mathcal{G}$ with n vertices. For a graph $G \in \mathcal{G}$ we denote by $I(G)$ the set of maximal independent sets of G and by

$$\mathcal{I}_n = \bigcup_{G \in \mathcal{G}_n} I(G) \times \{G\}$$

the system of all maximal independent sets of graphs of size n. More precisely, every maximal independent set J is *indexed* by the corresponding graph, this is formally done by taking pairs (J, G). Similarly, we denote by $M(G)$ the set of maximal matchings of G and by

$$\mathcal{M}_n = \bigcup_{G \in \mathcal{G}_n} M(G) \times \{G\}$$

the system of all maximal matchings of graphs of size n.

In this extended abstract, we present precise enumerative results on $\mathcal{I}_n$ and $\mathcal{M}_n$. In particular, we will apply our method to two important graph families: Cayley trees and series-parallel graphs. In principle our results can be extended to other graph classes that have a so-called *subcritical analytic structure*, we will make this more precise in Subsection 2.3 (for instance, cactus graphs and outerplanar graphs also satisfy this analytic scheme). For the mentioned graph classes we have the following universal structure in the asymptotic enumeration formula for the number of graphs on n vertices, for n large enough:

$$g_n = |\mathcal{G}_n| \sim c \, n^{-5/2} \rho^{-n} n!,$$

where $c > 0$ and ρ is the radius of convergence of the (exponential) generating function $G(x) = \sum_{n \geq 0} g_n \frac{x^n}{n!}$ associated to the graph class under study. The first result is an asymptotic estimate for both $|\mathcal{I}_n|$ and $|\mathcal{M}_n|$:

▶ **Theorem 1.** *Let $\mathcal{G}$ either be the class of vertex labelled trees, cactus graphs, outerplanar graphs or series-parallel graphs, and let ρ be the radius of convergence of the generating function $G(x)$ associated to $\mathcal{G}$. Then we have*

$$|\mathcal{I}_n| \sim A_1\, n^{-5/2}\rho_1^{-n}n! \quad and \quad |\mathcal{M}_n| \sim A_2\, n^{-5/2}\rho_2^{-n}n!,$$

where A_1, A_2, ρ_1, ρ_2 are positve constants with $0 < \rho_1 < \rho$ and $0 < \rho_2 < \rho$.

As a direct corollary we obtain:

▶ **Corollary 2.** *Let $\mathcal{G}$ be as in Theorem 1 and let AI_n be the average number of maximal independent sets in a graph of size n in $\mathcal{G}$ and AM_n be the average number of matchings in a graph of size n in $\mathcal{G}$. Then it holds that*

$$AI_n = \frac{|\mathcal{I}_n|}{g_n} \sim C \cdot \alpha^n \quad and \quad AM_n = \frac{|\mathcal{M}_n|}{g_n} \sim D \cdot \beta^n,$$

where C, D, α, β are positive constants and α and β are larger than 1.

The second main result concerns the distribution of the respective size of maximal independent sets and matchings. The following theorem shows that the limiting distribution follows a central limit theorem with linear expectation and variance:

▶ **Theorem 3.** *Let $\mathcal{G}$ either be the class of vertex labelled trees, cactus graphs, outerplanar graphs or series-parallel graphs. Furthermore, let SI_n denote the size of a uniformly randomly chosen maximal independent set in $\mathcal{I}_n$ and SM_n the size of a uniformly randomly chosen matching in $\mathcal{M}_n$. Then,*

$$\begin{aligned}
\mathbb{E}[SI_n] = \mu n + O(1), \quad & \mathbb{V}\mathrm{ar}[SI_n] = \sigma_1^2 n + O(1), \\
\mathbb{E}[SM_n] = \lambda n + O(1), \quad & \mathbb{V}\mathrm{ar}[SM_n] = \sigma_2^2 n + O(1),
\end{aligned}$$

for some constants $\mu, \lambda > 0$ and $\sigma_1^2, \sigma_2^2 > 0$. Moreover, SI_n and MI_n satisfy a central limit theorem:

$$\frac{SI_n - \mathbb{E}[SI_n]}{\sqrt{\mathbb{V}\mathrm{ar}[SI_n]}} \xrightarrow{d} \mathrm{N}(0,1) \quad and \quad \frac{SM_n - \mathbb{E}[SM_n]}{\sqrt{\mathbb{V}\mathrm{ar}[SM_n]}} \xrightarrow{d} \mathrm{N}(0,1).$$

Apart for constants C and D in Corollary 2, all the other appearing constants can be computed explicitly to any degree of precision. The following table lists some of them:

Family	α	μ	β	λ
Trees	1.273864	0.463922	1.313080	0.285910
Cactus graphs	1.282413	0.429472	1.371652	0.268268
Series-parallel graphs	1.430394	0.269206	1.470167	0.254122

Let us mention that in [13], Meir and Moon obtained the estimate of Theorem 1 and the expectation in Theorem 3 for maximal independent sets in Cayley trees, plane trees and binary trees. Our contribution generalises their work, providing a precise limiting distribution for the size of maximal independent sets in Cayley trees.

Finally, let us briefly discuss the extremal versions of those problems. In the literature, one can find two such directions. One of them, started by Wilf [17] who was motivated by the design of an algorithm to compute the chromatic number, consists in characterising the extremal instances of a given family of graphs containing the maximum number of maximal

independent sets (see [9], [15] and [18]), as well as maximum independent sets (see [19] and [12]). Furthermore, the maximum number of both maximal matchings [10] and maximum matchings [11] have been treated. The other direction consists in bounding the size of a maximum matching in a graph [3]. However, the problems discussed in this extended abstract seem to be of a different nature.

It is worth noticing that in [3], the authors also give tight bounds on the size of a maximal matching in 3-connected planar graphs and in graphs with bounded maximum degree.

Structure of the extended abstract

Section 2 introduces the necessary background, namely the language of generating functions and how they apply to graph decompositions in terms of their connectivity, as well as the analytic concepts needed in the context of subcritical graph classes. Later, in Section 3 we obtain a system of functional equations encoding maximal independent sets in subcritical graph classes. We then analyse it using complex analytic tools in Subsection 3.2. And in Section 4 we apply our results to the families of Cayley trees and series-parallel graphs. The reader will finally find the analoguous scheme for maximal matchings in an appendix at the end of the extended abstract.

2 Preliminaries

2.1 Generating functions

We follow the notation from [6]. A *labelled combinatorial class* is a set $\mathcal{A}$ together with a size measure, such that if $n \geq 0$, then the set of elements of size n, denoted by $\mathcal{A}_n$, is finite. Each element a of $\mathcal{A}_n$ is built from n atoms (typically, vertices in graph classes) assembled in a certain way, the atoms bearing distinct labels in the set $\{1, \ldots, n\}$. We always assume that a combinatorial class is stable under graph isomorphism, namely, $a \in \mathcal{A}$ if and only if all graphs a' isomorphic to a are also elements of $\mathcal{A}$.

In enumerative problems, it is often useful to use the exponential generating function (shortly the EGF) associated to the labelled class $\mathcal{A}$:

$$A(x) := \sum_{n \geq 0} \frac{|\mathcal{A}_n|}{n!} x^n, \qquad [x^n]A(x) = \frac{|\mathcal{A}_n|}{n!}.$$

In our setting, we use the (exponential) variable x to encode vertices.

We can root the elements of a class $\mathcal{A}$ by distinguishing one of the items and discounting it, which means that we reduce the size function by 1. Since we assume that our combinatorial class is stable under graph isomorphism, this procedure can be performed by taking the item with the largest label as the root. The corresponding new rooted class will be denoted by $\mathcal{A}^\circ$. Since every element of $\mathcal{A}$ correponds uniquely to an element of $\mathcal{A}^\circ$, but the corresponding term $x^n/n!$ in the generating function is replaced by $x^{n-1}/(n-1)!$ (for an element of size n), the correponding generating function satisfies

$$A^\circ(x) = A'(x).$$

Similarly, we can consider a pointed structure $\mathcal{A}^\bullet$ by distinguishing one of the items without discounting it. Since there are n different ways of choosing an item (for an element of size n), the corresponding term $x^n/n!$ in the generating function is replaced by $nx^n/n! = x^n/(n-1)!$ which leads to the relation

$$A^\bullet(x) = xA'(x).$$

Finally, we will deal with the set construction of classes: given a labelled combinatorial structure $\mathcal{A}$, the set construction $\mathrm{Set}(\mathcal{A})$ takes all possible sets of elements in $\mathcal{A}$. The corresponding generating function is then $\exp((A(x)))$, where $A(x)$ is the generating function associated to $\mathcal{A}$.

2.2 Graph decompositions

A *block* of a graph G is a maximal 2-connected subgraph of G. A graph class $\mathcal{G}$ is said to be *block-stable* if it contains the graph e, the unique connected graph with two labelled vertices, and satisfies that a connected graph G belongs to $\mathcal{G}$ if and only if any one of its blocks is in $\mathcal{G}$. The class $\mathcal{G}$ is also said to be *connected component-stable* when any graph G is in $\mathcal{G}$ if and only if all connected components of G belong to $\mathcal{G}$. For a graph class $\mathcal{G}$, we denote by $\mathcal{C}$ and $\mathcal{B}$ the families of connected and 2-connected graphs in $\mathcal{G}$, respectively. In particular, if $\mathcal{G}$ is a block-stable and connected-component stable class of graphs, then the following combinatorial decomposition holds:

$$\mathcal{G} = \mathrm{Set}(\mathcal{C}), \qquad\qquad\qquad \mathcal{C}^{\bullet} = \bullet \times \mathrm{Set}(\mathcal{B}^{\circ} \circ \mathcal{C}^{\bullet}).$$

The previous formulas read as follows: first, each graph in $\mathcal{G}$ is a set of elements in $\mathcal{C}$. Secondly, a pointed connected graph in $\mathcal{C}^{\bullet}$ can be decomposed as the root vertex, and a set of pointed blocks (the ones incident with the root vertex) where we substitute on each vertex a rooted connected graph. See [1, 4, 8] for details. These expressions translate into equations of EGF in the following way:

$$G(x) = \exp(C(x)), \qquad\qquad\qquad C^{\bullet}(x) = x \exp(B^{\circ}(C^{\bullet}(x))).$$

See [16] for further results on graph decompositions and connectivity on graphs.

2.3 Asymptotics for subcritical graph classes

We call a block-stable and vertex labelled graph class *subcritial* if $\eta B''(\eta) > 1$, where η denotes the radius of convergence of $B(x)$. In particular this is satisfied if $B''(x) \to \infty$ as $x \to \eta-$. Cayley trees, cactus graphs, outerplanar graphs and series-parallel graphs are subcritical. The main analytic property of subcritical graph classes is that they have many universal asymptotic behaviours, see [2, 5, 14, 7].

In our context, we will just use the fact that the property $\eta B''(\eta) > 1$ ensures that the functional equation $C^{\bullet}(x) = x \exp(B^{\circ}(C^{\bullet}(x)))$ has solution $C^{\bullet}(x)$ that has a squareroot singularty at its radius of convergence ρ and, thus, a local expansion of the form

$$C^{\bullet}(x) = xC'(x) = \underline{c}_0 + \underline{c}_1 \left(1 - \frac{x}{\rho}\right)^{1/2} + \underline{c}_2 \left(1 - \frac{x}{\rho}\right) + \underline{c}_3 \left(1 - \frac{x}{\rho}\right)^{3/2} + \cdots, \qquad (1)$$

where ρ is given by $\rho = \underline{c}_0 e^{-B'(\underline{c}_0)}$ and $0 < \underline{c}_0 = C^{\bullet}(\rho) < \eta$ is given by the equation $\underline{c}_0 B''(\underline{c}_0) = 1$. Furthermore $\underline{c}_1 < 0$. Note that the singular behaviour of $B(x)$ at its radius of convergence η is irrelevant for the singular behaviour of $C^{\bullet}(x) = xC'(x)$, we only make use of the (analytic) behaviour of $B'(x)$ around $x = \underline{c}_0 < \eta$.

From (1), and if we assume that the class is also connected component-stable, it follows that $C(x)$ and $G(x) = e^{C(x)}$ have the following singular behaviour around their common radius of convergence ρ:

$$C(x) = c_0 + c_2 \left(1 - \frac{x}{\rho}\right) + c_3 \left(1 - \frac{x}{\rho}\right)^{3/2} + \cdots,$$
$$G(x) = g_0 + g_2 \left(1 - \frac{x}{\rho}\right) + g_3 \left(1 - \frac{x}{\rho}\right)^{3/2} + \cdots,$$

where c_3 and g_3 are positive. If we further assume that $x = \rho$ is the only singularity on the circle of convergence $|x| = \rho$ which is satisfied for all our cases, and for proper positive constants c', c'', it then follows that (see for instance [6])

$$|\mathcal{C}_n| = n! \, [x^n] \, C(x) \sim c' n^{-5/2} \rho^{-n} n! \quad \text{and} \quad |\mathcal{G}_n| = n! \, [x^n] \, C(x) \sim c'' n^{-5/2} \rho^{-n} n!.$$

3   Counting in block-stable graph classes

In this section, we consider block-stable vertex labelled graph classes and set up functional equations for counting maximal independent subsets and maximal matchings. We use the notation $\mathcal{B}$ for the family of 2-connected blocks in a block-stable graph class $\mathcal{G}$ and $\mathcal{C}$ for the family of connected graphs in $\mathcal{G}$.

3.1   Maximal independent sets in block-stable graph classes

A *coloured block* is a pair (I, b) consisting of a block $b \in \mathcal{B}$ together with a distinguished independent set I of b (note that I can be any independent set of b and not only a maximal one). Let $B(x, y_0, y_1, y_2)$ be the generating function enumerating coloured-blocks, where the variable x marks vertices. The extra variables encodes the following: y_0 corresponds to vertices of I, y_1 corresponds to vertices adjacent to a vertex in I (i.e. at distance one from I), and y_2 corresponds to all other vertices, that is to vertices at distance at least two from I.

Similarly, a *pointed coloured block* is a pair (I, b°) consisting of a pointed block $b^\circ \in \mathcal{B}^\circ$ together with a distinguished independent set I of b. Let $B_i = B_i(x, y_0, y_1, y_2)$ be the generating function counting pointed coloured blocks, where the pointed vertex is at distance *exactly i* from I, for $i \in \{0, 1\}$, and at distance *at least* 2 (case $i = 2$). In those cases, the pointed-vertex must neither be encoded by x or by any y_i, for $i \in \{0, 1, 2\}$. Hence,

$$B_i = \frac{1}{x} \cdot \frac{\partial B}{\partial y_i}, \text{ for } i \in \{0, 1, 2\}.$$

A *coloured graph* (J, g) is a pair consisting of a connected graph $g \in \mathcal{C}$ and of a *maximal* independent set J of g. We can define pointed coloured graphs similarly to coloured blocks. Let $C = C(x, y_0, y_1)$ be the generating function counting coloured-graphs, where y_0 and y_1 have the same meaning as in coloured blocks. For $i \in \{0, 1\}$, let $C_i = C_i(x, y_0, y_1)$ be the generating functions enumerating pointed coloured-graphs, for which the pointed vertex is at distance *exactly i* from J. Those two generating functions are given by

$$C_i = \frac{1}{x} \cdot \frac{\partial C}{\partial y_i}, \text{ for } i \in \{0, 1\}. \tag{2}$$

We finally need an auxiliary class. A *special pointed coloured-graph* is a pair (J, g°) where J is an independent set of g which becomes maximal when adding the pointed vertex to J. In other words, a special pointed coloured-graph is obtained from a coloured-graph pointed at a vertex in J by removing it from J. We denote the corresponding counting formula by $C_2(x, y_0, y_1)$. Finally, observe that given a coloured-graph (J, g), the independent set J together with the vertices of g at distance one from J define a partition of $V(g)$. Hence, the following equalities hold:

$$\frac{\partial C}{\partial x} = \frac{y_0}{x} \frac{\partial C}{\partial y_0} + \frac{y_1}{x} \frac{\partial C}{\partial y_1} = y_0 C_0 + y_1 C_1. \tag{3}$$

Obviously we also have

$$G(x, y_0, y_1) = \exp(C(x, y_0, y_1)),$$
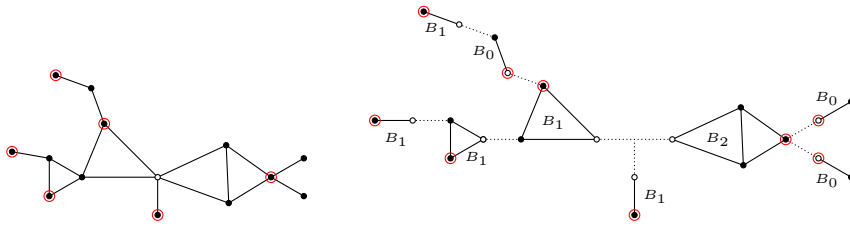
■ **Figure 1** Left is a connected series-parallel graph with a maximal independent set I (vertices circled in red) and pointed at a vertex at distance one from I. Right is its block-decomposition. Pointed vertices are coloured in white.

where $G(x, y_0, y_1)$ denotes the corresponding generating function of coloured graphs in $\mathcal{G}$. The following lemma describes connected structures in terms of their block-decomposition (see Figure 1 for an example). Thus, if we know $B(x, y_0, y_1, y_2)$ (or just $B_i(x, y_0, y_1, y_2)$, for each $i \in \{0, 1, 2\}$), then we can determine $\frac{\partial C}{\partial x}(x, y_0, y_1)$ and consequently $C(x, y_0, y_1)$ and $G(x, y_0, y_1)$.

▶ **Lemma 4.** *With the above notations, the following system of equations holds:*

$$\begin{aligned}
C_0 &= \exp(B_0(x, y_0 C_0, y_1(C_1 + C_2), y_1 C_1)), \\
C_1 &= (\exp(B_1(x, y_0 C_0, y_1(C_1 + C_2), y_1 C_1)) - 1) \cdot C_2, \\
C_2 &= \exp(B_2(x, y_0 C_0, y_1(C_1 + C_2), y_1 C_1)).
\end{aligned} \tag{4}$$

Proof. Let us start by finding an expression for C_0 and let (I, g°) be a pointed coloured-graph whose pointed vertex is in I. Following the decomposition of graphs into blocks, observe that the pointed vertex of g° determines a set of pointed coloured-blocks (J_i, b_i°) (with $i = 1, \ldots, k$ for a certain k) for which the root of each b_i° belongs to J_i, i.e. coloured-blocks with the pointed vertex in J_i (and hence, counted by B_0). Observe that the independent sets J_i can be extended to I by pasting pointed coloured-graphs on each of their vertices (and completing the graph to g°).

Without loss of generality, let us now fix a $j \in \{1, \ldots, k\}$ and analyse the pair (J_j, b_j°). First, to every vertex of b_j° in J_j must be attached a coloured-graph (L, h°) whose root is in L, i.e. a coloured-graph counted by C_0. In terms of generating functions, this translates to the substitution of y_0 by $y_0 C_0$. Second, to each vertex of b_j° at distance one from J_j, the root of the pointed coloured-graph (L, h°) attached to it can either be at distance one or more from L. This then translates to the substitution of y_1 by $y_1(C_1 + C_2)$. Finally, if a vertex of b_j° is at distance at least two from J_j, then the root of the coloured-graph (L, h°) attached to it must be at distance one from L, as we need to extend the independent set to a maximal one. This translates to the substitution of y_2 by $y_1 C_1$ and the first equation of (4) holds. The study of C_2 is obtained following the exact same arguments as in C_0.

Let us finally discuss the equation for C_1. Assume that (I, g°) is a pointed coloured-graph and that (J_i, b_i°) (for $i = 1, \ldots, k$) are the pointed coloured-blocks incident with the pointed vertex of g°. In particular, for each $i \in \{1, \ldots, k\}$ the pointed vertex of b_i° is either at distance one or at least two from J_i. Nevertheless, observe that there exists at least one of the pointed-blocks (J_j, b_j°) whose pointed vertex is at distance one from J_j. This gives us that

$$\begin{aligned}
C_1 &= \exp_{\geq 1}(B_1(x, y_0 C_0, y_1(C_1 + C_2), y_1 C_1)) \cdot \exp(B_2(x, y_0 C_0, y_1(C_1 + C_2), y_1 C_1)) \\
&= C_2 \left(\exp(B_1(x, y_0 C_0, y_1(C_1 + C_2), y_1 C_1)) - 1 \right).
\end{aligned}$$

Which concludes the argument.   ◀

3.2   Asymptotic Analysis

We study next the analytic structure of the solutions of the systems (4) and (9) provided that the functions B_i behave in a *proper way* that is similar to the behaviour of $B(x)$ in the case of sub-critical graph classes. Under these hypothesis (see Lemma 5) it is then very easy to prove Theorems 1 and 3 which will be done at the end of this subsection. For the sake of brevity we only discuss the system (4), the analysis of (9) runs along the same lines.

First we note that the functions $B_i(x, y_0, y_1, y_2)$ are actually functions in three variables since a monomial $x^n y_0^{k_0} y_1^{k_1} y_2^{k_2}$ can only appear if $k_0 + k_1 + k_2 = n$, that is, we have $B_i(x, y_0, y_1, y_2) = B_i(1, xy_0, xy_1, xy_2)$ or equivalently $B_i(x, y_0, y_1, y_2)$
$= B_i(xy_2, y_0/y_2, y_1/y_2, 1)$. However, it is more convenient to work with all four variables x, y_0, y_1, y_2. If y_0, y_1, y_2 are positive real numbers then the function $x \mapsto B(x, y_0, y_1, y_2)$ is a power series with non-negative coefficients. Hence the radius of convergence of this function coincides with its dominant singularity in x. We will denote this radius of convergence by $R(y_0, y_1, y_2)$. Similarly for the solution functions C_0, C_1, C_2 of the System (4) we denote by $\rho_i(y_0, y_1)$, $i = 0, 1, 2$, the radius of convergence with respect to x when y_0, y_1 are positive real numbers.

▶ **Lemma 5.** *Suppose that the function $R(y_0, y_1, y_2)$ extends to an analytic function $R(y_0, y_1, y_2)$ for a sufficiently small neighbourhood around the positive real numbers. Furthermore assume that for all positive real numbers y_0, y_1, y_2 we have*

$$\lim_{x \to R(y_0, y_1, y_2)-} \frac{\partial^2 B}{\partial y_i^2}(x, y_0, y_1, y_2) = \infty \tag{5}$$

for at least one of the $i \in \{0, 1, 2\}$. Then the solutions C_0, C_1, C_2 of the system (4) have the property that the functions $\rho_i(y_0, y_1)$, $i = 0, 1, 2$, coincide and extend to an analytic function $\rho(y_0, y_1)$ for a sufficiently small neighbourhood around the positive real numbers. Moreover, the dominant singularity is of squareroot type and we have a local expansion of the form

$$C_i(x, y_0, y_1) = \quad c_{i,0}(y_0, y_1) + c_{i,1}(y_0, y_1) \left(1 - \frac{x}{\rho_1(y_0, y_1)}\right)^{1/2}$$
$$+ c_{i,2}(y_0, y_1) \left(1 - \frac{x}{\rho_1(y_0, y_1)}\right) + \cdots, \tag{6}$$

where $c_{i,1}(y_0, y_1) < 0$ (for positive real y_0, y_1) and that extends to sufficiently small neighbourhood in x, y_0, y_1 around the positive real numbers.

Proof. We recall some basic facts on (positive) systems of functional equations that are taken from [4]. Suppose that we have a system of three equations of the form

$$\begin{aligned} C &= F(x, C, D, E), \\ D &= G(x, C, D, E), \\ E &= H(x, C, D, E), \end{aligned}$$

in unknown functions $C = C(x)$, $D = D(x)$, $E = E(x)$, where F, G, H are power series with non-negative coefficients. We also assume that the system is strongly connected which means that no subsystem can be solved before solving the whole system. We set

$$\Delta = \begin{vmatrix} 1 - F_C & -F_D & -F_E \\ -G_C & 1 - G_D & -G_E \\ -H_C & -H_D & 1 - H_E \end{vmatrix}$$

the functional determinant of the system $\{C - F = 0,\ D - G = 0,\ E - H = 0\}$ and let r be the spectral radius of the Jacobian matrix of the right hand-side of the system of equations.

Note that $r = 1$ implies that $\Delta = 0$. We also assume that there is a unique non-negative solution $C(0)$, $D(0)$, $E(0)$ for $x = 0$ with the property that $r < 1$, which also shows that $\Delta \neq 0$. Thus by iteration, the solution for $x = 0$ extends to power series solutions $C(x)$, $D(x)$, $E(x)$ with non-negative coefficients and a positive radius of convergence. By the strongly connectedness assumption, this radius of convergence ρ is the same for all three solutions functions $C(x)$, $D(x)$, $E(x)$. By the theory given in [4], this radius of convergence is determined by the condition $r = 1$ provided that we are still working within the region of convergence of F, G, and H. The condition $r = 1$ can be also witnessed by the condition $\Delta = 0$ or equivalently by the condition

$$\frac{F_D G_E H_C + F_E G_C H_C}{(1 - F_C)(1 - G_D)(1 - H_E)} + \frac{G_E H_D}{(1 - G_D)(1 - H_E)} + \frac{F_E H_C}{(1 - F_C)(1 - H_E)} + \frac{F_D G_C}{(1 - F_C)(1 - G_D)} = 1. \tag{7}$$

Note that the left hand-side is smaller than 1 for $x = 0$ and $C = C(0)$, $D = D(0)$, $E = E(0)$ and is strictly increasing in x. Thus, in order to find ρ we just have to find the value for which the left hand-side hits the value 1. If we are still inside the region of convergence of F, G, and H, then it follows that the solution functions $C(x)$, $D(x)$, $E(x)$ have a squareroot singularity of the form (1) at $x = \rho_1$.

In our special situation all the above assumptions concerning positivity, strongly connectedness etc. are satisfied. Now let us also observe that $\frac{\partial^2 B}{\partial y_0^2} \to \infty$ implies that $F_C \to \infty$, since $F(x) = \exp(B_0(x, y_0 C, y_1(D + E)), y_1 D)$ and $B_0 = \frac{1}{x} \frac{\partial B}{\partial y_0}$ (note the two different meanings of y_0). Similar observations hold for G_D and H_E. Thus, it is clear that (7) is satisfied inside the region of convergence of F, G and H. We recall the fact that the left hand-side of (7) is smaller than 1 for $x = 0$ and strictly increasing in x. ◄

Finally we show that under the hypothesis of Lemma 5, it is immediate to deduce our main results Theorem 1 and Theorem 3: from (6) and (3) it follows that $C(x, y_0, y_1)$ can be represented as

$$C(x, y_0, y_1) = c_0(y_0, y_1) + c_2(y_0, y_1) \left(1 - \frac{x}{\rho_1(y_0, y_1)} \right) + c_3(y_0, y_1) \left(1 - \frac{x}{\rho_1(y_0, y_1)} \right)^{3/2} + \cdots,$$

where $c_3(y_0, y_1) > 0$ for positive real y_0, y_1. Thus, if we set $y_0 = y_1 = 1$ and $\rho_1(1, 1) = \rho_1$, then we have

$$C(x, 1, 1) = c_0(1, 1) + c_2(1, 1) \left(1 - \frac{x}{\rho_1} \right) + c_3(1, 1) \left(1 - \frac{x}{\rho_1} \right)^{3/2} + \cdots,$$

and consequently

$$G(x, 1, 1) = \sum_{n \geq 0} |\mathcal{I}_n| \frac{z^n}{n!} = \exp(C(x, 1, 1))$$

$$= g_0(1, 1) + g_2(1, 1) \left(1 - \frac{x}{\rho_1} \right) + g_3(1, 1) \left(1 - \frac{x}{\rho_1} \right)^{3/2} + \cdots.$$

This directly implies Theorem 1 for the case of maximal independent sets by standard singularity analysis (see [6]). We just have to observe that $x_0 = \rho = \rho(1, 1)$ is the only singularity on the circle of convergence. However, this follows from the fact that there exists graphs of all sizes $n \geq 1$.

Finally, if we set $y_1 = 1$ then we have

$$G(x, y_0, 1) = \sum_{n \geq 0} \mathbb{E}[y_0^{SI_n}] |\mathcal{I}_n| \frac{z^n}{n!} = \exp(C(x, y_0, 1))$$

$$= g_0(y_0, 1) + g_2(y_0, 1) \left(1 - \frac{x}{\rho_1(y_0, 1)} \right) + g_3(1, 1) \left(1 - \frac{x}{\rho_1(y_0, 1)} \right)^{3/2} + \cdots$$

Hence, a direct application of [4, Theorem 2.35] implies a central limit theorem of the proposed form, as well as the asymptotic expansions for the expected value and variance. This proves Theorem 3 for the case of maximal independent sets.

What remains is to check condition (5). We work this out in details for trees and series-parallel graphs in Section 4. The other cases (cactus graphs and outerplanar graphs) can be handled in a similar way and this will be covered in the paper version of this extended abstract.

4   Applications

Our first application concerns the most basic subcritical graph class, namely Cayley trees. We note that the case of maximal independent sets was already discussed in [13]. We will then deal with the class of series-parallel graphs.

4.1   Maximal independent sets in trees

In both structures (maximal independent sets and maximal matchings), we proceed following the block-decomposition of trees, and we explicitly give the generating functions B_0, B_1 and B_2. Notice that in a tree, blocks are reduced to single edges. The computations of the constants given in Table 1 are obtained by computing the branch point of the corresponding system, using the explicit expressions for B_0, B_1 and B_2.

We first give the generating functions counting the rooted blocks carrying an independent set. From the possible choices of an independent set in a single edge, namely $B(x, y_0, y_1, y_2) = \frac{x^2}{2} \left(2y_0 y_1 + y_2^2 \right)$, we obtain that

$$B_0 = xy_1, \qquad\qquad B_1 = xy_0, \qquad\qquad B_2 = xy_2.$$

Thus, the following property holds:

$$\lim_{x \to \infty} \frac{\partial^2 B}{\partial y_2^2} = \lim_{x \to \infty} x = \infty.$$

So Lemma 5 applies in the case of maximal independent sets in trees, which completes the proof.

4.2   Maximal independent sets in series-parallel graphs

We are now concerned with the generating functions of the labelled series-parallel graphs carrying a maximal independent set. As above, the vertices of the graphs carrying an independent set I are said to be of type i ($i \in \{0, 1\}$), when they are at distance i from I, and of type 2 otherwise. We will now explicit the classical decomposition of graphs in terms of networks.

Series-parallel networks

A *series-parallel network* D_{ij} is a labelled graph with an oriented edge ij that is distinguished and whose endpoints, called the *poles*, are unlabeled and respectively of type i and j. Observe that by symmetry $D_{ij} = D_{ji}$, so we can restrict the range of the pairs of indexes ij to the set $\{00, 01, 02, 11, 12, 22\}$. The network D_{ij} is either the single rooted edge e_{ij}, where $e_{01} = e_{22} = y$ and $e_{ij} = 0$ otherwise, a series network counted by the generating function S_{ij}, or a parallel network counted by the generating function P_{ij}. We then specify those generating functions via the following positive system of 18 equations and 18 unknowns:

$$D_{ij} = e_{ij} + S_{ij} + P_{ij},$$
$$S_{ij} = D_{i0}xy_0(D_{0j} - S_{0j}) + (D_{i1} + D_{i2})xy_1(D_{1j} - S_{1j}) + (D_{i1}y_1 + D_{i2}y_2)x(D_{2j} - S_{2j}),$$
$$P_{00} = \exp_{\geq 2}(S_{00}),$$
$$P_{01} = y\exp_{\geq 1}(S_{01} + S_{02}) + \exp_{\geq 2}(S_{01}) + \exp_{\geq 1}(S_{01})\exp_{\geq 1}(S_{02}),$$
$$P_{02} = \exp_{\geq 2}(S_{02}),$$
$$P_{11} = \exp_{\geq 2}(S_{11}) + \exp_{\geq 1}(S_{11})(y\exp(2S_{12} + S_{22}) + \exp_{\geq 1}(2S_{12} + S_{22}))$$
$$\qquad + (1 + y)\exp_{\geq 1}(S_{12})^2\exp(S_{22}),$$
$$P_{12} = y\exp_{\geq 1}(S_{12})\exp(S_{22}) + \exp_{\geq 2}(S_{12}) + \exp_{\geq 1}(S_{12})\exp_{\geq 1}(S_{22}),$$
$$P_{22} = y\exp_{\geq 1}(S_{22}) + \exp_{\geq 2}(S_{22}).$$

In order to proceed further, we eliminate D_{ij} from this system to obtain a positive and strongly connected system of equations for $S_{ij} = S_{ij}(x, y, y_0, y_1, y_2)$ and $P_{ij} = P_{ij}(x, y, y_0, y_1, y_2)$, where the right hand-side consists of entire functions (note that for the equations defining S_{ij}, the term $D_{ij} - S_{ij} = e_{ij} + P_{ij}$, which makes the whole system positive). Thus, all functions have a common singular behaviour that is (again) of squareroot type:

$$S_{ij}(x, y, y_0, y_1, y_2) = s_{0;ij}(y, y_0, y_1, y_2) + s_{1;ij}(y, y_0, y_1, y_2)\left(1 - \frac{x}{\rho(y, y_0, y_1, y_2)}\right)^{1/2} + \cdots$$

and

$$P_{ij}(x, y, y_0, y_1, y_2) = p_{0;ij}(y, y_0, y_1, y_2) + p_{1;ij}(y, y_0, y_1, y_2)\left(1 - \frac{x}{\rho(y, y_0, y_1, y_2)}\right)^{1/2} + \cdots,$$

where $s_{1;ij}(y, y_0, y_1, y_2) < 0$ and $p_{1;ij}(y, y_0, y_1, y_2) < 0$ for positive y, y_0, y_1, y_2.

2-connected series-parallel graphs

The next step is to relate these network generating functions with the generating function $B(x, y, y_0, y_1, y_2)$ of independent sets in 2-connected series parallel graphs. Note that an added variable y takes into account the number of edges. In the (usual) counting procedure for series parallel graphs, we have the property that $\frac{\partial B}{\partial y} = \frac{x^2}{2}\exp(S(x, y))$, where $S(x, y)$ denotes the generating function of series networks (similarly to the above). The combinatorial property behind this relation is that an edge-rooted series-parallel graph (that corresponds to the generating function $\frac{\partial B}{\partial y}$) can be seen as a series-parallel network between the two vertices of the root-edge, consisisting of this edge and a collection of series-networks between the two vertices.

In our present situation we have a similar property, namely

$$\frac{\partial B}{\partial y} = x^2 y_0 y_1 \exp(S_{01} + S_{02}) + \frac{x^2}{2} y_2^2 \exp(S_{22}) + x^2 y_1 y_2 \exp_{\geq 1}(S_{12}) \exp(S_{22})$$

$$+ \frac{x^2}{2} y_1^2 \left(\exp(S_{11} + 2S_{12} + S_{22}) - 2\exp(S_{12} + S_{22}) + \exp(S_{22}) \right).$$

This is immediate by considering all possible situation for the rooted edge. Observe that, despite the negative terms, $\frac{\partial B}{\partial y}$ is in fact a positive function of the generating functions $\{S_{ij}\}$. Hence, $\frac{\partial B}{\partial y}$ has also a squareroot singularity:

$$\frac{\partial B}{\partial y} = b_0(y, y_0, y_1, y_2) + b_1(y, y_0, y_1, y_2) \left(1 - \frac{x}{R(y, y_0, y_1, y_2)} \right)^{1/2} + \cdots,$$

where $b_1(y, y_0, y_1, y_2) < 0$ for positive y, y_0, y_1, y_2. Next, by applying the proof method of [4, Lemma 2.28], we can integrate $\frac{\partial B}{\partial y}$ with respect to y and then take the derivative with respect to y_0 and obtain the same kind of squareroot singularity for $\frac{\partial B}{\partial y_0}$

$$\frac{\partial B}{\partial y_0} = b_{1,0}(y, y_0, y_1, y_2) + b_{1,1}(y, y_0, y_1, y_2) \left(1 - \frac{x}{R(y, y_0, y_1, y_2)} \right)^{1/2} + \cdots,$$

and consequently the following representation of $\frac{\partial^2 B}{\partial y_0^2}$:

$$\frac{\partial^2 B}{\partial y_0^2} = b_{2,-1}(y, y_0, y_1, y_2) \left(1 - \frac{x}{R(y, y_0, y_1, y_2)} \right)^{-1/2} + b_{2,1}(y, y_0, y_1, y_2) + \cdots,$$

which implies that (5) holds for $i = 0$. This completes the proof for maximal independent sets in series-parallel graphs.

References

1 François Bergeron, Gilbert Labelle, and Pierre Leroux. *Combinatorial species and tree-like structures*, volume 67. Cambridge University Press, 1998.

2 Nicla Bernasconi, Konstantinos Panagiotou, and Angelika Steger. The degree sequence of random graphs from subcritical classes. *Combinatorics, Probability and Computing*, 18(5):647–681, 2009.

3 Therese Biedl, Erik D. Demaine, Christian A. Duncan, Rudolf Fleischer, and Stephen G. Kobourov. Tight bounds on maximal and maximum matchings. *Discrete Mathematics*, 285:7–15, 2004.

4 Michael Drmota. *Random trees: an interplay between combinatorics and probability*. SpringerWienNewYork, 2009.

5 Michael Drmota, Éric Fusy, Mihyun Kang, Veronika Kraus, and Juanjo Rué. Asymptotic study of subcritical graph classes. *SIAM J. Discrete Math.*, 25(4):1615–1651, 2011.

6 Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, 2009.

7 Agelos Georgakopoulos and Stephan Wagner. Limits of subcritical random graphs and random graphs with excluded minors. Available on-line on `arXiv:1512.03572`.

8 Ian P. Goulden and David M. Jackson. *Combinatorial enumeration*. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1983.

9 Jerold R. Griggs, Charles M. Grinstead, and David R. Guichard. The number of maximal independent sets in a connected graph. *Discrete Mathematics*, 68:211–220, 1988.

10 Joanna Górska and Zdzisław Skupień. Trees with maximum number of maximal matchings. *Discrete Mathematics*, 307:1367–1377, 2007.

11 Clemens Heuberger and Stephan Wagner. The number of maximum matchings in a tree. *Discrete Mathematics*, 311(21):2512–2542, 2011.

12 Min-Jen Jou and Gerard J. Chang. The number of maximum independent sets in graphs. *Taiwanese Journal of Mathematics*, 4(4):685–695, 2000.

13 Amram Meir and John W. Moon. On maximal independent sets of nodes in trees. *Journal of Graph Theory*, 12(2):265–283, 1988.

14 Konstantinos Panagiotou, Benedikt Stufler, and Kerstin Weller. Scaling limits of random graphs from subcritical classes. *Ann. Probab.*, 44(5):3291–3334, 09 2016. `doi:10.1214/15-AOP1048`.

15 Bruce E. Sagan. A note on independent sets in trees. *SIAM Journal on Algebraic Discrete Methods*, 1(1):105–108, 1988.

16 William T. Tutte. *Connectivity in graphs*, volume 285. University of Toronto Press, 1966.

17 Herbert S. Wilf. The number of maximal independent sets in a tree. *SIAM Journal on Algebraic Discrete Methods*, 7(1):125–130, 1986.

18 Iwona Włoch. Trees with extremal numbers of maximal independent sets including the set of leaves. *Discrete Mathematics*, 308:4768–4772, 2008.

19 Jennifer Zito. The structure and maximum number of maximum independent sets in trees. *Journal of Graph Theory*, 15(2):207–221, 1991.

A Maximal matchings

A.1 Maximal matchings in block-stable classes of graphs

In this subsection we deal with the case of maximal matchings. Most of the definitions and concepts are the natural analogues of the ones developed in the case of maximal independent sets. Hence, we will skip unnecessary repetitions.

A *matched block* is a triple (I, M, b) with a block $b \in \mathcal{B}$, a matching M in b, and an independent set I of b, and where no element of I is incident to an edge in M. In other words, we split the set of vertices of b in three disjoint subsets: matched vertices, vertices in I, and the rest. A *pointed* matched block is a triple (I, M, b°), where $b^\circ \in \mathcal{B}^\circ$ and M and I are respectively a matching and an independent set of b, and where again no element of I is incident to any edge in M. Let $\overline{B}(x, z_0, z_1, z_2)$ be the generating function counting matched blocks, where the variable x marks vertices, z_0 marks vertices in I, z_1 marks vertices matched by M, and z_2 the remaining ones. For $i \in \{0, 1, 2\}$, let $\overline{B}_i = \overline{B}_i(x, z_0, z_1, z_2)$ be the generating function counting pointed matched blocks where the pointed vertex is either in I, is incident with M or none of the previous cases. In particular,

$$\overline{B}_i = \frac{1}{x} \cdot \frac{\partial \overline{B}}{\partial z_i}, \text{ for } i \in \{0, 1, 2\}.$$

A *matched graph* is a triple (I, M, g) consisting of a connected graph g in $\mathcal{C} \subseteq \mathcal{G}$, a matching M of g, and an independent set $I \subset V(g) \setminus V(M)$. Similarly, a *pointed* matched graph is a triple (M, I, g°) where now g° is a pointed graph. Let $\overline{C}(x, z_0, z_1, z_2)$ be the generating function counting matched graphs, where x, z_0, z_1 and z_2 respectively mark vertices, vertices incident with I, vertices incident with M, and the rest of the vertices. Notice that when $z_2 = 0$, $\overline{C} := \overline{C}(x, z_0, z_1) = \overline{C}(x, z_0, z_1, 0)$ encodes matched graphs where M is maximal. For each $i \in \{0, 1, 2\}$, let us define the following generating function

$$\overline{C}_i = \overline{C}_i(x, z_0, z_1) = \frac{1}{x} \cdot \frac{\partial \overline{C}}{\partial z_i}(x, z_0, z_1, 0).$$

Observe then that $\overline{C}_0$ counts pointed matched graphs, where the matching is maximal and the pointed vertex belongs to the independent set, $\overline{C}_1$ counts pointed matched graphs, where the matching is maximal and the pointed vertex belongs to the matching, whereas $\overline{C}_2$ counts pointed matched graphs, where the matching is not necessarily maximal and the pointed vertex does not belong to either the independent set or the matching. In the latter case, the matching is maximal except for possibly the pointed vertex, which might be unmatched and adjacent to other unmatched vertices. In particular, this implies that the generating function of pairs of connected graphs and maximal matchings is given by

$$\frac{\partial \overline{C}}{\partial x} = z_0 \overline{C}_0 + z_1 \overline{C}_1. \tag{8}$$

The main idea behind this encoding of the problem is that vertices in the independent set I play the role of vertices that will not be matched in the block decomposition. In particular, we exploit independence in order to ensure that the matching cannot be extended. On the other hand, the set of vertices that are unmatched and not in I will be matched by an attached block of the decomposition.

The following lemma relates all the previous generating functions. Note that the generating functions $\overline{C}(x, z_0, z_1, 0)$ and $\overline{G}(x, z_0, z_1) = \exp(\overline{C}(x, z_0, z_1))$ directly follow from the solution of the next system.

▶ **Lemma 6.** *The following equalities hold:*

$$\begin{aligned}
\overline{C}_0 &= \exp(\overline{B}_0(x, z_0\overline{C}_0, z_1\overline{C}_2, z_1\overline{C}_1)), \\
\overline{C}_1 &= \overline{C}_2\,\overline{B}_1(x, z_0\overline{C}_0, z_1\overline{C}_2, z_1\overline{C}_1), \\
\overline{C}_2 &= \exp(\overline{B}_2(x, z_0\overline{C}_0, z_1\overline{C}_2, z_1\overline{C}_1)).
\end{aligned} \tag{9}$$

Proof. Let (M, I, g°) be a pointed matched graph, with pointed vertex v. Suppose first that $v \in I$, i.e. the case counted by $\overline{C}_0$. It therefore is the pointed vertex of a (possibly empty) set of adjacent pointed blocks (I_j, M_j, b_j°), in which $v \in I_j$, and is not adjacent to any other pointed block. This means that all the pointed blocks adjacent to v are counted by $\overline{B}_0$. Suppose next that $v \in V(M)$, i.e. the case counted by $\overline{C}_1$. Then the edge of M incident with v must belong to a single pointed block whose pointed vertex (v) is incident to an edge of the respective matching. Hence, attached to v are this one block together with any number (possibly null) of pointed blocks counted by $\overline{B}_2$, since v is already incident to an edge of a matching. Suppose finally that we are in the case counted by $\overline{C}_2$. Then v is neither in I nor in $V(M)$. Therefore, any block attached to it must not have its pointed vertex in an independent set or incident to and edge of a matching. This means that v belongs to a (possibly empty) set of blocks counted by $\overline{B}_2$.

Let now $\{(I_i, M_i, b_i^\circ) : i = 1, \ldots k\}$ be the pointed blocks in the decomposition of (M, I, g°) and fix a $j \in \{1, \ldots k\}$. Then using the same arguments as just above, we see that to a vertex in I_j must be attached a pointed matched graph counted by $\overline{C}_0$, to a vertex in $V(M_j)$ one counted by $\overline{C}_2$ and to any other vertex must be attached a pointed matched graph counted by $\overline{C}_1$, as we need to extend the matching to maximality. ◀

A.2 Maximal matchings in trees

Observe that in this case $\overline{B}(x, z_0, z_1, z_2) = \frac{x^2}{2}\left(2z_0 z_2 + z_1^2 + z_2^2\right)$, which gives

$$\overline{B}_0 = x z_2, \qquad\qquad \overline{B}_1 = x z_1, \qquad\qquad \overline{B}_2 = x(z_0 + z_2).$$

Hence, we are in a similar situation as above and Lemma 5 applies. This completes the proof for maximal matchings in trees.

A.3   Maximal matchings in series-parallel graphs

We proceed similarly to Subsection 4.2. Let G be a series-parallel graph with a matching M and an independent set I such that $I \cap V(M) = \emptyset$. A vertex v of G is said to be of type 0 when $v \in I$, of type 1 when $v \in V(M)$ and of type 2 otherwise.

Series-parallel networks

Let $D_{ij}(x, y, z_0, z_1, z_2)$ be the exponential generating function counting matchings in series-parallel networks whose poles are of type i and j. As before, observe that $D_{ij} = D_{ji}$ and for $ij \in \{00, 01, 02, 11, 12, 22\}$, define S_{ij} and P_{ij} to be the generating functions counting matchings in networks that are respectively series and parallel.

The following system of 18 equations and 18 unknowns holds:

$$
\begin{aligned}
D_{ij} &= e_{ij} + S_{ij} + Pij, \\
S_{ij} &= (D_{i0} - S_{i0})xz_0 D_{0j} + (D_{i1} - S_{i1})xz_1 D_{2j} + (D_{i2} - S_{i2})x(z_1 D_{1j} + z_2 D_{2j}), \\
P_{00} &= \exp_{\geq 2}(S_{00}), \\
P_{01} &= S_{01}(y \exp(S_{02}) + \exp_{\geq 1}(S_{02})), \\
P_{02} &= y \exp_{\geq 1}(S_{02}) + \exp_{\geq 2}(S_{02}), \\
P_{11} &= (yS_{11} + (1 + y)S_{12}^2) \exp(S_{22}) + (y + S_{11}) \exp_{\geq 1}(S_{22}), \\
P_{12} &= S_{12}(y \exp(S_{22}) + \exp_{\geq 1}(S_{22})), \\
P_{22} &= y \exp_{\geq 1}(S_{22}) + \exp_{\geq 2}(S_{22}),
\end{aligned}
$$

where this time $e_{02} = e_{11} = e_{22} = y$ and $e_{ij} = 0$.

2-connected series-parallel graphs

It remains to check the relevent analytic properties of $B(x, y, z_0, z_1, z_2)$ in order to assure that Lemma 5 can be applied. Eliminating D_{ij} from the above system, we again get a positive and strongly connected system of equations for the set of generating functions $\{S_{ij}, P_{ij}\}$, where the right hand-side consists of entire functions. In particular, the functions S_{ij} and P_{ij} all have a common singular behaviour that is of squareroot type.

And we have that

$$
\begin{aligned}
\frac{\partial B}{\partial y} &= x^2 z_0 z_1 S_{01} \exp(S_{02}) + x^2 z_0 z_2 \exp(S_{02}) + x^2 z_1 z_2 S_{12} \exp(S_{22}) \\
&\quad + \frac{x^2}{2} z_2^2 \exp(S_{22}) + \frac{x^2}{2} z_1^2 \left(S_{11} + S_{12}^2 + 1\right) \exp(S_{22}).
\end{aligned}
$$

Finally, using the very same arguments as in the case of maximal independent sets, we show that (5) is satisfied in the context of maximal matchings in series-parallel graphs. Thus completing the proof.

The Number of Double Triangles in Random Planar Maps

Michael Drmota[1]

TU Wien, Institute of Discrete Mathematics and Geometry, Wiedner Hauptstrasse 8–10, 1040
Vienna, Austria
michael.drmota@tuwien.ac.at
(iD) https://orcid.org/0000-0002-6876-6569

Guan-Ru Yu[2]

TU Wien, Institute of Discrete Mathematics and Geometry, Wiedner Hauptstrasse 8–10, 1040
Vienna, Austria
guan-ru.yu@tuwien.ac.at

Abstract

The purpose of this paper is to provide a central limit theorem for the number of occurrences of
double triangles in random planar maps. This is the first result of this kind that goes beyond face
counts of given valency. The method is based on generating functions, an involved combinatorial
decomposition scheme that leads to a system of catalytic functional equations and an analytic
extension of the *Quadratic Method* to systems of equations.

1   Introduction

A *planar map* is a connected planar graph (loops and multiple edges are allowed) embedded
into the plane up to homeomorphism. A map is *rooted* if a vertex v is chosen from the map
and a half-edge e is chosen from all the edges incident to v, and called the *root vertex* and
root edge, respectively. Moreover, a planar map separates the plane into several connected
regions called *faces*. The *root face* in a rooted map is the face which is on the left side of e
(sometimes the root face is defined as the right side of e, but this does not make a principle
difference). Without loss of generality we may assume that the root face is the infinite (or
outer) face, in particular the root edge e is then adjacent to the outside face. In this paper,
all maps we consider are rooted and planar. By convenience we also include the trivial map
that consists just of one vertex and one face (which are also rooted). It is well known that
there are precisely $M_n = \dfrac{2 \cdot 3^n \binom{2n}{n}}{n(n+1)}$ different rooted planar maps with n edges [12]. In what

Figure 1 Degenerate cases of double triangles that are represented as bold edges.

follows we assume that for any fixed n every map with n edges is equally likely. Hence every parameter of rooted planar maps can be considered as a random variable related to random planar maps with n edges.

The main goal of this paper is to prove the following theorem:

▶ **Theorem 1.** *The number X_n of edges with valency 3 faces on both sides in a random planar map with n edges satisfies a central limit law, i.e.,*

$$\frac{X_n - \mathbb{E}[X_n]}{\mathrm{Var}[X_n]^{1/2}} \to \mathcal{N}(0,1), \tag{1}$$

where $\mathbb{E}[X_n] = \mu n + O(1)$ and $\mathrm{Var}[X_n] = \sigma^2 n + O(1)$, and μ, σ are positive constants.

▶ Remark. We cannot derive a simple analytic expression μ and σ since our analysis is implicitly based on an infinite system of equations. So they are definitely hard to compute, even in an approximate sense.

In an slight abuse of notation we will call the occurrence of an edge with valency 3 faces on both sides a *double triangle*. Namely there are some degenerate cases as Figure 1 shows (in the first case we identify vertices a and c and then also the edges ab and bc so that we havee two double triangles between two triangles; in the second case, we identify vertices b and d and then the edges ab and ad so that a bridge represents a double triangle).

The background of this result is a widely believed conjecture that the number of pattern occurrences in planar maps (and many related graph classes) obeys a central limit theorem. For (general) planar maps there are only very few results in this direction, see [7, 4] for the number of faces of given valency or [9] for triangulation patterns in 2-connects triangulations and quadrangulations patterns in simple quadrangulations. We also want to mention that the expected number of occurrences of a given pattern in a random planar map with n edges is asymptotically linear: $\mathbb{E} X_n \sim cn$ for some constant $c > 0$. This follows from the fact that random planar maps have a Benjamini-Schramm limit, see [8, 1, 10, 11]. As mentioned before it is expected that X_n satisfies a central limit theorem in all cases. However, it seems that this is out of reach at the moment. Even the simplest case beyond face-pattern that is considered in this paper requires a thorough and delicate analysis for the combinatorial part as well as for the analytic part. We use an approach that is in principle close to that of [7], namely we use generating functions, set up a system of catalytic functional equations (Section 2) and finally provide a proper analytic extension of the classical *Quadratic Method* [3, 12] (Section 3).

2   Combinatorics

Our goal is to set up a recursive structure of planar maps that is suitable to take occurrence of double triangles into account. For this purpose we distinguish between three different cases: the initial case (a map without any edge, denoted by $\bullet$), the bridgeable case (maps, where the root edge is a bridge, denoted by $\mathcal{D}^{(b)}$) and the non-bridgeable case (maps, where the root edge is not a bridge, denoted by $\mathcal{D}^{(n)}$):

$$\mathcal{D} = \bullet + \mathcal{D}^{(b)} + \mathcal{D}^{(n)}.$$

We let $D(z, u, w)$ be the ordinary generating function

$$D(z, u, w) = \sum_{n,k,\ell \geq 0} d_{n,k,\ell} z^n u^k w^\ell,$$

where $d_{n,k,\ell}$ is the number of planar maps with n edges, valency k on its root face and ℓ edges that represent double triangles, where edges on the root face are not considered. For the sake of brevity, we denote $D(z, u, w)$ by D and $D(z, 1, w)$ by $D(1)$. (The same *rule* will be applied to other generating functions.)

Clearly, the initial case corresponds to the generating function 1 and the bridgeable case to $zu^2 D^2$. The non-bridgeable case is split into two different classes: $\mathcal{D}_{\not\triangleright}$ denotes the class where the *second face* (the face on the right side of the root edge) has valency not equal to 3 and $\mathcal{D}_{\triangleright}$ denotes the class where the second face has valency 3. This means that we have $D = 1 + zu^2 D^2 + D_{\not\triangleright} + D_{\triangleright}$, where $D_{\not\triangleright}$ and $D_{\triangleright}$ are the corresponding generating functions of $\mathcal{D}_{\not\triangleright}$ and $\mathcal{D}_{\triangleright}$, respectively.

▶ **Lemma 2.** *The generating functions* $D = D(z, u, w)$, $D_{\not\triangleright} = D_{\not\triangleright}(z, u, w)$, *and* $D_{\triangleright} = D_{\triangleright}(z, u, w)$ *satisfy the following system of equations:*

$$
\begin{aligned}
D &= 1 + zu^2 D^2 + D_{\not\triangleright} + D_{\triangleright}, \\
D_{\not\triangleright} &= zu\frac{D(1) - uD}{1 - u} - zu^{-1}\left(D - 1 - u[u^1]D\right), \\
D_{\triangleright} &= zu^{-1}\left(D - 1 - u[u^1]D\right) \\
&\quad + (w - 1) \cdot \left[z^2 uD + (w + 1)\left(zu^{-1}D_{\triangleright} - z[u^1]D_{\triangleright}\right) - z^2 u(w - 1)DD_{\triangleright} \right. \\
&\quad \left. - (w - 1)\left(z^2\frac{D_{\triangleright}(1) - uD_{\triangleright}}{1 - u} - z^2 D_{\triangleright}(1) - z^2 u^{-2}\left(D_{\triangleright} - u[u^1]D_{\triangleright} - u^2[u^2]D_{\triangleright}\right)\right) \right].
\end{aligned}
\tag{2}
$$

▶ **Remark.** If $w = 1$ the system collapses to the well-known catalytic equation for the generating function $M(z, u) = D(z, u, 1)$ of planar maps:

$$M(z, u) = 1 + zu^2 M(z, u)^2 + zu\frac{M(z, 1) - uM(z, u)}{1 - u}. \tag{3}$$

Proof. We have already discussed the first equation of (2). Thus, we can concentrate on the non-bridgeable case. Here we relate the original map with the resulting map, where we have removed the root edge. Actually it is more transparent to consider the reverse process of adding a new root edge that *cuts across* the root face. This operation separates the root face into two faces. For instance, there are five possible situations of cutting across a root face of valency 4 as Figure 2 shows, and which have the following effect to the variable u:

$$u^4 \mapsto z(u^5 + u^4 + u^3 + u^2 + u).$$
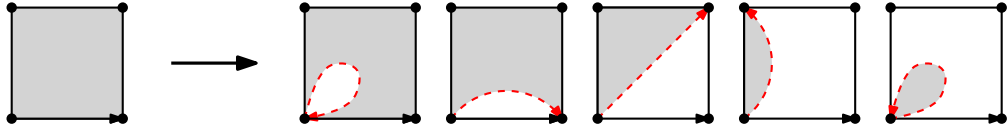
Figure 2 Cutting-across-process.
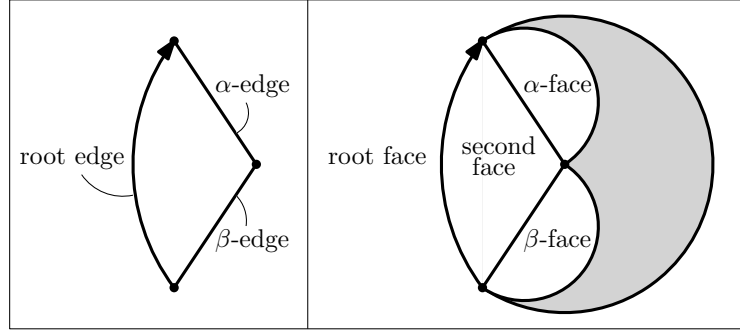
Figure 3 Definition of the α and β-edge and face.

When we consider $\mathcal{D}_{\not\triangleright}$, we have to discount the case where the second face has valency 3. In the cutting-across-process, we take out the situation that the new-appearing second face has valency 3. The corresponding effect with the root face of valency r is

$$u^r \mapsto z(u^{r+1} + u^r + \cdots + u^2 + u^1) + \begin{cases} -zu^{r-1} & \text{, if } r \geq 2 \\ 0 & \text{, if } r = 0 \text{ or } 1. \end{cases}$$

So the corresponding generating function of $\mathcal{D}_{\not\triangleright}$ is given by

$$D_{\not\triangleright} = zu\frac{D(1) - uD}{1 - u} - zu^{-1}\left(D - 1 - [u^1]D\right).$$

Next, we consider maps whose second face is of valency 3 and whose generating function is $D_{\triangleright}$. We introduce some notations. When the second face has valency 3, the edges following the root edge in clockwise order are called the α-*edge* and the β-*edge*. One side of the α-edge is the second face, we call the face on the other side the α-*face*. Similar to the α-face, the β-*face* is the face incident to the β-edge. Note that the α-face and the β-face might coincide (see Figure 3).

For describing the class $\mathcal{D}_{\triangleright}$, we consider four different cases: both the α-face and the β-edge are equal to the root face (denoted by $\mathcal{D}_{\triangleright}^{\alpha,\beta}$), only the α-face is equal to the root face (denoted by $\mathcal{D}_{\triangleright}^{\alpha}$), only the β-face is equal to the root face (denoted by $\mathcal{D}_{\triangleright}^{\beta}$) and neither the α-face nor the β-face is equal to the root face (denoted by $\mathcal{D}_{\trianglerighteq}$) (see Figure 4). Thus, we have $D_{\triangleright} = D_{\triangleright}^{\alpha,\beta} + D_{\triangleright}^{\alpha} + D_{\triangleright}^{\beta} + D_{\trianglerighteq}$.

The maps corresponding to the class $\mathcal{D}_{\triangleright}^{\alpha,\beta}$ can be divided into a triangle and three maps. Thus, we have $D_{\triangleright}^{\alpha,\beta} = z^3 u^3 D^3$.

The maps corresponding to the class $\mathcal{D}_{\triangleright}^{\alpha}$ and $\mathcal{D}_{\triangleright}^{\beta}$ can be divided into a map and a map stuck together with a triangle attached to an edge (see the left part of Figure 5). The structure of a map stuck together with a triangle attached to an edge has either the property that this edge corresponds to a double triangle or not (see the right of Figure 5).

If this edge is (resp. is not) a double triangle, we can think of it as adding two edges to a map which belong to $\mathcal{D}_{\triangleright}$ (resp. $\mathcal{D}_{\not\triangleright}$). The effect of these two additional edges is that the number of edges increased by 2 and the valency of the outside (root) face increased by 1. Hence, $D_{\triangleright}^{\alpha} = D_{\triangleright}^{\beta} = z^2 u \left(wD_{\triangleright} + D_{\not\triangleright}\right) D$.
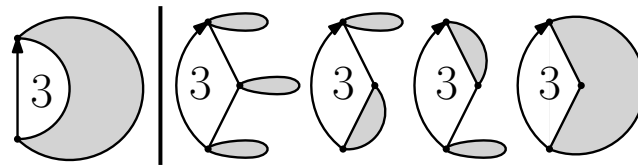
Figure 4 Four different cases of $\mathcal{D}_{\rhd}$.
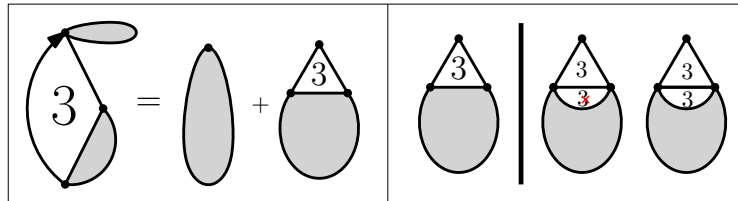
Figure 5 A map in $\mathcal{D}_{\rhd}^{\alpha}$ or $\mathcal{D}_{\rhd}^{\beta}$ can be divided into a map and a map stuck together with a triangle at an edge. This edge corresponds to a double triangle or not.

For the fourth class $\mathcal{D}_{\unrhd}$ we need to consider three different cases. The first one is when the α-edge is different from the β-edge but the α-face equals to the β-face (denoted by $\mathcal{D}_{\unrhd}^{\mathbb{U}}$). The second one is when the α-edge is different from the β-edge and the α-face is different from the β-face (denoted by $\mathcal{D}_{\unrhd}^{\varpi}$). The third one is when the α-edge equals to the β-edge. In this case, both the α-face and the β-face are equal to the second face (denoted by $\mathcal{D}_{\unrhd}^{\psi}$) (see Figure 6). By definition we have $D_{\unrhd} = D_{\unrhd}^{\mathbb{U}} + D_{\unrhd}^{\varpi} + D_{\unrhd}^{\psi}$.

When we deal with the maps in $\mathcal{D}_{\unrhd}^{\mathbb{U}}$, the α-face coincides with the β-face if both of them have valency 3, in particular, both the α-edge and the β-edge represent double triangles. Therefore, we have to take care of the valency of the α-face and of the β-face. For this purpose we consider the so-called "*border-(α,β)-path*", that starts from the α-edge, goes clockwise along the border of the α-face and finishes at the β-edge but dones not include the α-edge nor the β-edge. We distinguish between three different cases by considering the length of the border-(α,β)-path (denoted by $|(\alpha,\beta)|$): $|(\alpha,\beta)| = 0$, $|(\alpha,\beta)| = 1$ and $|(\alpha,\beta)| \geq 2$. The corresponding sets of maps are denoted by $\mathcal{D}_{\unrhd}^{\mathbb{U}_0}$, $\mathcal{D}_{\unrhd}^{\mathbb{U}_1}$, and $\mathcal{D}_{\unrhd}^{\mathbb{U}_{\geq 2}}$ respectively, see Figure 7. From the above relation, we have $D_{\unrhd}^{\mathbb{U}} = D_{\unrhd}^{\mathbb{U}_0} + D_{\unrhd}^{\mathbb{U}_1} + D_{\unrhd}^{\mathbb{U}_{\geq 2}}$ and (similar to the above considerations) they can be further decomposed which leads to the following relations:

$$
\begin{aligned}
D_{\unrhd}^{\mathbb{U}_0} &= z^3 u D \left[w^2 \left(w[u^1]D_{\rhd} + [u^1]D_{\not\rhd} \right) + \left(D(1) - [u^1]D \right) \right], \\
D_{\unrhd}^{\mathbb{U}_1} &= z^3 w^2 (w D_{\rhd} + D_{\not\rhd} + z u^2 D^2) + z^3 \left(D(1) - 1 \right) \left(D - 1 \right), \\
D_{\unrhd}^{\mathbb{U}_{\geq 2}} &= z^2 D(1) \left(z u \frac{D(1) - u D}{1 - u} - z(D - 1) - z u D \right),
\end{aligned}
\tag{4}
$$

The proof is given in the Appendix A.1.

Next, $\mathcal{D}_{\unrhd}^{\psi}$ is the class of maps that combines maps and an edge inside a loop. The edge inside the loop is a double triangle. Thus, we have $D_{\unrhd}^{\psi} = z^2 u w D$.

Finally, we discuss the class $\mathcal{D}_{\unrhd}^{\varpi}$. By distinguishing whether the α-face and the β-face have valency 3 we have to consider four different situations: neither the α-face nor the β-face has valency 3 (denoted by $\mathcal{D}_{\unrhd}^{\otimes}$), only the β-face has valency 3 (denoted by $\mathcal{D}_{\unrhd}^{\beta}$), only the α-face has valency 3 (denoted by $\mathcal{D}_{\unrhd}^{\alpha}$) and both the α-face and the β-face have valency 3 (denoted by $\mathcal{D}_{\unrhd}^{\alpha,\beta}$): $D_{\unrhd}^{\varpi} = D_{\unrhd}^{\otimes} + D_{\unrhd}^{\beta} + D_{\unrhd}^{\alpha} + D_{\unrhd}^{\alpha,\beta}$.
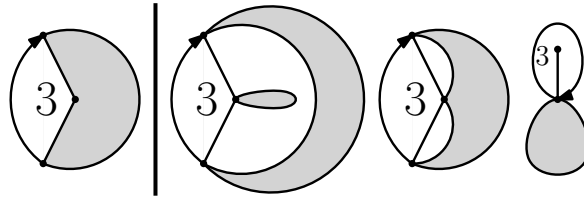
Figure 6 Three different cases of $\mathcal{D}_{\unrhd}$.
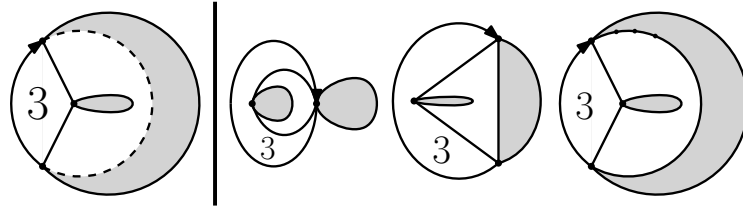
Figure 7 Three different cases of the length of the border-(α,β)-path of the maps in $\mathcal{D}_{\unrhd}^{\mathbb{U}}$.

When we study the class $\mathcal{D}_{\unrhd}^{\otimes}$, we need to build up maps, where the second face has valency 3 and neither the α-face nor the β-face has valency 3. We start with $\mathcal{D}_{\not\unrhd}$ and do the cutting across process that adds an edge starting from the end point of the root edge of the map. In cutting across process (see Figure 2) we always keep the second face valency different from 3 and the outside face valency greater than 1 (in order to make sure that the new α-edge and the new β-edge exist). In a second step we add an edge to complete the construction (see the left of Figure 8).

We have to be careful in the cutting across process. For example, if the root face valency equals r before we start the process, we have to avoid the case, where the root face valency would get $r - 1$ in cutting across process. This means that the cases $r = 0, 1, 2$ have to be considered separately. If $r = 0$ or $r = 1$ the root face valency $r - 1$ in cutting across process can not appear, and when $r = 2$ the resulting root face of valency $r - 1 = 1$ is also excluded. The effect on the variable u is therefore

$$u^r \mapsto z(u^{r+1} + u^r + \cdots + u^2 + u^1) - zu^1 + \begin{cases} -zu^{r-1} & \text{, if } r \geq 3 \\ 0 & \text{, if } r = 0 \text{ or } 1 \text{ or } 2. \end{cases}$$

After adding an edge in second step we obtain the following relations for the corresponding generating function: $\mathcal{D}_{\unrhd}^{\otimes}$ is

$$D_{\unrhd}^{\otimes} = zu^{-1}\left(zu\frac{D_{\not\unrhd}(1) - uD_{\not\unrhd}}{1-u} - zuD_{\not\unrhd}(1) - zu^{-1}\left(D_{\not\unrhd} - u[u^1]D_{\not\unrhd} - u^2[u^2]D_{\not\unrhd} \right) \right).$$

By using similar ideas (by using $\mathcal{D}_{\rhd}$ instead) and by observing that the new β-edge will be a double triangle (see the right of Figure 8) we obtain

$$D_{\unrhd}^{\beta} = zu^{-1}w\left(zu\frac{D_{\rhd}(1) - uD_{\rhd}}{1-u} - zuD_{\rhd}(1) - zu^{-1}\left(D_{\rhd} - u[u^1]D_{\rhd} - u^2[u^2]D_{\rhd} \right) \right).$$

By symmetry we have $D_{\unrhd}^{\alpha} = D_{\unrhd}^{\beta}$.

In order to describe the class $\mathcal{D}_{\unrhd}^{\alpha,\beta}$ we need to *adjust* both the root face and the face (we call this face *clockwise-face*) on the right of the clockwise-edge have valency 3 (see Figure 9), where the *clockwise-edge* is the edge in clockwise direction of the root edge on the outside face. Suppose that $\mathcal{D}_{\mathfrak{B}}$ is the class of maps, where both the root face and the clockwise-face have valency 3, we have $D_{\unrhd}^{\alpha,\beta} = zu^{-1}w^2 D_{\mathfrak{B}}$.
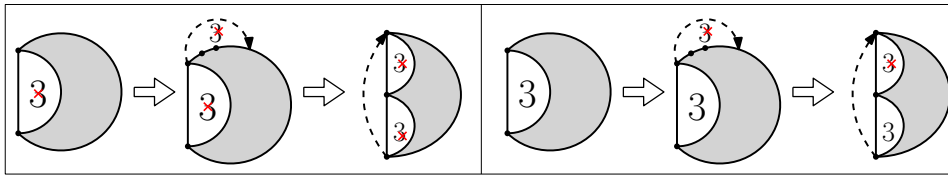
Figure 8 Construction of a map contained in $\mathcal{D}_{\unrhd}^{\otimes}$ and $\mathcal{D}_{\unrhd}^{\beta}$, respectively.
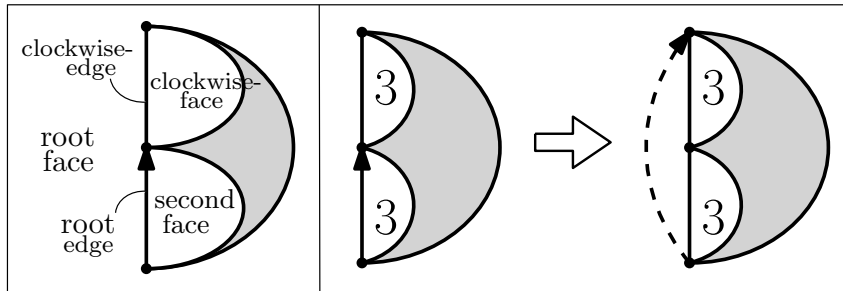
Figure 9 Definition of clockwise-edge and clockwise-face. Relation between $\mathcal{D}_{\mathfrak{B}}$ and $\mathcal{D}_{\unrhd}^{\alpha,\beta}$.

The class $\mathcal{D}_{\mathfrak{B}}$ is a subclass of $\mathcal{D}_{\rhd}$. Hence, we can get $D_{\mathfrak{B}}$ by eliminating some cases of $\mathcal{D}_{\rhd}$. When we consider the clockwise-edge and the clockwise-face of $\mathcal{D}_{\rhd}$, we have five different cases. The first three cases where the clockwise-edge is not a bridge, and first, where the clockwise-face has valency 3, second, where the clockwise-face has valency not equal to 3 and third, where the clockwise-face is equal to the second face. In the fourth case the clockwise-edge is a bridge and in the last the clockwise-edge does not exist (see Figure 10).

The first case of $\mathcal{D}_{\rhd}$ is precisely $\mathcal{D}_{\mathfrak{B}}$.

The second case of $\mathcal{D}_{\rhd}$ (clockwise-face has valency not equal to 3) corresponds precisely to the first step of the construction of $\mathcal{D}_{\unrhd}^{\beta}$ in Figure 8. Hence, the corresponding generating function is given by

$$zu\frac{D_{\rhd}(1) - uD_{\rhd}}{1 - u} - zuD_{\rhd}(1) - zu^{-1}\left(D_{\rhd} - u[u^1]D_{\rhd} - u^2[u^2]D_{\rhd}\right).$$

The only difference to $D_{\unrhd}^{\beta}$ is the factor $zu^{-1}w$.

In the third case of $\mathcal{D}_{\rhd}$ we have to consider several subcases that lead to the following generating function:

$$z^2u^2wD[u^1]D_{\rhd} + z^2u^2D[u^1]D_{\not\rhd} + z^2uwD_{\rhd} + z^2uD_{\not\rhd} + z^3u^3D^2.$$

In the fourth case of $\mathcal{D}_{\rhd}$ the second face has valency 3 and the clockwise-edge is a bridge. Thus, it corresponds to the generating function $zu^2DD_{\rhd}$.

Finally, in the last case the root face valency equals 1 and the second face has valency 3. Consequently its corresponding generating function is $u[u^1]D_{\rhd}$.

Summing up, the generating function of $\mathcal{D}_{\mathfrak{B}}$ is given by

$$D_{\mathfrak{B}} = D_{\rhd} - zu^2DD_{\rhd} - u[u^1]D_{\rhd}$$
$$- \left(zu\frac{D_{\rhd}(1) - uD_{\rhd}}{1 - u} - zuD_{\rhd}(1) - zu^{-1}\left(D_{\rhd} - u[u^1]D_{\rhd} - u^2[u^2]D_{\rhd}\right)\right)$$
$$- \left(z^2u^2D[u^1]D_{\not\rhd} + z^2u^2wD[u^1]D_{\rhd} + z^2uD_{\not\rhd} + z^2uwD_{\rhd} + z^3u^3D^2\right).$$

By collecting all these parts and by applying some simplifications (that are described in the Appendix A.2) we obtain the third equation of the system (2). ◀
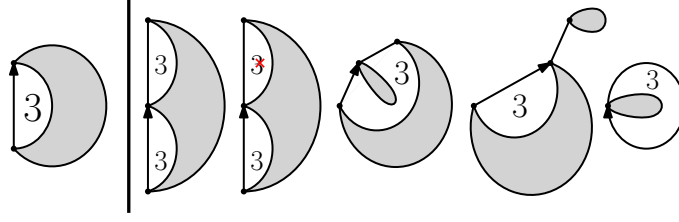
Figure 10 Five different cases of $\mathcal{D}_{\triangleright}$.

3 Asymptotic analysis

In order to analyze the system of equations (2) we apply a 2-step procedure that is in principle close to that of [7]. In the first step we eliminate the terms $[u^1]D$, $[u^1]D_{\triangleright}$, and $[u^2]D_{\triangleright}$ so that the system (2) is transferred into a catalytic system of equations that will be solved then in a second step.

▶ **Lemma 3.** *Suppose that* $D = D(z, u, w)$, $D_{\not\triangleright} = D_{\not\triangleright}(z, u, w)$, *and* $D_{\triangleright} = D_{\triangleright}(z, u, w)$ *are the solution functions of the system (2). Then there exist analytic functions* $K_{ij}(z, w, x_0, x_1, x_2)$ *(for* $|z| < \frac{1}{2}$, $|x_0| < 2$, $|x_1| < 2$, $|x_2| < 2$, *and* $|w - 1| < \eta$ *for some sufficiently small* $\eta > 0$*),* $i \in \{0, 1, 2\}$, $j \in \{1, 2\}$ *such that for* $j \in \{1, 2\}$

$$[u^j]D(z, u, w) = K_{0,j}\left(z, w, D(z, 1, w), D_{\not\triangleright}(z, 1, w), D_{\triangleright}(z, 1, w)\right),$$
$$[u^j]D_{\not\triangleright}(z, u, w) = K_{1,j}\left(z, w, D(z, 1, w), D_{\not\triangleright}(z, 1, w), D_{\triangleright}(z, 1, w)\right),$$
$$[u^j]D_{\triangleright}(z, u, w) = K_{2,j}\left(z, w, D(z, 1, w), D_{\not\triangleright}(z, 1, w), D_{\triangleright}(z, 1, w)\right).$$

Proof. We rewrite the system (2) into an equivalent one. We substitute in all instances $D_{\not\triangleright} = D_1 - zu^{-1}(D - 1 - u[u^1]D)$ and $D_{\triangleright} = D_2 + zu^{-1}(D - 1 - u[u^1]D)$ so that we obtain a system of the form

$$D = 1 + zu^2 D^2 + D_1 + D_2, \quad D_1 = zu\frac{D(1) - uD}{1 - u}, \quad D_2 = (w - 1)H, \tag{5}$$

where H is equal to

$$z^2 u D + (w + 1)\left(zu^{-1}\left(D_2 - u[u^1]D_2\right) + z^2 u^{-2}\left(D - 1 - u[u^1]D - u^2[u^2]D\right)\right)$$

$$+ (w - 1)\left(-z^2 u D D_2 - z^3 D(D - 1 - u[u^1]D) - z^2 u\frac{D_2(1) - D_2}{1 - u} - z^3\frac{uD(1) - D}{1 - u} - z^3\right.$$

$$\left. + z^2 u^{-2}\left(D_2 - u[u^1]D_2 - u^2[u^2]D_2\right) + z^3 u^{-3}\left(D - 1 - u[u^1]D - u^2[u^2]D - u^3[u^3]D\right)\right).$$

Next we consider the functions D, D_1, D_2 as power series in u:

$$D = 1 + \sum_{\ell \geq 1} d_\ell u^\ell, \quad D_1 = \sum_{\ell \geq 1} d_{1,\ell} u^\ell, \quad D_2 = \sum_{\ell \geq 1} d_{2,\ell} u^\ell,$$

and rewrite the system (5) into an infinite system of equations:

$$d_\ell = z\sum_{j=0}^{\ell-2} d_j d_{\ell-2-j} + zD(1) - z\sum_{j=0}^{\ell-2} d_j + (w - 1)[u^\ell]H,$$

$$d_{1,\ell} = zD(1) - z\sum_{j=0}^{\ell-2} d_j, \tag{6}$$

$$d_{2,\ell} = (w - 1)[u^\ell]H,$$

where $\ell \geq 1$, $d_0 := 1$, and $[u^\ell]H$ is equal to

$$z^2 d_{\ell-1} + (w+1)(z d_{2,\ell+1} + z^2 d_{\ell+2}) + (w-1)\left[-z^2 \sum_{i=0}^{\ell-2} d_i d_{2,\ell-1-i} - z^3 \sum_{i=0}^{\ell-2} d_i d_{\ell-i} \right.$$

$$\left. - z^2 \left(D_2(1) - \sum_{i=0}^{\ell-1} d_{2,i} \right) - z^3 \left(D(1) - \sum_{i=0}^{\ell} d_i \right) + z^2 d_{2,\ell+2} + z^3 d_{\ell+3} \right].$$

Note that we have not substituted $D(1)$, $D_1(1)$, and $D_2(1)$. In a final step we use the substitutions $y_{0,\ell} = d_\ell v^\ell$, $y_{1,\ell} = d_{1,\ell} v^\ell$, $y_{2,\ell} = d_{2,\ell} v^\ell$, $\ell = 1, 2, \ldots$ (and $y_{0,0} = 1$) for some parameter $v > 0$ to rewrite (6) to

$$y_{0,\ell} = zv^2 \sum_{j=0}^{\ell-2} y_{0,j} y_{0,\ell-2-j} + zD(1)v^\ell - zv^2 \sum_{j=0}^{\ell-2} y_{0,j} v^{\ell-2-j} + (w-1)H_\ell, \tag{7}$$

$$y_{1,\ell} = zD(1)v^\ell - zv^2 \sum_{j=0}^{\ell-2} y_{0,j} v^{\ell-2-j}, \qquad y_{2,\ell} = (w-1)H_\ell,$$

where

$$H_\ell = z^2 v y_{0,\ell-1} + (w+1)(zv^{-1} y_{2,\ell+1} + z^2 v^{-2} y_{0,\ell+2})$$

$$+ (w-1)\left[-z^2 v \sum_{i=0}^{\ell-2} y_{0,i} y_{2,\ell-1-i} - z^3 \sum_{i=0}^{\ell-2} y_{0,i} y_{0,\ell-i} + z^2 v^{-2} y_{2,\ell+2} + z^3 v^{-3} y_{0,\ell+3} \right.$$

$$\left. - z^2 \left(v^\ell D_2(1) - v \sum_{i=0}^{\ell-1} y_{2,i} v^{\ell-1-i} \right) - z^3 \left(v^\ell D(1) - \sum_{i=0}^{\ell} y_{0,i} v^{\ell-i} \right) \right].$$

Now we consider $D(1)$, $D_1(1)$, and $D_2(1)$ as new variables x_0, x_1, and x_2 and rewrite the system (7) into a new system

$$y_{0,\ell} = zv^2 \sum_{j=0}^{\ell-2} y_{0,j} y_{0,\ell-2-j} + zx_0 v^\ell - zv^2 \sum_{j=0}^{\ell-2} y_{0,j} v^{\ell-2-j} + (w-1)\tilde{H}_\ell, \tag{8}$$

$$y_{1,\ell} = zx_0 v^\ell - zv^2 \sum_{j=0}^{\ell-2} y_{0,j} v^{\ell-2-j}, \qquad y_{2,\ell} = (w-1)\tilde{H}_\ell,$$

where $\tilde{H}_\ell$ results from H_ℓ by this substitution. The solution functions $y_{i,\ell} = y_{i,\ell}(z, w, x_0, x_1, x_2)$ are now considered as functions in z, w, x_0, x_1, x_2 and in a next step we will show that these functions are actually analytic in these variables (in a certain range). Of course, if we have proved this assertion then we can obtain, for example,

$$d_\ell = d_\ell(z, w) = y_{0,\ell}(z, w, D(z, 1, w), D_1(z, 1, w), D_2(z, 1, w))v^{-\ell}$$

as an analytic function in $z, w, D(z, 1, w), D_1(z, 1, w), D_2(z, 1, w)$. This also proves the lemma after re-substituting $D_{\not\triangleright}$ and $D_\triangleright$ in terms of D, D_1, and D_2.

The idea of solving (8) is to consider it as a fixed point equation in a complete metric space and to solve it with the help of Banach's fixed point theorem. For this purpose we have to adjust the parameter $v > 0$ so that the right hand side of (8) is a contraction. More precisely we set $\mathbf{y}_0 = (y_{0,\ell})_{\ell \geq 1}$, $\mathbf{y}_1 = (y_{1,\ell})_{\ell \geq 1}$, $\mathbf{y}_2 = (y_{2,\ell})_{\ell \geq 1}$, and $\mathbf{y} = (\mathbf{y}_0, \mathbf{y}_1, \mathbf{y}_2)$ and consider the ℓ^1 norm $\|\mathbf{y}\|_1 = \|\mathbf{y}_0\|_1 + \|\mathbf{y}_1\|_1 + \|\mathbf{y}_2\|_1$, where

$$\|\mathbf{y}_j\|_1 = \sum_{\ell \geq 1} |y_{j,\ell}|, \qquad j \in \{0, 1, 2\}.$$

Furthermore we define the mapping $\mathbf{T} : \ell^1(\mathbb{C})^3 \to \ell^1(\mathbb{C})^3$ by $\mathbf{T}(\mathbf{y}) = (\mathbf{T}_0(\mathbf{y}), \mathbf{T}_1(\mathbf{y}), \mathbf{T}_2(\mathbf{y}))$, where

$$\mathbf{T}_0(\mathbf{y}) = \left(zv^2 \sum_{j=0}^{\ell-2} y_{0,j} y_{0,\ell-2-j} + zx_0 v^\ell - zv^2 \sum_{j=0}^{\ell-2} y_{0,j} v^{\ell-2-j} + (w-1)\tilde{H}_\ell \right)_{\ell \geq 1},$$

$$\mathbf{T}_1(\mathbf{y}) = \left(zx_0 v^\ell - zv^2 \sum_{j=0}^{\ell-2} y_{0,j} v^{\ell-2-j} \right)_{\ell \geq 1}, \qquad \mathbf{T}_2(\mathbf{y}) = \left((w-1)\tilde{H}_\ell \right)_{\ell \geq 1},$$

where z, w are considered as complex parameters and $v > 0$ will be chosen in a proper way. Clearly, a fixed point of $\mathbf{T}$ is a solution of (8).

By definition it follows that

$$\|\mathbf{T}_0(\mathbf{y})\|_1 \leq v^2 |z| \left(1 + \|\mathbf{y}_0\|_1\right)^2 + \frac{v}{1-v} |zx_0| + \frac{v^2}{1-v} |z| \left(1 + \|\mathbf{y}_0\|_1\right)$$

$$+ \frac{|w-1|}{v^2} P_0 \left(|z|, |x_0|, |x_1|, |x_2|, \|\mathbf{y}_0\|_1, \|\mathbf{y}_1\|_1, \|\mathbf{y}_2\|_1, v, \frac{1}{1-v} \right),$$

$$\|\mathbf{T}_1(\mathbf{y})\|_1 \leq \frac{v}{1-v} |zx_0| + \frac{v^2}{1-v} |z| \left(1 + \|\mathbf{y}_0\|_1\right),$$

$$\|\mathbf{T}_2(\mathbf{y})\|_1 \leq \frac{|w-1|}{v^2} P_0 \left(|z|, |x_0|, |x_1|, |x_2|, \|\mathbf{y}_0\|_1, \|\mathbf{y}_1\|_1, \|\mathbf{y}_2\|_1, v, \frac{1}{1-v} \right),$$

where P_0 is some polynomial with non-negative coefficients. Similarly we get

$$\|\mathbf{T}_0(\mathbf{y}) - \mathbf{T}_0(\mathbf{z})\|_1 \leq \left(v^2 |z| \left(2 + \|\mathbf{y}_0\|_1 + \|\mathbf{z}_0\|_1\right) + \frac{v^2 |z|}{1-v} \right) \|\mathbf{y}_0 - \mathbf{z}_0\|_1$$

$$+ \frac{|w-1|}{v^2} \tilde{P}_0 \left(|z|, |x_0|, |x_1|, |x_2|, \|\mathbf{y}_0\|_1, \|\mathbf{y}_1\|_1, \|\mathbf{y}_2\|_1, v, \frac{1}{1-v} \right) \|\mathbf{y} - \mathbf{z}\|_1,$$

$$\|\mathbf{T}_1(\mathbf{y}) - \mathbf{T}_1(\mathbf{z})\|_1 \leq \frac{v^2 |z|}{1-v} \|\mathbf{y}_0 - \mathbf{z}_0\|_1,$$

$$\|\mathbf{T}_2(\mathbf{y}) - \mathbf{T}_2(\mathbf{z})\|_1 \leq$$

$$\frac{|w-1|}{v^2} \tilde{P}_0 \left(|z|, |x_0|, |x_1|, |x_2|, \|\mathbf{y}_0\|_1, \|\mathbf{y}_1\|_1, \|\mathbf{y}_2\|_1, v, \frac{1}{1-v} \right) \|\mathbf{y} - \mathbf{z}\|_1,$$

where $\tilde{P}_0$ is another polynomial with non-negative coefficients.

Thus, given upper bounds Z, X_0, X_1, X_2, and Y for $|z|$, $|x_0|$, $|x_1|$, $|x_2|$, and $\|\mathbf{y}\|_1$ it is easy to choose $v > 0$ and $\eta > 0$ such that for $|w - 1| \leq \eta$ the mapping $\mathbf{T}$ maps the set $\{\mathbf{y} \in \ell^1(\mathbb{C})^3 : \|\mathbf{y}\|_1 \leq Y\}$ into itself and is a contraction, too. This shows that (8) has a unique solution that can be obtained as the uniform limit of the iterations $\mathbf{T}^k(\mathbf{0})$. By definition it is clear that all components of $\mathbf{T}^k(\mathbf{0})$ are analytic functions in z, w, x_0, x_1, x_2. Hence, the limits are analytic, too. This completes the proof of the lemma. ◀

We now go back to the original system (2) and substitute $[u^1]D$, $[u^1]D_\rhd$, and $[u^2]D_\rhd$ by the analytic functions K_{ij} given by Lemma 3 so that it can be rewritten as

$$D = 1 + zu^2 D^2 + D_{\not\rhd} + D_\rhd,$$

$$D_{\not\rhd} = Q_1(z, u, w, D, D(1), D_{\not\rhd}, D_{\not\rhd}(1), D_\rhd, D_\rhd(1)),$$

$$D_\rhd = Q_2(z, u, w, D, D(1), D_{\not\rhd}, D_{\not\rhd}(1), D_\rhd, D_\rhd(1))$$

with proper functions Q_1, Q_2. This is a catalytic system of three equations. In order to make our analysis a little bit easier we eliminate D and $D(1)$ by using the first equation. By substituting D (and similarly $D(1)$) by

$$D = \frac{1 - \sqrt{1 - 4zu^2(1 + D_{\not\triangleright} + D_{\triangleright})}}{2zu^2}$$

in the second and the third equation we finally obtain a system of two equations that we represent in the form

$$P_1(z, u, w, D_{\not\triangleright}, D_{\not\triangleright}(1), D_{\triangleright}, D_{\triangleright}(1)) = 0, \quad P_2(z, u, w, D_{\not\triangleright}, D_{\not\triangleright}(1), D_{\triangleright}, D_{\triangleright}(1)) = 0$$

for proper functions P_1, P_2 (that are by the way non-linear in $D_{\not\triangleright}, D_{\not\triangleright}(1), D_{\triangleright}, D_{\triangleright}(1)$).

We recall now a method by Bousquet-Mélou and Jehanne [2] on catalytic equations of the form

$$P(z, u, M(z, u), M_1(z)) = 0,$$

where $M_1(z)$ is usually $M(z, 1)$ or $M(z, 0)$ and $P = P(z, u, x_0, x_1)$ is usually a polynomial (however, the method also works with proper regularity conditions for P). The first step is to find functions $u(z)$, $y(z)$, and $f(z)$ that satisfy the system of equations

$$\begin{aligned} P(z, u(z), y(z), f(z)) &= 0, \\ P_u(z, u(z), y(z), f(z)) &= 0, \\ P_{x_0}(z, u(z), y(z), f(z)) &= 0, \end{aligned} \tag{9}$$

where P_u and P_{x_0} denote the partial derivatives $\frac{\partial P}{\partial u}$ and $\frac{\partial P}{\partial x_0}$, respectively. Then we can set $M_1(z) = f(z)$ and can recover $M(z, u)$ – if necessary – from the equation

$$P(z, u, M(z, u), f(z)) = 0. \tag{10}$$

This method generalizes the classical *Quadratic Method* and can be extended in various ways. It is also possible to guarantee unique power series solutions etc., for details we refer to [2].

We emphasize here some further extensions. First we can directly add a parameter w or several parameters $\mathbf{w} = (w_1, \ldots, w_k)$ into the equation without any change of the method. From $P(z, u, \mathbf{w}, M(z, u, \mathbf{w}), M_1(z, \mathbf{w})) = 0$ we, thus, obtain the solutions $M_1(z, \mathbf{w})$ and $M(z, u, \mathbf{w})$.

It was shown in [7] and [6] that the solution function $M_1(z)$ of a catalytic equation (10) that is singular at $z = \rho$ has usually a singularity of the form

$$M_1(z) = g(z) + h(z)\left(1 - \frac{z}{\rho}\right)^{3/2}, \tag{11}$$

where $g(z)$ and $h(z)$ are analytic at $z = \rho$. This is in particular true for the generating function $M(z, 1)$ that counts planar maps and is the solution of the catalytic equation (3):

$$M(z, 1) = \frac{18z - 1 + (1 - 12z)^{3/2}}{54z^2}.$$

Here $\rho = 1/12$ is the radius of convergence of $M(z, 1)$. Since $M(z, 1) = D(z, 1, 1)$ it also follows that $D(z, 1, 1)$ and consequently the functions $D_{\not\triangleright}(z, 1, 1)$ and $D_{\triangleright}(z, 1, 1)$ have the same kind of singularity at $z = 1/12$. What we show next (and which is actually the main

property that will be used to prove the central limit theorem) is that we have the same kind of singular behavior if we add some parameters. In particular we will show that $D(z, 1, w)$ can be represented as

$$D(z, 1, w) = g_D(z, w) + h_D(z, w) \left(1 - \frac{z}{\rho(w)}\right)^{3/2}, \tag{12}$$

where g_D, h_D and ρ are analytic at $z = 1/12$ and $w = 1$.

We will first consider one catalytic equation and will then generalize it to a system.

▶ **Lemma 4.** *Suppose that $M(z, u, \mathbf{w})$ and $M_1(z, \mathbf{w})$ are the solutions of the catalytic equation $P(z, u, \mathbf{w}, M(z, u, \mathbf{w}), M_1(z, \mathbf{w})) = 0$, where the function $P(z, u, \mathbf{w}, x_0, x_1)$ is analytic and $M_1(z, \mathbf{1})$ has a singularity at $z = \rho_0$ of form (11) with $g(\rho_0) \neq 0$, $h(\rho_0) \neq 0$ such that for $z = \rho_0$, $u = u_0$, $x_0 = M(\rho_0, u_0, \mathbf{1})$, $x_1 = M_1(\rho_0, \mathbf{1})$, and $\mathbf{w} = \mathbf{1}$ we have[3]*

$$P = 0, \quad P_u = 0, \quad P_{x_0} = 0, \quad P_{x_1} \neq 0, \quad P_{x_0 x_0} P_{uu} = P_{x_0 u}^2.$$

Furthermore, let $z = \rho(\mathbf{w})$, $u = u_0(\mathbf{w})$, $x_0 = x_0(\mathbf{w})$, $x_1 = x_1(\mathbf{w})$ for $\mathbf{w}$ close to $\mathbf{1}$ be defined by $\rho(\mathbf{1}) = \rho_0$, $u_0(\mathbf{1}) = u_0$, $x_0(\mathbf{1}) = M(\rho_0, u_0, \mathbf{1})$, $x_1(\mathbf{1}) = M_1(\rho_0, \mathbf{1})$ and by the system

$$P = 0, \quad P_u = 0, \quad P_{x_0} = 0, \quad P_{x_0 x_0} P_{uu} = P_{x_0 u}^2.$$

Then for $\mathbf{w}$ close to $\mathbf{1}$ the function $M_1(z, \mathbf{w})$ has a local singular representation of the form

$$M_1(z, \mathbf{w}) = \bar{g}(z, \mathbf{w}) + \bar{h}(z, \mathbf{w}) \left(1 - \frac{z}{\rho(\mathbf{w})}\right)^{3/2}, \tag{13}$$

where $\bar{g}(z, \mathbf{w})$, $\bar{h}(z, \mathbf{w})$ are analytic at $z = \rho_0$ and $\mathbf{w} = \mathbf{1}$ and satisfy $\bar{g}(\rho_0, \mathbf{1}) = g(\rho_0) \neq 0$, $\bar{h}(\rho_0, \mathbf{1}) = h(\rho_0) \neq 0$.

The **Proof** is an adaption of the methods of [7]. The essential step is to represent (with the help of the Weierstrass preparation theorem) the function P locally around $z = \rho_0$, $u = u_0$, $x_0 = M(\rho_0, u_0, \mathbf{1})$, $x_1 = M_1(\rho_0, \mathbf{1})$, and $\mathbf{w} = \mathbf{1}$ by

$$P(z, u, \mathbf{w}, x_0, x_1) = K(z, u, \mathbf{w}, x_0, x_1) \left((x_0 - G(z, u, \mathbf{w}, x_1))^2 - H(z, u, \mathbf{w}, x_1)\right),$$

where all appearing functions are analytic and we have $K(\rho_0, u_0, \mathbf{1}, M(\rho_0, u_0, \mathbf{1}), M_1(\rho_0, \mathbf{1})) \neq 0$, $G(\rho_0, u_0, \mathbf{1}, M_1(\rho_0, \mathbf{1})) = M(\rho_0, u_0, \mathbf{1})$ and $H(\rho_0, u_0, \mathbf{1}, M_1(\rho_0, \mathbf{1})) = 0$. The system (9) translates into a smaller system of the form $H(z, \mathbf{w}, u(z, \mathbf{w}), f(z, \mathbf{w})) = 0$, $H_u(z, \mathbf{w}, u(z, \mathbf{w}), f(z, \mathbf{w})) = 0$ which is suitable to extract the singular behavior of the form (13). In particular the condition $P_{x_0 x_0} P_{uu} = P_{x_0 u}^2$ is equivalent to $H_{uu} = 0$. Now we proceed as in [7], observe the singular expansion for $M_1(z, \mathbf{w})$ of the form (13) and by comparing it with (11) we also get the properties $\bar{g}(\rho_0, \mathbf{1}) = g(\rho_0) \neq 0$, $\bar{h}(\rho_0, \mathbf{1}) = h(\rho_0) \neq 0$.

In the case of a system of two catalytic equations $P_1 = 0$, $P_2 = 0$ (in unknown functions $M(z, u, \mathbf{w})$, $M_1(z, \mathbf{w})$, $N(z, u, \mathbf{w})$, $N_1(z, \mathbf{w})$) we apply an elimination procedure to reduce it to a single catalytic equation so that Lemma 4 can be applied. We consider first the second equation and replace $M(z, u, \mathbf{w})$, $M_1(z, \mathbf{w})$ by two new variables v_0, v_1:

$$P_2(z, u, \mathbf{w}, v_0, v_1, N, N_1) = 0$$

---

[3] The notation P_x denotes the partial derivative with respect to x – and similarly for partial derivatives with respect to other variables or for higher order derivatives.

and solve this catalytic equation in order to obtain solution functions $N = \overline{N}(z, u, \mathbf{w}, v_0, v_1)$ and $N_1 = \overline{N}_1(z, \mathbf{w}, v_0, v_1)$. Then we substitute these solutions into the first equation and obtain a single catalytic equation for $M = M(z, u, \mathbf{w})$, $M_1 = M_1(z, \mathbf{w})$:

$$P_1(z, u, \mathbf{w}, M, M_1, \overline{N}(z, u, \mathbf{w}, M, M_1), \overline{N}_1(z, \mathbf{w}, M, M_1)) = 0.$$

Finally we apply Lemma 4 and obtain the proposed singular representation. The only thing that has to be checked is that $P_{2,NN}P_{2,uu} \neq P_{2,Nu}^2$ and $P_{2,N_1} \neq 0$ so that the functions $N = \overline{N}(z, u, \mathbf{w}, v_0, v_1)$ and $N_1 = \overline{N}_1(z, \mathbf{w}, v_0, v_1)$ are analytic in the region of interest. In our special situation this is easy to check. With this method we obtain singular representations for $D_{\not\triangleright}(z, 1, w)$ and $D_{\triangleright}(z, 1, w)$ and consequently (12) for $D(z, 1, w)$.

The **Proof of Theorem 1** is now almost immediate. Let Y_n denote the number of edges in a random planar map with n edges that represent double triangles but are not on the root face. Then we have

$$D(z, 1, w) = \sum_{n \geq 0} M_n \mathbb{E}[w^{Y_n}] z^n,$$

where $M_n = [z^n]M(z, 1)$ denotes the number of planar maps with n edges. By a direct application of [5, Theorem 2.35] it follows that Y_n satisfies a central limit theorem of the form (1) with expected value and variance asymptotically proportional to n. The only difference between X_n and Y_n is the number of edges on the root face that represent a double triangle. However, if X_n and Y_n are different then the root face has valency 3 which means that the difference between X_n and Y_n is at most 3. Hence, the central limit theorem (as well as asymptotics for expected value and variance) of Y_n transfers directly into a corresponding central limit theorem for X_n which completes the proof of Theorem 1.

References

1 O. Angel and O. Schramm. Uniform infinite planar triangulations. *Commun. Math. Phys.*, 241(2-3):191–213, 2003. `doi:10.1007/s00220-003-0932-3`.

2 M. Bousquet-Mélou and A. Jehanne. Polynomial equations with one catalytic variable, algebraic series and map enumeration. *J. Comb. Theory, Ser. B*, 96(5):623–672, 2006. `doi:10.1016/j.jctb.2005.12.003`.

3 W.G. Brown and W.T. Tutte. On the enumeration of rooted non-separable planar maps. *Can. J. Math.*, 16:572–577, 1964. `doi:10.4153/CJM-1964-058-7`.

4 G. Collet, M. Drmota, and L. Klausner. Limit Laws of Planar Maps with Prescribed Vertex Degrees. *Combinatorics Probability Computing*, to appear.

5 M. Drmota. *Random trees. An interplay between combinatorics and probability.* Wien: Springer, 2009. `doi:10.1007/978-3-211-75357-6`.

6 M. Drmota and M. Noy. Universal exponents and tail estimates in the enumeration of planar maps. In *Extended abstracts of the sixth European conference on combinatorics, graph theory and applications, EuroComb 2011, Budapest, Hungary, August 29 – September 2, 2011*, pages 309–317. Amsterdam: Elsevier, 2011.

7 M. Drmota and K. Panagiotou. A central limit theorem for the number of degree-k vertices in random maps. *Algorithmica*, 66(4):741–761, 2013. `doi:10.1007/s00453-013-9751-x`.

8 M. Drmota and B. Stufler. Benjamini-schramm convergence of random planar maps. *arXiv preprint arXiv:1801.10007*, 2018.

9 Z. Gao and N. C. Wormald. Asymptotic normality determined by high moments, and submap counts of random maps. *Probab. Theory Relat. Fields*, 130(3):368–376, 2004. `doi:10.1007/s00440-004-0356-9`.

10   M. Krikun. Local structure of random quadrangulations. *arXiv preprint math/0512304*, 2005.

11   R. Stephenson. Local convergence of large critical multi-type Galton-Watson trees and applications to random maps. *J. Theor. Probab.*, 31(1):159–205, 2018. `doi:10.1007/s10959-016-0707-3`.

12   W. T. Tutte. A census of planar maps. *Can. J. Math.*, 15:249–271, 1963. `doi:10.4153/CJM-1963-029-x`.

A   Appendix

A.1   Proof of the relations (4)

In the first part of the Appendix, we present a proper decomposition of the sets $\mathcal{D}_{\triangleright}^{\mathbb{U}_0}$, $\mathcal{D}_{\triangleright}^{\mathbb{U}_1}$, and $\mathcal{D}_{\triangleright}^{\mathbb{U}_{\geq 2}}$ that translate into the system (4).

In order to represent $\mathcal{D}_{\triangleright}^{\mathbb{U}_0}$ (and consequently the generating function of $\mathcal{D}_{\triangleright}^{\mathbb{U}_0}$) which corresponds to the case $|(\alpha, \beta)| = 0$, the main argument will focus on the valency of the α-face (that equals to that of the β-face) which depends on the outside (root) face valency of the map between (or inside) the α-edge and the β-edge. If this map has root face valency 1, then the α-face has valency 3 which means the α-edge and the β-edge are both double triangles. Moreover, in case this map has root face valency 1, if this map belongs to $\mathcal{D}_{\triangleright}$ (the second face has valency 3), then the root edge of this map will become a double triangle after putting this map into the chink between the α-edge and the β-edge and vice versa. Therefore, we have $z^3 u D w^2 \left(w[u^1]D_{\triangleright} + [u^1]D_{\not\triangleright} \right)$. Contrarily, if this map has root face valency not equal to 1, then the valency of the α-face is not equal to 3, it corresponds to $z^3 u D \left(D(1) - [u^1]D \right)$. Thus, the corresponding generating function is $D_{\triangleright}^{\mathbb{U}_0} = z^3 u D \left[w^2 \left(w[u^1]D_{\triangleright} + [u^1]D_{\not\triangleright} \right) + \left(D(1) - [u^1]D \right) \right].$

If $|(\alpha, \beta)| = 1$ which corresponds to the class $\mathcal{D}_{\triangleright}^{\mathbb{U}_1}$ the border-(α,β)-path is just an edge and the valency of the α-face (and of the β-face) is three (because of the α-edge, the β-edge and the border-(α,β)-path) plus the outside (root) face valency of the map inside this triangle.

If the map inside the triangle has no edge (which means the corresponding generating function of the map is 1), then the α-face has valency 3 which means that both the α-edge and the β-edge represent double triangles. And whether the edge that equals to the border-(α,β)-path corresponds to a double triangle or not depends on the other incident face of this edge. The face on the other side may or may not have valency 3 and also may equal to the outside face (see the above case of Figure 11). Hence this part corresponds to $z^3 w^2 (w D_{\triangleright} + D_{\not\triangleright} + z u^2 D^2)$. If the map (inside the triangle) has some edges (corresponding to the generating function $D(1) - 1$), then the valency of α-face is not equal to 3 which means that neither the α-edge nor the β-edge correspond to a double triangle, and the edge that equals to the border-(α,β)-path must not correspond to a double triangle. We also have to distinguish between three different coases fo the other incident face (see the below case of Figure 11). This part corresponds to $z^3 \left(D(1) - 1 \right) \left(D_{\triangleright} + D_{\not\triangleright} + z u^2 D^2 \right)$ which can be simplified to $z^3 \left(D(1) - 1 \right) \left(D - 1 \right)$ by the first equation of (2). Summing up we get the corresponding generating function of $\mathcal{D}_{\triangleright}^{\mathbb{U}_1}$ as follows:

$$D_{\triangleright}^{\mathbb{U}_1} = z^3 w^2 (w D_{\triangleright} + D_{\not\triangleright} + z u^2 D^2) + z^3 \left(D(1) - 1 \right) \left(D - 1 \right).$$

Finally $\mathcal{D}_{\triangleright}^{\mathbb{U}_{\geq 2}}$ is easier to describe, since the α-face (that is equal to the β-face) has valency not equal to 3. So we do not have to care about whether the α-edge and the β-edge are double triangles. We can directly decompose the map into two parts: one is a map with second face valency greater than 3 (the length of the border-(α,β)-path greater than 2 and plus the root edge), the other one is a map with plus edges (see Figure 12).
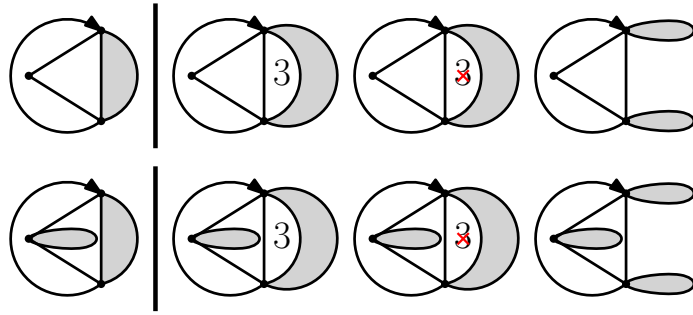
■ **Figure 11** Decomposition of $\mathcal{D}_{\rhd}^{\mathbb{U}^1}$: In the first (upper) case the the map (inside the triangle) has no edge, whereas in the second (below) case this map is non-trivial. In both case we have the right side face of the border-(α,β)-path is different to the root face and its valency is either equal to 3 or not, or it equals the root face.
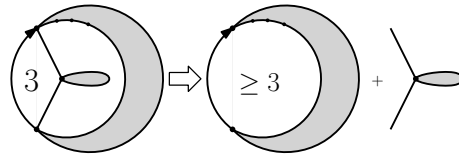
■ **Figure 12** Decompose a map that belongs to $\mathcal{D}_{\rhd}^{\mathbb{U}^{\geq 2}}$ into two parts.

The first map class can be counted with the help of a cutting across process (see Figure 2) where we have take out the situation where the new-appearing second face has valency 1 or 2. The corresponding effect to u^r is

$$u^r \mapsto z(u^{r+1} + u^r + \cdots + u^2 + u^1) - z(u^{r+1} + u^r) + \begin{cases} 0 & \text{, if } r \geq 1 \\ zu^0 & \text{, if } r = 0 \end{cases}$$

which leads to $zu\frac{D(1)-uD}{1-u} - z(D-1) - zuD$. After combining this with a map plus two edges (which is counted by $z^2 D(1)$) we have,

$$D_{\rhd}^{\mathbb{U}^{\geq 2}} = z^2 D(1) \left(zu\frac{D(1)-uD}{1-u} - z(D-1) - zuD \right)$$

which completes the proof of (4).

A.2 Simplification of the representation of $D_{\rhd}$

In the second part of the Appendix we prove that $D_{\rhd}$ can be simplified into the form that is given in (2).

After collecting all parts of $D_\triangleright$ that are described in the Proof of Lemma 2 we obtain

$$
\begin{aligned}
D_\triangleright ={}& z^3 u^3 D^3 + 2z^2 uD\left(wD_\triangleright + D_{\not\triangleright}\right) \\
&+ z^3 uD\left[w^2\left(w[u^1]D_\triangleright + [u^1]D_{\not\triangleright}\right) + \left(D(1) - [u^1]D\right)\right] \\
&+ z^3 w^2(wD_\triangleright + D_{\not\triangleright} + zu^2 D^2) + z^3\left(D(1) - 1\right)(D-1) \\
&+ z^2 D(1)\left(zu\frac{D(1) - uD}{1-u} - z(D-1) - zuD\right) + z^2 uwD \\
&+ zu^{-1}\left(zu\frac{D_{\not\triangleright}(1) - uD_{\not\triangleright}}{1-u} - zuD_{\not\triangleright}(1) - zu^{-1}\left(D_{\not\triangleright} - u[u^1]D_{\not\triangleright} - u^2[u^2]D_{\not\triangleright}\right)\right) \\
&+ 2zu^{-1}w\left(zu\frac{D_\triangleright(1) - uD_\triangleright}{1-u} - zuD_\triangleright(1) - zu^{-1}\left(D_\triangleright - u[u^1]D_\triangleright - u^2[u^2]D_\triangleright\right)\right) \\
&+ zu^{-1}w^2 D_\triangleright - z^2 uw^2 DD_\triangleright - zw^2[u^1]D_\triangleright \\
&- zu^{-1}w^2\left(zu\frac{D_\triangleright(1) - uD_\triangleright}{1-u} - zuD_\triangleright(1) - zu^{-1}\left(D_\triangleright - u[u^1]D_\triangleright - u^2[u^2]D_\triangleright\right)\right) \\
&- zw^2\left(z^2 uD[u^1]D_{\not\triangleright} + z^2 uwD[u^1]D_\triangleright + z^2 D_{\not\triangleright} + z^2 wD_\triangleright + z^3 u^2 D^2\right).
\end{aligned}
$$

We use the first two terms of the 2nd line and the first three terms of the 3rd line to cancel the last line. We also cancel the third term of the 2nd line and the third term of the 4th line. Moreover, we cancel part of the last term of the 3rd line and the second term of the 4th line.

$$
\begin{aligned}
D_\triangleright ={}& z^3 u^3 D^3 + z^2 uD\left(wD_\triangleright + D_{\not\triangleright}\right) + z^2 uD\left(wD_\triangleright + D_{\not\triangleright}\right) \\
&- z^3 uD[u^1]D - z^3(D-1) + z^2 D(1)\left(zu\frac{D(1) - uD}{1-u}\right) + z^2 uwD \\
&+ zu^{-1}\left(zu\frac{D_{\not\triangleright}(1) - uD_{\not\triangleright}}{1-u} - zuD_{\not\triangleright}(1) - zu^{-1}\left(D_{\not\triangleright} - u[u^1]D_{\not\triangleright} - u^2[u^2]D_{\not\triangleright}\right)\right) \\
&+ 2zu^{-1}w\left(zu\frac{D_\triangleright(1) - uD_\triangleright}{1-u} - zuD_\triangleright(1) - zu^{-1}\left(D_\triangleright - u[u^1]D_\triangleright - u^2[u^2]D_\triangleright\right)\right) \\
&+ zu^{-1}w^2 D_\triangleright - z^2 uw^2 DD_\triangleright - zw^2[u^1]D_\triangleright \\
&- zu^{-1}w^2\left(zu\frac{D_\triangleright(1) - uD_\triangleright}{1-u} - zuD_\triangleright(1) - zu^{-1}\left(D_\triangleright - u[u^1]D_\triangleright - u^2[u^2]D_\triangleright\right)\right).
\end{aligned}
$$

We now rewrite $D_\triangleright$ according to the appearing power of w and separate as follows:

$$
\begin{aligned}
D_\triangleright &= A_0 + wA_1 + w^2 A_2 \\
&= A_0 + A_1 + A_2 + (w-1)A_1 + (w^2 - 1)A_2 \\
&= (A_0 + A_1 + A_2) + (w-1)\left(A_1 + (w+1)A_2\right)
\end{aligned}
$$

where A_0, A_1, A_2 are explicit functions in $z, u, D, D_{\not\triangleright}, D_\triangleright, [u^1]D, [u^2]D, [u^1]D_{\not\triangleright}, [u^2]D_{\not\triangleright}, [u^1]D_\triangleright, [u^2]D_\triangleright$.

In order to show that this representation can be simplified to the form in (2) we first have to show that $A_0 + A_1 + A_2 = zu^{-1}\left(D - 1 - u[u^1]D\right)$. By summing up the expressions

of $A_0 + A_1 + A_2$ (and cancelling already two terms) we get

$$z^3 u^3 D^3 + z^2 uD \left(D_\triangleright + D_{\not\triangleright} \right) + z^2 uD \left(D_\triangleright + D_{\not\triangleright} \right)$$

$$- z^3 uD[u^1]D - z^3(D-1) + z^2 D(1) \left(zu \frac{D(1) - uD}{1 - u} \right) + z^2 uD$$

$$+ zu^{-1} \left(zu \frac{D_{\not\triangleright}(1) - uD_{\not\triangleright}}{1 - u} - zuD_{\not\triangleright}(1) - zu^{-1} \left(D_{\not\triangleright} - u[u^1]D_{\not\triangleright} - u^2[u^2]D_{\not\triangleright} \right) \right)$$

$$+ zu^{-1} \left(zu \frac{D_\triangleright(1) - uD_\triangleright}{1 - u} - zuD_\triangleright(1) - zu^{-1} \left(D_\triangleright - u[u^1]D_\triangleright - u^2[u^2]D_\triangleright \right) \right)$$

$$+ zu^{-1}D_\triangleright - z^2 uDD_\triangleright - z[u^1]D_\triangleright.$$

Now by using the relation $D_\triangleright + D_{\not\triangleright} = D - 1 - zu^2 D^2$, we can deduce two properties:

$$F_1: \quad [u^1]D_\triangleright + [u^1]D_{\not\triangleright} = [u^1](D - 1 - zu^2 D^2) = [u^1]D,$$
$$F_2: \quad [u^2]D_\triangleright + [u^2]D_{\not\triangleright} = [u^2](D - 1 - zu^2 D^2) = [u^2]D - z.$$

We combine the 3$^{\text{rd}}$ and 4$^{\text{th}}$ line by applying F_1 and F_2 and use the last term of it to cancel z^3 in the 2$^{\text{nd}}$ line. Then, applying the relation $D = 1 + zu^2 D^2 + D_{\not\triangleright} + D_\triangleright$ in the 1$^{\text{st}}$ line, we obtain

$$z^2 uD(D-1) + z^2 uD \left(D_\triangleright + D_{\not\triangleright} \right)$$

$$- z^3 uD[u^1]D - z^3 D + z^2 D(1) \left(zu \frac{D(1) - uD}{1 - u} \right) + z^2 uD$$

$$+ z^2 \frac{uD(1) - uD + zu^3 D^2 - zuD(1)^2}{1 - u} - z^2 u^{-2} \left(D - 1 - zu^2 D^2 - u[u^1]D - u^2[u^2]D \right)$$

$$+ zu^{-1}D_\triangleright - z^2 uDD_\triangleright - z[u^1]D_\triangleright.$$

We cancel some terms from 1$^{\text{st}}$, 2$^{\text{nd}}$, and 4$^{\text{th}}$. Next, We introduce the notation $K := zu \frac{D(1) - uD}{1 - u}$ and use it in the 2$^{\text{nd}}$ and 3$^{\text{rd}}$ line:

$$z^2 uD^2 + z^2 uDD_{\not\triangleright} - z^3 uD[u^1]D - z^3 D + z^2 D(1)K$$

$$- z^2 D(1) + zu^{-1}K - z^2(uD + D(1))K - z^2 u^{-2} \left(D - 1 - zu^2 D^2 - u[u^1]D - u^2[u^2]D \right)$$

$$+ zu^{-1}D_\triangleright - z[u^1]D_\triangleright.$$

After canceling some terms from the first two lines and applying $D_\triangleright = D - 1 - zu^2 D^2 - D_{\not\triangleright}$ in the 3$^{\text{rd}}$ line we obtain

$$z^2 uD^2 + z^2 uDD_{\not\triangleright} - z^3 uD[u^1]D - z^3 D$$

$$- z^2 D(1) + zu^{-1}K - z^2 uDK - z^2 u^{-2} \left(D - 1 - zu^2 D^2 - u[u^1]D - u^2[u^2]D \right)$$

$$+ zu^{-1}(D - 1 - zu^2 D^2 - D_{\not\triangleright}) - z([u^1]D - [u^1]D_{\not\triangleright}).$$

We apply $D_{\not\triangleright} = K - zu^{-1}(D - 1 - u[u^1]D)$ in the 1$^{\text{st}}$ and 3$^{\text{rd}}$ line and apply $[u^1]D_{\not\triangleright} = [u^1]K - [u^1]zu^{-1} \left(D - 1 - u[u^1]D \right) = [u^1]K - z[u^2]D$ in the 3$^{\text{rd}}$. After simplifying we obtain

$$z^2 uD^2 - z^2 D(1) + zu^{-1}K + zu^{-1}(D - 1 - zu^2 D^2 - K) - z([u^1]D - [u^1]K).$$

We replace now K by $D - 1 - zu^2 D^2$ and apply $[u^1]K = [u^1]D$ so that we have

$$z^2 uD^2 - z^2 D(1) + zu^{-1} \left(D - 1 - zu^2 D^2 \right)$$

which reduces to $zu^{-1}(D - 1 - zuD(1))$.

Finally we observe that we have relation

$$[u^1]D = [u^1](1 + zu^2D^2 + zu\frac{D(1) - uD}{1 - u}) = z[u^0]\frac{D(1) - uD}{1 - u} = zD(1)$$

which implies that we actually end up with

$$A_0 + A_1 + A_2 = zu^{-1}(D - 1 - zuD(1)) = zu^{-1}(D - 1 - u[u^1]D)$$

as proposed.

Finally we apply some simplifications to A_1 and A_2. Recall that the second term of $D_\triangleright$ is $(w - 1)(A_1 + (w + 1)A_2)$. It is clear that

$$A_1 = 2z^2uDD_\triangleright + z^2uD + 2zu^{-1}P(D_\triangleright)$$
$$A_2 = zu^{-1}D_\triangleright - z^2uDD_\triangleright - z[u^1]D_\triangleright - zu^{-1}P(D_\triangleright)$$

where

$$P(D_\triangleright) = \left(zu\frac{D_\triangleright(1) - uD_\triangleright}{1 - u} - zuD_\triangleright(1) - zu^{-1}\left(D_\triangleright - u[u^1]D_\triangleright - u^2[u^2]D_\triangleright\right)\right).$$

After canceling some terms we finally get that $A_1 + (w + 1)A_2$ is equal to

$$z^2uD + (w + 1)\left(zu^{-1}D_\triangleright - z[u^1]D_\triangleright\right) - z^2u(w - 1)DD_\triangleright - zu^{-1}(w - 1)P(D_\triangleright)$$

which completes the proof.

Fixed Partial Match Queries in Quadtrees

Amalia Duch
Universitat Politècnica de Catalunya
duch@cs.upc.edu
 https://orcid.org/0000-0003-4371-1286

Gustavo Lau
Universitat Politècnica de Catalunya
glau@cs.upc.edu
 https://orcid.org/0000-0002-3460-9186

Conrado Martínez
Universitat Politècnica de Catalunya
conrado@cs.upc.edu
 https://orcid.org/0000-0003-1302-9067

──── **Abstract** ────

Several recent papers in the literature have addressed the analysis of the cost $\mathcal{P}_{n,\mathbf{q}}$ of partial match search for a given fixed query $\mathbf{q}$ –that has s out of K specified coordinates– in different multidimensional data structures. Indeed, detailed asymptotic estimates for the main term in the expected cost $P_{n,\mathbf{q}} = \mathbb{E}\{\mathcal{P}_{n,\mathbf{q}}\}$ in standard and relaxed K-d trees are known (for any dimension K and any number s of specified coordinates), as well as stronger distributional results on $\mathcal{P}_{n,\mathbf{q}}$ for standard 2-d trees and 2-dimensional quadtrees. In this work we derive a precise asymptotic estimate for the main order term of $P_{n,\mathbf{q}}$ in quadtrees, for any values of K and s, $0 < s < K$, under the assumption that the limit of $P_{n,\mathbf{q}}/n^{\alpha}$ when $n \to \infty$ exists, where α is the exponent of n in the expected cost of a *random* partial match query with s specified coordinates in a random K-dimensional quadtree.

2012 ACM Subject Classification Theory of computation → Data structures design and analysis, Theory of computation → Design and analysis of algorithms

Keywords and phrases Quadtree, Partial match queries, Associative queries, Multidimensional search, Analysis of algorithms

Digital Object Identifier 10.4230/LIPIcs.AofA.2018.20

1 Introduction

One of the fundamental features of any hierarchical multidimensional data structure such as quadtrees is to efficiently support partial match (PM) queries. These queries are as follows. Given a collection F of K-dimensional ($K \geq 2$) tuples of the form $\mathbf{x} = (x_0, \ldots, x_{K-1})$, with each x_i ($0 \leq i < K$) belonging to a totally ordered domain $\mathcal{D}_i$, and a query $\mathbf{q} = (q_0, \ldots, q_{K-1})$

such that $q_i \in D_i \cup \{*\}$ $(0 \leq i < K)$, the goal of a PM query is to find all those tuples in F such that x_i matches q_i whenever $q_i \neq *$. Coordinates such that $q_i \neq *$ are called *specified*, otherwise they are called *unspecified*; we assume that the number s of specified coordinates satisfies $0 < s < K$.

The average-case analysis of PM queries in random quadtrees and other multidimensional data structures has a long history. In the case of quadtrees, a fundamental milestone was the paper by Flajolet, Gonnet, Puech, and Robson [7] where the authors proved that the expected cost of random PM queries with s specified coordinates in random K-dimensional quadtrees of n nodes is $\beta_{s,K} \, n^{\alpha(s/K)} + l.o.t.$ for some constant $\beta_{s,K}$; and $\alpha = \alpha(s/K)$ the unique real solution in $[0, 1]$ of the indicial equation

$$(\alpha + 2)^s (\alpha + 1)^{K-s} = 2^K. \tag{1}$$

The exponent α turns out to be exactly the same as in the expected cost of random PM queries in standard K-d trees. It was not until 2003 that Chern and Hwang [2] obtained an explicit expression for $\beta_{s,K}$, for general s and K, this is:

$$\beta_{s,K} = \frac{1}{(2^{K-s} - 1)\Gamma(\alpha + 1)^{K-s}\Gamma(\alpha + 2)^s} \prod_{2 \leq j \leq K} \frac{\Gamma(\alpha - \alpha_j)}{\Gamma(-\alpha_j)}, \tag{2}$$

for $0 < s < K$ and $K \geq 2$ and where Γ is the Gamma function and the α_j's are the roots of equation (1) and $\alpha = \alpha_1 > \Re(\alpha_2) \geq \cdots \geq \Re(\alpha_K)$. Note that Chern and Hwang [2] used the indicial equation for $\alpha + 1$ so they gave a formula for $\beta_{s,K}$ as a function of $\alpha_j' = \alpha_j + 1$, $j = 1, \ldots, K - 1$.

In 2011 fixed PM queries were studied for the first time in 2-dimensional quadtrees by Curien and Joseph [3] where the authors computed the expected cost $\mathbb{E}\{\mathcal{P}_{n,\mathbf{q}}\}$ of a fixed PM query in 2-dimensional quadtrees. In particular, they showed that if $\mathbf{q} = (q, *)$, then $P_{n,\mathbf{q}} = \mathbb{E}\{\mathcal{P}_{n,\mathbf{q}}\} \sim \nu_{1,2} \cdot (q \cdot (1 - q))^{\alpha/2} \cdot n^\alpha$, where $\alpha = \alpha(1/2) = (\sqrt{17} - 3)/2$ is the same exponent as in the expected cost for random PM queries [7], and $\nu_{1,2} = \frac{\Gamma(2\alpha+2)\Gamma(\alpha+2)}{2\Gamma^3(\alpha+1)\Gamma^2(\frac{\alpha}{2}+1)}$. The asymptotic distribution was obtained for this particular case by Broutin, Neininger and Sulzbach in 2012 [1].

In this work, we extend the results of [3] to give a precise asymptotic estimate of the expected cost of a fixed PM query in random K-dimensional quadtrees, for general K and s. In particular, we show that this cost is of the form

$$\nu_{s,K} \cdot \left(\prod_{i:q_i \neq *} q_i (1 - q_i) \right)^{\alpha/2} \cdot n^\alpha + \text{l.o.t.},$$

where $\nu_{s,K}$ is a constant that depends on s, K and the particular query $\mathbf{q}$ and $\alpha = \alpha(s/K)$ is the same as for random PM queries (see above).

The paper is organised as follows. In Section 2 we give some preliminaries. We explain our methodology in Section 3 through the simplest case $K = 2$ (Subsection 3.1). We continue with the general case of arbitrary s and K (Subsection 3.2). To complete the analysis one needs to solve an integral equation; that is the subject of Subsection 3.3. Section 4 contains some final remarks as well as some future lines of work.

2 Preliminaries

Let F be a collection of n multidimensional records, each one endowed with a K-dimensional key $\mathbf{x} = (x_0, \ldots, x_{K-1})$, with coordinate x_j drawn from a totally ordered domain $\mathcal{D}_j$. For convenience, here we will assume that, for all $0 \leq j < K$, $\mathcal{D}_j = [0, 1]$.
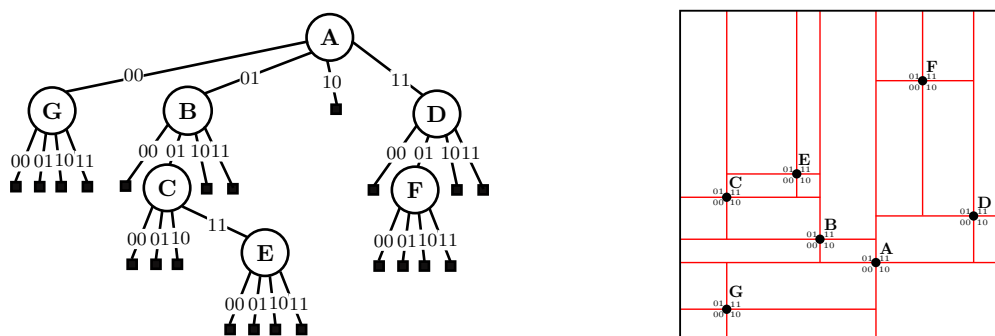
Figure 1 A 2-dimensional quadtree of file $F = \{A, B, C, D, E, F, G\}$ and the partition that it induces of the space. In this example $F_{00} = \{G\}$, $F_{01} = \{B, C, E\}$ and $F_{0*} = \{B, C, E, G\}$.

▶ **Definition 1.** A quadtree T of size n is a 2^K-ary tree storing a collection F of n K-dimensional records. T is either empty (when $n = 0$) or each one of its n nodes holds a key from F, such that the root node of T stores a record with key $\mathbf{x}$ and pointers to 2^K subtrees, that hold the remaining $n - 1$ records of F. Every subtree of T, let say $T_{\mathbf{w}}$, is associated to a bitstring $\mathbf{w} = w_0 w_1 \ldots w_{K-1} \in \{0, 1\}^K$, in such a way that $T_{\mathbf{w}}$ is a quadtree, and for any key $y \in T_{\mathbf{w}}$, it holds that $y_j \leq x_j$ if $w_j = 0$ and $y_j > x_j$ if $w_j = 1$, for all $0 \leq j < K$.

Any quadtree of size n induces a partition of the domain into $(2^K - 1)n + 1$ regions, each corresponding to a leaf (or equivalently empty subtree) in the quadtree. An example of a quadtree and the partition of the space that it induces is shown in Figure 1. To build a quadtree starting from an empty tree, each insertion of a new record with key $\mathbf{x}$ follows a path from the root to a leaf; at each step, we compare $\mathbf{x}$ and the key at the current node to determine in which of the 2^K subtrees the insertion should continue recursively, and the process ends when a leaf is reached and it is replaced by a new node containing $\mathbf{x}$ and 2^K empty subtrees. The region associated to the substituted leaf is called the *bounding box* of the subtree rooted at $\mathbf{x}$. Following the same convention used for the names of the subtrees, we will denote by $B_{\mathbf{w}}$ the bounding boxes of subtrees $T_{\mathbf{w}}$ associated to the tree rooted at $\mathbf{x}$ and by $F_{\mathbf{w}}$ the subset of data points of F that fall inside $B_{\mathbf{w}}$.

Consider a string $\mathbf{v}$ over the alphabet $\Sigma = \{0, 1, *\}$. We define as $\mathcal{L}(\mathbf{v})$ the set of binary strings matching $\mathbf{v}$; that is, where each occurrence of the symbol $*$ stands for a 0 or a 1. For instance, $\mathcal{L}(001) = \{001\}$, $\mathcal{L}(0*1) = \{001, 011\}$ and $\mathcal{L}(1**00) = \{10000, 10100, 11000, 11100\}$. With this notation let us define the following extension of the notion of bounding box $B_{\mathbf{v}} = \bigcup_{\mathbf{w} \in \mathcal{L}(\mathbf{v})} B_{\mathbf{w}}$.

Likewise $F_{\mathbf{v}}$ is the union of the (disjoint) $F_{\mathbf{w}}$'s with $\mathbf{w}$ matching $\mathbf{v}$. For example, in two dimensions $B_{**} = [0, 1]^2$ is the bounding box of the root of the quadtree, F_{0*} is the subset of all those keys with first coordinate smaller than the first coordinate of the root, that is, the ones stored in T_{00} and T_{01} (see Figure 1).

To perform a PM search with query $\mathbf{q}$, the quadtree is recursively explored as follows. First, we check whether the root $\mathbf{x}$ matches $\mathbf{q}$ or not, to report it in the former case. Then, we make recursive calls in all the 2^{K-s} subtrees $T_{\mathbf{w}}$ such that the first s bits of $\mathbf{w}$ are such that $w_i = 0$ whenever $q_i \neq *$ and $q_i \leq x_i$, and $w_i = 1$ whenever $q_i \neq *$ and $q_i > x_i$, $0 \leq i < s$, and the remaining $K - s$ bits can be either 0 or 1.

One key observation about the PM search in quadtrees (or similar data structures) is that, except for eventual matches, only the relative ranks of the coordinates matter. Let

us call the *rank vector* of a query $\mathbf{q}$ the vector $\mathbf{r}(\mathbf{q}) = (r_0, \ldots, r_{K-1})$ such that $r_i = *$, if $q_i = *$, and r_i is the number of records $\mathbf{x}$ in the collection F such that $x_i \le q_i$ $(0 \le r_i \le n)$, if $q_i \ne *$. Then for any two given queries $\mathbf{q}$ and $\mathbf{q}'$ with equal rank vectors $\mathbf{r}(\mathbf{q}) = \mathbf{r}(\mathbf{q}')$ the PM procedure described above will visit exactly the same set of nodes of the tree. In our analysis, we shall be using rank vectors instead of the queries themselves (as done in [6]) and consider, for instance, the cost $\mathcal{P}_{n,\mathbf{r}}$ of a PM query with given rank vector $\mathbf{r}$ in a random quadtree of size n. The probability model for random quadtrees that we will use throughout this work is that the tree is built by inserting in any order n keys drawn independently at random (coordinate by coordinate) from a continuous distribution. For the sake of simplicity, we can safely assume that the distribution is Uniform$(0, 1)$. Because of the symmetry of the model we can also assume that the s specified coordinates of $\mathbf{q}$ are the first s coordinates, $0 < s < K$, and therefore that $\mathbf{q} = (q_0, \ldots, q_{s-1}, *, \ldots, *)$ and $\mathbf{r} = (r_0, \ldots, r_{s-1}, *, \ldots, *)$. We shall write hence $\mathbf{q} = (q_0, \ldots, q_{s-1})$ and $\mathbf{r} = (r_0, r_1, \ldots, r_{s-1})$ with the convention that the implicit $K - s$ remaining components are all $*$'s.

3   Analysis

Our goal in this section is to find the expected cost $P_{n,\mathbf{r}} = \mathbb{E}\{\mathcal{P}_{n,\mathbf{r}}\}$, measured as the number of visited nodes, of a PM query with a fixed rank vector $\mathbf{r}$ in a random quadtree of n nodes.

In order to show our methodology and to give some intuition on the problem we are going to start our analysis with the easiest case $K = 2$ in Subsection 3.1. Afterwards, in Subsection 3.2, we analyze the general case.

In both subsections we are going to obtain a recurrence for $P_{n,\mathbf{r}}$. Then, in order to solve the general recurrence, we translate it into an integral equation whose solution will give us the leading term in the asymptotic estimate for $P_{n,\mathbf{r}}$. The solution of the integral equation is given in Subsection 3.3.

3.1   The case $K = 2$

Given a 2-dimensional quadtree T, its root splits the space into four rectangles: B_{00} (south-west of the root), B_{01} (north-west of the root), B_{10} (south-east of the root) and B_{11} (north-east of the root). These four rectangles are the corresponding *bounding boxes* of the four subtrees T_{00}, T_{01}, T_{10} and T_{11} from Definition 1. Recall also that $B_{0*} = B_{00} \cup B_{01}$ and $B_{*0} = B_{00} \cup B_{10}$ are, respectively, the rectangles west and south of the root. For any string $\mathbf{u} \in \{0, 1, *\}^2$, the number of data points in $B_{\mathbf{u}}$ (equivalently, the cardinality of $F_{\mathbf{u}}$) will be denoted $\mathcal{N}_{\mathbf{u}}$. For a random quadtree the $\mathcal{N}_{\mathbf{u}}$'s are random variables.

Let us now address the recurrence for $P_{n,\mathbf{r}}$, and to simplify let us write P_{n,r_0}, as $\mathbf{r} = (r_0, *)$. The basis of recursion is trivially $P_{0,r_0} = 0$. If $n > 0$, let $\mathbf{j} = (j_0, j_1)$ be the rank vector of the root. Since $\mathbf{q}$ contains only one specified coordinate, the relation between j_0 and r_0 determines whether the query intersects either B_{0*} or B_{1*}. If $r_0 \le j_0$, then the query intersects B_{0*}; otherwise it intersects B_{1*}. In our recurrence for P_{n,r_0} the value $j_0 = \mathcal{N}_{0*} = |F_{0*}|$ run from r_0 to $n - 1$, leading to a non-empty intersection of B_{0*} and the query, or from 0 to $r_0 - 1$, leading to a non-empty intersection of B_{1*} and the query. Because of the randomness assumptions, each possible value of $\mathcal{N}_{0*}$ has probability $1/n$ and hence this factor will weight the expected cost of the PM query conditioned to $\mathcal{N}_{0*} = j_0$.

The number of data points in B_{0*} is j_0 by definition, and the number of data points in B_{1*} is $n - 1 - j_0$. If the query intersects B_{0*} then the rank of the query with respect to B_{0*} is still r_0, but if it intersects B_{1*} then its rank with respect to B_{1*} is $r_0 - 1 - j_0$. So the contribution to P_{n,r_0} coming from the recursive traversal of B_{0*} involves a set of j_0
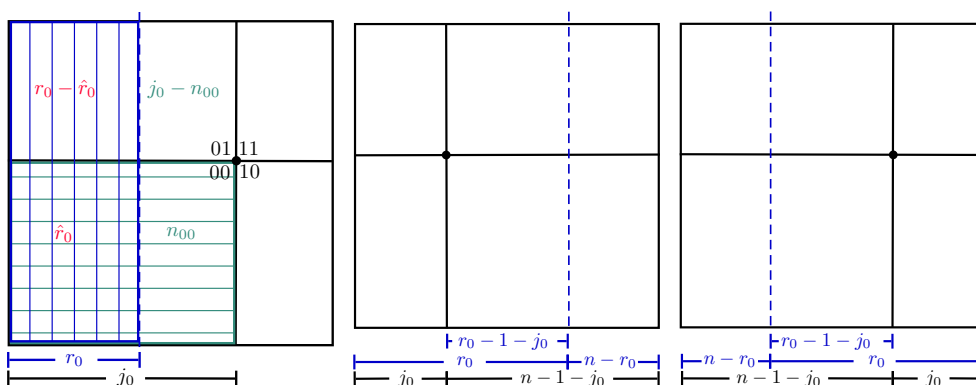
Figure 2 A partial match in a two-dimensional quadtree. The first diagram shows the case $r_0 \leq j_0$, the second one the case $j_0 < r_0$ and the third one how the east-west symmetry converts the second case into the first one.

points and the rank of the query is r_0 while the contribution coming from B_{1*} involves a set $n - 1 - j_0$ points and, because of the symmetry $P_{n,r_0} = P_{n,n-r_0}$, the rank of the query is $n - r_0$. Hence, we can reduce the case $j_0 < r_0$ to the case $r_0 \leq j_0$, see Figure 2.

In the general case we would have to consider 2^s regions $B_{\mathbf{w}}$ described by bitstrings $\mathbf{w} = w_0 \cdots w_{s-1} * \cdots *$, where each w_i is 0 or 1 depending on whether $r_i \leq j_i$ or not; as we consider all possible $\mathbf{j}$, the query will intersect these 2^s different regions, and we will be able to use these "east-west" symmetry considerations to reduce their analysis to the analysis of one of them, say, $B_{00\ldots 0*\ldots *}$.

Let us come back to $K = 2$. The region B_{0*} is the union of the two bounding boxes B_{00} and B_{01} (in general we will consider regions $B_{\mathbf{w}}$ that contain 2^{K-s} bounding boxes) and our goal is to use further symmetries to reduce the analysis of the cost of traversing both bounding boxes to the analysis of just traversing one of them, say, B_{00}.

Let Q_{j_0,r_0} be the contribution to the expected cost of a PM query due to the recursive call in T_{00}, when the query has rank r_0 in the first coordinate and given that there are $j_0 \geq r_0$ nodes to the west of the root.

Suppose that $\mathcal{N}_{00} = n_{00}$. The rank vector of the query in the recursive call to T_{00} will be $(\hat{r}_0, *)$, and the contribution to the expected cost will then be $P_{n_{00},\hat{r}_0}$. So it only remains to determine: a) the probability that $\mathcal{N}_{00} = n_{00}$, given the rank vector of the root $\mathbf{j}$ and, b) the probability that the rank vector of the query with respect to B_{00} is $(\hat{r}_0, *)$. Let us define the subsets of data points $F'_{\mathbf{v}}$ and the corresponding bounding boxes $B'_{\mathbf{v}}$ like $F_{\mathbf{v}}$ and $B_{\mathbf{v}}$, but with respect to the given query, instead of the root. The value $\hat{r}_0$ is the number of data points in the intersection between B_{00} and B'_{0*}, see Figure 2. We will use $\mathcal{R}_{\langle \mathbf{0} \rangle} := |F_{00} \cap F'_{0*}|$.

In general, $\langle \mathbf{i} \rangle := *^i 0 *^{K-1-i}$, so using this convention, we can also write $\mathcal{N}_{\langle \mathbf{0} \rangle} = j_0$ and $|F'_{\langle \mathbf{0} \rangle}| = r_0$. Conditioned on the sizes of F_{00}, $F_{\langle \mathbf{0} \rangle}$ and $F'_{\langle \mathbf{0} \rangle}$, the random variable $\mathcal{R}_{\langle \mathbf{0} \rangle}$ obeys a hypergeometric distribution:

$$\Pr\left\{ \mathcal{R}_{\langle \mathbf{0} \rangle} = \hat{r}_0 \,|\, \mathcal{N}_{00} = n_{00}, \mathcal{N}_{\langle \mathbf{0} \rangle} = j_0, |F'_{\langle \mathbf{0} \rangle}| = r_0 \right\} = \frac{\binom{n_{00}}{\hat{r}_0}\binom{j_0 - n_{00}}{r_0 - \hat{r}_0}}{\binom{j_0}{r_0}}.$$

Now if we look at the contribution to the expected cost due to the traversal of T_{01}, we have that $\mathcal{N}_{01} = j_0 - n_{00}$ and the rank of the query with respect to B_{01} is $(r_0 - \hat{r}_0, *)$. The fact that the second coordinate is unspecified allow us to do the analysis above with n_{01} instead of n_{00} and we would have obtained symmetric formulas. We can exploit this

north-south symmetry that will give us a factor of 2. Taking into account the visit to the root and our discussion so far we can write

$$P_{n,r_0} = 1 + \frac{2}{n}\left(\sum_{j_0=0}^{r_0-1} Q_{n-1-j_0,n-r_0} + \sum_{j_0=r_0}^{n-1} Q_{j_0,r_0}\right), \tag{3}$$

where, for $n_{0*} \geq r$, we have

$$Q_{j_0,r_0} = \sum_{n_{00}=0}^{j_0} \Pr\{\mathcal{N}_{00} = n_{00} \,|\, \mathcal{N}_{\langle\mathbf{0}\rangle} = j_0\} \sum_{\hat{r}_0=0}^{r_0}\left(\frac{\binom{n_{00}}{\hat{r}_0}\binom{j_0-n_{00}}{r_0-\hat{r}_0}}{\binom{j_0}{r_0}} P_{n_{00},\hat{r}_0}\right). \tag{4}$$

To complete the recurrence for P_{n,r_0} we need only to obtain the probability that $\mathcal{N}_{00} = n_{00}$, conditioned on $\mathcal{N}_{\langle\mathbf{0}\rangle} = j_0$. Since $\mathcal{N}_{\langle\mathbf{1}\rangle}$ can take any value in $[0..n-1]$ with identical probability, the number of points in B_{00} will take any value between 0 and j_0 with identical probability $1/(j_0+1)$. Plugging this probability and (4) into (3) yields to the desired recurrence for P_{n,r_0}.

An asymptotic estimate of the main term of P_{n,r_0} follows by deriving an integral equation for $f(z_0) := \lim_{n\to\infty} P_{n,z_0 n}/n^\alpha$ and solving that integral equation. We give the details of the derivation of the integral equation in the case of $K = 2$ in Lemma 4.

3.2   The general case

Let $\mathbf{r} = (r_0, r_1, \ldots, r_{s-1})$ be the query rank vector and let $\mathbf{j} = (j_0, \ldots, j_{s-1})$ be the first s coordinates of the rank vector for the root of the random quadtree. Thus we have that j_i is the value of $|F_{\langle\mathbf{i}\rangle}| = \mathcal{N}_{\langle\mathbf{i}\rangle}$. These K strings of the form $\langle\mathbf{i}\rangle$ constitute a "basis" in the sense that we can obtain any region $B_{\mathbf{w}}$ by complementation ($B_{*^i 1*^{K-1-i}} = B_{*\cdots*} \setminus B_{\langle\mathbf{i}\rangle}$) and intersection of the appropriate $B_{\langle\mathbf{i}\rangle}$'s.

Like we did for $K = 2$ our goal is to use the symmetries of the problem to reduce the whole analysis to the analysis of the contribution to the total cost of one particular subtree, namely, T_{0^s}. Again, call $Q_{\mathbf{j},\mathbf{r}}$ the contribution of the recursive call in T_{0^s}, conditioned to $r_i \leq j_i$ for all i, $0 \leq i < s$. This condition guarantees that the PM search will recursively continue in that subtree.

Then, because of the $K - s$ symmetries on unspecified coordinates (like the north-south symmetry of the case $K = 2$) and because of the s symmetries for specified coordinates (like the east-west symmetry when $K = 2$), we can express $P_{n,\mathbf{r}}$ in terms of $Q_{\mathbf{j},\mathbf{r}}$'s. In particular, considering all the possibilities for $\mathbf{j}$ gives a factor $1/n^s$, and a summation over all bitstrings $\mathbf{w}$ of length s to cover the cases where the query intersects $B_{\mathbf{w}}$. Finally the factor 2^{K-s} stems from the 2^{K-s} bounding boxes that each $B_{\mathbf{w}}$ contains. Hence,

$$P_{n,\mathbf{r}} = 1 + \frac{2^{K-s}}{n^s} \sum_{\mathbf{w}\in\{0,1\}^s} \sum_{j_0} \cdots \sum_{j_{s-1}} Q_{\mathbf{j}'_{\mathbf{w}}(\mathbf{j}),\mathbf{r}'_{\mathbf{w}}(\mathbf{r})}, \tag{5}$$

where the summation ranges are $r_i \leq j_i \leq n-1$ if $w_i = 0$, and $0 \leq j_i \leq r_i - 1$ if $w_i = 1$, and the rank vectors $\mathbf{j}'_{\mathbf{w}} = (j'_0, \ldots, j'_{s-1})$ and $\mathbf{r}'_{\mathbf{w}} = (r'_0, \ldots, r'_{s-1})$ are defined as follows: if $w_i = 0$ then $j'_i = j_i$ and $r'_i = r_i$, otherwise if $w_i = 1$ then $j'_i = n - 1 - j_i$ and $r'_i = n - r_i$.

For any i, $0 \leq i < K$, we will denote $\mathbf{0}^i$ the string $0^i *^{K-i}$, that is, a string of length K consisting of i zeros, followed by $K - i$ *'s.

The method to obtain a formula for $Q_{\mathbf{j},\mathbf{r}}$ consists of the following steps: 1) First we use Lemma 5 to obtain the probability distribution of the number of data points $\mathcal{N}_{\mathbf{0}^s}$ in the

"corner" hyperrectangle, by intersecting the sets $F_{\langle 0 \rangle}$, $F_{\langle 1 \rangle}$, ..., $F_{\langle s-1 \rangle}$, with sizes j_0, ..., j_{s-1}, respectively. This will be expressed by $s-1$ "hypergeometric" sums that will give us the probability that $\mathcal{N}_{\mathbf{0}^s} = \ell_s$; 2) Given that the last $K - s$ coordinates are unspecified, and conditioned on $j_i = \mathcal{N}_{\langle \mathbf{i} \rangle}$, $0 \le i < s$, all the potential sizes of $\mathcal{N}_{\langle \mathbf{i} \rangle} = |F_{\langle \mathbf{i} \rangle}|$, $s \le i < K$, are equiprobable. This will be expressed by $K - s$ "uniform" sums that will allow us to derive the probability distribution for $\mathcal{N}_{\mathbf{0}^K}$, and 3) Now conditioning on $\mathcal{N}_{\mathbf{0}^K} = |F_{\mathbf{0}^K}|$, and given $\mathbf{r}$ we intersect $F_{\mathbf{0}^K}$ with each of $F'_{\langle 0 \rangle}$, $F'_{\langle 1 \rangle}$, ..., $F'_{\langle \mathbf{s}-\mathbf{1} \rangle}$ to obtain the components of $\mathbf{r}_{\mathbf{0}^K} = (\hat{r}_0, \ldots, \hat{r}_{s-1})$. We will denote $\mathcal{R}_{\langle \mathbf{i} \rangle} = |F_{\mathbf{0}^K} \cap F'_{\langle \mathbf{i} \rangle}|$ the random variable that gives the i-th component of $\mathbf{r}_{\mathbf{0}^K}$. As in the case $K = 2$, the probability distribution of the $\mathcal{R}_{\langle \mathbf{i} \rangle}$'s is hypergeometric and it will lead to s additional "hypergeometric" sums.

Therefore the general formula for $Q_{\mathbf{j},\mathbf{r}}$ is:

$$Q_{\mathbf{j},\mathbf{r}} = \sum_{\ell_s=0}^{j_{s-1}} \Pr\left\{\mathcal{N}_{\mathbf{0}^s} = \ell_s \,\middle|\, \bigwedge_{i=0}^{s-1} \mathcal{N}_{\langle \mathbf{i} \rangle} = j_i\right\} \times \sum_{\ell_K=0}^{\ell_s} \Pr\left\{\mathcal{N}_{\mathbf{0}^K} = \ell_K \,\middle|\, \mathcal{N}_{\mathbf{0}^s} = \ell_s\right\}$$

$$\times \sum_{\mathbf{r}_{\mathbf{0}^K}=(\hat{r}_0,\ldots,\hat{r}_{s-1})} \Pr\left\{\bigwedge_{i=0}^{s-1} \mathcal{R}_{\langle \mathbf{i} \rangle} = \hat{r}_i \,\middle|\, \mathcal{N}_{\mathbf{0}^K} = \ell_K, \bigwedge_{i=0}^{s-1} |F'_{\langle \mathbf{i} \rangle}| = r_i\right\} \times P_{\ell_K, \mathbf{r}_{\mathbf{0}^K}}. \quad (6)$$

We can expand this last expression as:

$$Q_{\mathbf{j},\mathbf{r}} = \sum_{\ell_s=0}^{j_{s-1}} \cdots \sum_{\ell_2=0}^{j_1} \left(\frac{\binom{j_0}{\ell_2}\binom{n-1-j_0}{j_1-\ell_2}}{\binom{n-1}{j_1}} \cdots \frac{\binom{\ell_{s-1}}{\ell_s}\binom{n-1-\ell_{s-1}}{j_{s-1}-\ell_s}}{\binom{n-1}{j_{s-1}}} \right)$$

$$\times \frac{1}{\ell_s + 1} \sum_{\ell_{s+1}=0}^{\ell_s} \cdots \frac{1}{\ell_{K-1}+1} \sum_{\ell_K=0}^{\ell_{K-1}}$$

$$\sum_{\hat{r}_0=0}^{\ell_K \wedge r_0} \frac{\binom{\ell_K}{\hat{r}_0}\binom{j_0-\ell_K}{r_0-\hat{r}_0}}{\binom{j_0}{r_0}} \cdots \sum_{\hat{r}_{s-1}=0}^{\ell_K \wedge r_{s-1}} \frac{\binom{\ell_K}{\hat{r}_{s-1}}\binom{j_1-\ell_K}{r_{s-1}-\hat{r}_{s-1}}}{\binom{j_{s-1}}{r_{s-1}}} P_{\ell_K,(\hat{r}_0,\ldots,\hat{r}_{s-1})}, \quad (7)$$

where we have used $x \wedge y = \min(x, y)$ to stress the intersections that are involved in each case, e.g. $\hat{r}_i$ ranges from 0 to $\ell_K \wedge r_i$ since the number of data points is given by $|F_{\mathbf{0}^K} \cap F'_{\langle \mathbf{i} \rangle}|$; with $|F_{\mathbf{0}^K}| = \mathcal{N}_{\mathbf{0}^K} = \ell_K$ and $F'_{\langle \mathbf{i} \rangle} = r_i$.

To derive the integral equation corresponding to the recurrence above we can use arguments similar to those in the case $K = 2$. We give all the details of this derivation, as well as other necessary technical lemmas in Apprendix A.

▶ **Lemma 2.** *If $f(z_0, \ldots, z_{s-1}) = \lim_{n \to \infty} \frac{P_{n,\mathbf{r}}}{n^\alpha}$ exists, with $\alpha = \alpha(s/K)$ the solution of the indicial equation (1) and $z_i = \lim_{n \to \infty} r_i/n$, $0 < z_i < 1$, for all i, $0 \le i < s$, then $f(z_0, \ldots, z_{s-1})$ is the unique solution of*

$$f(z_0, \ldots, z_{s-1}) = \left(\frac{2}{\alpha+1}\right)^{K-s} \times \sum_{\mathbf{w} \in (0+1)^s} \Bigg\{$$

$$\int_{I_{w_0}(z_0)} \cdots \int_{I_{w_{s-1}}(z_{s-1})} f\Big(\varphi_{w_0}(z_0, u_0), \ldots, \varphi_{w_{s-1}}(z_{s-1}, u_{s-1})\Big)$$

$$\cdot \Big(\psi_{w_0}(u_0) \cdots \psi_{w_{s-1}}(u_{s-1})\Big)^\alpha \, du_{s-1} \cdots du_0 \Bigg\}, \quad (8)$$

where $I_0(z) = [0, z]$, $I_1(z) = [z, 1]$, $\psi_0(u) = 1 - u$, $\psi_1(u) = u$, $\varphi_0(z, u) = (1-z)/(1-u)$ and $\varphi_1(z, u) = z/u$, which satisfies the following boundary conditions:

1. $f(z_0, \ldots, z_{s-1})$ *is symmetric on all variables, that is, for any i and j,*

$$f(z_0, \ldots, z_i, \ldots, z_j, \ldots, z_{s-1}) = f(z_0, \ldots, z_j, \ldots, z_i, \ldots, z_{s-1}).$$

2. *For any $z_i \in (0, 1)$, $0 \le i < s$, f is symmetric with respect to the axis $z_i = 1/2$, that is,*

$$f(z_0, \ldots, z_i, \ldots, z_{s-1}) = f(z_0, \ldots, 1 - z_i, \ldots, z_{s-1}).$$

3. *For any i, $0 \le i < s$,*

$$\lim_{z_i \to 0^+} f(z_0, \ldots, z_i, \ldots, z_{s-1}) = \lim_{z_i \to 1^-} f(z_0, \ldots, z_i, \ldots, z_{s-1}) = 0.$$

4.

$$\int_0^1 \int_0^1 \cdots \int_0^1 f(z_0, \ldots, z_{s-1}) \, dz_0 \cdots dz_{s-1} = \beta_{s,K}.$$

Proof. We will follow a procedure similar to the one in the proof of Lemma 4, which covers the case $K = 2$.

The steps that we will give to obtain the integral equation for general K are:

1. Apply Lemma 6 to (7) s times in the s hypergeometric sums (the last sums over the $\hat{r}_i$'s)

2. Convert the $K - s$ uniform sums (the middle sums over the ℓ_i's, $s < i \le K$) into the corresponding integral by passing to the limit. That gives $K - s$ factors $1/(\alpha + 1)$.

3. Apply Lemma 7 once to the first $s - 1$ hypergeometric sums (over the ℓ_i's, $2 \le i \le s$).

4. Convert all the sums in (5) into integrals by passing to the limit.

Here, we use ℓ_i to denote the values that the random variables $\mathcal{N}_{\mathbf{0}^i}$ can take, like we did in subsection 3.2, and in particular in (6) and successive.

Defining $f\left(\frac{r_0}{n}, \ldots, \frac{r_{s-1}}{n}\right) := P_{n,\mathbf{r}}/n^\alpha$, where α is the solution of the indicial equation for quadtrees, we get:

$$
\frac{Q_{\mathbf{j},\mathbf{r}}}{n^\alpha} = \sum_{\ell_s=0}^{j_{s-1}} \cdots \sum_{\ell_2=0}^{j_1} \left(\frac{\binom{j_0}{\ell_2}\binom{n-1-j_0}{j_1-\ell_2}}{\binom{n-1}{j_1}} \cdots \frac{\binom{\ell_{s-1}}{\ell_s}\binom{n-1-\ell_{s-1}}{j_{s-1}-\ell_s}}{\binom{n-1}{j_{s-1}}} \right)
$$
$$
\times \frac{1}{\ell_s + 1} \sum_{\ell_{s+1}=0}^{\ell_s} \cdots \frac{1}{\ell_{K-1} + 1} \sum_{\ell_K=0}^{\ell_{K-1}}
$$
$$
\sum_{\hat{r}_0=0}^{\ell_K \wedge r_0} \frac{\binom{\ell_K}{\hat{r}_0}\binom{j_0-\ell_K}{r_0-\hat{r}_0}}{\binom{j_0}{r_0}} \cdots \sum_{\hat{r}_{s-1}=0}^{\ell_K \wedge r_{s-1}} \frac{\binom{\ell_K}{\hat{r}_{s-1}}\binom{j_1-\ell_K}{r_{s-1}-\hat{r}_{s-1}}}{\binom{j_{s-1}}{r_{s-1}}} \times f\left(\frac{\hat{r}_0}{\ell_K}, \ldots, \frac{\hat{r}_{s-1}}{\ell_K}\right) \left(\frac{\ell_K}{n}\right)^\alpha.
$$

Hence, defining $u_{0^i} = \lim_{n \to \infty}(\ell_i/n)$ for $s \le i \le K$, $z_i = \lim_{n \to \infty}(r_i/n)$ and $u_i =$

$\lim_{n \to \infty}(j_i/n)$ for $0 \le i < K$ and applying Lemma 6 s times:

$$
\begin{aligned}
\lim_{n\to\infty} \frac{Q_{\mathbf{j},\mathbf{r}}}{n^\alpha} &= \lim_{n\to\infty} \sum_{\ell_s=0}^{j_{s-1}} \cdots \sum_{\ell_2=0}^{j_1} \left(\frac{\binom{j_0}{\ell_2}\binom{n-1-j_0}{j_1-\ell_2}}{\binom{n-1}{j_1}} \cdots \frac{\binom{\ell_{s-1}}{\ell_s}\binom{n-1-\ell_{s-1}}{j_{s-1}-\ell_s}}{\binom{n-1}{j_{s-1}}} \right) \\
&\quad \times \frac{1}{\ell_s+1} \sum_{\ell_{s+1}=0}^{\ell_s} \cdots \frac{1}{\ell_{K-1}+1} \sum_{\ell_K=0}^{\ell_{K-1}} f\left(\frac{r_0}{j_0},\dots,\frac{r_{s-1}}{j_{s-1}}\right)\left(\frac{\ell_K}{n}\right)^\alpha \\
&= \lim_{n\to\infty} \sum_{\ell_s=0}^{j_{s-1}} \cdots \sum_{\ell_2=0}^{j_1} \left(\frac{\binom{j_0}{\ell_2}\binom{n-1-j_0}{j_1-\ell_2}}{\binom{n-1}{j_1}} \cdots \frac{\binom{\ell_{s-1}}{\ell_s}\binom{n-1-\ell_{s-1}}{j_{s-1}-\ell_s}}{\binom{n-1}{j_{s-1}}} \right) \\
&\quad \times \frac{1}{u_{0^s}} \int_0^{u_{0^s}} \cdots \frac{1}{u_{0^{K-1}}} \int_0^{u_{0^{K-1}}} f\left(\frac{z_0}{u_0},\dots,\frac{z_{s-1}}{u_{s-1}}\right) u_{0^K}^\alpha\, du_{0^K} \dots du_{0^{s+1}} \\
&= \lim_{n\to\infty} \sum_{\ell_s=0}^{j_{s-1}} \cdots \sum_{\ell_2=0}^{j_1} \left(\frac{\binom{j_0}{\ell_2}\binom{n-1-j_0}{j_1-\ell_2}}{\binom{n-1}{j_1}} \cdots \frac{\binom{\ell_{s-1}}{\ell_s}\binom{n-1-\ell_{s-1}}{j_{s-1}-\ell_s}}{\binom{n-1}{j_{s-1}}} \right) \\
&\quad \times f\left(\frac{z_0}{u_0},\dots,\frac{z_{s-1}}{u_{s-1}}\right) \frac{u_{0^s}^\alpha}{(\alpha+1)^{K-s}}.
\end{aligned}
$$

Replacing u_{0^s} by ℓ_s/n and applying Lemma 7 once to the first $s-1$ hypergeometric sums we obtain:

$$
\lim_{n\to\infty} \frac{Q_{\mathbf{j},\mathbf{r}}}{n^\alpha} = \frac{1}{(\alpha+1)^{K-s}} f\left(\frac{z_0}{u_0},\dots,\frac{z_{s-1}}{u_{s-1}}\right) \prod_{i=0}^{s-1} u_i^\alpha. \tag{9}
$$

Finally, introduce the following notation: $I_0(z) = [0,z]$, $I_1(z) = [z,1]$, $\varphi_0(z,u) = (1-z)/(1-u)$ and $\varphi_1(z,u) = z/u$. Plugging (9) into (5)) and passing to the limit (the fourth step in the procedure that we have described) yields the stated integral equation. ◄

Conditions 1 and 2 in the lemma follow from the combinatorics of the problem. By symmetry, $P_{n,\mathbf{r}} = P_{n,\mathbf{r}'}$ for any permutation $\mathbf{r}'$ of the rank vector $\mathbf{r}$. Likewise, if $\mathbf{r} = (r_0,\dots,r_i,\dots,r_{s-1})$ and $\mathbf{r}' = (r_0,\dots,r_{i-1},n-r_i,r_{i+1},\dots,r_{s-1})$ then $P_{n,\mathbf{r}} = P_{n,\mathbf{r}'}$. Condition 3 needs an inductive argument in the number of non-extreme ($z_i \ne 0$ and $z_i \ne 1$) coordinates. When all specified coordinates are extreme, say, $z_0 = z_1 = \dots = z_{s-1} = 0$ we must have $f = 0$; indeed, it is very easy to prove that $P_{n,(0,\dots,0)} = o(n^\alpha)$. We do not give here a complete and detailed analysis when $s_0 \le s$ specified coordinates are extreme; the computations and the reasoning is analogous to that carried out in [6] for K-d trees. Last but not least, Condition 4 follows by summing the expected cost $P_{n,\mathbf{r}}$ over all possible rank vectors $\mathbf{r}$ and dividing by $(n+1)^s$: it must yield the known expected cost of a random partial match query $\beta_{s,K} n^\alpha + o(n^\alpha)$. In terms of f, we must integrate f in the domain $[0,1]^s$ to obtain $\beta_{s,K}$. For a detailed justification the reader can refer to [6]: it is straightforward to adapt the discussion there to the case of quadtrees.

3.3 Solving the integral equation

From the integral equation (8) in Lemma 2 we can obtain an equivalent partial differential equation (PDE) by application of the differential operators

$$
\Phi_j(f) = z_j(1-z_j)\frac{\partial^2 f}{\partial z_j^2} + \alpha(2z_j-1)\frac{\partial f}{\partial z_j} - \alpha(\alpha+1)f.
$$

Indeed, if we define the operator

$$I_i(f) = z_i^{\alpha+1} \int_{z_i}^1 f(z_0, \ldots, z_{i-1}, u_i, z_{i+1}, \ldots, z_{s-1}) \frac{du_i}{u_i^{\alpha+2}} +$$

$$(1 - z_i)^{\alpha+1} \int_0^{z_i} f(z_0, \ldots, z_{i-1}, v_i, z_{i+1}, \ldots, z_{s-1}) \frac{dv_i}{(1-v_i)^{\alpha+2}}$$

then the integral equation (8) in Lemma 2 can be written as

$$f = \left(\frac{2}{\alpha+1}\right)^{K-s} I_0(I_1(\cdots(I_{s-1}(f)\cdots)),$$

using the changes of variables $u_i := z_i/u_i$ and $v_i := (1 - z_i)/(1 - u_i)$.

Then, as

$$\Phi_i(I_j(g)) = \Psi_i(g) = (2z_i - 1)\frac{\partial g}{\partial z_i} - 2\alpha g$$

it follows that

$$\Phi_0(\Phi_1(\cdots(\Phi_{s-1}(f))\cdots) = \left(\frac{2}{\alpha+1}\right)^{K-s} \Phi_0(\Phi_1(\cdots(\Phi_{s-1}(I_0(I_1(\cdots(I_{s-1}(f)\cdots))\cdots).$$

Now, since Φ_i's and Ψ_i's commute – $\Phi_i(\Phi_j(g)) = \Phi_j(\Phi_i(g))$, $\Psi_i(\Psi_j(g)) = \Psi_j(\Psi_i(g))$ – and $\Phi_i(\Psi_j(g)) = \Psi_j(\Phi_i(g))$ for any $i \neq j$, we can manipulate the equation above to get

$$\Phi_0(\Phi_1(\cdots(\Phi_{s-1}(f))\cdots) = \left(\frac{2}{\alpha+1}\right)^{K-s} \Psi_0(\Psi_1(\cdots(\Psi_{s-1}(f)\cdots)$$

or

$$\left(\Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{s-1} - \left(\frac{2}{\alpha+1}\right)^{K-s} \Psi_0 \circ \Psi_1 \circ \cdots \circ \Psi_{s-1}\right)(f) = 0, \tag{10}$$

which is the sought PDE, succinctly expressed in terms of the linear differential operators Φ_i and Ψ_i, $i = 0, \ldots, s - 1$.

The resulting PDE is homogeneous and linear, hence it is natural to try to solve it by separation of variables. The shape of equation (10) also cries out to try a solution in separated variables. Therefore, we will assume that the solution to the integral equation (8) is a function: $f(z_0, z_1, \ldots, z_{s-1}) = \phi_0(z_0) \cdot \phi_1(z_1) \cdots \phi_{s-1}(z_{s-1})$.

Given that the function f is symmetric with respect to any permutation of its arguments, we can also safely assume that all the functions $\phi_0, \phi_1, \cdots, \phi_{s-1}$ are the same function ϕ. Rather than working with the PDE itself, we may use our assumption to rewrite equation (8) as:

$$\phi(z_0) \cdot \phi(z_1) \cdots \phi(z_{s-1}) = \left(\frac{2}{\alpha+1}\right)^{K-s} \prod_{i=0}^{s-1} \left(\int_0^{z_i} \phi\left(\frac{1-z_i}{1-u_i}\right)(1-u_i)^\alpha du_i\right.$$

$$\left. + \int_{z_i}^1 \phi\left(\frac{z_i}{u_i}\right)u_i^\alpha du_i\right). \tag{11}$$

If ϕ is a solution of the following equation

$$\phi(z) = \left(\frac{2}{\alpha+1}\right)^{\frac{K-s}{s}} \left(\int_0^z \phi\left(\frac{1-z}{1-u}\right)(1-u)^\alpha du + \int_z^1 \phi\left(\frac{z}{u}\right)u^\alpha du\right), \tag{12}$$

then it would be a solution of equation (11). As shown in [4],

$$\phi(z) = \mu\big(z(1-z)\big)^{\delta-1}, \qquad \delta = \left(\frac{2}{\alpha+1}\right)^{\frac{K-s}{s}},$$

is such a solution, where μ is an arbitrary constant and we have discarded additional terms in the general solution based on symmetry considerations.

Because the exponent α is a solution to the indicial equation (1) it follows that $\delta = \frac{\alpha}{2}+1$ and hence the solution to (8) is:

$$f(z_0,\ldots,z_{s-1}) = \nu_{s,K} \cdot \prod_{i=0}^{s-1}\big(z_i(1-z_i)\big)^{\alpha/2},$$

where $\nu_{s,K}$ is a constant that depends on s and K only. To finish our derivation and to obtain the value of $\nu_{s,K}$ we replace f by the expression above in Condition 4 of Lemma 2 and we get:

$$\nu_{s,K}\left(\int_0^1 \big(z(1-z)\big)^{\alpha/2} dz\right)^s = \nu_{s,K}\left(\frac{\Gamma(\alpha/2+1)^2}{\Gamma(\alpha+2)}\right)^s = \beta_{s,K},$$

so we can use the expression for $\beta_{s,K}$ in Equation (2) to find an explicit formula for $\nu_{s,K}$.

To argue unicity of the solution, we should begin noticing that the linear homogeneous PDE satisfied by the function f has all real-analytic coefficients in the domain $(0,1)^s$, because the coefficients of the operators Ψ_i and Φ_i are analytic too in that domain and the PDE results from the composition of such operators.

Moreover, the highest derivative in the PDE is $\partial^{2s}f/\partial z_0^2\cdots\partial z_{s-1}^2$ and its coefficient $\prod_{0\le i<s} z_i(1-z_i)$ is clearly always positive in $(0,1)^s$, hence, the PDE is elliptic. Then, by Holmgren's theorem, any solution is real-analytic; and from Cauchy-Kovalevskaya theorem it follows that it must be unique, since this last theorem guarantees that there is a unique real-analytic solution (see for instance [8, 11]). Altogether, these results tell us that the solution that we have found, starting from the *ansatz* that it admitted a representation in separable variables, is unique.

It remains to verify by direct substitution that $P_{n,\mathbf{r}} = f(\mathbf{r}/n)n^\alpha$ is a solution of recurrence (5) replacing the independent term by $o(1)$, which is the error resulting from approximating the summations by integrals. With this our main result follows.

▶ **Theorem 3.** *If $\lim_{n\to\infty}\frac{P_{n,\mathbf{r}}}{n^\alpha}$ exists then the expected cost $P_{n,\mathbf{r}}$ of a PM query with given rank vector $\mathbf{r}$ such that $r_i = z_i n + o(n)$ for some $z_i \in (0,1)$, $0 \le i < s$, in a random K-dimensional quadtree of size n is*

$$P_{n,\mathbf{r}} = \nu_{s,K}\left(\prod_{i=0}^{s-1} z_i(1-z_i)\right)^{\alpha/2} n^\alpha + o(n^\alpha),$$

where α is the unique solution in $(0,1)$ of

$$(\alpha+2)^s(\alpha+1)^{K-s} = 2^K,$$

$$\nu_{s,K} = \frac{1}{(2^{K-s}-1)\Gamma(\alpha+1)^{K-s}\Gamma(\alpha/2+1)^{2s}}\prod_{2\le j\le K}\frac{\Gamma(\alpha-\alpha_j)}{\Gamma(-\alpha_j)},$$

and the α_j's, with $\alpha = \alpha_1 > \Re(\alpha_2) \ge \cdots \ge \Re(\alpha_K)$, are the roots of the indicial equation above.
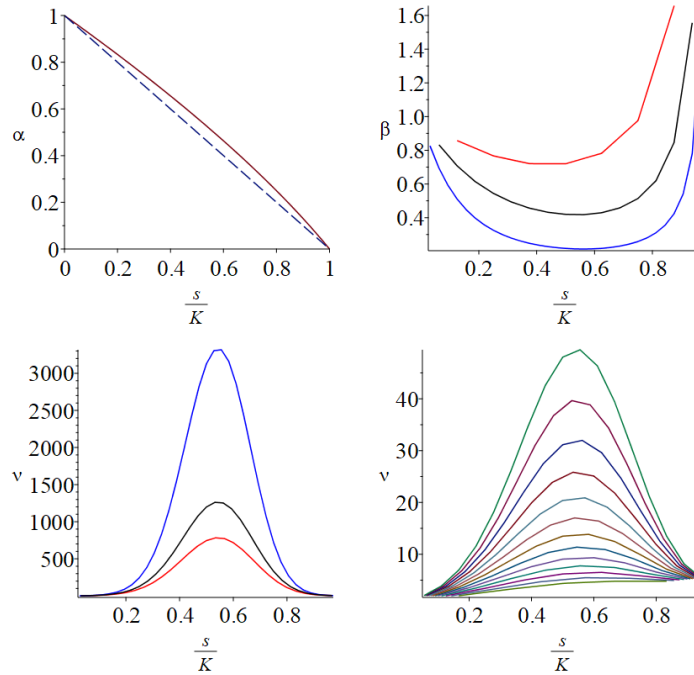
Figure 3 Variation of the exponent $\alpha(s/K)$ (top-left), $\beta(s, K)$ for $K \in \{8, 16, 32\}$ (top-right) and $\nu(s, K)$ for $K \in \{30, 32, 36\}$ (bottom-left), as well as $\nu(s, K)$ for all $6 \leq K \leq 18$ (bottom-right).

Figure 3 depicts how the exponent $\alpha = \alpha(s/K)$, and the constants $\beta(s, K)$ and $\nu(s, K)$ vary with respect to s and K. In all cases, the x-axis is s/K to ease the comparison – α is a function of s/K alone, but β and ν depend on both s and K. In the graphs for $\beta(s, K)$ and $\nu(s, K)$ we have drawn three curves in each case, corresponding to $K = 8$ (red), $K = 16$ (black) and $K = 32$ (blue) in the graph for $\beta(s, K)$, and $K = 30$ (red), $K = 32$ (black) and $K = 36$ (blue) in the graph for $\nu(s, K)$. Moreover in the graph of $\alpha(s/K)$ we have also plotted $1 - s/K$ (dashed line) for reference. For fixed K, $\beta(s, K)$ is a convex function with a minimum close to $s = K/2$ but slowly shifted to the right. Likewise, for fixed K, $\nu(s, K)$ is a bell-shaped function with a single global maximum near $s = K/2$ but also slightly shifted to the right ($\nu(s, K)$ is not defined for $s = K$). If we denote $\nu^*(K) = \max_{0<s<K} \nu(s, K)$ the graph shows that $\nu^*(K)$ grows with K. On the other hand, the graph and further numerical computations suggest that there is a limiting curve $\beta_\infty(x) = \lim_{K \to \infty} \beta(\lfloor xK \rfloor, K)$ that is a lower bound for any $\beta(s, K)$ as $K \to \infty$.

When $s = 0$ (no coordinate is specified), we have $\alpha(0) = \beta(0, K) = \nu(0, K) = 1$, despite all these constant are not well defined when $s = 0$. Notice that for $s = 0$ the partial match degenerates to a full traversal of the quadtree and visits its n nodes.

In the opposite situation, when all coordinates are specified, $s = K$, β and ν are undefined, and $\alpha(1) = 0$. The expected cost of a partial match is not $\Theta(1) = \Theta(n^0)$ but $\Theta(\log n)$ as it is actually an exact search.

4   Conclusions and Future Work

Our main result, Theorem 3, gives the main order term of the expected cost $P_{n,\mathbf{r}}$ of a PM search with a fixed query of rank vector $\mathbf{q}$, for quadtrees of any dimension K and any number

of specified coordinates. It can be easily translated to an equivalent result in terms of the coordinates q_i of the query, namely,

$$P_{n,\mathbf{q}} = \nu_{s,K} \cdot \left(\prod_{i:q_i \neq *} q_i(1 - q_i) \right)^{\alpha/2} \cdot n^\alpha + \text{l.o.t.}$$

under the assumption of uniformity of the coordinates of the data points (see, for instance, [6]).

We show that quadtrees behave qualitatively as standard and relaxed K-d trees [6]. There we conjectured that the form of the expected cost of a PM search with fixed query would have the same "shape" for a wide variety of multidimensional data structures, excluding those producing very balanced partitions of the space (e.g., quadtries, squarish K-d trees). Duch and Lau [5] have disproved the conjecture, in its broadest terms, as it does not apply to locally balanced K-d trees. However, it seems that the conjecture might hold for hierarchical multidimensional data structures where: 1) no balancing of subtrees occurs; 2) the partition at each node follows a fixed rule independent of the current data point.

From the methodological viewpoint, we systematically exploit the many symmetries that appear in the problem to simplify its formulation and to make its mathematical manipulation feasible.

Several open problems remain. To begin with, the existence of $\lim_{n\to\infty} \frac{P_{n,\mathbf{r}}}{n^\alpha}$, which has been rigorously proved for $K = 2$ in [3] (also in [1]); our result in that case coincides with the previous ones. We are currently working in the proof of the existence of the required limit for general K; meanwhile, our results follow from the – yet unproven – assumption that such limit exists. We shall mention that there is compelling evidence that this is the case. On the other hand, the existence of a limiting distribution for $\mathcal{P}_{n,\mathbf{r}}/n^\alpha$ has been shown only for the case of standard 2-d trees and 2-dimensional quadtrees, but not for other data structures or larger dimensions, and this is a question worth of further study.

Another goal for future research, more technical in nature but also more ambitious, is to develop tools that would allow a straightforward, (semi-)automatic derivation of the recurrences or distributional equations, the proof of the existence of the limiting distribution, the corresponding integral equations for the expectation and other higher order moments, etc. This kind of techniques would ease the obtainment of results, such as the ones in previous literature and the ones in this paper, for many other multidimensional data structures and it might also open the door for "universality" results such as the ones conjectured in [6].

References

1. Nicolas Broutin, Ralph Neininger, and Henning Sulzbach. Partial match queries in random quadtrees. In Yuval Rabani, editor, *Proc. of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1056–1065, 2012.

2. H.-H. Chern and H.-K. Hwang. Partial match queries in random quadtrees. *SIAM J. Comput.*, 32:904–915, 2003.

3. N. Curien and A. Joseph. Partial match queries in two-dimensional quadtrees: A probabilistic approach. *Advances in Applied Probability*, 43:178–194, 2011.

4. A. Duch, R. M. Jiménez, and C. Martínez. Selection by rank in k-dimensional binary search trees. *Random Structures and Algorithms*, 2012. `doi:10.1002/rsa.20476`.

5. Amalia Duch and Gustavo Lau. Partial match queries in relaxed K-dt trees. In *Proc. of the Fourteenth ACM-SIAM Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 131–138, 2017. `doi:10.1137/1.9781611974775.13`.

6   Amalia Duch, Gustavo Lau, and Conrado Martínez.   On the cost of fixed partial match queries in k-d trees. *Algorithmica*, 75(4):684–723, 2016. `doi:10.1007/s00453-015-0097-4`.

7   Philippe Flajolet, Gaston Gonnet, Claude Puech, and John Michael Robson. Analytic variations on quad trees. *Algorithmica*, 10:473–500, 1993.

8   Gerald B. Folland. *Introduction to Partial Differential Equations.* Princeton University Press, 2nd edition, 1995.

9   Steven G. Krantz and Harold R. Parks. *A Primer of Real Analytic Functions.* Birkhäuser, 2nd edition, 2002.

10 S. Ross. *A First Course in Probability.* Prentice Hall, Upper Saddle River, New Jersey, 8th edition, 2010.

11 Daniel Zwillinger. *Handbook of Differential Equations.* Academic Press, 3rd edition, 1997.

A   Technical Lemmas

▶ **Lemma 4.** *If $f(z) = \lim_{n\to\infty} \frac{P_{n,r}}{n^\alpha}$ exists, with $\alpha = \alpha(1/2)$ the solution of the indicial equation* (1) *when $s = 1$ and $K = 2$, and $z = \lim_{n\to\infty} r/n$, $0 < z < 1$, then*

$$f(z) = \frac{2}{\alpha+1}\left(\int_0^z f\left(\frac{1-z}{1-u}\right)(1-u)^\alpha du + \int_z^1 f\left(\frac{z}{u}\right)u^\alpha du\right). \tag{13}$$

The symmetry $P_{n,r_0} = P_{n,n-r_0}$ implies that in general $f(z) = f(1-z)$ and in particular $f\left(\frac{1-z}{1-u}\right) = f\left(1 - \frac{1-z}{1-u}\right)$ from where it follows that equation (13) *is the same as the one for standard 2d-trees (see [4]).*

Proof. Let $f(r_0/n) := P_{n,r_0}/n^\alpha$. Then we have that

$$\frac{P_{a,(b,*)}}{n^\alpha} = f\left(\frac{b}{a}\right)\left(\frac{a}{n}\right)^\alpha$$

and therefore, substituting into (4)

$$\frac{Q_{j_0,r_0}}{n^\alpha} = \frac{1}{j_0+1}\sum_{n_{00}=0}^{j_0}\sum_{\hat{r}_0=0}^{r_0}\left(\frac{\binom{n_{00}}{\hat{r}_0}\binom{j_0-n_{00}}{r_0-\hat{r}_0}}{\binom{j_0}{r_0}}f\left(\frac{\hat{r}_0}{n_{00}}\right)\left(\frac{n_{00}}{n}\right)^\alpha\right)$$

$$= \frac{1}{j_0+1}\sum_{n_{00}=0}^{j_0}\sum_{\hat{r}_0=0}^{r_0}\left(\frac{\binom{n_{00}}{\hat{r}_0}\binom{j_0-n_{00}}{r_0-\hat{r}_0}}{\binom{j_0}{r_0}}f\left(\frac{\hat{r}_0}{j_0}\frac{j_0}{n}\frac{n}{n_{00}}\right)\left(\frac{n_{00}}{n}\right)^\alpha\right)$$

The last sum is the expected value of a function of a hypergeometric random variable. Passing to the limit when $n \to \infty$, Lemma 6 allows us to exchange the expected value and the function. Therefore passing to the limit when $n \to \infty$, with $z = \lim_{n\to\infty}(r/n)$, $u_{0*} = \lim_{n\to\infty}(j_0/n)$, $u_{00} = \lim_{n\to\infty}(n_{00}/n)$, and assuming that f is real analytic in Lemma 6 we can apply it to get:

$$\lim_{n\to\infty}\frac{Q_{j_0,r_0}}{n^\alpha} = \frac{1}{u_{0*}}\int_0^{u_{0*}} f\left(\frac{u_{00}}{u_{0*}}\frac{z}{u_{0*}}\frac{u_{0*}}{u_{00}}\right)u_{00}^\alpha du_{00} = \frac{1}{u_{0*}}\int_0^{u_{0*}} f\left(\frac{z}{u_{0*}}\right)u_{00}^\alpha du_{00}$$

$$= \frac{1}{\alpha+1}f\left(\frac{z}{u_{0*}}\right)u_{0*}^\alpha.$$

and similarly

$$\lim_{n\to\infty}\frac{Q_{n-1-j_0,n-r_0}}{n^\alpha} = \frac{1}{1-u_{0*}}\int_0^{1-u_{0*}} f\left(\frac{1-z}{1-u_{0*}}\right)u_{00}^\alpha du_{00}$$

$$= \frac{1}{\alpha+1}f\left(\frac{1-z}{1-u_{0*}}\right)(1-u_{0*})^\alpha$$

Since $j_0 = 0 \implies u_{0*} = 0$, $j_0 = r_0 \implies u_{0*} = z_0$ and in the limit $j_0 = r_0 - 1 \implies u_{0*} = z_0$, $j_0 = n - 1 \implies u_{0*} = 1$ and $\frac{\Delta j_0}{n} \to du_{0*}$ replacing in (3) and passing to the limit we obtain this integral equation:

$$
\begin{aligned}
f(z_0) &= 2 \int_0^{z_0} \frac{1}{1 - u_{0*}} f\left(\frac{1 - z_0}{1 - u_{0*}}\right) \int_0^{1 - u_{0*}} u_{00}^\alpha \, du_{00} \, du_{0*} \\
&\quad + 2 \int_{z_0}^1 \frac{1}{u_{0*}} f\left(\frac{z_0}{u_{0*}}\right) \int_0^{u_{0*}} u_{00}^\alpha \, du_{00} \, du_{0*} \\
&= 2 \int_0^{z_0} \frac{1}{1 - u_{0*}} f\left(\frac{1 - z_0}{1 - u_{0*}}\right) \frac{(1 - u_{0*})^{\alpha+1}}{\alpha + 1} \, du_{0*} + 2 \int_{z_0}^1 \frac{1}{u_{0*}} f\left(\frac{z_0}{u_{0*}}\right) \frac{u_{0*}^{\alpha+1}}{\alpha + 1} \, du_{0*}.
\end{aligned}
$$

Replacing now in (3)), passing to the limit $n \to \infty$ and, to simplify, replacing u_{0*} by u we get the integral equation (13) in the statement of the Lemma. ◄

▶ **Lemma 5.** *Given a random K dimensional quadtree with n data points the conditional probability that $\mathcal{N}_{\mathbf{0}^K} = \ell_K$ given that $\mathcal{N}_{\langle \mathbf{i} \rangle} = n_{\langle \mathbf{i} \rangle}$ for $0 \le i \le K - 1$ is:*

$$
\Pr\left\{ \mathcal{N}_{\mathbf{0}^K} = \ell_K \,\middle|\, \bigwedge_{i=0}^{K-1} \mathcal{N}_{\langle \mathbf{i} \rangle} = n_{\langle \mathbf{i} \rangle} \right\} = \sum_{\ell_{K-1}=0}^{n_{\langle \mathbf{i} \rangle \mathbf{K-2}}} \cdots \sum_{\ell_3=0}^{n_{\langle \mathbf{2} \rangle}} \sum_{\ell_2=0}^{n_{\langle \mathbf{1} \rangle}} \left(\frac{\binom{\ell_1}{\ell_2}\binom{n-1-\ell_1}{n_{\langle \mathbf{1} \rangle}-\ell_2}}{\binom{n-1}{n_{\langle \mathbf{1} \rangle}}} \right.
$$
$$
\left. \frac{\binom{\ell_2}{\ell_3}\binom{n-1-\ell_2}{n_{\langle \mathbf{2} \rangle}-\ell_3}}{\binom{n-1}{n_{\langle \mathbf{2} \rangle}}} \cdots \frac{\binom{\ell_{K-2}}{\ell_{K-1}}\binom{n-1-\ell_{K-2}}{n_{\langle \mathbf{K-2} \rangle}-\ell_{K-1}}}{\binom{n-1}{n_{\langle \mathbf{K-2} \rangle}}} \frac{\binom{\ell_{K-1}}{\ell_K}\binom{n-1-\ell_{K-1}}{n_{\langle \mathbf{K-1} \rangle}-\ell_K}}{\binom{n-1}{n_{\langle \mathbf{K-1} \rangle}}} \right). \quad (14)
$$

Proof. In the base case $K = 2$ given n, $\mathcal{N}_{\langle \mathbf{0} \rangle} \equiv \mathcal{N}_{0*} = n_{0*}$ and $\mathcal{N}_{\langle \mathbf{1} \rangle} \equiv \mathcal{N}_{*0} = n_{*0}$, the probability that the intersection of the rectangles $B_{\langle \mathbf{0} \rangle} = B_{0*}$ and $B_{\langle \mathbf{1} \rangle} = B_{*0}$ contains $\ell_2 = n_{00}$ nodes is the probability of having $\ell_2 = n_{00}$ successes in n_{*0} draws without replacement from a population of size $n - 1$ that contains n_{0*} successes. It is $n - 1$ instead of n because the root cannot be in the intersections. Therefore the distribution is hypergeometric:

$$
\Pr\left\{ \mathcal{N}_{00} = n_{00} \,\middle|\, \mathcal{N}_{0*} = n_{0*}, \mathcal{N}_{*0} = n_{*0} \right\} = \frac{\binom{n_{0*}}{n_{00}}\binom{n-1-n_{0*}}{n_{*0}-n_{00}}}{\binom{n-1}{n_{*0}}}.
$$

Assume that the lemma is true for K dimensions. We can do the inductive step based on writing the intersection of $K + 1$ sets as an intersection of K sets followed by the intersection of two sets:

$$
\bigcap_{i=0}^{K} F_{*^i 0 *^{K-i}} = \left(\bigcap_{i=0}^{K-1} F_{*^i 0 *^{K-i}} \right) \cap F_{*^K 0} = F_{0^K *} \cap F_{*^K 0} = F_{0^{K+1}}.
$$

Taking into account all the possible values of $\mathcal{N}_{0^K *}$, we have:

$$
\Pr\left\{ \mathcal{N}_{0^{K+1}} = n_{0^{K+1}} \,\middle|\, \bigwedge_{i=0}^{K} \mathcal{N}_{*^i 0 *^{K-i}} = n_{*^i 0 *^{K-i}} \right\}
$$
$$
= \sum_{n_{0^K *}=0}^{n_{*^{K-1} 0 *}} \left(\Pr\left\{ \mathcal{N}_{0^K *} = n_{0^K *} \,\middle|\, \bigwedge_{i=0}^{K-1} \mathcal{N}_{*^i 0 *^{K-i}} = n_{*^i 0 *^{K-i}} \right\} \right.
$$
$$
\left. \times \Pr\left\{ \mathcal{N}_{0^{K+1}} = n_{0^{K+1}} \,\middle|\, \mathcal{N}_{0^K *} = n_{0^K *}, \mathcal{N}_{*^K 0} = n_{*^K 0} \right\} \right)
$$
$$
= \sum_{n_{0^K *}=0}^{n-1} \left(\Pr\left\{ \mathcal{N}_{0^K *} = n_{0^K *} \,\middle|\, \bigwedge_{i=0}^{K-1} \mathcal{N}_{*^i 0 *^{K-i}} = n_{*^i 0 *^{K-i}} \right\} \times \frac{\binom{n_{0^K *}}{n_{0^{K+1}}}\binom{n-1-n_{0^K *}}{n_{*^K 0}-n_{0^{K+1}}}}{\binom{n-1}{n_{*^K 0}}} \right),
$$

applying the inductive hypothesis (14) (adding a $*$ to the end of each string) completes the proof. Notice that we have used ℓ_i instead of $n_{0^i *^{K-i}}$ and $n_{\langle \mathbf{i} \rangle} = n_{*^i 0 *^{K-1-i}}$ in the statement of the theorem. ◄

▶ **Lemma 6.** *Given a random two dimensional quadtree let* $\mathcal{N}_{0*}$, $\mathcal{N}_{*0}$ *and* $\mathcal{N}_{00}$ *be respectively the random variables of the number of nodes west, south and south-west of the root. If* f *is a real analytic function [9] in* $(0,1)$, $\lim_{n \to \infty} n_{0*}/n = u_{0*}$ *and* $\lim_{n \to \infty} n_{*0}/n = u_{*0}$, *where* $u_{0*}, u_{*0} \in (0,1)$, *then*

$$\lim_{n \to \infty} \mathbb{E}\left\{ f\left(\frac{\mathcal{N}_{00}}{n}\right) \bigg| \mathcal{N}_{0*} = n_{0*}, \mathcal{N}_{*0} = n_{*0} \right\} = \lim_{n \to \infty} \sum_{n_{00}=0}^{n_{*0}} \left(\frac{\binom{n_{0*}}{n_{00}}\binom{n-n_{0*}}{n_{*0}-n_{00}}}{\binom{n}{n_{*0}}} f\left(\frac{n_{00}}{n}\right) \right)$$

$$= \lim_{n \to \infty} \sum_{n_{00}=0}^{n_{*0}} \left(\frac{\binom{n_{0*}}{n_{00}}\binom{n-n_{0*}}{n_{*0}-n_{00}}}{\binom{n}{n_{*0}}} f\left(\frac{n_{00}}{n}\right) \right)$$

$$= f(u_{0*} u_{*0}). \tag{15}$$

Proof. For simplicity, in the hypergeometric probability formulas we have replaced $n-1$ by n as in the limit they are the same.

Since f is real analytic all derivatives of f exist in $(0,1)$ and we can write, for some $x_0 \in (0,1)$,

$$f(x) = \sum_{i=0}^{\infty} a_i (x - x_0)^i = \sum_{i=0}^{\infty} a_i \sum_{k=0}^{i} \binom{i}{k} (-x_0)^{i-k} x^k.$$

Since the series on the right side converges we can use the linearity of expectations:

$$\mathbb{E}\left\{ f(x) \right\} = \sum_{i=0}^{\infty} a_i \sum_{k=0}^{i} \binom{i}{k} (-x_0)^{i-k} \mathbb{E}\left\{ x^k \right\}.$$

Therefore we only need to prove the lemma for $f(x) = x^k$. If $X_{n,m,N}$ is a hypergeometric random variable with parameters n, m, and N then [10]:

$$\mathbb{E}\left\{ X_{n,m,N}^k \right\} = \frac{nm}{N} \mathbb{E}\left\{ (X_{n-1,m-1,N-1} + 1)^{k-1} \right\}.$$

Based on that it is easy to prove by induction that for every $k \in \mathbb{N}$ there are integers $c_{k,i}$, with $c_{k,k} = 1$, such that:

$$\mathbb{E}\left\{ X_{n,m,N}^k \right\} = \sum_{i=0}^{k} c_{k,i} \frac{n^{\underline{i}} m^{\underline{i}}}{N^{\underline{i}}}.$$

Therefore if $f(x) = x^k$:

$$\mathbb{E}\left\{ f\left(\frac{\mathcal{N}_{00}}{n}\right) \bigg| \mathcal{N}_{0*} = n_{0*}, \mathcal{N}_{*0} = n_{*0} \right\} = \mathbb{E}\left\{ \frac{\mathcal{N}_{00}^k}{n^k} \bigg| \mathcal{N}_{0*} = n_{0*}, \mathcal{N}_{*0} = n_{*0} \right\}$$

$$= \frac{\sum_{i=0}^{k} c_{k,i} \frac{n_{0*}^{\underline{i}} n_{*0}^{\underline{i}}}{n^{\underline{i}}}}{n^k} = \sum_{i=0}^{k} c_{k,i} \frac{n_{0*}^{\underline{i}} n_{*0}^{\underline{i}}}{n^{\underline{i}} n^k}.$$

In the last sum the only term that does not go to zero as $n \to \infty$ is the last one, where $i = k$. Given that $c_{k,k} = 1$, we have:

$$\lim_{n \to \infty} \mathbb{E}\left\{ \frac{\mathcal{N}_{00}^k}{n^k} \bigg| \mathcal{N}_{0*} = n_{0*}, \mathcal{N}_{*0} = n_{*0} \right\} = \lim_{n \to \infty} \frac{\frac{n_{0*}^{\underline{k}} n_{*0}^{\underline{k}}}{n^{\underline{k}}}}{n^k} = \lim_{n \to \infty} \left(\frac{n_{0*}}{n}\right)^k \left(\frac{n_{*0}}{n}\right)^k = u_{0*}^k u_{*0}^k.$$

That proves the lemma for $f(x) = x^k$. ◄

The lemma can be generalised to any dimension K using mathematical induction on the number of dimensions, again assuming that the function f is real analytic (in several variables).

▶ **Lemma 7.** *Given a random quadtree let $\mathcal{N}_{\langle \mathbf{i} \rangle}$ be the random variable of the number of data points that have their i-th coordinate less than the i-th coordinate of the root and the rest of the coordinates undetermined and let $\mathcal{N}_{0^K}$ be the random variable of the size of the cuboid where all the coordinates have values lower than the respective coordinates of the root. If f is real analytic in $(0,1)^K$, $\lim_{n \to \infty} n_{\langle \mathbf{i} \rangle}/n = u_i$ for $0 \le i < K$, where $u_i \in (0,1)$, then*

$$\lim_{n \to \infty} \mathbb{E} \left\{ f\left(\frac{\mathcal{N}_{0^K}}{n} \right) \middle| \bigwedge_{i=0}^{K-1} \mathcal{N}_{\langle \mathbf{i} \rangle} = n_{\langle \mathbf{i} \rangle} \right\} = f\left(\prod_{i=0}^{K-1} u_i \right).$$

Proof. The base case $K = 2$ has been proved. Assume that the lemma is true for K dimensions. Then:

$$\lim_{n \to \infty} \mathbb{E} \left\{ f\left(\frac{\mathcal{N}_{0^{K+1}}}{n} \right) \middle| \bigwedge_{i=0}^{K} \mathcal{N}_{*^i 0 *^{K-i}} = n_{*^i 0 *^{K-i}} \right\}$$

$$= \lim_{n \to \infty} \sum_{n_{0^{K+1}}=0}^{n-1} \Pr \left\{ \mathcal{N}_{0^{K+1}} = n_{0^{K+1}} \middle| \bigwedge_{i=0}^{K} \mathcal{N}_{*^i 0 *^{K-i}} = n_{*^i 0 *^{K-i}} \right\} f\left(\frac{n_{0^{K+1}}}{n} \right)$$

$$= \lim_{n \to \infty} \sum_{n_{0^{K+1}}=0}^{n-1} \sum_{n_{0^K *}=0}^{n-1} \left(\Pr \left\{ \mathcal{N}_{0^K *} = n_{0^K *} \middle| \bigwedge_{i=0}^{K-1} \mathcal{N}_{*^i 0 *^{K-i}} = n_{*^i 0 *^{K-i}} \right\} \right.$$

$$\left. \times \Pr \left\{ \mathcal{N}_{0^{K+1}} = n_{0^{K+1}} \middle| \mathcal{N}_{0^K *} = n_{0^K *}, \mathcal{N}_{*^K 0} = n_{*^K 0} \right\} f\left(\frac{n_{0^{K+1}}}{n} \right) \right.$$

$$= \lim_{n \to \infty} \sum_{n_{0^K *}=0}^{n-1} \Pr \left\{ \mathcal{N}_{0^K *} = n_{0^K *} \middle| \bigwedge_{i=0}^{K-1} \mathcal{N}_{*^i 0 *^{K-i}} = n_{*^i 0 *^{K-i}} \right\}$$

$$\times \mathbb{E} \left\{ f\left(\frac{\mathcal{N}_{0^{K+1}}}{n} \right) \middle| \mathcal{N}_{0^K *} = n_{0^K *}, \mathcal{N}_{*^K 0} = n_{*^K 0} \right\}$$

$$= \lim_{n \to \infty} \sum_{n_{0^K *}=0}^{n-1} \Pr \left\{ \mathcal{N}_{0^K *} = n_{0^K *} \middle| \bigwedge_{i=0}^{K-1} \mathcal{N}_{*^i 0 *^{K-i}} = n_{*^i 0 *^{K-i}} \right\} \times f\left(\lim_{n \to \infty} \frac{n_{0^K *} n_{*^K 0}}{(n-1)n} \right)$$

$$= \lim_{n \to \infty} \mathbb{E} \left\{ f\left(\lim_{n \to \infty} \frac{\mathcal{N}_{0^K *} n_{*^K 0}}{(n-1)n} \right) \middle| \bigwedge_{i=0}^{K-1} \mathcal{N}_{*^i 0 *^{K-i-1} 0} = n_{*^i 0 *^{K-i-1} 0} \right\}.$$

Replacing $n-1$ by n, because in the limit they are equivalent, and using the induction hypothesis (adding 0 at the end of each string) we have:

$$\lim_{n \to \infty} \mathbb{E} \left\{ f\left(\frac{\mathcal{N}_{0^{K+1}}}{n} \right) \middle| \bigwedge_{i=0}^{K} \mathcal{N}_{*^i 0 *^{K-i}} = n_{*^i 0 *^{K-i}} \right\} = f\left(\lim_{n \to \infty} \left(\prod_{i=0}^{K-1} \frac{n_{*^i 0 *^{K-i-1} 0}}{n_{*^K 0}} \right) \frac{n_{*^K 0}}{n} \right)$$

$$= f\left(\lim_{n \to \infty} \prod_{i=0}^{K} \frac{n_{*^i 0 *^{K-i}}}{n} \right) = f\left(\prod_{i=0}^{K} u_{*^i 0 *^{K-i}} \right).$$

◀

▶ **Lemma 8.** *The real function $f(x) = x^a (1-x)^a$ is real analytic, i. e. it is infinitely differentiable and agrees with its Taylor series, in the interval $(0,1)$ for any real number a.*

Proof. By the binomial series, or Newton's generalized binomial theorem, $f_1(x) = (1 - x)^a$ is real analytic in $(-1, 1)$ and $f_2(x) = x^a = (1 + (x - 1))^a$ is real analytic in $(0, 2)$. Therefore their product, $f(x)$, is real analytic in $(0, 1)$. ◀

On the Tails of the Limiting QuickSort Density

James Allen Fill

Department of Applied Mathematics and Statistics, The Johns Hopkins University, 3400
N. Charles Street, Baltimore, MD 21218-2682, USA
jimfill@jhu.edu

Wei-Chun Hung

Department of Applied Mathematics and Statistics, The Johns Hopkins University, 3400
N. Charles Street, Baltimore, MD 21218-2682, USA
whung6@jhu.edu

──── **Abstract** ────

We give upper and lower asymptotic bounds for the left tail and for the right tail of the continuous limiting `QuickSort` density f that are nearly matching in each tail. The bounds strengthen results from a paper of Svante Janson (2015) concerning the corresponding distribution function F. Furthermore, we obtain similar upper bounds on absolute values of derivatives of f of each order.

1 Introduction

Let X_n denote the (random) number of comparisons when sorting n distinct numbers using the algorithm `QuickSort`. Clearly $X_0 = 0$, and for $n \geq 1$ we have the recurrence relation

$$X_n \stackrel{\mathcal{L}}{=} X_{U_n-1} + X^*_{n-U_n} + n - 1,$$

where $\stackrel{\mathcal{L}}{=}$ denotes equality in law (i.e., in distribution); $X_k \stackrel{\mathcal{L}}{=} X^*_k$; the random variable U_n is uniformly distributed on $\{1, \ldots, n\}$; and $U_n, X_0, \ldots, X_{n-1}, X^*_0, \ldots, X^*_{n-1}$ are all independent. It is well known that

$$\mathbb{E}X_n = 2\,(n+1)\,H_n - 4n,$$

where H_n is the nth harmonic number $H_n := \sum_{k=1}^n k^{-1}$ and (from a simple exact expression) that $\text{Var}\,X_n = (1 + o(1))(7 - \frac{2\pi^2}{3})n^2$. To study distributional asymptotics, we first center and scale X_n as follows:

$$Z_n = \frac{X_n - \mathbb{E}X_n}{n}.$$

Using the Wasserstein d_2-metric, Rösler [8] proved that Z_n converges to Z weakly as $n \to \infty$. Using a martingale argument, Régnier [7] proved that the slightly renormalized $\frac{n}{n+1}Z_n$ converges to Z in L^p for every finite p, and thus in distribution; equivalently, the same conclusions hold for Z_n. The random variable Z has everywhere finite moment generating function with $\mathbb{E}Z = 0$ and $\text{Var}\,Z = 7 - \left(2\pi^2/3\right)$. Moreover, Z satisfies the distributional identity

$$Z \stackrel{\mathcal{L}}{=} UZ + (1-U)Z^* + g(U).$$

On the right, $Z^* \stackrel{\mathcal{L}}{=} Z$; U is uniformly distributed on $(0,1)$; U, Z, Z^* are independent; and

$$g(u) := 2u \ln u + 2(1-u)\ln(1-u) + 1.$$

Further, the distributional identity together with the condition that $\mathbb{E}Z$ (exists and) vanishes characterizes the limiting `Quicksort` distribution; this was first shown by Rösler [8] under the additional condition that $\text{Var}\,Z < \infty$, and later in full by Fill and Janson [1].

Fill and Janson [2] derived basic properties of the limiting `QuickSort` distribution $\mathcal{L}(Z)$. In particular, they proved that $\mathcal{L}(Z)$ has a (unique) continuous density f which is everywhere positive and infinitely differentiable, and for every $k \geq 0$ that $f^{(k)}$ is bounded and enjoys superpolynomial decay in both tails, that is, for each $p \geq 0$ and $k \geq 0$ there exists a finite constant $C_{p,k}$ such that $\left|f^{(k)}(x)\right| \leq C_{p,k}|x|^{-p}$ for all $x \in \mathbb{R}$.

In this paper, we study asymptotics of $f(-x)$ and $f(x)$ as $x \to \infty$. Janson [3] concerned himself with the corresponding asymptotics for the distribution function F and wrote this: "Using non-rigorous methods from applied mathematics (assuming an as yet unverified regularity hypothesis), Knessl and Szpankowski [4] found very precise asymptotics of both the left tail and the right tail." Janson specifies these Knessl–Szpankowski asymptotics for F in his equations (1.6)–(1.7). But Knessl and Szpankowski actually did more, producing asymptotics for f, which were integrated by Janson to get corresponding asymptotics for F. We utilize the same abbreviation $\gamma := (2 - \frac{1}{\ln 2})^{-1}$ as Janson [3]. With the *same* constant c_3 as in (1.6) of [3], the density analogues of (1.6) (omitting the middle expression) and (1.7) of [3] are that, as $x \to \infty$, Knessl and Szpankowski [4] find

$$f(-x) = \exp\left[-e^{\gamma x + c_3 + o(1)}\right] \tag{1}$$

for the left tail and

$$f(x) = \exp[-x \ln x - x \ln \ln x + (1 + \ln 2)x + o(x)] \tag{2}$$

for the right tail.

We will come as close to these non-rigorous results for the density as Janson [3] does for the distribution function and we also obtain corresponding asymptotic upper bounds for absolute values of derivatives of the density. Although our asymptotics for f imply the asymptotics for F in Janson's Theorem 1.1, it is important to note that in the case of upper bounds (but not lower bounds) we use his results in the proofs of ours.

The next two theorems are our main results.

▶ **Theorem 1.1.** *Let* $\gamma := (2 - \frac{1}{\ln 2})^{-1}$. *As* $x \to \infty$, *the limiting* `QuickSort` *density function* f *satisfies*

$$\exp\left[-e^{\gamma x + \ln \ln x + O(1)}\right] \leq f(-x) \leq \exp\left[-e^{\gamma x + O(1)}\right], \tag{3}$$

$$\exp[-x \ln x - x \ln \ln x + O(x)] \leq f(x) \leq \exp[-x \ln x + O(x)]. \tag{4}$$

▶ **Theorem 1.2.** *Given an integer $k \geq 0$, as $x \to \infty$ the kth derivative of the limiting* QuickSort *density function f satisfies*

$$\left| f^{(k)}(-x) \right| \leq \exp\left[-e^{\gamma x + O(1)} \right], \tag{5}$$

$$\left| f^{(k)}(x) \right| \leq \exp[-x \ln x + O(x)]. \tag{6}$$

▶ **Remark.** The non-rigorous arguments of Knessl and Szpankowski [4] suggest that the following asymptotics as $x \to \infty$ obtained by repeated formal differentiation of (1)–(2) are correct for every $k \geq 0$:

$$f^{(k)}(-x) = \exp\left[-e^{\gamma x + c_3 + o(1)} \right], \tag{7}$$

$$f^{(k)}(x) = (-1)^k \exp[-x \ln x - x \ln \ln x + (1 + \ln 2)x + o(x)]. \tag{8}$$

But these remain conjectures for now. Unfortunately, for $k \geq 1$ we don't even know how to identify rigorously the asymptotic signs of $f^{(k)}(\mp x)$! Concerning $k = 1$, it has long been conjectured that f is unimodal. This would of course imply that $f'(-x) > 0$ and $f'(x) < 0$ for sufficiently large x.

As already mentioned, Fill and Janson [2] proved that or each $p \geq 0$ and $k \geq 0$ there exists a finite constant $C_{p,k}$ such that $\left| f^{(k)}(x) \right| \leq C_{p,k} |x|^{-p}$ for all $x \in \mathbb{R}$. Our technique for proving the upper bounds in Theorems 1.1 and 1.2 is to use explicit bounds on the constants $C_k := C_{0,k}$ together with the Landau–Kolmogorov inequality (see, for example, [9]).

Our extended abstract is organized as follows. In Section 2 we deal with preliminaries: We restate (to render this extended abstract self-contained) the asymptotic results of Janson [3, Theorem 1.1], bound C_k explicitly in terms of k, review the Landau–Kolmogorov inequality, and recall an integral equation for f that is the starting point for our lower-bound results. In Section 3 we establish the left-tail upper bounds on $|f^{(k)}|$ for $k \geq 0$ claimed in (3) and (5). In Section 4, we establish the right-tail upper bounds on $|f^{(k)}|$ for $k \geq 0$ claimed in (4) and (6). Sections 5 and 6 derive the stated lower bounds on the left and right tails, respectively, of f using an iterative approach similar to that of Janson [3] for the distribution function.

2 Preliminaries

2.1 Janson's asymptotic bounds on F

The upper bounds in the following main Theorem 1.1 of Janson [3] are used in our proof of the upper bounds in our Theorems 1.1 and 1.2.

▶ **Proposition 2.1.** *Let $\gamma := (2 - \frac{1}{\ln 2})^{-1}$. As $x \to \infty$, the limiting* QuickSort *distribution function F satisfies*

$$\exp\left[-e^{\gamma x + \ln \ln x + O(1)} \right] \leq F(-x) \leq \exp\left[-e^{\gamma x + O(1)} \right], \tag{9}$$

$$\exp[-x \ln x - x \ln \ln x + O(x)] \leq 1 - F(x) \leq \exp[-x \ln x + O(x)]. \tag{10}$$

2.2 Explicit constant bounds for absolute derivatives

We also make use of the following two results extracted from [2, Theorem 2.1 and (3.3)].

▶ **Lemma 2.2.** *Let ϕ denote the characteristic function corresponding to f. Then for every real $p \geq 0$ we have*

$$|\phi(t)| \leq 2^{p^2 + 6p} |t|^{-p} \quad \text{for all } t \in \mathbb{R}.$$

▶ **Lemma 2.3.** *For every integer $k \geq 0$ we have*

$$\sup_{x \in \mathbb{R}} |f^{(k)}(x)| \leq \frac{1}{2\pi} \int_{t=-\infty}^{\infty} |t|^k |\phi(t)| \, dt.$$

Using these two results, it is now easy to bound $f^{(k)}$.

▶ **Proposition 2.4.** *For every integer $k \geq 0$ we have*

$$\sup_{x \in \mathbb{R}} |f^{(k)}(x)| \leq 2^{k^2 + 10k + 17}.$$

Proof. For every integer $k \geq 0$ we have

$$
\begin{aligned}
\sup_{x \in \mathbb{R}} |f^{(k)}(x)| &\leq \frac{1}{2\pi} \int_{t=-\infty}^{\infty} |t|^k |\phi(t)| \, dt \\
&\leq \frac{1}{2\pi} \left[\int_{|t|>1} |t|^k |\phi(t)| \, dt + \int_{|t| \leq 1} |t|^k |\phi(t)| \, dt \right] \\
&\leq \frac{1}{2\pi} \left[\int_{|t|>1} 2^{(k+2)^2 + 6(k+2)} t^{-2} \, dt + \int_{|t| \leq 1} |t|^k \, dt \right] \\
&\leq \frac{1}{\pi} \left[2^{k^2 + 10k + 16} + \frac{1}{k+1} \right] \leq 2^{k^2 + 10k + 17},
\end{aligned}
$$

as desired. ◀

2.3 Landau–Kolmogorov inequality

For an overview of the Landau–Kolmogorov inequality, see [6, Chapter 1]. Here we state a version of the inequality well-suited to our purposes; see [5] and [9, display (21) and the display following (17)].

▶ **Lemma 2.5.** *Let $n \geq 2$, and suppose $h : (0, \infty) \to \mathbb{R}$ has n derivatives. If h and $h^{(n)}$ are both bounded, then for $1 \leq k < n$ so is $h^{(k)}$. Moreover, there exist constants $c_{n,k}$ (not depending on h) such that the supremum norm $\| \cdot \|$ satisfies*

$$\|h^{(k)}\| \leq c_{n,k} \|h\|^{1-(k/n)} \|h^{(n)}\|^{k/n}, \quad 1 \leq k < n.$$

Further, for $1 \leq k \leq n/2$ the best constants $c_{n,k}$ satisfy

$$c_{n,k} \leq n^{(1/2)[1-(k/n)]} (n-k)^{-1/2} \left(\frac{e^2 n}{4k} \right)^k \leq \left(\frac{e^2 n}{4k} \right)^k.$$

2.4 An integral equation for f

Fill and Janson [2, Theorem 4.1 and (4.2)] produced an integral equation satisfied by f, namely,

$$f(x) = \int_{u=0}^{1} \int_{z \in \mathbb{R}} f(z) f\left(\frac{x - g(u) - (1-u)z}{u} \right) \frac{1}{u} \, dz \, du. \tag{11}$$

3 Left Tail Upper Bound for Absolute Derivatives

The left-tail upper bound (5) in Theorem 1.2 can be written in the equivalent form that, for each fixed integer $k \geq 0$, we have

$$\limsup_{x \to \infty} \left(\gamma x - \ln \left[-\ln \left| f^{(k)}(-x) \right| \right] \right) < \infty, \tag{12}$$

just as Janson's left-tail upper-bound on F in (9) can be written

$$\limsup_{x \to \infty} \left(\gamma x - \ln \left[-\ln F(-x) \right] \right) < \infty. \tag{13}$$

In this section we prove (5) $\equiv$ (12) in the strengthened form LHS(3.1) $\leq$ LHS(3.2), for which the following proposition is clearly sufficient.

▶ **Proposition 3.1.** *For each fixed $k \geq 0$ we have*

$$\limsup_{x \to \infty} \left(-\ln \left[-\ln \left| f^{(k)}(-x) \right| \right] + \ln[-\ln F(-x)] \right) \leq 0. \tag{14}$$

Proof. Choosing any x and applying the Landau–Kolmogorov inequality Lemma 2.5 to the function h defined for $t \geq 0$ by $h(t) := F(-x - t)$, we find for $0 \leq k \leq (n/2) - 1$ that

$$\left| f^{(k)}(-x) \right| \leq \sup_{t \geq x} \left| f^{(k)}(-t) \right|$$

$$\leq \left[\frac{e^2 n}{4(k+1)} \right]^{k+1} [F(-x)]^{1-[(k+1)/n]} \left[\sup_{t \geq x} \left| f^{(n-1)}(-t) \right| \right]^{(k+1)/n}.$$

For $n \geq 2$ we can bound the last supremum using Proposition 2.4 simply by

$$2^{(n-1)^2 + 10(n-1) + 17} = 2^{n^2 + 8n + 8} \leq 2^{7n^2}. \tag{15}$$

Thus the argument of the lim sup in (14) can be bounded above by

$$-\ln \left[1 - \frac{k+1}{n} - \frac{\ln a_k + (k+1)(7n \ln 2 + \ln n)}{-\ln F(-x)} \right],$$

with $a_k := [e^2/(4(k+1))]^{k+1}$. Letting $n \equiv n(x) \to \infty$ with $n(x) = o(e^{\gamma x})$ and again using the upper bound from (9), the claimed inequality follows. ◀

4 Right Tail Upper Bound for Absolute Derivatives

In this section we establish the next proposition, a right-tail analogue of Proposition 3.1, which [by Janson's right-tail upper bound on F in (10)] implies the following strengthened form of (6):

$$\limsup_{x \to \infty} x^{-1} \left(x \ln x + \ln \left| f^{(k)}(x) \right| \right) \leq \limsup_{x \to \infty} x^{-1} \left(x \ln x + \ln [1 - F(x)] \right) < \infty.$$

▶ **Proposition 4.1.** *For each fixed $k \geq 0$ we have*

$$\limsup_{x \to \infty} x^{-1} \left(\ln \left| f^{(k)}(x) \right| - \ln [1 - F(x)] \right) \leq 0. \tag{16}$$

Proof. Proceeding as in the proof of Proposition 3.1, for any x and any $0 \leq k \leq (n/2) - 1$ we have

$$\left| f^{(k)}(x) \right| \leq \left[\frac{e^2 n}{4(k+1)} \right]^{k+1} [1 - F(x)]^{1 - [(k+1)/n]} \left[\sup_{t \geq x} \left| f^{(n-1)}(t) \right| \right]^{(k+1)/n};$$

we again bound the third factor by (15).

Thus the argument of the $\limsup$ in (16) can be bounded above by

$$x^{-1} \left[\frac{k+1}{n} (- \ln[1 - F(x)]) + \ln a_k + (k+1)(7n \ln 2 + \ln n) \right],$$

again with $a_k := [e^2 / (4(k+1))]^{k+1}$. Letting $n \equiv n(x)$ satisfy $n(x) = \omega(\log x)$ and $n(x) = o(x)$, and now using the right-tail *lower* bound on F from (10), the claimed inequality follows. ◄

5   Left Tail Lower Bound on f

Our iterative approach to finding the left tail lower bound on f is similar to the method used by Janson [3] for F. The following lemma gives us an inequality that is essential in this section; as we shall see, it is established from a recurrence inequality. For $z \geq 0$ define

$$m_z := \left(\min_{x \in [-z, 0]} f(x) \right) \wedge 1.$$

▶ **Lemma 5.1.** *Given $\epsilon \in (0, 1/10)$, let $a \equiv a(\epsilon) := -g\left(\frac{1}{2} - \epsilon \right) > 0$. Then for any integer $k \geq 2$ we have*

$$m_{ka} \geq \left(2\epsilon^3 m_{2a} \right)^{2^{k-2}}.$$

We delay the proof of Lemma 5.1 in order to show next how the lemma leads us to the desired lower bound in (3) on the left tail of f by using the same technique as in [3] for F.

▶ **Proposition 5.2.** *As $x \to \infty$ we have*

$$\ln f(-x) \geq -e^{\gamma x + \ln \ln x + O(1)}.$$

Proof. By Lemma 5.1, for $x > a$ we have

$$f(-x) \geq m_x \geq m\left(\left\lceil \frac{x}{a} \right\rceil a \right) \geq \left(2\epsilon^3 m_{2a} \right)^{2^{\lceil x/a \rceil - 2}} \geq \left(2\epsilon^3 m_{2a} \right)^{2^{x/a}},$$

provided ϵ is sufficiently small that $2\epsilon^3 m_{2a} < 1$. The same as Janson [3], we pick $\epsilon = x^{-1/2}$ and, setting $\gamma = (2 - \frac{1}{\ln 2})^{-1}$, get $\frac{1}{a} = \frac{\gamma}{\ln 2} + O(x^{-1})$ and

$$\begin{aligned}
\ln f(-x) &\geq 2^{\frac{\gamma}{\ln 2} x + O(1)} \cdot \ln \left(2\epsilon^3 m_{2a} \right) \\
&= e^{\gamma x + O(1)} \cdot \left(-\frac{3}{2} \ln x + \ln m_{2a} + \ln 2 \right) \\
&\geq -e^{\gamma x + \ln \ln x + O(1)}.
\end{aligned}$$
◄

Now we go back to prove Lemma 5.1:

Proof of Lemma 5.1. By the integral equation (11) satisfied by f (and symmetry in u about $u = 1/2$), for arbitrary z and a we have

$$f(-z - a) = 2 \int_{u=0}^{1/2} \int_{y \in \mathbb{R}} f(y) f\left(\frac{-z - a - g(u) - (1 - u)y}{u} \right) \frac{1}{u} \, dy \, du. \tag{17}$$

Since f is everywhere positive, we can get a lower bound on $f(-z-a)$ by restricting the range of integration in (17). Therefore,

$$f(-z-a) \geq 2 \int_{u=\frac{1}{2}-\frac{\epsilon}{2}}^{1/2} \int_{y=-z}^{-z+\epsilon^2} f(y) f\left(\frac{-z-a-g(u)-(1-u)y}{u}\right) \frac{1}{u}\, dy\, du. \tag{18}$$

We claim that in this integral region, we have $\frac{-z-a-g(u)-(1-u)y}{u} \geq -z$, which is equivalent to $y+z \leq \frac{-a-g(u)}{1-u}$. Here is a proof. Observe that when ϵ is small enough and $u \in [\frac{1}{2}-\frac{\epsilon}{2}, \frac{1}{2}]$, we have

$$\frac{-a-g(u)}{1-u} \geq \frac{g\left(\frac{1}{2}-\epsilon\right)-g\left(\frac{1}{2}-\frac{\epsilon}{2}\right)}{\frac{1}{2}+\frac{\epsilon}{2}}$$

$$\geq \frac{\frac{\epsilon}{2}\left|g'\left(\frac{1}{2}-\frac{\epsilon}{2}\right)\right|}{\frac{1}{2}+\frac{\epsilon}{2}} = \frac{\epsilon}{1+\epsilon}\left|2\ln\left(1-\frac{2\epsilon}{1+\epsilon}\right)\right|$$

$$\geq \frac{4\epsilon^2}{(1+\epsilon)^2} \geq \epsilon^2.$$

Also, in this integral region we have $y+z \leq \epsilon^2$. So we conclude that $y+z \leq \frac{-a-g(u)}{1-u}$.

Next, we claim that $\frac{-z-a-g(u)-(1-u)y}{u} \leq 0$ in this integral region if z is large enough. Here is a proof. Let $\frac{-z-a-g(u)-(1-u)y}{u} = -z+\delta$ with $\delta \geq 0$. Then in the integral region we have $0 \leq y+z = \frac{-a-g(u)-u\delta}{1-u}$. Therefore

$$\delta \leq \frac{-a-g(u)}{u} \leq \frac{-a-g\left(\frac{1}{2}\right)}{\frac{1}{2}-\frac{\epsilon}{2}} = \frac{2}{1-\epsilon}\left[g\left(\frac{1}{2}-\epsilon\right)-g\left(\frac{1}{2}\right)\right]$$

$$\leq \frac{2\epsilon}{1-\epsilon}\left|2\ln\left(1-\frac{4\epsilon}{1+2\epsilon}\right)\right|$$

$$\leq 19\epsilon^2,$$

where the last inequality can be verified to hold for $\epsilon < 1/10$. That means if we pick z large enough, for example, $z \geq 20\epsilon^2$, then $\frac{-z-a-g(u)-(1-u)y}{u} = -z+\delta$ will be negative. It can also be verified that $a \geq 30\epsilon^2$ for $\epsilon < 1/10$.

Now consider $\epsilon < 1/10$, an integer $k \geq 3$, $z \in [(k-2)a, (k-1)a]$, and $x = z+a \in [(k-1)a, ka]$. Noting $z \geq a \geq 30\epsilon^2 > 20\epsilon^2$, by (18) we have

$$f(-x) \geq 2 \cdot \frac{\epsilon}{2} \cdot m_z^2 \cdot \epsilon^2 \cdot 2 \geq 2\epsilon^3 m_{(k-1)a}^2.$$

Further, for $x \in [0, (k-1)a]$ we have

$$f(-x) \geq m_{(k-1)a} > 2\epsilon^3 m_{(k-1)a}^2$$

since $2\epsilon^3 < 1$ and $m_{(k-1)a} \leq 1$ by definition. Combine these two facts, we can conclude that for $x \in [0, ka]$ we have $f(-x) \geq 2\epsilon^3 m_{(k-1)a}^2$. This implies the recurrence inequality

$$m_{ka} \geq 2\epsilon^3 m_{(k-1)a}^2.$$

The desired inequality follows by iterating:

$$m_{ka} \geq \left(2\epsilon^3\right)^{2^{k-2}-1} m_{2a}^{2^{k-2}} \geq \left(2\epsilon^3 \cdot m_{2a}\right)^{2^{k-2}}. \qquad \blacktriangleleft$$

6   Right Tail Lower Bound on f

Once again we use an iterative approach to derive our right-tail lower bound. The following key lemma is established from a recurrence inequality. Define

$$c := 2[F(1) - F(0)] \in (0, 2)$$

and

$$m_z := \min_{x \in [0,z]} f(x), \quad z \geq 0.$$

▶ **Lemma 6.1.** *Suppose $b \in [0, 1)$ and that $\delta \in (0, 1/2)$ is sufficiently small that $g(\delta) \geq b$. Then for any integer $k \geq 1$ satisfying*

$$2 + (k - 1)b \leq [g(\delta) - b]/\delta$$

we have

$$m_{2+kb} \geq (c\delta)^{k-1} m_3.$$

We delay the proof of Lemma 6.1 in order to show next how the lemma leads us to the desired lower bound in (4) on the right tail of f.

▶ **Proposition 6.2.** *As $x \to \infty$ we have*

$$f(x) \geq \exp[-x \ln x - x \ln \ln x + O(x)].$$

Proof. Given $x \geq 3$ suitably large, we will show next that we can apply Lemma 6.1 for suitably chosen $b > 0$ and δ and $k = \lceil (x - 2)/b \rceil \geq 2$. Then, by the lemma,

$$f(x) \geq m_{2+kb} \geq (c\delta)^{k-1} m_3 \geq (c\delta)^{(x-2)/b} m_3, \tag{19}$$

and we will use (19) to establish the proposition.

We make the same choices of δ and b as in [3, Sec. 4], namely, $\delta = 1/(x \ln x)$ and $b = 1 - (2/\ln x)$. To apply Lemma 6.1, we need to check that $g(\delta) \geq b$ and $2 + (k-1)b \leq [g(\delta) - b]/\delta$, for the latter of which it is sufficient that $x \leq [g(\delta) - b]/\delta$. Indeed, if x is sufficiently large, then

$$g(\delta) \geq 1 + 3\delta \ln \delta = 1 - \frac{3}{x \ln x}(\ln x + \ln \ln x) \geq 1 - \frac{4}{x},$$

where the elementary first inequality is (4.1) in [3], and so

$$g(\delta) - b \geq \frac{2}{\ln x} - \frac{4}{x} \geq \frac{1}{\ln x} > 0$$

and

$$\frac{g(\delta) - b}{\delta} \geq \frac{1/\ln x}{1/(x \ln x)} = x.$$

Finally, we use (19) to establish the proposition. Indeed,

$$
\begin{aligned}
-\ln f(x) &\leq \tfrac{x-2}{b} \ln(\tfrac{1}{c\delta}) - \ln m_3 \\
&\leq \tfrac{x}{1-(2/\ln x)}[\ln(x \ln x) + \ln(\tfrac{1}{c})] - \ln m_3 \\
&= \tfrac{x}{1-(2/\ln x)} \ln(x \ln x) + O(x).
\end{aligned}
$$

But

$$\frac{x}{1 - (2/\ln x)} \ln(x \ln x)$$

$$= x \left[1 + \frac{2}{\ln x} + O\left(\frac{1}{(\log x)^2}\right) \right] (\ln x + \ln \ln x)$$

$$= (x \ln x) \left[1 + \frac{2}{\ln x} + O\left(\frac{1}{(\log x)^2}\right) \right] \left(1 + \frac{\ln \ln x}{\ln x} \right)$$

$$= (x \ln x) \left[1 + \frac{\ln \ln x}{\ln x} + \frac{2}{\ln x} + \frac{2 \ln \ln x}{(\ln x)^2} + O\left(\frac{1}{(\log x)^2}\right) \right]$$

$$= x \ln x + x \ln \ln x + 2x + \frac{2x \ln \ln x}{\ln x} + O\left(\frac{x}{\log x}\right)$$

$$= x \ln x + x \ln \ln x + O(x).$$

So

$$- \ln f(x) \le x \ln x + x \ln \ln x + O(x),$$

as claimed. ◀

Now we go back to prove Lemma 6.1, but first we need two preparatory results.

▶ **Lemma 6.3.** *Suppose $z \ge 2$, $b \ge 0$, and $\delta \in (0, 1/2)$ satisfy $g(\delta) \ge b$ and $z \le [g(\delta) - b]/\delta$. Then*

$$f(z + b) \ge c\,\delta\,m_z.$$

Proof. By the integral equation (11) satisfied by f (and symmetry in u about $u = 1/2$), for arbitrary z and b we have

$$f(z + b) = 2 \int_{u=0}^{1/2} \int_{y \in \mathbb{R}} f(y) f\left(\frac{z + b - g(u) - (1 - u)y}{u} \right) \frac{1}{u}\, dy\, du.$$

Since f is positive everywhere, a lower bound on $f(z + b)$ can be achieved by shrinking the region of integration:

$$f(z + b) \ge 2 \int_{u=0}^{\delta} \int_{y=0}^{z} f(y) f\left(\frac{z + b - g(u) - (1 - u)y}{u} \right) \frac{1}{u}\, dy\, du$$

$$\ge 2 m_z \int_{u=0}^{\delta} \int_{y=0}^{z} f\left(\frac{z + b - g(u) - (1 - u)y}{u} \right) \frac{1}{u}\, dy\, du$$

$$= 2 m_z \int_{u=0}^{\delta} \int_{\xi = z + \frac{b-g(u)}{u}}^{\frac{z+b-g(u)}{u}} f(\xi) \frac{1}{1 - u}\, d\xi\, du. \tag{20}$$

The equality comes from a change of variables. We next claim that the integral of integration for ξ contains $(0, z - 1)$, and then the desired result follows. Indeed, if $u \in (0, \delta)$ and $\xi \in (0, z - 1)$ then

$$\xi < z - 1 < \tfrac{z-1}{u} \le \tfrac{z+b-g(u)}{u},$$

where the last inequality holds because $b \ge 0$ and $g(u) \le 1$; and, because $g(u) \ge g(\delta)$ and $g(\delta) \ge b$ and $z \le [g(\delta) - b]/\delta$, we have

$$\xi > 0 = z + \tfrac{b-g(u)}{u} - [z + \tfrac{b-g(u)}{u}] \ge z + \tfrac{b-g(u)}{u} - [z + \tfrac{b-g(\delta)}{u}]$$

$$\ge z + \tfrac{b-g(u)}{u} - [z + \tfrac{b-g(\delta)}{\delta}] \ge z + \tfrac{b-g(u)}{u}. \qquad ◀$$

▶ **Lemma 6.4.** *Suppose $b \geq 0$ and that $\delta \in (0, 1/2)$ is sufficiently small that $g(\delta) \geq b$. Then for any integer $k \geq 2$ satisfying*

$$2 + (k-1)b \leq [g(\delta) - b]/\delta$$

we have

$$m_{2+kb} \geq c\,\delta\,m_{2+(k-1)b}.$$

Proof. For $y \in [2 + (k-1)b, 2 + kb]$, application of Lemma 6.3 with $z = y - b$ yields

$$f(y) \geq c\,\delta\,m_{y-b} \geq c\,\delta\,m_{2+(k-1)b}.$$

Also, for $y \in [0, 2 + (k-1)b]$ we certainly have

$$f(y) \geq m_{2+(k-1)b} > c\,\delta\,m_{2+(k-1)b}.$$

The result follows. ◀

We are now ready to complete this section by proving Lemma 6.1.

Proof of Lemma 6.1. By iterating the recurrence inequality of Lemma 6.4, it follows that

$$m_{2+kb} \geq (c\,\delta)^{k-1}m_{2+b}.$$

Lemma 6.1 then follows since $b < 1$. ◀

───── **References** ─────

1 James Allen Fill and Svante Janson. A characterization of the set of fixed points of the Quicksort transformation. *Electron. Comm. Probab.*, 5:77–84, 2000. `doi:10.1214/ECP.v5-1021`.

2 James Allen Fill and Svante Janson. Smoothness and decay properties of the limiting quicksort density function. *Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities, Trends in Mathematics*, pages 53–64, 2000.

3 Svante Janson. On the tails of the limiting quicksort distribution. *Electronic Communications in Probability*, 20, 2015.

4 Charles Knessl and Wojciech Szpankowski. Quicksort algorithm again revisited. *Discrete Mathematics and Theoretical Computer Science*, 3(2), 1999.

5 A. P. Matorin. On inequalities between the maxima of the absolute values of a function and its derivatives on a half-line. *Ukrain. Mat. Ž.*, 7:262–266, 1955.

6 D. S. Mitrinović, J. E. Pečarić, and A. M. Fink. *Inequalities involving functions and their integrals and derivatives*, volume 53 of *Mathematics and its Applications (East European Series)*. Kluwer Academic Publishers Group, Dordrecht, 1991. `doi:10.1007/978-94-011-3562-7`.

7 Mireille Régnier. A limiting distribution for quicksort. *RAIRO-Theoretical Informatics and Applications*, 23(3):335–343, 1989.

8 Uwe Rösler. A limit theorem for "quicksort". *RAIRO-Theoretical Informatics and Applications*, 25(1):85–100, 1991.

9 Sergei Borisovich Stechkin. Inequalities between the upper bounds of the derivatives of an arbitrary function on the half-line. *Mathematical notes of the Academy of Sciences of the USSR*, 1(6):442–447, 1967.

Stationary Distribution Analysis of a Queueing Model with Local Choice

Plinio S. Dester[1]

INRIA Paris, 2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12, France
plinio.santini-dester@polytechnique.edu

Christine Fricker

Département d'informatique de l'ENS, École normale supérieure, CNRS, PSL Research University, 75005 Paris, France,
INRIA Paris, 2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12, France
Christine.Fricker@inria.fr

Hanene Mohamed[2]

Université Paris Nanterre, Modal'X, UPL, 92000 Nanterre, France
hanene.mrad@u-parisnanterre.fr

─── Abstract ───

The paper deals with load balancing between one-server queues on a circle by a local choice policy. Each one-server queue has a Poissonian arrival of customers. When a customer arrives at a queue, he joins the least loaded queue between this queue and the next one, ties solved at random. Service times have exponential distribution. The system is stable if the arrival-to-service rate ratio called load is less than one. When the load tends to zero, we derive the first terms of the expansion in this parameter for the stationary probabilities that a queue has 0 to 3 customers. We investigate the error, comparing these expansion results to numerical values obtained by simulations. Then we provide the asymptotics, as the load tends to zero, for the stationary probabilities of the queue length, for a fixed number of queues. It quantifies the difference between policies with this local choice, no choice and the choice between two queues chosen at random.

1 Introduction

1.1 A load balancing policy

The paper deals with the impact of choice between two neighbors in a large set of queues. Load balancing is present in a wide literature and includes various policies as choice, offloading, redundancy or work stealing ([6], [17], [8] and others) for example. The two-choice policy is a well-known distributed way to improve load balancing. See [14] and [12] for one-server

---

29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018).
Editors: James Allen Fill and Mark Daniel Ward; Article No. 22; pp. 22:1–22:18
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

queues. For this policy, the arriving customers choose two queues at random and join the shortest one, ties being solved at random. The paper focuses on the case where only local choice can be processed. This case occurs in many applications with geographical constraints, like vehicle-sharing systems or cloud computing.

1.2   The model

The model we present is called *local choice model*. It consists in a set of N one-server queues with infinite capacity where customers arrive at each queue according to independent Poisson processes with rate λ, which means that inter-arrival times are independent with exponential distribution with parameter λ. When a customer arrives at queue i, $1 \leq i \leq N$, he chooses between queues i and $i + 1$ the least loaded one and joins it. By convention, queue $N + 1$ is queue 1. If queues i and $i + 1$ have the same number of customers, he joins one of these queues with probability $1/2$. The service times are iid with exponential distribution with parameter μ. When the customer is served, he leaves the system. All inter-arrival and service times are independent. The load ρ is by definition λ/μ.

1.3   The problem

The main issue addressed in the paper concerns the marginal distribution of the number of customers in one queue at equilibrium for the *local choice model*. We investigate the asymptotics of the stationary probabilities for one queue as the load tends to zero. The number N of queues is fixed throughout the paper. We compare them with the same quantities for the *random choice model*, where an arriving customer chooses two queues at random and joins the least loaded one and *the no choice model*, where a customer who arrives at queue i is served at this queue.

The *no choice model* is simply N independent $M/M/1$ queues. The *random choice model* is classical, see [14] and [12]. For $\rho < 1$, the limiting stationary tail probability, i.e. the limit as N gets large of the stationary probability that a queue has more than k customers, is doubly exponentially decreasing, more precisely is ρ^{2^k-1}, $k \geq 0$. This doubly exponential decrease is known in the literature as *the power of choice*. Indeed it is much smaller than the tail probability ρ^k, $k \geq 0$ in the *no choice model* as in the $M/M/1$ queue, the queue length stationary distribution is geometric with parameter ρ. What is this tail probability for the *local choice model*?

1.4   The results

They concern the *local choice model* previously described. In the paper, N is fixed and $\rho < 1$ to ensure the ergodicity of the queue length process. We consider the stationary probabilities as analytical functions of parameter ρ. Based on some crucial arguments (see Lemmas 2 and 3), an induction procedure provides all the terms of the power series expansion. We apply this procedure to find the first terms explicitly. Then, in the study of the marginal distribution for one queue, it gives the first terms (at order 6) of the expansion in ρ of the stationary probability that a queue has m customers, for small m ($m \leq 3$). This expansion is an approximation for the stationary probability at light traffic, which is compared to simulations.

The main result of the paper gives the asymptotics as ρ tends to 0 of the stationary probability that a queue has m customers, for any m. It is claimed in Proposition 8 that these asymptotics are $2\rho^{2m-1}$ for $N = 2$ and $12(\rho/2)^{2m-1}$ for $N \geq 3$. It gives the rate of

decay in parameter ρ of the stationary queue length at light traffic, which is ρ^2 for $N = 2$ and $(\rho/2)^2$ for $N \geq 3$. In other words, compared to the N independent $M/M/1$ queues, the *local choice model* does not lead an improvement as large as in the *random choice model* which is doubly exponential.

1.5 Related work

The choice between two queues at random among N one-server queues is well understood for N large via mean-field method for the late 90's with [14] and [12] and knows a great interest in literature. Nevertheless, local choice is a quite challenging open problem in queueing theory. As far as we know, very few papers investigate the problem. For this model, where the underlying graph is linear, more precisely a circle, and more generally for a graph $G = (V, E)$, [7] gives an approximation of the steady-state queue length distribution which seems numerically accurate compared to simulations. This approximation, called pair-approximation, is obtained from the empirical measure on pairs of neighbors. It is a mean field limit as the graph gets large. But this limit, solution of an ODE, is hard to study analytically. In [7], the expression of the ODE is explicitly given, but its equilibrium point is investigated by numerical simulations.

The series expansion of the stationary probabilities in parameter ρ is the key tool in [2] for the study of the *JSQ model*. It is the classical model of N queues, where arriving customers join the shortest queue among all the queues. The paper gives asymptotics in light traffic for the mean and the variance of the total number of customers at equilibrium. Nevertheless the method to obtain them is quite different.

1.6 Related models

Some papers deal with such models, but without departure. They are called urn models in computer science literature, and deposition models or crystal growth models in statistical physics. The problems addressed in both cases are quite different.

THE URN MODEL. Urns are put at vertices of a finite graph $G = (V, E)$ with $|V| = N$. Arrival of balls are associated to edges. For each ball, an edge is chosen at random and the ball is put in the least loaded of the two end-points of the edge. The problem of the maximum number of balls per urn for N balls in N urns is investigated. The conclusion is that the *power of choice* does not hold for d-regular graphs, d constant, as this maximum is not in $\log \log N$ (see [10], also [3] and references therein). But the main difference with our study is that we deal with the stationary regime. The poor load balancing result in the urn problem might come from the fact that with N balls in N urns, the equilibrium is not reached.

THE CRYSTAL GROWTH MODEL. In this model, consider N sites $1, \ldots, N$. There is also no departure. Particles arrive at each site, say i, at rate λ. If the two (respectively just one or none) neighboring sites $i - 1$ and $i + 1$ have more particles than site i, the arrival rate at the site i is β_2, β_1 and β_0, respectively. [9, 1, 5] give ergodic conditions for the shape process, which is Markov. Our arrival process is a variant of this model in the special case where $\beta_0 = 0$ and $\beta_2 = 2\beta_1$ (see Section 2 for details). Note that if we extend the *local choice model*, to the case where the customers, arriving at queue i, choose between the two neighboring queues i and $i + 1$ with some probability α and do not choose otherwise, it will still fit in this framework as a variant, but with $\beta_0 \neq 0$.

1.7 Outline

The paper is organized as follows. Section 2 gives the model description and the notations. Section 3 leads to the induction procedure to obtain the power series expansion of the stationary state probabilities of the model. Section 4 deals with the results for the marginal distribution for one queue.

2 Model description and notations

Consider a system of N queues with infinite capacities, each of them served by one server at rate μ. In all the following, queue $N + 1$ means queue 1. The arrival rate at each queue is λ but the arriving customer at queue i joins the least loaded queue between queues i and $i + 1$, ties being solved at random. All inter-arrival and service times are independent with exponential distribution. The i.i.d. Bernoulli variables with parameter $1/2$ introduced to solve the ties are independent of the previous random variables. By definition, $\rho = \lambda/\mu$.

2.1 The state process

For $1 \leq i \leq N$, let $X_i(t)$ be the number of customers at queue i at time t and $X(t) = (X_i(t))_{1 \leq i \leq N}$. The queue length process $(X(t))_{t \geq 0}$ is a Markov process on state space $\mathbb{N}^N$ with Q-matrix Q, given for $n = (n_1, \ldots, n_N)$ here and in all the following, by its non-negative components, for $1 \leq i \leq N$,

$$Q(n, n + e_i) = \lambda c_i(n)$$
$$Q(n, n - e_i) = \mu \mathbb{1}_{n_i > 0}$$

where $c : \mathbb{N} \times \mathbb{N}^N \longrightarrow \mathbb{R}_+$, called the *contribution function*, quantifies the amount of arrivals at the different queues and $(e_i)_{1 \leq i \leq N}$ is the canonical basis of $\mathbb{R}^N$.

For our *local choice model*, this contribution function is called *local choice function* and is denoted by c^{lc}. Function c^{lc} at queue i, depends only on the state of this queue and the two neighboring queues $i - 1$ and $i + 1$ and is defined by

$$c_i^{lc}(n) = d(n_i, n_{i+1}) + d(n_i, n_{i-1}) \text{ where } d(k, l) = \frac{1}{2} \mathbb{1}_{\{k=l\}} + \mathbb{1}_{\{k<l\}} \tag{1}$$

with, by convention, $n_0 = n_N$ and $n_{N+1} = n_1$. Dispatching function d is the basis of our *local choice model* since it implements the load balancing policy: join the least loaded among two neighboring queues.

▶ Remark. The local choice function c^{lc} can also be defined by

$$c_i^{lc}(n) = \omega(\Delta_{i-1}n, -\Delta_i n),$$

in terms of the shape function Δ defined by $n \mapsto \Delta n = (\Delta_1 n, \ldots, \Delta_N n)$ where $\Delta_j n = n_j - n_{j+1}$, $1 \leq j \leq N$ and the so-called deposition function ω given by

$$\omega(a, b) = \frac{1}{2} \left(\mathbb{1}_{\{a=0\}} + \mathbb{1}_{\{b=0\}} \right) + \mathbb{1}_{\{a>0\}} + \mathbb{1}_{\{b>0\}}, \ a, b \in \mathbb{Z}. \tag{2}$$

Note that the Gates-Wescott process studied in [5] is the shape process $(\Delta X(t))$ for the model without departure associated to the following deposition function

$$\omega(a, b) = \beta_{\mathbb{1}_{\{a>0\}} + \mathbb{1}_{\{b>0\}}}, \ a, b \in \mathbb{Z}, \tag{3}$$

with β_0, β_1 and $\beta_2 > 0$.

3 An algorithm to compute the stationary distribution

In this section we study, for N fixed, the queue length process $(X(t))$ at stationarity. We prove first that $(X(t))$ is ergodic for $\rho < 1$ if $c = c^{lc}$. See Proposition 1. For a general c, if $(X(t))$ is ergodic, it has a unique invariant measure $y = (y_n, \ n \in \mathbb{N}^N)$ on $\mathbb{N}^N$, solution of the global balance equations

$$\sum_{n' \in \mathbb{N}^N} y(n')Q(n', n) = 0, \quad n \in \mathbb{N}^N. \tag{4}$$

Our aim is not to solve these equations but rather look for an analytical solution for y of the form

$$y_n(\rho) = \sum_{k \geq 0} \alpha_k(n)\rho^k, \ n \in \mathbb{N}^N.$$

Assuming the existence of $\varepsilon > 0$ such that the solution of the global balance equations $y_n(\rho)$ has a serie expansion for $0 < \rho < \varepsilon$, we prove that each α_k, $k \geq 0$, has a finite support. See Lemma 4. Then we explain the algorithm to obtain by induction on $k \geq 0$ the explicit expressions of α_k and compute explicitely the first terms.

3.1 Ergodicity for c^{lc}

For local choice, contribution function c^{lc} is given by equation (1). The following result gives us the necessary and sufficient condition for ergodicity of the Markov state process $(X(t))$ in this case.

▶ **Proposition 1** (Ergodicity). *For $c = c^{lc}$, the Markov process $(X(t))_{t \geq 0}$ is ergodic if $\rho < 1$ and transient if $\rho > 1$.*

The proof based on Foster's criterion is postponed in Appendix.

3.2 Power series expansion in ρ of the stationary probabilities

For ρ such that $(X(t))$ is ergodic, let $y(\rho) = (y_n(\rho), n \in \mathbb{N}^N)$ be its invariant measure, the unique solution of the global balance equations

$$\left(\sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} + \rho \sum_{i=1}^{N} c_i(n) \right) y_n(\rho) =$$

$$\sum_{i=1}^{N} y_{n+e_i}(\rho) + \rho \sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} c_i(n - e_i) y_{n-e_i}(\rho), \quad n \in \mathbb{N}^N \tag{5}$$

obtained by plugging the expression of Q in equation (4).

We look for an invariant measure $(y_n, n \in \mathbb{N}^N)$ satisfying the following condition.

($\mathbf{H_0}$) There exists $\varepsilon > 0$, such that, for $\rho \in [0, \varepsilon[$ and $n \in \mathbb{N}^N$,

$y_n(\rho)$ can be written as a series expansion of the form

$$y_n(\rho) = \sum_{k \geq 0} \alpha_k(n)\rho^k. \tag{6}$$

▶ Remark. In all the following, (H_0) will be assumed. This question of analyticity of stationary probabilities of a family of Markov chains depending on one parameter is the major issue addressed by [16, Chapter IV] (see also [15, Chapter 7]). The main tool for proving such analyticity is the Lyapunov function in Foster's criterion for ergodicity. We get a quadratic Lyapunov function to prove the ergodicity in Proposition 1 (see the proof in Appendix A). But the dynamics of our model do not allow to apply the results of [16, Chapter IV] or [15, Chapter 7], due to the contribution function part. This question is the object of future work.

Moreover the following technical assumption

$$\sum_{k \geq 0} \sum_{n \in \mathbb{N}^N} \alpha_k(n)\rho^k = \sum_{n \in \mathbb{N}^N} \sum_{k \geq 0} \alpha_k(n)\rho^k, \quad \rho < \varepsilon \tag{7}$$

is used.

▶ Remark. According to Proposition 1, for $c = c^{lc}$, as analyticity requires the existence of the stationary measure, thus implicitly the ergodicity of process $(X(t))$, it holds that $\varepsilon \leq 1$. Note that assumption (H_0) could have been written with 1 instead of ε. We introduce ε in (H_0) of this form because some results in the following apply for more general c than c^{lc}, where the ergodicity condition can be written $\rho < \varepsilon$.

Under assumption (H_0), for each $n \in \mathbb{N}^N$, $\rho \mapsto y_n(\rho)$ is C^∞ on $[0, \varepsilon[$ and $\alpha_k(n) = y_n^{(k)}(0)/k!$. Taking the k-th derivative in the global balance equations (5) with respect to ρ and evaluating it at $\rho = 0$, it holds that, for any $n \in \mathbb{N}^N$ and $k \in \mathbb{N}^*$,

$$\left(\sum_{i=1}^N \mathbb{1}_{\{n_i > 0\}}\right) \alpha_k(n) =$$

$$\sum_{i=1}^N \alpha_k(n + e_i) + \sum_{i=1}^N \mathbb{1}_{\{n_i > 0\}} c_i(n - e_i)\alpha_{k-1}(n - e_i) - \left(\sum_{i=1}^N c_i(n)\right)\alpha_{k-1}(n). \tag{8}$$

3.3 Some crucial lemmas

Equation (8) allows us to prove that, for k fixed, α_k has a finite support. It is the purpose of Lemma 4. For that, we need to prove the two following technical lemmas. Lemma 2, formulated with α for sake of simplicity, will be applied for each $\alpha_k, k \geq 1$. Before that, let us introduce the following set

$$\mathcal{A}_k \stackrel{def}{=} \{n \in \mathbb{N}^N, n_1 + n_2 + \ldots + n_N = k\}, \ k \in \mathbb{N}. \tag{9}$$

▶ **Lemma 2.** *Let* $\alpha : \mathbb{N}^N \longrightarrow \mathbb{R}$ *and* $k_0 \in \mathbb{N}^*$ *be such that, for* $n = (n_1, \ldots, n_N)$ *with* $|n| = n_1 + \cdots + n_N > k_0$,
  (i) $\alpha(n) \geq 0$,
  (ii) *the following recurrence equation holds,*

$$\left(\sum_{i=1}^N \mathbb{1}_{\{n_i > 0\}}\right) \alpha(n) = \sum_{i=1}^N \alpha(n + e_i), \tag{10}$$

  (iii) $\sum_{n, |n| > k_0} \alpha(n) < \infty$
then, for all n *such that* $|n| > k_0$, $\alpha(n) = 0$.

Proof. Let $k_0 \in \mathbb{N}^*$ be fixed. First, we claim that, for any $k > k_0$,

$$\sum_{n \in \mathcal{A}_k} \sum_{i=1}^{N} \alpha(n + e_i) = \sum_{n \in \mathcal{A}_{k+1}} \left(\sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} \right) \alpha(n). \tag{11}$$

Indeed, for $n \in \mathcal{A}_{k+1}$, for i such that $n_i \neq 0$, $\alpha(n)$ can be written as $\alpha(\hat{n} + e_i)$, for a unique $\hat{n} \in \mathcal{A}_k$. The number of elements in $\mathcal{A}_k$ that can generate n when we add them to e_i is exactly equal to the number of non-zero coordinates n_i of n, $1 \leq i \leq N$. Therefore, equation (11) holds.

Then we replace $\sum_{i=1}^{N} \alpha(n + e_i)$ in the left-hand side of (11) by the left-hand side of equation (10). It yields, for any $k > k_0$,

$$\sum_{n \in \mathcal{A}_k} \left(\sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} \right) \alpha(n) = \sum_{n \in \mathcal{A}_{k+1}} \left(\sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} \right) \alpha(n).$$

Thus, for any $k > k_0$,

$$\sum_{n \in \mathcal{A}_k} \left(\sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} \right) \alpha(n) = C \tag{12}$$

where C is non-negative due to (i) and independent of k. As $\sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} \leq N$,

$$\sum_{n \in \mathcal{A}_k} N \alpha(n) \geq C.$$

If $C > 0$, $\sum_{k > k_0} \sum_{n \in \mathcal{A}_k} \alpha(n)$ diverges, since $\sum_{n \in \mathcal{A}_k} \alpha(n) \geq \dfrac{C}{N} > 0$. But, this contradicts the fact that $\sum_{n, |n| > k_0} \alpha(n) < \infty$. Thus $C = 0$. Using the fact that $\alpha(n) \geq 0$ in equation (12), for all n such that $|n| > k_0$, $\alpha(n) = 0$.   ◀

The following lemma is a key argument for both computing the $\alpha_k(n)$ (see Section 3.4) and in the proof of Lemma 4.

▶ **Lemma 3.** *The following property holds:*

$$\sum_{n \in \mathbb{N}^N} \alpha_k(n) = 0, \quad k > 0. \tag{13}$$

Proof. Permuting the sums, by equation (7), it holds that, for $\rho < \varepsilon$,

$$\sum_{k \geq 0} \left(\sum_{n \in \mathbb{N}^N} \alpha_k(n) \right) \rho^k = \sum_{n \in \mathbb{N}^N} \left(\sum_{k \geq 0} \alpha_k(n) \rho^k \right) = \sum_{n \in \mathbb{N}^N} y_n(\rho) = 1.$$

as $y(\rho) = (y_n(\rho), n \in \mathbb{N}^N)$ is a probability measure. The left-hand side of this equation is a power series whose all the terms except the first one are null. It ends the proof.   ◀

We can now prove the following result.

▶ **Lemma 4.** *Let $k \in \mathbb{N}$. For all n, $|n| > k$, $\alpha_k(n) = 0$.*

Proof. We prove this assertion by induction on k. Take $k = 0$. From equation (6), $y_n(0) = \alpha_0(n)$. As the invariant measure for $\rho = 0$ (no arrival) is $y_n(0) = \delta_{0_N}$, the Dirac mass at $(0, \ldots, 0) \in \mathbb{N}^N$ denoted by 0_N, the assertion is true for $k = 0$. Let $k \in \mathbb{N}$ be fixed. If we suppose that the assertion holds for $k' \leq k$, then Lemma 2, applied to $\alpha = \alpha_{k+1}$ and $k_0 = k + 1$, guarantees that the assertion is true for $k' = k + 1$. Indeed let us check assertions (i), (ii) and (iii) for all n with $|n| > k + 1$. Let such a n be fixed. In equation (8), $\alpha_{(k+1)-1}(n) = \alpha_{(k+1)-1}(n - e_i) = 0$ since $|n| > k$ and $|n - e_i| > k$ and induction assumption. Therefore equation (8) is rewritten as equation (10), giving (ii). Moreover, by induction assumption, in equation (6), $\alpha_{k+1}(n)$ represents the first possible non-zero coefficient for $y_n(\rho)$. This coefficient $\alpha_{k+1}(n) \geq 0$, because otherwise, as

$$y_n(\rho) \sim_{\rho \to 0} \alpha_{k+1}(n)\rho^{k+1},$$

it would exist ρ such that $y_n(\rho) < 0$, which is false as $y(\rho)$ is a probability measure. It gives (i). Eventually, by equation (13), $\sum_{n=(n_1, \ldots, n_N)} \alpha_{k+1}(n) = 0$ and, as $\sum_{|n| \leq k+1} \alpha_{k+1}(n)$ is finite, then $\sum_{|n| > k+1} \alpha_{k+1}(n)$ is finite too, which is (iii). ◀

3.4   Induction procedure

The algorithm to obtain all the coefficients $\alpha_k(n)$ is an induction procedure on $k \geq 0$. We use that $\alpha_0 = \delta_{0_N}$ and key equation (8). For $k \geq 1$, assume that we know the coefficients $\alpha_{k-1}(n)$, for all $n \in \mathbb{N}^N$ and find the coefficients $\alpha_k(n)$, $n \in \mathbb{N}^N$. First, by Lemma 4, $\alpha_k(n) = 0$ for $|n| > k$. Second we derive each coefficient $\alpha_k(n)$ for $n \in \mathcal{A}_k$, defined by equation (9), as the left-hand side of equation (8). Indeed, in the right-hand side of the same equation, the first term is null due to Lemma 4. The other terms are known as coefficients for $k - 1$. By the same procedure, we compute the $\alpha_k(n)$ for $n \in \mathcal{A}_{k-1}$: Since $n + e_i \in \mathcal{A}_k$, we still know also the first term of the right-hand side of equation (8). Then we determine the coefficients for $n \in \mathcal{A}_{k-2}$, $n \in \mathcal{A}_{k-3}$ and so on, until $n \in \mathcal{A}_1$. It remains to compute the last coefficient $\alpha_k(0_N)$. It is given by the additional equation (13) in Lemma 3.

▶ Remark. For Lemma 4 and the previous induction procedure, we do not use the specific expression (1) of contribution function c. We just choose ρ in the domain of analyticity of the $y_n, n \in \mathbb{N}^N$. What follows remains valid for a general contribution function c satisfying the following additional assumptions

(**H$_1$**) For $n \in \mathbb{N}^N$, $c_1(n) + \cdots + c_N(n) = N$.

(**H$_2$**) c is invariant by cyclic permutation or reflection (reverse order).

More precisely, the second assumption means that, for such a permutation σ on $\{1, 2, \ldots, N\}$, for $n \in \mathbb{N}^N$ and $1 \leq i \leq N$, $c_{\sigma(i)}(\sigma(n)) = c_i(n)$. These assumptions are obviously true for the local choice function $c = c^{lc}$ defined by equation (1).

3.5   Deriving the first terms

Let us derive the coefficients until order 3 under (H_0), (H_1) and (H_2). It is given by the following proposition.

▶ **Proposition 5.** *For $k = 0$,*

$$\alpha_0(n) = \mathbb{1}_{\{n=0_N\}}. \tag{14}$$

For $k = 1$,

$$\begin{cases} \alpha_1(0_N) & = -N, \\ \alpha_1(e_i) & = 1,\ 1 \le i \le N \\ \alpha_1(n) & = 0\ otherwise. \end{cases} \tag{15}$$

For $k = 2$, for $i, j \in \{1, 2, ..., N\}$,

$$\begin{cases} \alpha_2(0_N) & = \frac{1}{2}(N^2 - Nc_1(e_1)), \\ \alpha_2(e_i) & = -N, \\ \alpha_2(e_i + e_j) & = c_i(e_j), \\ \alpha_2(n) & = 0\ otherwise. \end{cases} \tag{16}$$

For $k = 3$, for all $i, j, l \in \{1, 2, ..., N\}$, $i \ne j$, $j \ne l$ and $l \ne i$,

$$\begin{cases} \alpha_3(0_N) & = -\sum_{n \ne 0_N} \alpha_3(n) \\ \alpha_3(e_i) & = \frac{1}{2}(N^2 - Nc_1(e_1)) \\ \alpha_3(e_i + e_j) & = \frac{1}{2}\left(\sum_{v=1}^{N} \alpha_3(e_i + e_j + e_v) - 3Nc_i(e_j)\right), \\ \alpha_3(2e_i) & = \frac{1}{2}\left(\sum_{v=1}^{N} c_i(e_v)c_i(e_i + e_v) - 3Nc_i(e_i)\right), \\ \alpha_3(e_i + e_j + e_l) & = \frac{1}{3}(c_i(e_j)c_l(e_i + e_j) + c_j(e_l)c_i(e_j + e_l) \\ & \quad + c_l(e_i)c_j(e_l + e_i)), \\ \alpha_3(2e_i + e_j) & = \frac{1}{2}\left(c_i(e_j)c_i(e_i + e_j) + c_i(e_i)c_j(2e_i)\right), \\ \alpha_3(3e_i) & = c_1(e_1)c_1(2e_1) \\ \alpha_3(n) & = 0\ otherwise. \end{cases} \tag{17}$$

Proof. For $\rho = 0$, the solution is $y_n(0) = \mathbb{1}_{\{n=0_N\}}$, which gives the coefficients for $k = 0$. For $k = 1, 2$ and 3, we use the method previously described and assumptions (H_1) and (H_2). ◀

It is interesting to notice that, for $k = 0$ and 1, the coefficients $\alpha_k(n)$ do not depend on the choice function c. It means that, for ρ sufficiently small, the choice policy does not influence the system. For $k \ge 4$, the expressions become huge, which is not a problem if performed numerically.

4 Marginal distribution for one queue

Our objective is to study the expansion of the stationary probability that queue i, $1 \le i \le N$, has $m \in \mathbb{N}$ customers assuming an analytical solution for y. We give the series expansion at order 6, for small m ($m \le 3$), for the local choice contribution function. Moreover we investigate the accuracy of this expansion, compared to numerical values obtained by simulations. Then the main result of the section provides the first term of the expansion for every $m \ge 1$.

As our system is invariant by cyclic permutation, by assumption (H_2), for $m \in \mathbb{N}$ and $i \in \{1, \ldots, N\}$, the probability that queue i has m customers does not depend on i. This probability, denoted by $\pi_m(\rho)$, is given by

$$\pi_m(\rho) = \sum_{n=(n_1,\ldots,n_N),\, n_1=m} y_n(\rho). \tag{18}$$

Under assumption (H_0) that $y_n(\rho)$ is analytical on $[0, \varepsilon[$, $\pi_m(\rho)$ has a series expansion, that can be written as

$$\pi_m(\rho) = \sum_{k \geq 0} \phi_k(m)\rho^k, \quad 0 \leq \rho < \varepsilon \tag{19}$$

where $\phi_k(m) = \pi_m^{(k)}(0)/k!$ is given from equation (6) by

$$\phi_k(m) = \mathbb{1}_{\{m \leq k\}} \sum_{n_2+n_3+...+n_N \leq k-m} \alpha_k(m, n_2, n_3, ..., n_N). \tag{20}$$

4.1   Expansion for a general contribution function

Note that, in equation (20), $\phi_m(m)$ is the first possibly non-null coefficient of the expansion of $\pi_m(\rho)$. This follows directly from Lemma 4. Moreover this coefficient is derived in the following proposition, which also gives the third order expansion of the π_m's.

▶ **Proposition 6.** *If the choice function c satisfies (H_0), (H_1) and (H_2), then*

$$\begin{cases} \pi_0(\rho) = 1 - \rho, \\ \pi_1(\rho) = \rho - c_1(e_1)\rho^2 + \left(Nc_1(e_1) - \sum_{j=1}^N c_1(e_1 + e_j)c_1(e_j)\right)\rho^3 + \mathcal{O}(\rho^4), \\ \pi_2(\rho) = c_1(e_1)\rho^2 - \left(Nc_1(e_1) - \sum_{j=2}^N c_1(e_1 + e_j)c_1(e_j)\right)\rho^3 + \mathcal{O}(\rho^4), \\ \pi_m(\rho) = \left(\prod_{j=1}^{m-1} c_1(je_1)\right)\rho^m + \mathcal{O}(\rho^{m+1}), \quad m \geq 3 \end{cases} \tag{21}$$

where ρ tends to 0.

Proof. Equation (21) comes straightforwardly from equation (19) and two intermediate results, Lemma 9 and Lemma 10, postponed in Appendix. Note that, as at equilibrium the rates of incoming and outgoing customers are the same, i.e., $N\lambda = N\mu(1 - \pi_0)$, it gives another way to obtain that $\pi_0(\rho) = 1 - \rho$. ◀

4.2   Expansions for the local choice contribution function

Equation (21) can be rewritten in the case of the local choice function $c = c^{lc}$ defined by equation (1). It gives the following result.

▶ **Corollary 7.** *For the local choice function c^{lc}, for $N \geq 3$*

$$\pi_0(\rho) = 1 - \rho, \qquad\qquad \pi_1(\rho) = \rho - \frac{3}{2}\rho^3 + \mathcal{O}(\rho^4),$$

$$\pi_2(\rho) = \frac{3}{2}\rho^3 + \mathcal{O}(\rho^4), \qquad\qquad \pi_m(\rho) = \mathcal{O}(\rho^{m+1}), \quad m > 2.$$

For $N = 2$, the coefficient $3/2$ of ρ^3 is replaced by 2.

The main point is that $\phi_m(m)$ is null in this case. The aim will be to find the first non vanishing term of the expansion of $\pi_m(\rho)$ for every $m \geq 1$. It is the purpose of Section 4.6. Let us begin by giving more terms in the previous series expansion.
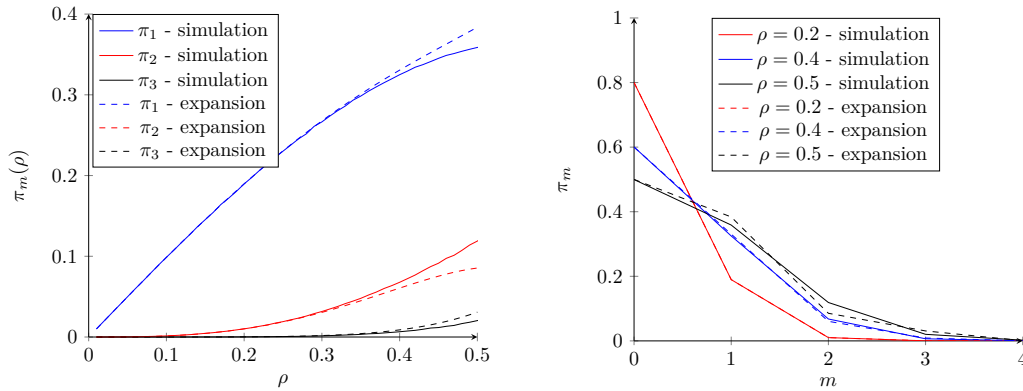
(a) For $m = 1$, 2 and 3, as a function of ρ.   **(b)** For $\rho = 0.2$, 0.4 and 0.5.

Figure 1 Invariant distribution $\pi = (\pi_m, m \in \mathbb{N})$ of the number of customers in one queue for a set of N queues with local choice.

4.3 Further expansions for the local choice function

As the amount of cases to analyze grows exponentially with k, it is rather difficult to obtain further series expansions. The following expansions are obtained with help of mathematical software. For that, we observe the following property, which remains to be proved, that for each $k \in \mathbb{N}$, there exists $N_0(k)$ such that if $N > N_0(k)$, for each $m \geq 1$, $\phi_k(m)$ does not depend on N. For small values of k, it is easy to see that this property holds, given the recurrence equation and the local choice function. Using this, from global balance equations (5) for some N sufficiently large, the following result holds. For $\rho < 1$ tending to 0,

$$\pi_0(\rho) = 1 - \rho$$
$$\pi_1(\rho) = \rho - \frac{3}{2}\rho^3 + \frac{11}{8}\rho^4 - \frac{7}{3}\rho^5 + \frac{10727}{2880}\rho^6 + \mathcal{O}(\rho^7)$$
$$\pi_2(\rho) = \frac{3}{2}\rho^3 - \frac{11}{8}\rho^4 + \frac{47}{24}\rho^5 - \frac{1583}{320}\rho^6 + \mathcal{O}(\rho^7)$$
$$\pi_3(\rho) = \frac{3}{8}\rho^5 + \frac{11}{9}\rho^6 + \mathcal{O}(\rho^7)$$
$$\pi_i(\rho) = \mathcal{O}(\rho^7), \;\; i > 3.$$

4.4 Validation by simulation

In Figure 1, we investigate numerically the accuracy of the previous expansion. Recall that π is the stationary queue length distribution of any queue in this symmetric system of N queues. In figure 1a, we plot π_m for $m = 0$, 1, 2 and 3 as a function of ρ given by simulation and by the series expansion at order 6. The conclusion is that the previous series expansion gives a quite good approximation for small values of ρ ($\rho \leq 0.3$), reasonable for $\rho \leq 0.4$. Figure 1b gives the distribution for different small values of ρ. It indicates that, as ρ increases, the distribution deviates from a geometric distribution. Moreover, the series expansion gives a quite good approximation for $\rho \leq 0.4$.

4.5 No choice policy: the case of independent queues

For the case where each queue receives independently customers at rate λ and serves them at rate μ, the contribution function becomes $c_i(n) = 1$, $n \in \mathbb{N}^N$ and $i \in \mathbb{N}$. We can easily

verify that

$$\alpha_k(n) = (-1)^{k-|n|} \binom{N}{k-|n|} \mathbb{1}_{|n| \leq k}$$

satisfies equation (8), where $|n| = n_1 + \ldots + n_N$. Using equation (20), we have for any $r \in \mathbb{N}$, $0 \leq r \leq k$,

$$\phi_k(k-r) = (-1)^r \sum_{i=0}^{r} (-1)^i \binom{N-2+i}{i} \binom{N}{r-i}$$

$$= (-1)^r \mathbb{1}_{\{r \leq 1\}}. \tag{22}$$

The term $\binom{N-2+i}{i}$ comes from the fact that we need to distribute the remaining i customers in the remaining $N-1$ queues. The last equality, of the form $a_r = b_r$ for all $r \in \mathbb{N}$, is obtained proving that the generating functions $\sum_{r \geq 0} a_r z_r$ and $\sum_{r \geq 0} b_r z_r$ are equal by developing the product

$$1 + z = (1+z)^N \frac{1}{(1+z)^{N-1}}.$$

With straightforward algebra, plugging equation (22) in equation (19), we retrieve that the stationary distribution $\pi(\rho) = (\pi_m(\rho), m \in \mathbb{N})$ for one queue is the geometric distribution with parameter ρ, as each queue is a $M/M/1$ queue with arrival-to-service-rate ratio $\rho = \lambda/\mu$.

4.6   Main result: Asymptotics for the stationary queue length distribution in light traffic

Let us then present the main result.

▶ **Proposition 8.** *For the local choice function c^{lc} defined by equation* (1) *and under assumption* (H_0), *for $m \geq 2$, the stationary probability $\pi_m(\rho)$ that a queue has m customers verifies*

$$\pi_m(\rho) = \begin{cases} 12 \left(\dfrac{\rho}{2}\right)^{2m-1} + \mathcal{O}(\rho^{2m}) & \text{if } N \geq 3 \\ 2\rho^{2m-1} + \mathcal{O}(\rho^{2m}) & \text{if } N = 2 \end{cases}$$

when ρ tends to zero.

Proposition 8 guarantees that, for ρ sufficiently small, the probability of having m customers in the queue follows a geometric decay of parameter $\rho^2/4$ as m grows. The following table illustrates where the local choice is situated.

■ **Table 1** Comparison of asymptotics for the stationary probability that a queue has more than k customers in light traffic (as parameter ρ tends to 0) for different allocation policies, $N \geq 3$.

Allocation policy	$u_k = \sum_{k \geq k} \pi_m$
No-choice	$\sim \rho^m$
Local choice	$\sim (\rho/2)^{2k-1}$
Random choice	$\sim \rho^{2^k-1}$

As expected, the performance of *local choice* policy is between the other two policies. However, for light traffic, its behavior is closer to *no choice* than to *random choice*. Indeed, the two first asymptotics are exponential while the third one is double exponential in ρ.

The light-traffic asymptotics obtained in this paper are for the limit when t tends to $+\infty$ first and then N tends to $+\infty$, since the asymptotic result is independent of N for $N \geq 3$, while from mean-field approximation for the *random choice model* the limit is when N first and then t tends to $+\infty$. The comparison we made is rigorous and justified by the interchange of the order of these two limits, see [14].

References

1 E. Andjel, M. Menshikov, and V. Sisko. Positive recurrence of processes associated to crystal growth models. *Annals of Applied Probability*, pages 1059–1085, 2006.

2 Johannes Pieter Cornelis Blanc. The power-series algorithm applied to the shortest-queue model. *Operations Research*, 40(1):157–167, 1992.

3 P. Bogdan, T. Sauerwald, A. Stauffer, and H. Sun. Balls into bins via local search. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 16–34. SIAM, 2013.

4 F. Ezanno. *Systèmes de particules en interaction et modèles de déposition aléatoire*. PhD thesis, Aix-Marseille Université, 2012.

5 F. Ezanno. Some results about ergodicity in shape for a crystal growth model. *Electronic Journal of Probability*, 18(33):1–20, 2013.

6 Ayalvadi Ganesh, Sarah Lilienthal, D Manjunath, Alexandre Proutiere, and Florian Simatos. Load balancing via random local search in closed and open systems. In *ACM SIGMETRICS Performance Evaluation Review*, volume 38, pages 287–298. ACM, 2010.

7 N. Gast. The power of two choices on graphs: the pair-approximation is accurate? *ACM SIGMETRICS Performance Evaluation Review*, 43(2):69–71, 2015.

8 Nicolas Gast and Gaujal Bruno. A mean field model of work stealing in large-scale systems. In *ACM SIGMETRICS Performance Evaluation Review*, volume 38, pages 13–24. ACM, 2010.

9 D. Gates and M. Westcott. Markov models of steady crystal growth. *Annals of Applied Probability*, pages 339–355, 1993.

10 K. Kenthapadi and R. Panigrahy. Balanced allocation on graphs. In *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pages 434–443. Society for Industrial and Applied Mathematics, 2006.

11 John Lamperti. Criteria for the recurrence or transience of stochastic process. i. *Journal of Mathematical Analysis and applications*, 1(3-4):314–330, 1960.

12 M. Mitzenmacher. *The Power of Two Choices in Randomized Load Balancing*. PhD thesis, Berkeley, 1996.

13 P. Robert. *Stochastic networks and queues*, volume 52. Springer, 2013.

14 N. Vvedenskaya, R. Dobrushin, and F. Karpelevich. Queueing system with selection of the shortest of two queues: An asymptotic approach. *Problemy Peredachi Informatsii*, 32(1):20–34, 1996.

15 G. Fayolle, V. Malyshev and M. Menshikov. *Topics in the constructive theory of countable Markov chains*, Cambridge university press, 1995.

16 V. A. Malyshev and M. Menshikov. Ergodicity, continuity and analyticity of countable Markov chains. *Trudy Moskovskogo Matematicheskogo Obshchestva*, 39, 3-48, 1979.

17 K. Gardner, M. Harchol-Balter, A. Scheller-Wolf, M. Velednitsky and S. Zbarsky. Redundancy-d: The power of d choices for redundancy. *Operations Research*, 65(4): 1078–1094, 2017.

A   Proof of Proposition 1

Proof. Assume that $\rho < 1$. We prove ergodicity by Foster's criterion for Markov processes based on a Lyapunov function (see for example [13, Proposition 8.14]). Here the Lyapunov function f is quadratic, given by $f(n) = n_1^2 + \cdots + n_N^2$, $n = (n_1, \ldots, n_N)$.

Let us denote $|n| = \sum_{i=1}^N n_i$. The infinitesimal generator is given by

$$
\begin{aligned}
Lf(n) &= \sum_{n' \in \mathbb{N}^N} Q(n, n')(f(n') - f(n)) \\
&= \sum_{i=1}^N \lambda c_i(n)(f(n + e_i) - f(n)) + \mathbb{1}_{n_i > 0}\mu(f(n - e_i) - f(n)),
\end{aligned}
\tag{23}
$$

for $f : \mathbb{N}^N \to \mathbb{R}$ with finite support. With straightforward algebra, using equation (1), it holds that

$$
n_1 c_1^{lc}(n) + \cdots + n_N c_N^{lc}(n) \le |n| \text{ and } c_1^{lc}(n) + \cdots + c_N^{lc}(n) = N.
\tag{24}
$$

This gives

$$
\begin{aligned}
L(f)(n) &= \lambda \sum_{i=1}^N c_i^{lc}(n)((n_i + 1)^2 - n_i^2) + \mu \sum_{i=1}^N \mathbb{1}_{n_i > 0}((n_i - 1)^2 - n_i^2) \\
&= 2\lambda \sum_{i=1}^N c_i^{lc}(n)n_i + \lambda \sum_{i=1}^N c_i^{lc}(n) - 2\mu \sum_{i=1}^N n_i + \mu \sum_{i=1}^N \mathbb{1}_{n_i > 0} \\
&\le (\lambda + \mu)N - 2(\mu - \lambda)|n|.
\end{aligned}
\tag{25}
$$

By the equivalence of norms in $\mathbb{R}^N$, there is a constant $C > 0$ such that, for all n, $\sqrt{f(n)} \le C^{-1}|n|$ where $|n| = n_1 + \cdots + n_N$. Thus, if $f(n) > K$ then $|n| \ge C\sqrt{K}$. As $\rho = \lambda/\mu < 1$, K can be chosen large enough to get $\gamma = -(\lambda + \mu)N + 2(\mu - \lambda)C\sqrt{K} > 0$.

Thus, by equation (25), if $f(n) > K$ then $L(f)(n) \le -\gamma$. Moreover the set $F = \{n \in \mathbb{N}^N, f(n) \le K\}$ is finite and the random variables $\sup_{0 \le s \le 1} f(X(s))$ and $\int_0^1 L(f)(X(s))ds$ are integrable. Indeed,

$$
\sup_{0 \le s \le 1} f(X(s)) \le C^{-2} \sup_{0 \le s \le 1} |X(s)|^2 \le C^{-2}(\mathcal{N}_{\lambda N}([0, 1[)^2
$$

where the arrival process in the system, denoted by $\mathcal{N}_{\lambda N}$, is a Poisson process with intensity $\lambda N ds$, as the sum of the N independent Poisson processes with parameter λ of arrivals at the N queues. Using again equation (25),

$$
\int_0^1 L(f)(X(s))ds \le (\lambda + \mu)N.
$$

Thus, the Markov process $(X(t))_{t \ge 0}$ is ergodic if $\rho < 1$.

If $\rho > 1$, we apply [13, Theorem 8.10], a simplified version of a Lamperti's result, to prove the transience of the embedded Markov chain (M_n) at jump times of $(X(t))$. It gives the transience of $(X(t))$. Let g be defined by $g(n) = n_1 + \cdots + n_N$. Using that $c_1^{lc}(n) + \ldots + c_N^{lc}(n) = N$, see equation (24), for all $n \in \mathbb{N}^N$,

$$
\mathbb{E}_n(g(M_1) - g(n)) = Lg(n) = \lambda \sum_{i=1}^N c_i(n) - \mu \sum_{i=1}^N \mathbb{1}_{\{n_i > 0\}} \ge (\lambda - \mu)N > 0.
$$

Moreover, for all $n \in \mathbb{N}^N$,

$$\mathbb{E}_n(|g(M_1) - g(n)|^2) = \sum_{n' \in \mathbb{N}^N} Q(n, n')|g(n') - g(n)|^2$$

$$= \lambda \sum_{i=1}^{N} c_i(n) + \mu \sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} \leq (\lambda + \mu)N,$$

thus $\sup_{n \in \mathbb{N}^N} \mathbb{E}_n(|g(M_1) - g(n)|^2) < \infty$. The sufficient conditions for applying [13, Theorem 8.10] hold. It ends the proof. ◀

B Two lemmas

▶ **Lemma 9.** *For integer k, $0 \leq k \leq 3$, the coefficients $\phi_k(m)$, $m \in \mathbb{N}$, are given by*

$$\phi_0(0) = 1, \text{ and } \phi_0(m) = 0, \; m > 0,$$
$$\phi_1(0) = -1, \; \phi_1(1) = 1 \text{ and } \phi_1(m) = 0, \; m > 1,$$
$$\phi_2(0) = 0, \; \phi_2(1) = -c_1(e_1), \; \phi_2(2) = c_1(e_1) \text{ and } \phi_2(m) = 0, \; m > 2,$$

$$\phi_3(0) = 0, \; \phi_3(1) = -\phi_3(2) = Nc_1(e_1) - \sum_{j=1}^{N} c_1(e_1 + e_j)c_1(e_j),$$

$$\phi_3(3) = c_1(e_1)c_1(2e_1) \text{ and } \phi_3(m) = 0, \; m > 3.$$

Proof. We use, for $k \leq 3$, the expressions of α_k given by Proposition 5 to compute ϕ_k. ◀

▶ **Lemma 10.** *For $k \geq 1$, $\phi_k(k) = \alpha_k(ke_1) = \prod_{j=1}^{k-1} c_1(je_1)$.*

Proof. For $k \in \mathbb{N}^*$, by equation (20), $\phi_k(k) = \alpha_k(ke_1)$. Taking $n = ke_1$ in equation (8),

$$\alpha_k(ke_1) = \sum_{i=1}^{N} \alpha_k(ke_1 + e_i) + c_1((k-1)e_1)\alpha_{k-1}((k-1)e_1) - N\alpha_{k-1}(ke_1).$$

By Lemma 4, for any i, $1 \leq i \leq N$, $\alpha_k(ke_1 + e_i) = 0$ and $\alpha_{k-1}(ke_1) = 0$. It gives that

$$\phi_k(k) = c_1((k-1)e_1)\phi_{k-1}(k-1).$$

This recurrence equation in $\phi_k(k)$ leads to the desired result, since $\phi_1(1) = 1$. ◀

C Proof of Proposition 8

Proof. In the proof, the following definition will be used.

▶ **Definition 11.** *The state $n = (n_1, \ldots, n_N)$ exists at order k if and only if, in equation (6), $\alpha_k(n) \neq 0$.*

First step. To prove Proposition 8, the first step is to obtain that, for a state $n = (n_1, \ldots, n_N)$ existing at order k, the maximum possible queue length is $\lceil k/2 \rceil$. Indeed, by Lemma 4, n exists at order k only if $|n| \leq k$. Moreover, the following result holds.

▶ **Lemma 12.** *Let $k \in \mathbb{N}$ and $n = (n_1, \ldots, n_N) \in \mathbb{N}^N$. If $|n| \leq k$ and $n_1 > \lceil k/2 \rceil$ then $\alpha_k(n) = 0$.*

Proof. The following assertion is proved by induction on $p \geq 0$.

($\mathcal{B}_p$) For $k \in \mathbb{N}$ and $n = (n_i)_{1 \le i \le N}$, if $|n| = k - p$ and $n_1 > \lceil k/2 \rceil$ then $\alpha_k(n) = 0$.

Let us prove ($\mathcal{B}_0$). Let $k \in \mathbb{N}$ and n such that $|n| = k$ and $n_1 > \lceil k/2 \rceil$. As $|n| = k$, by Lemma 4, for each i, $1 \le i \le N$, $\alpha_k(n + e_i) = \alpha_{k-1}(n) = 0$. Thus equation (8) is rewritten as

$$\left(\sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} \right) \alpha_k(n) = \sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} c_i^{lc}(n - e_i) \alpha_{k-1}(n - e_i). \qquad (26)$$

As $|n| \le k$ and $n_1 > \lceil k/2 \rceil$, $n_2 + n_N \le k - n_1 < k - \lceil k/2 \rceil < n_1$. Thus $n_2 + n_N \le k - \lceil k/2 \rceil - 1 < n_1 - 1$.

It means that each neighboring queue of queue 1 has strictly less than $n_1 - 1$ customers. Thus the contribution on queue 1 for our local choice function c^{lc} defined by equation (1) gives $c_1^{lc}(n - e_1) = 0$ and equation (26) can be rewritten

$$\left(\sum_{i=1}^{N} \mathbb{1}_{\{n_i > 0\}} \right) \alpha_k(n) = \sum_{i=2}^{N} \mathbb{1}_{\{n_i > 0\}} c_i^{lc}(n - e_i) \alpha_{k-1}(n - e_i). \qquad (27)$$

Therefore, state n exists at order k only if there is $i_1 \ne 1$ such that $\alpha_{k-1}(n - e_{i_1}) \ne 0$. But $|n - e_{i_1}| = k - 1$ and we can repeat the previous arguments for $k - 1$ instead of k and $n - e_{i_1}$ instead of n, with $(n - e_{i_1})_1 > \lceil k/2 \rceil \ge \lceil (k-1)/2 \rceil$, and so on until we obtain $n_1 e_1$. In conclusion, n exists at order k only if $\alpha_{n_1}(n_1 e_1) \ne 0$. It contradicts Lemma 10. Therefore $\alpha_k(n) = 0$.

Assume now, for $p \ge 1$, that ($\mathcal{B}_{p-1}$) is true, and prove ($\mathcal{B}_p$). For that, let $k \in \mathbb{N}$ and n be such that $|n| = k - p$ and $n_1 > \lceil k/2 \rceil$. By induction assumption ($\mathcal{B}_{p-1}$), applied to k and $n + e_i$ as $|n + e_i| = k - (p - 1)$, then to $k - 1$ and n as $|n| = k - 1 - (p - 1)$, it holds that $\alpha_k(n + e_i) = \alpha_{k-1}(n) = 0$. Then the arguments used for ($\mathcal{B}_0$) give that $\alpha_k(n) = 0$. It ends the proof.   ◀

One can then deduce easily the following result.

▶ **Lemma 13.** *Let m be in $\mathbb{N}^*$. The first possibly non vanishing term of the expansion when ρ tends to zero of the stationary probability $\pi_m(\rho)$ that a queue has m customers is $\phi_{2m-1}(m) \rho^{2m-1}$.*

Proof. For $m \in \mathbb{N}$, by definition, see (19), $\pi_m(\rho) = \sum_{k \ge 0} \phi_k(m) \rho^k$ with

$$\phi_k(m) = \sum_{\substack{n = (m, n_2, \ldots, n_N) \\ |n| \le k}} \alpha_k(m, n_2, \ldots, n_N).$$

If $k < 2m - 1$ then, for $n = (m, n_2, \ldots, n_N)$ such that $|n| \le k$, $n_1 = m > \lceil k/2 \rceil$. Thus, by Lemma 12, all the $\alpha_k(m, n_2, \ldots, n_N)$ in the right-hand side of the previous equation are null for $k < 2m - 1$. It ends the proof.   ◀

Second step. Moreover the states which exist at order $k = 2m - 1$ with one queue with the maximum value m correspond just to two neighboring queues with m and $j < m$. It is given by the following lemma.

▶ **Lemma 14.** *If $|n| \le k = 2m - 1$ (k odd), $n_1 = m$ and there exists two distinct j and l, different from 1, such that $n_j > 0$ and $n_l > 0$ then $\alpha_k(n) = 0$.*

Proof. The following assertion is proved by induction on $p \ge 0$

($\mathcal{B}_p$) For $k = 2m - 1$, $m \in \mathbb{N}$, for n such that $|n| = k - p$, $n_1 = m$, $n_j > 0$ and $n_l > 0$ with j and l distinct, different from 1, then $\alpha_k(n) = 0$.

Let us prove ($\mathcal{B}_0$). Let $k = 2m - 1$ and n chosen as indicated. As $|n| = k$, by Lemma 4, for each i, $1 \leq i \leq N$, one gets $\alpha_k(n + e_i) = \alpha_{k-1}(n) = 0$. As before, using Lemma 2, equation (26) holds. By assumption, as in the proof of Lemma 12, it holds that each neighboring queue of queue 1 has strictly less than $n_1 - 1$ customers, which yields $c_1(n - e_1) = 0$. Thus equation (26) can be rewritten equation (27). We conclude as in the proof of Lemma 12. ◄

Step 3. We distinguish two cases:

Case 1: $N \geq 3$. From equation (20) and applying Lemma 14,

$$\phi_{2m-1}(m) = \sum_{i=1}^{m-1} \alpha_{2m-1}(m, i, 0, ...0) + \sum_{i=1}^{m-1} \alpha_{2m-1}(m, 0, ..., 0, i) + \alpha_{2m-1}(m, 0, ..., 0)$$

then by symmetry,

$$\phi_{2m-1}(m) = 2 \sum_{i=1}^{m-1} \alpha_{2m-1}(m, i, 0, ...0) + \alpha_{2m-1}(m, 0, 0, ...0). \tag{28}$$

This means that only these terms are non null. The rest of the proof consists in deriving them.

Let n_1 and n_2 be chosen as follows: $n_1 = (k + 1)/2$ and $n_2 = (k - 1)/2$. Using the same arguments as in Lemma 12, equation (8) gives, for $k = 2m - 1$ with m integer and $m \geq 2$,

$$2\alpha_{2m-1}(m, m - 1, 0, ..., 0) = \frac{1}{2}\alpha_{2(m-1)}(m - 1, m - 1, 0, ..., 0). \tag{29}$$

Let $k = 2m$, and $n = (m, m, 0, ..., 0)$. For $m \in \mathbb{N}$, $m \geq 2$, as $c_2^{lc}(m, m - 1, 0, ..., 0) = 1$, equation (8) gives

$$2\alpha_{2m}(m, m, 0, ..., 0) = 2\alpha_{2m-1}(m, m - 1, 0, ..., 0). \tag{30}$$

Combining equations (29) and (30), for $m \geq 3$,

$$\alpha_{2m-1}(m, m - 1, 0, ..., 0) = \frac{1}{2^2}\alpha_{2m-3}(m - 1, m - 2, 0, ..., 0)$$

and then, using equation (17) to show that $\alpha_3(2, 1, 0, ..., 0) = 3/8$, for $m \geq 3$,

$$\alpha_{2m-1}(m, m - 1, 0, ..., 0) = \frac{1}{2^{2(m-2)}}\alpha_3(2, 1, 0, ..., 0) = \frac{3}{2^{2m-1}}. \tag{31}$$

Then, for $n = (m, i, 0, ..., 0)$, for $0 < i < m - 1$, from equation (8),

$$2\alpha_{2m-1}(m, i, 0, ..., 0) = \alpha_{2m-1}(m, i + 1, 0, ..., 0).$$

By induction and using equation (31), for $0 < i < m - 1$,

$$\alpha_{2m-1}(m, i, 0, ..., 0) = \frac{1}{2^{m-1-i}}\alpha_{2m-1}(m, m - 1, 0, ..., 0) = \frac{3}{2^{2m-1}}\frac{1}{2^{m-1-i}}. \tag{32}$$

With similar arguments and then using equation (32) for $i = 1$,

$$\alpha_{2m-1}(m, 0, 0, ..., 0) = \alpha_{2m-1}(m, 1, 0, ..., 0) + \alpha_{2m-1}(m, 0, ..., 0, 1) \tag{33}$$

$$= \frac{6}{2^{2m-1}}\frac{1}{2^{m-2}}.$$

Plugging equations (31), (32) and (33) in equation (28),

$$\phi_{2m-1}(m) = 2\frac{3}{2^{2m-1}} + 2\sum_{i=1}^{m-2}\frac{3}{2^{2m-1}}\frac{1}{2^{m-1-i}} + \frac{6}{2^{2m-1}}\frac{1}{2^{m-2}}$$

$$= \frac{6}{2^{2m-1}}\left(\sum_{i=1}^{m-1}\frac{1}{2^{m-1-i}} + \frac{1}{2^{m-2}}\right) = \frac{12}{2^{2m-1}}$$

Using it in Lemma 13 gives the result.

Case 2: $N = 2$. With similar arguments, equation (28) is rewritten in this case

$$\phi_{2m-1}(m) = \sum_{i=1}^{m-1}\alpha_{2m-1}(m,i) + \alpha_{2m-1}(m,0).$$

while equations (29) and (30) become $2\alpha_{2m-1}(m,m-1) = \alpha_{2(m-1)}(m-1,m-1)$ and $2\alpha_{2m}(m,m) = 2\alpha_{2m-1}(m,m-1)$. Following exactly the same lines as in case 1, one gets

$$\phi_{2m-1}(m) = \sum_{i=1}^{m-1}\frac{1}{2^{m-1-i}} + \frac{1}{2^{m-2}} = 2.$$

It ends the proof.   ◀

Refined Asymptotics for the Number of Leaves of Random Point Quadtrees

Michael Fuchs[1]

Department of Applied Mathematics, National Chiao Tung University, 300 Hsinchu, Taiwan
mfuchs@math.nctu.edu.tw

Noela S. Müller

Institute for Mathematics, Goethe University, 60054 Frankfurt a.M., Germany
nmueller@math.uni-frankfurt.de

Henning Sulzbach

University of Birmingham, School of Mathematics, B15 2TT Birmingham, United Kingdom
h.sulzbach@bham.ac.uk

──── **Abstract** ────

In the early 2000s, several phase change results from distributional convergence to distributional non-convergence have been obtained for shape parameters of random discrete structures. Recently, for those random structures which admit a natural martingale process, these results have been considerably improved by obtaining refined asymptotics for the limit behavior. In this work, we propose a new approach which is also applicable to random discrete structures which do not admit a natural martingale process. As an example, we obtain refined asymptotics for the number of leaves in random point quadtrees. More applications, for example to shape parameters in generalized m-ary search trees and random gridtrees, will be discussed in the journal version of this extended abstract.

1 Introduction and Result

In this extended abstract, we investigate shape parameters of random discrete structures whose distributional behavior is known to undergo a phase change as a structural characteristic of the structure varies. Several such phase change phenomena, in particular with a change from distributional convergence to distributional non-convergence, have been found in the early 2000s. We start by recalling a particular nice and surprising result in this direction which was obtained by Janson in [11]: the phase change of the number of nodes with depth in a fixed congruent class in random recursive trees.

First, we recall the definition of random recursive trees. Starting from a root, nodes are added consecutively where the n-th node is attached uniformly at random as left-most child

---

■ **Table 1** A summary of shape parameters and discrete structures for which the distributional behavior changes from normal to non-convergence.

structure	parameter	non-convergence	refined asymptotics
recursive trees	nodes with depths divisible by m	$m \geq 7$	[16, 17]
m-ary search trees	size	$m \geq 27$	[15]
d-dimensional quadtrees	number of leaves	$d \geq 9$	this paper

to one of the existing nodes. In such a tree with n nodes, let M_n denote the number of nodes with depth (distance from the root) divisible by m where $m \geq 2$ is fixed. Set

$$\omega_n = \begin{cases} n, & \text{if } 6 \nmid m; \\ n \log n, & \text{if } 6 \mid m. \end{cases}$$

Then, in [11], the following result was proved: if $2 \leq m \leq 6$, we have

$$\frac{M_n - n/m}{\sqrt{\omega_n}} \xrightarrow{d} N(0, \sigma_m^2), \tag{1}$$

where $\sigma_m > 0$; for all other m, we have that M_n with the standard normalization, i.e., $(M_n - n/m)/\sqrt{\mathrm{Var}(M_n)}$, does not converge to a fixed limit law.

A similar result holds if the depths of nodes are required to fall into another residue class. Moreover, the same phase change phenomenon is present in random binary search trees, too; see [11]. Also, several other shape parameters in diverse families of random trees have been proved to exhibit a similar phase change behavior from distributional convergence to distributional non-convergence, e.g., the size of m-ary search trees proved by Chern and Hwang [4] (see also Mahmoud and Pittel [14] and Lew and Mahmoud [13] for preliminary results) and the number of leaves in random point quadtrees proved by Chern, Fuchs, and Hwang [2]; see Table 1 for a summary of these results and [2, 4] for many more examples.

After the above results have been published, subsequent research has focused on clarifying the stochastic behavior in the non-convergence regime; e.g. see [2], Chern, Fuchs, Hwang, and Neininger [3], Chauvin and Pouyanne [1], Fill and Kapur [6], and [11]. This line of research has recently culminated in the realization that subtracting a sufficiently large number of suitable random variables leads to a central limit theorem. To give some more details, consider again the above random variable M_n. Set $r = \lfloor (m-1)/6 \rfloor$ and

$$\zeta_k := \cos\left(\frac{2\pi k}{m}\right) \qquad \text{and} \qquad \eta_k := \sin\left(\frac{2\pi k}{m}\right).$$

Following a technique developed by Neininger [18] in a refined analysis of the complexity of the randomised Quicksort algorithms, it was proved by the second author of this extended abstract and Neininger [16, 17] that there exist complex random variables $\Xi_1, \ldots, \Xi_r$ such that

$$\frac{1}{\sqrt{\omega_n}} \left(M_n - \frac{n}{m} - 2 \sum_{k=1}^{r} \Re\left(\Xi_k n^{i\eta_k}\right) n^{\zeta_k} \right) \xrightarrow{d} N(0, \sigma_m^2)$$

with $\sigma_m > 0$. Note that this result yields (1) as a special case.

The proof of the above result made use of a natural martingale process related to random recursive trees. Moreover, another proof method (also using the martingale process) was

■ **Table 2** Value of p in (3) for small values of d.

d	$1, \ldots, 8$	$9, \ldots, 17$	$18, \ldots, 26$	$27, \ldots, 34$
p	0	1	2	3

proposed by the second author in [15], where the above result was extended to generalized Pólya urns; see Janson [10] for background. The latter result contains both the above result and a similar result for m-ary search trees; see Table 1.

The purpose of this work is to propose yet another approach which does not make use of the martingale process (the possibility of such an approach was already announced in [17]). The advantage of such a method is that it can be applied to random discrete structures which do not admit such a process. This is for instance the case for random point quadtrees which we use in this work as guiding example. Other applications of our approach in the context of, e.g., generalized m-ary search trees and gridtrees (where there are again no natural martingale processes) will be discussed in the journal version of this extended abstract.

We first recall the definition of random point quadtrees (which for brevity will be called random quadtrees in the sequel). Fix a dimension d and consider an infinite sequence of stochastically independent points chosen uniformly at random from the d-dimensional unit cube. Then, the first point is stored in the root which has 2^d subtrees that correspond to the 2^d quadrants into which the d-dimensional unit cube is split by the first point. These subtrees contain the points which fall into these quadrants respectively. Moreover, subtrees are built recursively via the same process. The resulting tree after n steps is called random quadtree of size n.

In such a tree of size n, let L_n denote the number of leaves. Then, in [2], the following phase change result was proved: if $1 \leq d \leq 8$, then

$$\frac{L_n - \kappa_d n}{\sqrt{n}} \xrightarrow{d} N(0, \sigma_d^2),$$

where $\sigma_d > 0$ and

$$\kappa_d = 1 - \frac{2}{d} \xi'(1), \tag{2}$$

where $\xi(s)$ is given in (4); for all other d, we have that L_n with the standard normalization does not converge to a fixed limit law. (For $d = 1$, the result goes back to Devroye [5].)

The main result of this extended abstract is the following extension of this result which gives an asymptotic expansion of the limit behavior in the style of [15, 16, 17].

▶ **Theorem 1.** *Let $d \geq 1$. Then, there exist complex random variables $\mathcal{Z}_1, \ldots, \mathcal{Z}_p$ such that*

$$\frac{1}{\sqrt{n}} \left(L_n - \kappa_d n - 2 \sum_{k=1}^{p} \Re(\mathcal{Z}_k n^{i\beta_k}) n^{\alpha_k} \right) \xrightarrow{d} N(0, \sigma_d^2), \tag{3}$$

where $\sigma_d > 0$. Here,

$$\alpha_k := 2 \cos\left(\frac{2\pi k}{d}\right) - 1 \qquad and \qquad \beta_k := 2 \sin\left(\frac{2\pi k}{d}\right)$$

and p is the largest number in $\{0, \ldots, \lfloor d/2 \rfloor\}$ with $\alpha_k > 1/2$; see Table 2.

We conclude the introduction with a discussion of the proof of Theorem 1 and an outline of the manuscript. Following [18, 16], the proof relies on the following three steps:

(i) the construction of the limiting random variables $\mathcal{Z}_1, \ldots, \mathcal{Z}_p$,

(ii) an expansion of the variance of the residual $L_n - \kappa_d n - 2 \sum_{k=1}^{p} \Re(\mathcal{Z}_k n^{i\beta_k}) n^{\alpha_k}$, and

(iii) general techniques to deduce the asymptotic normality (3) from (ii) from a distributional recurrence for the sequence of residuals.

In the literature, step (iii) in the present context has been carried out relying on two different techniques which both apply with straightforward modifications in our setting: the contraction method [18, 16] and the method of moments [9]. As this part does not involve significantly new arguments, we refrain from discussing the details in this extended abstract and refer the reader to the journal version of this work (to come).

The remainder of the manuscript is organized as follows. In Section 2, we give an explicit construction of the quadtree sequence and state known asymptotic expansions for the mean number of leaves. Section 3 is dedicated to step (i) and uses contraction arguments; the proofs are found in Appendix A.

The most technical part of the work, step (ii), crucially relies on a recursive distributional decomposition of the residual sequence and asymptotic transfer theorems developed in Chern, Fuchs, and Hwang [2] for general parameters in quadtrees. This part, worked out in Section 4, is based on conceptually novel ideas since second moments cannot be computed by direct means exploiting a martingale structure. Proofs of technical lemmas required here are collected in the Appendix B.

2   Preliminaries

Let us start with an explicit construction of the quadtrees. To this end, let $Y^{(i)}, i \geq 1$ be a sequence of independent random variables following the uniform distribution on $[0, 1]^d$. We define a sequence of 2^d-ary trees $T_0, T_1, \ldots$ where T_i stores the values $Y^{(1)}, \ldots, Y^{(i)}$ as follows: initially, we start with an empty tree T_0 consisting of a placeholder associated with the unit cube. $Y^{(1)}$ replaces the placeholder thereby creating a tree T_1 consisting of a root node to which we associate 2^d placeholders which are assigned the 2^d rectangular regions in which the components of $Y^{(1)}$ partition the unit cube. (In computer science, these placeholders are often called external nodes.) Inductively, having constructed the tree T_n storing $Y^{(1)}, \ldots, Y^{(n)}$ with $1 + (2^d - 1)n$ placeholders corresponding to $1 + (2^d - 1)n$ regions partitioning the unit cube, we obtain the tree T_{n+1} by storing $Y^{(n+1)}$ in the placeholder associated with the rectangle containing $Y^{(n+1)}$ and adding 2^d placeholders which are assigned the rectangular regions in which $Y^{(n+1)}$ partitions the aforementioned rectangle.

We let L_n denote the number of leaves in the random quadtree T_n. Set $\mu_n := \mathbb{E}[L_n]$. To describe the asymptotic behavior of μ_n, it is necessary to introduce some terminology from [7]: first, for $s \in \mathbb{C} \setminus \{0\}$, let $[s] := 1 - \frac{2^d}{s^d}$. Then, for $n \in \mathbb{N}, n \geq 3$, we define the d-analogue of the factorial as

$$[n]! := [3] \cdot [4] \cdots [n] \qquad \text{and} \qquad [2]! := 1.$$

Let $A := \{2\omega^k - j : k \in \{0, \ldots, d-1\}, j \in \mathbb{N}\}$. The definition of $[n]!$ extends holomorphically to complex numbers $s \in \mathbb{C} \setminus A$ through

$$[s]! := \prod_{j=1}^{\infty} \frac{[j+2]}{[j+s]}, \qquad \text{and} \quad [\infty]! := [3] \cdot [4] \cdot [5] \cdots.$$

Flajolet et al. [7], showed that, for all $n \geq 2$,

$$\mu_n = \sum_{k=0}^{n} \binom{n}{k}(-1)^k \mu_k^* \quad \text{with} \quad \mu_k^* = \begin{cases} 0, & k = 0 \\ -1, & k = 1 \\ -\sum_{j=2}^{k} \frac{[k]!}{[j]!}, & k \geq 2. \end{cases}$$

From here, standard techniques relying on Nörlund-Rice integrals for meromorphic functions arising in the analysis of finite differences such as, e.g. [8, Theorem 2], allow to derive asymptotic expansions of μ_n (as $n \to \infty$) of arbitrary precision. In particular, following the notation in [7], with

$$\xi(s) := \frac{s-1}{[\infty]!} + \sum_{j=2}^{\infty} \left(\frac{1}{[j]!} - \frac{1}{[s+j-1]!} \right), \tag{4}$$

one finds

$$\mu_n = \kappa_d n - 2 \sum_{1 \leq k \leq \lfloor d/2 \rfloor, \alpha_k > 0} \Re\left(\gamma_k n^{i\beta_k}\right) n^{\alpha_k} + \mathcal{O}(1), \tag{5}$$

where κ_d is given in (2) and, with $\lambda_k = \alpha_k + i\beta_k$ for $k = 1, \ldots, \lfloor d/2 \rfloor$ with $\alpha_k > 0$,

$$\gamma_k = -\frac{\lambda_k + 1}{d} \Gamma(-\lambda_k)\xi(\lambda_k)[\lambda_k + 1]!.$$

Here, $\Gamma(\cdot)$ denotes the Gamma function. The details of the argument show that $\gamma_k \neq 0$ for all $k \geq 1, \alpha_k > 0$, so no term in the asymptotic expansion (5) vanishes. For later purposes, note that $\alpha_k \neq 1/2$ for all $k = 1, \ldots, d-1$, since the converse would imply the existence of a k-th root of unity with real part $3/4$. But any rational real part of a root of unity takes values in the set $\{0, -1/2, 1/2, 1, -1\}$ since, with $\omega := \alpha_1 + i\beta_1$, the value $2\Re(\omega^k) = \omega^k + \omega^{d-k}$ is an algebraic integer for any $1 \leq k \leq d$.

3 A family of limiting random variables

As opposed to the applications discussed in [18, 16, 17], there is no natural martingale process associated with the sequence $L_n, n \geq 1$. Therefore, it is necessary to construct the limiting random variables $\mathcal{Z}_1, \ldots, \mathcal{Z}_p$ in Theorem 1 in an ad-hoc way guided by the recursive distributional decomposition of L_n. In this section, we give the details of this construction.

Let $\mathbb{T}$ be the complete infinite 2^d-ary tree represented in standard Ulam-Harris notation by

$$\mathbb{T} = \bigcup_{i \geq 0} \{0, \ldots, 2^d - 1\}^i.$$

Through the canonical embedding of the sequence $T_0, T_1, \ldots$ of increasing trees into $\mathbb{T}$, to any $v \in \mathbb{T}$, we shall associate a random integer $\ell(v) \geq 1$ such that $Y^{(\ell(v))}$ is stored at node v. (Clearly, as the fill-up level of T_n grows to infinity, $\ell(v)$ exists for all nodes $v \in \mathbb{T}$.) For $\ell \geq 1$, let I_ℓ be the rectangle corresponding to the placeholder in $T_{\ell-1}$ which contains $Y^{(\ell)}$. Define $\tilde{\mathcal{Y}}^{(\ell)}$ as the vector of components of $Y^{(\ell)}$ relative to the boundaries of I_ℓ. Formally, with $I_\ell = [i_1^\wedge, i_1^\vee] \times \cdots \times [i_d^\wedge, i_d^\vee]$, we set

$$\tilde{\mathcal{Y}}_k^{(\ell)} = \frac{Y_k^{(\ell)} - i_k^\wedge}{i_k^\vee - i_k^\wedge}, \quad k = 1, \ldots, d.$$

Finally, for $v \in \mathbb{T}$, let

$$U^{(v)} := \tilde{\mathcal{Y}}^{(\ell(v))}.$$

By construction, $\{U^{(v)} : v \in \mathbb{T}\}$ is a family of independent random variables with the uniform distribution on $[0, 1]^d$. While the placeholders associated with the nodes in the tree T_n give rise to a partition of the unit cube, the construction of the limiting random variables relies on different decompositions of the unit cube traversing $\mathbb{T}$ level-wise. To this end, to every $v \in \mathbb{T}$ and $0 \leq j \leq 2^d - 1$, writing $j = \sum_{k=1}^{d} \varepsilon_k 2^{k-1}$ with $\varepsilon_1, \ldots, \varepsilon_d \in \{0, 1\}$, we associate the random variables $\Delta_j^{(v)} := V_1^{(v)} \cdots V_d^{(v)}$, where

$$V_k^{(v)} := \begin{cases} U_k^{(v)}, & \text{if } \varepsilon_k = 0 \\ 1 - U_k^{(v)}, & \text{if } \varepsilon_k = 1. \end{cases}$$

Note that $\sum_{j=0}^{2^d-1} \Delta_j^{(v)} = 1$. Subsequently, write $\Delta^{(v)} = (\Delta_0^{(v)}, \ldots, \Delta_{2^d-1}^{(v)})$.

Let $k \in \{1, \ldots, d-1\}$ with $\alpha_k > 1/2$ and define a family of random variables $\{\mathcal{Z}_{n,k}^{(v)} : n \geq 0, v \in \mathbb{T}\}$ as follows: first, set $\mathcal{Z}_{0,k}^{(v)} = \gamma_k$ for all $v \in \mathbb{T}$. Then, for $n \geq 1$ and $v \in \mathbb{T}$, we recursively define

$$\mathcal{Z}_{n,k}^{(v)} := \sum_{j=0}^{2^d-1} \left(\Delta_j^{(v)} \right)^{\lambda_k} \cdot \mathcal{Z}_{n-1,k}^{(vj)}.$$

Note that, for all $n \geq 0$, we have $\mathcal{Z}_{n,d-k}^{(v)} = \overline{\mathcal{Z}_{n,k}^{(v)}}$. Let $\Pi_\emptyset := 1$, and, recursively, for $v \in \mathbb{T}$ and $j = 0, \ldots, 2^d - 1$,

$$\Pi_{vj} = \Delta_j^{(v)} \Pi_v.$$

Then, we have the following forward expression for $\mathcal{Z}_{n,k}^{(\emptyset)}$:

$$\mathcal{Z}_{n,k}^{(\emptyset)} = \gamma_k \sum_{|v|=n} \Pi_v^{\lambda_k}.$$

Analogous expansions can be stated for $\mathcal{Z}_{n,k}^{(v)}, v \in \mathbb{T}$. Let $\mathcal{F}_{-1}$ be the trivial σ-field, and, for $n \geq 0$, set $\mathcal{F}_n = \sigma \left(U^{(v)} : v \in \mathbb{T}, |v| \leq n \right)$. It follows immediately from the previous display that $\mathcal{Z}_{n,k}^{(\emptyset)}, n \geq 0$ is a martingale with respect to the filtration $\mathcal{F}_n, n \geq -1$.

This martingale has the following important property.

▶ **Proposition 2.** *For all $v \in \mathbb{T}$ there exists a random variable $\mathcal{Z}_k^{(v)}$ such that, almost surely and with respect to all moments,*

$$\mathcal{Z}_{n,k}^{(v)} \to \mathcal{Z}_k^{(v)}. \tag{6}$$

We have
(i) the random variables $\mathcal{Z}_k^{(v)}, v \in \mathbb{T}$ are identically distributed,
(ii) $\mathcal{Z}_k^{(v0)}, \ldots, \mathcal{Z}_k^{(v(2^d-1))}, \Delta^{(v)}$ are stochastically independent and

$$\mathcal{Z}_k^{(v)} = \sum_{j=0}^{2^d-1} \left(\Delta_j^{(v)} \right)^{\lambda_k} \cdot \mathcal{Z}_k^{(vj)},$$

(iii) the law of $\mathcal{Z}_k^{(\emptyset)}$ is the unique distribution satisfying $\mathbb{E}[\mathcal{Z}_k^{(\emptyset)}] = \gamma_k, \mathbb{E}[|\mathcal{Z}_k^{(\emptyset)}|^2] < \infty$ and

$$\mathcal{Z}_k^{(\emptyset)} \stackrel{d}{=} \sum_{j=0}^{2^d-1} \left(\Delta_j^{(\emptyset)}\right)^{\lambda_k} \cdot \mathcal{Z}_k^{*,(j)}, \tag{7}$$

where $\mathcal{Z}_k^{*,(0)}, \ldots, \mathcal{Z}_k^{*,(2^d-1)}$ are independent copies of $\mathcal{Z}_k^{*,(\emptyset)}$, independent of $\Delta^{(\emptyset)}$.
In the remainder of the manuscript, we agree to drop the upper index $\emptyset$ when referring to the quantities $\mathcal{Z}_k^{(\emptyset)}, k = 1, \ldots, p$ and $\Delta_j^{(\emptyset)}, j = 0, \ldots, 2^d - 1$ and $U_j^{(\emptyset)}, j = 1, \ldots, d$.

Below, we will need the following property of the $\mathcal{Z}_k$'s which follows from Leckey [12]; see Appendix A.

▶ **Proposition 3.** *Let $1 \le k \le p$. The vector $(\Re(\mathcal{Z}_k), \Im(\mathcal{Z}_k))$ has a Schwartz density f on $\mathbb{R}^2$, that is, f is infinitely differentiable, where f and all its derivatives decay faster to zero at infinity than any polynomial.*

4   The variance of the residual

In the final chapter of the manuscript, we discuss the techniques to prove step (ii) outlined in the introduction. Let

$$L_n^* := L_n - \kappa_d n - 2 \sum_{k=1}^{p} \Re\left(\mathcal{Z}_k n^{i\beta_k}\right) n^{\alpha_k} + \delta_n,$$

where δ_n is deterministic such that $\mathbb{E}[L_n^*] = 0$. (Exact scaling simplifies arguments in the following.) By (5), we have $\delta_n = \mathcal{O}\left(n^{\max(\alpha_{p+1}, 0)}\right)$. (One actually has $\alpha_{p+1} > 0$ for all $d > 11$.) For $j = 0, \ldots, 2^d - 1$, let N_j be the size of the j-th subtree of the root and $L_n^{(j)}$ be the number of leaves it contains. Given $\Delta_0, \ldots, \Delta_{2^d-1}$, the vector $(N_0, \ldots, N_{2^d-1})$ has the multinomial distribution with parameter $(n - 1; \Delta_0, \ldots, \Delta_{2^d-1})$. We now set up a distributional recurrence for L_n^*. As $\mathcal{Z}_k = \sum_{j=0}^{2^d-1} \Delta_j^{\lambda_k} \mathcal{Z}_k^{(j)}$ it follows that

$$L_n^* = \sum_{j=0}^{2^d-1} \left(L_n^{(j)} - \kappa_d N_j + \delta_{N_j} - 2 \sum_{k=1}^{p} \Re\left(\mathcal{Z}_k^{(j)} N_j^{i\beta_k}\right) N_j^{\alpha_k} \right) + r_n + D_n$$

$$=: \sum_{j=0}^{2^d-1} \mathbb{L}_n^{(j)} + r_n + D_n, \tag{8}$$

where

$$r_n := \delta_n - \sum_{j=0}^{2^d-1} \delta_{N_j} - \kappa_d, \quad \text{and } D_n := 2 \sum_{j=0}^{2^d-1} \sum_{k=1}^{p} \Re\left(\mathcal{Z}_k^{(j)} \left(N_j^{\lambda_k} - (\Delta_j n)^{\lambda_k}\right)\right).$$

By the construction of the quadtree, $\left(\mathbb{L}_n^{(0)}, \ldots, \mathbb{L}_n^{(2^d-1)}\right) \stackrel{d}{=} \left(\bar{L}_{N_0}^{(0)*}, \ldots, \bar{L}_{N_{2^d-1}}^{(2^d-1)*}\right)$, where $\left(\bar{L}_k^{(0)*}\right)_{k\ge0}, \ldots, \left(\bar{L}_k^{(2^d-1)*}\right)_{k\ge0}$ are independent copies of the process $(L_k^*)_{k\ge0}$. Note that $(N_0, \ldots, N_{2^d-1})$ and $\{U^{(v)} : v \in \mathbb{T} \setminus \{\emptyset\}\}$ are independent. Note however, that D_n and $\left(\mathbb{L}_n^{(0)}, \ldots, \mathbb{L}_n^{(2^d-1)}\right)$ are not stochastically independent, not even given $(\Delta, N_0, \ldots, N_{2^d-1})$, since both quantities involve $\mathcal{Z}_k^{(j)}, j = 0, \ldots, 2^d - 1, k = 1, \ldots, p$.

4.1 An asymptotic expansion for the variance

The remainder of this extended abstract is devoted to the proof of the following proposition.

▶ **Proposition 4.** *There exists $0 < \sigma_d < \infty$ such that, as $n \to \infty$,*

$$\mathrm{Var}(L_n^*) = \sigma_d n + o(n).$$

Of course, as $\delta_n = o(\sqrt{n})$, the same asymptotic expansion applies to the variance of the residual sequence $L_n^* - \delta_n$. To prove the proposition, note that, from (8), straightforward calculations reveal that, with $a(n) := \mathbb{E}\left[(L_n^*)^2\right]$, we have

$$a(n) = 2^d \mathbb{E}[a(N_0)] + \mathbb{E}\left[r_n^2\right] + \mathbb{E}\left[D_n^2\right] + 2\mathbb{E}\left[D_n r_n\right] + 2\mathbb{E}\left[D_n \sum_{j=0}^{2^d-1} \mathbb{L}_n^{(j)}\right]$$

$$=: 2^d \mathbb{E}[a(N_0)] + b(n). \tag{9}$$

This is the quadtree recurrence (see Lemma 5 below). Our aim is to apply the asymptotic transfer theorems for it developed in Chern, Fuchs, and Hwang[2]. To this end, we need to understand the asymptotic behavior of the additive sequence $b(n)$ in the last display. In particular, we would like to use the following result from [2].

▶ **Theorem 5** ([2], Theorem 2(i)). *Consider the quadtree recurrence*

$$a_n = b_n + 2^d \sum_{0 \le j < n} \pi_{n,j} a_j, \qquad (n \ge 1),$$

where $a_0 = 0$ and

$$\pi_{n,j} = \mathbb{P}\left(N_0 = j\right) = \binom{n-1}{j} \int_0^1 u^j (1-u)^{n-1-j} \frac{(-\log u)^{d-1}}{(d-1)!} \mathrm{d}u.$$

If $b_n = o(n)$ and the series $\sum_{n \ge 1} b_n/n^2$ converges, then $a_n = \kappa n + o(n)$ for some $\kappa \in \mathbb{R}$.

For infinite sum representations of the limiting constant κ, we refer to [2]. The theorem does not exclude the case that $\kappa = 0$, which explains the necessity of the following lemma, whose proof is deferred to the Appendix B.

▶ **Lemma 6.** *In the set-up of the previous theorem, assume that*
(a) *(i) b_n is non-negative for all n, and (ii) b_n is positive for at least one n, or*
(b) *(i) a_n is non-negative for all n, and (ii) b_n is positive for all n large enough.*
Then, $a_n = \Omega(n)$.

We also need the following two lemmas, where the first is a straightforward implication of the multivariate central limit theorem for $(N_0, \ldots, N_{2^d-1})$, while the technical proof of the second lemma is given in the Appendix B.

▶ **Lemma 7** (Multivariate central limit theorem). *Let $z \in \mathbb{C}$ with $1/2 \le \Re(z) < 1$. In distribution, in $\mathbb{C}^{2^d}$,*

$$\left(\frac{N_0^z - (\Delta_0 n)^z}{n^{z-1/2}}, \ldots, \frac{N_{2^d-1}^z - (\Delta_{2^d-1} n)^z}{n^{z-1/2}} \right) \to X,$$

where $X_i = z \cdot \Delta_i^{z-1} Y_i$ with $Y = \Sigma^{1/2} \cdot \mathcal{N}$, where $\mathcal{N} = (\mathcal{N}_0, \ldots, \mathcal{N}_{2^d-1})$ has the standard multivariate normal distribution, $(\Delta_0, \ldots, \Delta_{2^d-1})$ and $\mathcal{N}$ are stochastically independent, and the covariance matrix Σ satisfies

$$\Sigma_{i,j} = \begin{cases} \Delta_i(1 - \Delta_i) & \text{if } i = j, \\ -\Delta_i \Delta_j & \text{if } i \ne j. \end{cases}$$

▶ **Lemma 8.** *We have the following asymptotic expansions:*

(i) *for any $z \in \mathbb{C}$ with $0 < \Re(z) < 1$ and $\varepsilon > 0$, we have, as $n \to \infty$,*

$$\mathbb{E}[N_0^z] = \mathbb{E}\left[\Delta_0^z\right] n^z + \frac{z(z-1)}{2} \mathbb{E}\left[(1 - \Delta_0) \Delta_0^{z-1}\right] n^{z-1} + \mathcal{O}(n^{\varepsilon-1}).$$

(ii) *For any $z \in \mathbb{C}$ with $1/2 < \Re(z) < 1$ and fixed $p \in \mathbb{N} \setminus \{0\}$, we have*

$$\left\|N_0^z - (\Delta_0 n)^z\right\|_p = |z| \left\|\Delta_0^{\Re(z)/2} \sqrt{1 - \Delta_0}\right\|_p \|\mathcal{N}_0\|_p \, n^{\Re(z)-1/2} + o(n^{\Re(z)-1/2}).$$

(iii) *For any $z \in \mathbb{C}$ with $0 < \Re(z) < 1/2$ and fixed $p \in \mathbb{N} \setminus \{0\}$, we have*

$$\left\|N_0^z - (\Delta_0 n)^z\right\|_p = \mathcal{O}(1).$$

The first step to show Proposition 4 is to verify that the contribution of the mixed term in $b(n)$ is asymptotically negligible.

▶ **Lemma 9.** *As $n \to \infty$, we have $\mathbb{E}\left[D_n \sum_{j=0}^{2^d-1} \mathbb{L}_n^{(j)}\right] = \mathcal{O}(n^{\alpha_1 - 1/2})$.*

Proof. First of all, note that $\mathbb{E}\left[r_n^2\right] = \mathcal{O}(n^{2 \max(\alpha_{p+1}, 0)})$ since $\delta_n = \mathcal{O}\left(n^{\max(\alpha_{p+1}, 0)}\right)$ and $N_j \leq n$ for all $j = 0, \ldots, 2^d - 1$. As $\mathcal{Z}_k^{(j)}$ and (N_j, Δ_j) are stochastically independent, it follows from part *(ii)* of the previous lemma that

$$\mathbb{E}\left[\left|\mathcal{Z}_k^{(j)}\left(N_j^{\lambda_k} - (\Delta_j n)^{\lambda_k}\right)\right|^2\right] = \mathbb{E}\left[|\mathcal{Z}_k|^2\right] \mathbb{E}\left[\left|N_0^{\lambda_k} - (\Delta_0 n)^{\lambda_k}\right|^2\right] = \mathcal{O}(n^{2\alpha_k - 1}).$$

A standard application of the Cauchy-Schwarz inequality shows that $\mathbb{E}\left[D_n^2\right] = \mathcal{O}(n^{2\alpha_1 - 1})$. Next, by independence of quantities defined in subtrees, we obtain

$$\mathbb{E}\left[D_n \sum_{j=0}^{2^d-1} \mathbb{L}_n^{(j)}\right] = 2 \sum_{j=0}^{2^d-1} \mathbb{E}\left[\mathbb{L}_n^{(j)} \sum_{k=1}^p \Re\left(\mathcal{Z}_k^{(j)}\left(N_j^{\lambda_k} - (\Delta_j n)^{\lambda_k}\right)\right)\right]. \tag{10}$$

Conditionally on $\{N_0 = n_0, \ldots, N_{2^d-1} = n_{2^d-1}\}$ where $n_0 + \cdots + n_{2^d-1} = n - 1$, we have

(i) the random variables $(\Delta_0, \ldots, \Delta_{2^d-1}), (\mathcal{Z}_1^{(j)}, \ldots, \mathcal{Z}_p^{(j)}, \mathbb{L}_n^{(j)})$ are stochastically independent, and

(ii) $(\mathcal{Z}_1^{(j)}, \ldots, \mathcal{Z}_p^{(j)}, \mathbb{L}_n^{(j)})$ is distributed like $(\mathcal{Z}_1, \ldots, \mathcal{Z}_p, L_{n_j}^*)$.

To estimate (10), consider the terms

$$\mathbb{E}\left[\mathbb{L}_n^{(j)} \Re(\mathcal{Z}_k^{(j)}(N_j^{\lambda_k} - (\Delta_j n)^{\lambda_k}))\right] = \sum_{\ell=0}^{n-1} \mathbb{P}\left(N_j = \ell\right) \mathbb{E}\left[L_\ell^* \Re(\mathcal{Z}_k)\right] \mathbb{E}\left[\Re\left(\ell^{\lambda_k} - (\Delta_j n)^{\lambda_k}\right)\right]$$
$$- \sum_{\ell=0}^{n-1} \mathbb{P}\left(N_j = \ell\right) \mathbb{E}\left[L_\ell^* \Im(\mathcal{Z}_k)\right] \mathbb{E}\left[\Im\left(\ell^{\lambda_k} - (\Delta_j n)^{\lambda_k}\right)\right].$$

By the trivial bound $\mathbb{E}[(L_n^*)^2] = \mathcal{O}(n^2)$, it follows from the Cauchy-Schwarz inequality that there exists a constant $C > 0$ such that

$$\max\left\{\mathbb{E}\left[L_n^* \Re(\mathcal{Z}_k)\right], \mathbb{E}\left[L_n^* \Im(\mathcal{Z}_k)\right]\right\} \leq Cn.$$

Therefore,

$$\left|\mathbb{E}\left[\mathbb{L}_n^{(j)} \Re(\mathcal{Z}_k^{(j)}(N_j^{\lambda_k} - (\Delta_j n)^{\lambda_k}))\right]\right| \leq 2Cn \left|\mathbb{E}\left[N_j^{\lambda_k} - (\Delta_j n)^{\lambda_k}\right]\right|. \tag{11}$$

From part *(i)* of the previous lemma, it follows that the right hand side of (11) grows at most of the order n^{α_k}. Overall, this shows that

$$\mathbb{E}\left[D_n \sum_{j=0}^{2^d-1} \mathbb{L}_n^{(j)}\right] = \mathcal{O}(n^{\alpha_1}).$$

Combining the bounds on $\mathbb{E}\left[r_n^2\right], \mathbb{E}\left[D_n^2\right]$ and the last display, Theorem 5 yields $\mathrm{Var}(L_n^*) = \mathcal{O}(n)$. Repeating the last steps using this improved bound concludes the proof. ◀

The previous proposition suggests that the order of magnitude of the additive term in (9) is $\max\{n^{2\max(\alpha_{p+1},0)}, n^{2\alpha_1-1}\}$. For most values of d, we have $2\alpha_1 - 1 > 2\alpha_{p+1}$. Indeed, for $9 \leq d \leq 10,000$, there exist only 31 values ranging from $d = 15$ to $d = 8598$ for which the converse is true. It is important to note that, for all $d \geq 9$, we have $2\alpha_1 - 1 \neq 2\alpha_{p+1}$ since the contrary would imply that $\omega + \omega^{d-1} - \omega^{p+1} - \omega^{d-p-1} = 1/2$ which is impossible since the left hand side is an algebraic integer. In particular, in light of Theorem 5 and Lemma 6, the following two propositions verifying that $b(n) \to \infty$ are the missing pieces to conclude the proof of Proposition 4.

▶ **Proposition 10.** *Let $\alpha_{p+1} > 0$, that is, $d > 11$ and*

$$W := \sum_{i=0}^{2^d-1} \Delta_i^{\lambda_{p+1}} = \prod_{i=1}^{d}\left(U_i^{\lambda_{p+1}} + (1 - U_i)^{\lambda_{p+1}}\right).$$

For $x \in \mathbb{R}$, let

$$\Phi(x) := 2\Re\left(\gamma_{p+1}^2 \mathbb{E}\left[(1 - W)^2\right] e^{2i\beta_{p+1}x}\right) + 2|\gamma_{p+1}|^2 \mathbb{E}\left[|1 - W|^2\right].$$

Φ is a smooth periodic function with period π/β_{p+1}, amplitude $2|\gamma_{p+1}|^2 |\mathbb{E}\left[(1 - W)^2\right]|$ and

$$\min_{x\in\mathbb{R}} \Phi(x) = 2|\gamma_{p+1}|^2 \left[\mathbb{E}\left[|1 - W|^2\right] - |\mathbb{E}\left[(1 - W)^2\right]|\right] > 0.$$

As $n \to \infty$,

$$\mathbb{E}\left[r_n^2\right] = \Phi(\log n)n^{2\alpha_{p+1}} + \mathcal{O}(n^{\alpha_{p+1}+\alpha_{p+2}}).$$

▶ **Proposition 11.** *Let (Δ, Y) be as in Lemma 7 and stochastically independent of $\mathcal{Z}_1^{(0)}, \ldots, \mathcal{Z}_1^{(2^d-1)}$. Set*

$$\mathcal{W} = \sum_{j=0}^{2^d-1} \lambda_1 \mathcal{Z}_1^{(j)} \Delta_j^{\lambda_1-1} Y_j.$$

For $x \in \mathbb{R}$, define

$$\Psi(x) := 2\Re\left(\mathbb{E}\left[\mathcal{W}^2\right] e^{2i\beta_{p+1}x}\right) + 2\mathbb{E}\left[|\mathcal{W}|^2\right].$$

Ψ is a smooth periodic function with period π/β_{p+1}, amplitude $2|\mathbb{E}\left[\mathcal{W}^2\right]|$ and

$$\min_{x\in\mathbb{R}} \Psi(x) = 2\left(\mathbb{E}\left[|\mathcal{W}|^2\right] - |\mathbb{E}\left[\mathcal{W}^2\right]|\right) > 0.$$

As $n \to \infty$, we have

$$\mathbb{E}[D_n^2] = \Psi(\log n)n^{2\alpha_1-1} + o(n^{2\alpha_1-1}).$$

The proofs of these propositions are very similar and we only present the proof of Proposition 11 which is more involved.

Proof of Proposition 11. By definition, Ψ has period π/β_{p+1}. Next, for any $z \in \mathbb{C}$, it is easy to see that the global maximum and minimum of the function $x \mapsto \Re(z \exp(ix))$ are $|z|$ and $-|z|$. This implies the remaining claims on the shape of Ψ. $\min_{x \in \mathbb{R}} \Psi(x) > 0$ follows from triangle inequality upon verifying that $\arg(\mathcal{W})$ is not almost surely constant. This, in turn follows from that fact that, for any given (affine) line $L \subseteq \mathbb{C}$, we have $\mathbb{P}(\mathcal{Z}_1 \in L) = 0$. This is an immediate corollary of the fact that $(\Re(\mathcal{Z}_1), \Im(\mathcal{Z}_1))$ admits a density on $\mathbb{R}^2$ (see Proposition 3). For the asymptotic expansion of D_n, note that, following the steps involving the Cauchy-Schwarz inequality and the bounds stated in the proof of Proposition 9, it is straightforward to verify that

$$\mathbb{E}\left[D_n^2\right] = 4\mathbb{E}\left[\left(\sum_{j=0}^{2^d-1} \Re\left(\mathcal{Z}_1^{(j)}\left(N_j^{\lambda_1} - (n\Delta_j)^{\lambda_1}\right)\right)\right)^2\right] + \mathcal{O}(n^{\alpha_1 + \alpha_2 - 1}).$$

By the multivariate central limit theorem stated in Lemma 7, the first term is asymptotically equivalent to $\Psi(\log n) n^{2\alpha_1 - 1}$ which proves the expansion. ◀

---- **References** ----

1 B. Chauvin and N. Pouyanne. m-ary search trees when $m > 27$: a strong asymptotics for the space requirements. *Random Struct. Algor.*, 24(2):133–154, 2004.

2 H.-H. Chern, M. Fuchs, and H.-K. Hwang. Phase changes in random point quadtrees. *ACM Trans. Algorithms*, 3(3):51 pages, 2007.

3 H.-H. Chern, M. Fuchs, H.-K. Hwang, and R. Neininger. Dependence and phase changes in random m-ary search trees. *Random Struct. Algor.*, 50(3):353–379, 2017.

4 H.-H. Chern and H.-K. Hwang. Phase changes in random m-ary search trees and generalized quicksort. *Random Struct. Algor.*, 19(3-4):316–358, 2001.

5 L. Devroye. Limit laws for local counters in random binary search trees. *Random Struct. Algor.*, 2(3):303–315, 1991.

6 J. A. Fill and N. Kapur. The space requirement of m-ary search trees: distributional asymptotics for $m > 27$. *In Proceedings of the 7th Iranian Statistical Conference*, 2004. URL: http://www.ams.jhu.edu/~fill/papers/periodic.pdf.

7 P. Flajolet, G. Labelle, L. Laforest, and B. Salvy. Hypergeometrics and the cost structure of quadtrees. *Random Struct. Algor.*, 7(2):117–144, 1995.

8 P. Flajolet and R. Sedgewick. Mellin transforms and asymptotics: finite differences and Rice's integrals. *Theoretical Computer Science*, 144(1-2):101–124, 1995.

9 M. Fuchs. A note on the quicksort asymptotics. *Random Struct. Algor.*, 46(4):677–687, 2015.

10 S. Janson. Functional limit theorems for multitype branching processes and generalised Pólya urns. *Stoch. Process. Appl.*, 110:177–245, 2004.

11 S. Janson. Congruence properties of depths in some random trees. *Alea Lat. Am. J. Probab. Math. Stat.*, 1:347–366, 2006.

12 K. Leckey. On densities for solutions to stochastic fixed point equations. *Preprint*, 2016. URL: https://arxiv.org/abs/1604.05787.

13 W. Lew and H. M. Mahmoud. The joint distribution of elastic buckets in multiway search trees. *SIAM J. Comput.*, 23(5):1050–1074, 1994.

14 H. M. Mahmoud and B. Pittel. Analysis of the space of search trees under the random insertion algorithm. *J. Algorithms*, 10(1):52–75, 1989.

15   N. S. Müller. Central limit theorem analogues for multicolour urn models. *Preprint*, 2016. URL: `https://arxiv.org/abs/1604.02964`.

16   N. S. Müller and R. Neininger. The CLT analogue for cyclic urns. *Analytic Algorithmics and Combinatorics (ANALCO)*, pages 121–127, 2016.

17   N. S. Müller and R. Neininger. Refined asymptotics for the composition of cyclic urns. *Submitted*, 2016. URL: `https://arxiv.org/abs/1612.08930`.

18   R. Neininger. Refined quicksort asymptotics. *Random Struct. Algor.*, 46(2):346–361, 2015.

A   Appendix A

Proof of Proposition 2. These arguments are well-known. By construction,

$$
\mathcal{Z}_{n+1,k}^{(v)} - \mathcal{Z}_{n,k}^{(v)} = \sum_{j=0}^{2^d-1} \left(\Delta_j^{(v)} \right)^{\lambda_k} \cdot \left(\mathcal{Z}_{n,k}^{(vj)} - \mathcal{Z}_{n-1,k}^{(vj)} \right),
$$

and therefore

$$
\begin{aligned}
\Delta_n^{(v)} &:= \mathbb{E}\left[\left| \mathcal{Z}_{n+1,k}^{(v)} - \mathcal{Z}_{n,k}^{(v)} \right|^2 \right] \\
&= \mathbb{E}\left[\left| \mathcal{Z}_{n,k}^{(v)} - \mathcal{Z}_{n-1,k}^{(v)} \right|^2 \right] \sum_{j=0}^{2^d-1} \mathbb{E}\left[\left(\Delta_j^{(v)} \right)^{2\alpha_k} \right] =: q \cdot \Delta_{n-1}^{(v)},
\end{aligned}
$$

where we used that the claims of part (i) and part (ii) of this proposition also hold if $Z_k^{(v)}$ is replaced by $Z_{n,k}^{(v)}$ (from construction). As $0 < q < 1$, it immediately follows that $\mathbb{E}\big[|\mathcal{Z}_{n,k}^{(v)}|^2\big], n \geq 1$ is a bounded sequence. Since $\mathcal{Z}_{n,k}^{(v)}, n \geq 1$ is a martingale, the sequence converges almost surely and in L_2 by the L_2-convergence theorem for martingales. This shows (6).

(i) and *(ii)* follow from the construction (see the comment succeeding the above display).

(iii) follows from a standard contraction argument for probability measures on $\mathbb{C}$ with mean γ_k and finite second moment. Convergence of p-th moments is proved inductively using $p = 2$ as base case; details will be given in the journal version of this paper.   ◄

Proof of Proposition 3. Leckey [12] recently established a set of conditions under which solutions of fixed-point equations such as (7) admit Schwartz densities. More precisely, since we have already seen that $\mathcal{Z}_k$ has finite moments of all orders, applying [12, Theorem 4.2] in conjunction with Remark 4.9 only requires to verify conditions (A1) - (A5) from Definition 4.1. The only condition which is not trivially satisfied is (A4): the support of $\mathcal{Z}_k$ ought to be in general position, that is, contain three points z_1, z_2, z_3 which do not lie on a line. For all $x \in [0,1]$, the vector $(x, 1-x, 0, 0, \ldots, 0)$ lies in the support of Δ. Therefore, $(x^{\lambda_k}, (1-x)^{\lambda_k}, 0, 0, \ldots, 0)$ lies in the support of Δ^{λ_k}. Hence, for any z in the support of $\mathcal{Z}_k$, also $(x^{\lambda_k} + (1-x)^{\lambda_k})z$ lies in the support of $\mathcal{Z}_k$. As the support of $\mathcal{Z}_k$ contains a non-zero element and $\beta_k \neq 0$, this concludes the proof.   ◄

B   Appendix B

Proof of Lemma 6. We start with part (a). Let n_0 be the first index such that $b_{n_0} > 0$. Set

$$
\tilde{b}_n = \begin{cases} 0, & \text{if } 1 \leq n \leq n_0 \\ b_n + 2^d \pi_{n,n_0} b_{n_0}, & \text{if } n \geq n_0 + 1 \end{cases}
$$

and denote by $\tilde{a}_n$ the corresponding sequence. Obviously, $a_n \geq \tilde{a}_n$ and thus it suffices to prove the claim for the sequence $\tilde{a}_n$. Note that by the above definition

$$\tilde{b}_n \geq c \frac{\log^d n}{n}, \qquad (n \geq n_0 + 1)$$

for some positive $c > 0$ since

$$\pi_{n,j} = \frac{1}{d!} \cdot \frac{\log^d n}{n} \left(1 + \mathcal{O}\left(\frac{1}{\log n} \right) \right)$$

for fixed j (see Lemma 4 in [7]). We now claim that

$$\tilde{a}_n \geq d \left(n + \frac{1}{2^d - 1} \right), \qquad (n \geq n_0 + 1) \tag{12}$$

for some $d > 0$ which will be chosen below. We prove this claim by induction. Clearly, the claim is true for $n = n_0 + 1$. Next, in order to prove the induction step, plug the above claim into the recurrence for $\tilde{a}_n$. This yields

$$
\begin{aligned}
\tilde{a}_n &\geq d2^d \sum_{0 \leq j < n} \pi_{n,j} \left(j + \frac{1}{2^d - 1} \right) - d2^d \sum_{0 \leq j < n_0 + 1} \pi_{n,j} \left(j + \frac{1}{2^d - 1} \right) + c \frac{\log^d n}{n} \\
&\geq d \left(n - 1 + \frac{2^d}{2^d - 1} \right) + (c - dK) \frac{\log^d n}{n} \\
&\geq d \left(n + \frac{1}{2^d - 1} \right),
\end{aligned}
$$

where in the second estimate we used

$$2^d \sum_{0 \leq j < n} j \cdot \pi_{n,j} = \mathbb{E}\left(\sum_{\ell=0}^{2^d - 1} N_\ell \right) = n - 1$$

and

$$2^d \sum_{0 \leq j < n_0 + 1} \pi_{n,j} \left(j + \frac{1}{2^d - 1} \right) \leq K \frac{\log^d n}{n}$$

which follows from (12). Moreover, the last estimate follows if d is chosen such that $0 < d \leq c/K$. This concludes the induction step and thus also the proof.

(b) Assume that $b_n > 0$ for all $n \geq n_0$. The claim follows from part (a) by setting

$$\tilde{b}_n = \begin{cases} 0, & \text{if } 1 \leq n < n_0, \\ b_n, & \text{if } n \geq n_0 \end{cases}$$

and noting that the corresponding sequence $\tilde{a}_n$ satisfies $a_n \geq \tilde{a}_n$. ◀

Proof of Lemma 8. Throughout the proof, let $\alpha = \Re(z)$. Further, here, and subsequently, we write $\text{Bin}(n - 1, u)$ for a random variable with binomial distribution with parameters $n - 1$ and u.

(i) By construction, Δ_0 is distributed as $\exp(-\Gamma^*(d))$, where $\Gamma^*(d)$ is a random variable with the Gamma distribution with density $((d - 1)!)^{-1} t^{d-1} \exp(-t)$ for $t > 0$. It follows that Δ_0 has density $((d - 1)!)^{-1}(-\log t)^{d-1}$ for $t \in (0, 1)$. Hence,

$$\left| \mathbb{E}\left[N_0^z \mathbf{1}_{[0, n^{-1+\varepsilon}]}(\Delta_0) \right] \right| \leq \mathbb{E}\left[\text{Bin}(n, n^{-1+\varepsilon}) \right]^\alpha \mathbb{P}\left(\Delta_0 \leq n^{-1+\varepsilon} \right) \leq C n^{-1+\varepsilon} \log n.$$

Next, by part *(i)* of the (well-known) postponed Lemma 12 below, we have

$$\mathbb{E}\left[N_0^z \mathbf{1}_{[n^{-1+\varepsilon},1]}(\Delta_0)\right] = \mathbb{E}\left[\Delta_0^z \mathbf{1}_{[n^{-1+\varepsilon},1]}(\Delta_0)\right] n^z$$
$$+ \frac{z(z-1)}{2} \mathbb{E}\left[(1-\Delta_0)\Delta_0^{z-1}\mathbf{1}_{[n^{-1+\varepsilon},1]}(\Delta_0)\right] n^{z-1} + \mathcal{O}(n^{\varepsilon-1}).$$

Dropping the indicators on the right hand side only adds a negligible error term as $\left|\mathbb{E}\left[\Delta_0^z\mathbf{1}_{[0,n^{-1+\varepsilon})}(\Delta_0)\right]\right| \leq n^{(-1+\varepsilon)\Re(z)-1+\varepsilon}\log n$ with a similar computation for the second summand.

(ii) We have

$$\mathbb{E}\left[\left|N_0^z - (\Delta_0 n)^z\right|^p\right]$$
$$\leq 2^p((d-1)!)^{-1}(\mathbb{E}\left[\text{Bin}(n-1,1/n)^{\alpha p}\right]+1)\int_0^{1/n}(-\log u)^{d-1}du$$
$$+ ((d-1)!)^{-1}\int_{1/n}^1(-\log u)^{d-1}\mathbb{E}\left[\left|(\text{Bin}(n-1,u))^z - (un)^z\right|^p\right]du.$$

Part *(ii)* of Lemma 12 below shows that the integral in the second summand is bounded by

$$C\int_{1/n}^\infty(-\log u)^{d-1}\left((un)^{p(\alpha-1/2)} + (un)^{p\alpha}e^{-Cun}\right)du = \mathcal{O}(n^{p(\alpha-1/2)}).$$

As

$$\int_0^{1/n}(-\log u)^{d-1}du = \frac{1}{n}(\log n)^{d-1}(1+\mathcal{O}((\log n)^{-1})),$$

it follows that

$$\left\|N_0^z - (\Delta_0 n)^z\right\|_p = \mathcal{O}\left(n^{\alpha-1/2}\right).$$

This shows that the marginals in the mutivariate central limit theorem stated in Lemma 7 converge with respect to all moments. This shows *(ii)*. *(iii)* follows along similar lines. ◀

▶ **Lemma 12.** *Let $z \in \mathbb{C}$ with $0 < \alpha := \Re(z) < 1$. We have the following asymptotic expansions:*

(i) *for any $\varepsilon > 0$ sufficiently small, as $n \to \infty$, uniformly in $n^{-1+\varepsilon} \leq u \leq 1$,*

$$\mathbb{E}[\text{Bin}(n,u)^z] = (nu)^z + \frac{z(z-1)}{2}(1-u)(nu)^{z-1} + \mathcal{O}(n^{\varepsilon-1}).$$

(ii) *For $p \in \mathbb{N}\setminus\{0\}$, there exists a constant $C > 0$ such that*

$$\mathbb{E}\left[\left|\text{Bin}(n,u)^z - (nu)^z\right|^p\right]$$
$$\leq C\left(\mathbf{1}_{(0,1/n)}(u) + \mathbf{1}_{[1/n,1]}(u)\left((un)^{p(\alpha-1/2)} + (un)^{p\alpha}e^{-Cun}\right)\right).$$

Proof. *(i)* On $[1/2,3/2]$, we have

$$x^z = 1 + z(x-1) + \frac{z(z-1)}{2}(x-1)^2 + \gamma(x)(x-1)^3,$$

for some function γ which is bounded on $[1/2, 3/2]$. Let $A = \{\mathrm{Bin}(n, u)/(nu) \in [1/2, 3/2]\}$. Plugging $x = \mathrm{Bin}(n, u)\mathbf{1}_A/(nu)$ into the last display and taking the expectation gives

$$\frac{\mathbb{E}[\mathrm{Bin}(n, u)^z \mathbf{1}_A]}{(nu)^z} = 1 + z\mathbb{E}\left[\frac{\mathrm{Bin}(n, u)\mathbf{1}_A}{nu} - 1\right] + \frac{z(z-1)}{2}\mathbb{E}\left[\left(\frac{\mathrm{Bin}(n, u)\mathbf{1}_A}{nu} - 1\right)^2\right]$$
$$+ \mathcal{O}\left(\mathbb{E}\left[\left(\frac{\mathrm{Bin}(n, u)\mathbf{1}_A}{nu} - 1\right)^3\right]\right).$$

By Chernoff's inequality, since $u \geq n^{-1+\varepsilon}$, we have $\mathbb{P}(A) \leq C_1 \exp(-C_2 n^\varepsilon)$ for some universal constants $C_1, C_2 > 0$. Hence, dropping the indicator $\mathbf{1}_A$ in all expectations in the last display adds a negligible error term.

(ii) For $u \leq 1/n$, we can bound $\mathbb{E}\left[\mathrm{Bin}(n, u)^{\alpha k}\right] \leq \mathbb{E}\left[\mathrm{Bin}(n, 1/n)^{\alpha k}\right] \to \mathbb{E}\left[P^{\alpha k}\right]$ as $n \to \infty$. (Here, P denotes a random variable with the Poisson distribution and mean one.) Obviously, $(nu)^{\alpha k} \leq 1$. This shows one part of the inequality. For the more interesting case $u \geq 1/n$, first observe that

$$\mathbb{E}\left[\left|\mathrm{Bin}(n, u)^z - (nu)^z\right|^p\right]$$
$$\leq 2^k \left(\mathbb{E}\left[|\mathrm{Bin}(n, u)^\alpha - (nu)^\alpha|^p\right] + \mathbb{E}\left[\left|\mathrm{Bin}(n, u)^\alpha \cdot \log\frac{\mathrm{Bin}(n, u)}{nu}\right|^p\right]\right)$$
$$=: 2^k(f_1(u, n) + f_2(u, n)).$$

Set $E_n = \{\mathrm{Bin}(n, u) > (nu)/2\}$ and define

$$f_1(u, n) = \mathbb{E}\left[|\mathrm{Bin}(n, u)^\alpha - (nu)^\alpha|^p \, \mathbf{1}_{E_n}\right] + \mathbb{E}\left[|\mathrm{Bin}(n, u)^\alpha - (nu)^\alpha|^p \, \mathbf{1}_{E_n^c}\right]$$
$$=: g_1(u, n) + h_1(u, n),$$

and

$$f_2(u, n) = \mathbb{E}\left[\left|\mathrm{Bin}(n, u)^\alpha \cdot \log\frac{\mathrm{Bin}(n, u)}{nu}\right|^p \mathbf{1}_{E_n}\right]$$
$$+ \mathbb{E}\left[\left|\mathrm{Bin}(n, u)^\alpha \cdot \log\frac{\mathrm{Bin}(n, u)}{nu}\right|^p \mathbf{1}_{E_n^c}\right]$$
$$=: g_2(u, n) + h_2(u, n),$$

We now give bounds on g_1, g_2, h_1 and h_2. Let $\varrho(t) = (1+t)^\alpha$. Then, $|\varrho'(t)| \leq \alpha 2^{1-\alpha}$ for all $t \geq -1/2$. Thus, by the postponed Lemma 13 below,

$$g_1(u, n) = (nu)^{\alpha p}\mathbb{E}\left[\left|\varrho\left(\frac{\mathrm{Bin}(n, u) - nu}{nu}\right) - 1\right|^p \mathbf{1}_{E_n}\right]$$
$$\leq (\alpha 2^{1-\alpha})^p (nu)^{p(\alpha-1)}\mathbb{E}\left[|\mathrm{Bin}(n, u) - nu|^p\right] \leq C(un)^{p(\alpha-1/2)}$$

for some $C > 0$. Next, we consider g_2. Let $\psi(t) = t^\alpha |\log t|$. As ψ' is bounded on $[1/2, \infty)$, by, say $C_1 > 0$, we have

$$g_2(u, n) = (nu)^{\alpha p}\mathbb{E}\left[\psi\left(\frac{\mathrm{Bin}(n, u) - nu}{nu}\right)^p \mathbf{1}_{E_n}\right]$$
$$\leq C_1(nu)^{p(\alpha-1)}\mathbb{E}\left[|\mathrm{Bin}(n, u) - nu|^p\right] \leq C_2(un)^{p(\alpha-1/2)}.$$

for some $C_2 > 0$. Next, by the Cauchy-Schwarz inequality and the postponed Lemma 13 below,

$$h_1(u, n) \leq \mathbb{E}\left[|\mathrm{Bin}(n, u) - nu|^{2p\alpha}\right]^{1/2} \mathbb{P}\left(E_n^c\right)^{1/2}$$
$$\leq C_1 (nu)^{p\alpha/2} e^{-Cun}$$

for some $C > 0$. Since ψ in bounded on $[0, 1]$ by, say $C > 0$, we also have

$$h_2(u, n) \leq (nu)^{\alpha p} \mathbb{E}\left[\psi\left(\frac{\mathrm{Bin}(n, u)}{nu}\right)^p \mathbf{1}_{E_n^c}\right] \leq (nu)^{\alpha p} e^{-Cun}.$$

This concludes the proof.   ◀

▶ **Lemma 13.** *For any real $r \geq 1$ there exists a constant $C > 0$ such that, for all $n \geq 1$ and $u \in [0, 1]$, we have*

$$\mathbb{E}\left[|\mathrm{Bin}(n, u) - nu|^r\right] \leq C(nu)^{r/2}.$$

Proof. By Jensen's inequality, we may restrict ourselves to the case of integer r. Using Bernstein's inequality, we obtain

$$\mathbb{E}\left[|\mathrm{Bin}(n, u) - nu|^r\right] = r \int_0^\infty y^{r-1} \mathbb{P}\left(|\mathrm{Bin}(n, u) - nu| \geq y\right) dy$$
$$\leq r \int_0^\infty y^{r-1} \exp\left(-\frac{y^2}{2nu + 2y/3}\right) dy$$
$$\leq r \int_0^{6nu} y^{r-1} \exp\left(-\frac{y^2}{6np}\right) dy + r \int_{6nu}^\infty y^{r-1} e^{-y} dy.$$

Sustituting $x = y/\sqrt{6nu}$, one finds that the first term is bounded by $C(nu)^{k/2}$ for all $n \geq 1$ and $u \in [0, 1]$. The second summand is $\mathcal{O}(\exp(-\alpha np))$ for any $\alpha < 6$.   ◀

Slow Convergence of Ising and Spin Glass Models with Well-Separated Frustrated Vertices

David Gillman
Department of Computer Science, New College of Florida, Sarasota, FL, 34243, USA
dgillman@ncf.edu

Dana Randall[1]
School of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332, USA
randall@cc.gatech.edu

─── **Abstract** ───

Many physical models undergo phase transitions as some parameter of the system is varied. This phenomenon has bearing on the convergence times for local Markov chains walking among the configurations of the physical system. One of the most basic examples of this phenomenon is the ferromagnetic Ising model on an $n \times n$ square lattice region Λ with mixed boundary conditions. For this spin system, if we fix the spins on the top and bottom sides of the square to be $+$ and the left and right sides to be $-$, a standard Peierls argument based on energy shows that below some critical temperature t_c, any local Markov chain $\mathcal{M}$ requires time exponential in n to mix.

Spin glasses are magnetic alloys that generalize the Ising model by specifying the strength of nearest neighbor interactions on the lattice, including whether they are ferromagnetic or antiferromagnetic. Whenever a face of the lattice is bounded by an odd number of edges with ferromagnetic interactions, the face is considered *frustrated* because the local competing objectives cannot be simultaneously satisfied. We consider spin glasses with exactly four well-separated frustrated faces that are symmetric around the center of the lattice region under 90 degree rotations. We show that local Markov chains require exponential time for all spin glasses in this class. This class includes the ferromagnetic Ising model with mixed boundary conditions described above, where the frustrated faces are on the boundary. The standard Peierls argument breaks down when the frustrated faces are on the interior of Λ and yields weaker results when they are on the boundary of Λ but not near the corners. We show that there is a universal temperature T below which $\mathcal{M}$ will be slow for all spin glasses with four well-separated frustrated faces. Our argument shows that there is an exponentially small cut indicated by the *free energy*, carefully exploiting both entropy and energy to establish a small bottleneck in the state space to establish slow mixing.

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains

Keywords and phrases Mixing time, spin glass, Ising model, mixed boundary conditions, frustration

Digital Object Identifier 10.4230/LIPIcs.AofA.2018.24

---

1 Introduction

The celebrated Ising model on the Cartesian lattice is a fundamental model for ferromagnetism and one of the simplest models demonstrating an order-disorder phase transition. Each configuration σ in the state space $\Omega = \{-1, +1\}^{n^2}$ consists of an assignment of a $+$ or $-$ spin to each of the vertices, and the *Gibbs (or Boltzmann) distribution* assigns weight

$$\pi(\sigma) = e^{-\beta H(\sigma)}/Z(\beta),$$

where

$$H(\sigma) = - \sum_{(i,j) \in E} \sigma_i \sigma_j$$

is the *Hamiltonian* (or energy) of the system, $\beta = 1/T$ is inverse temperature, and $Z(\beta) = \sum_{\sigma \in \Omega} e^{-\beta H(\sigma)}$ is the normalizing constant known as the *partition function*. In Sections 3 and 4 it will be convenient to write the probability of a configuration in terms of $\lambda = e^{2\beta} = e^{2/T}$, where λ can be seen as the weight assigned to edges whose endpoints are assigned like spins.

Physicists characterize when there is a phase transition in a physical model by asking whether there is a unique limiting conditional distribution on finite subregions as the lattice size grows. The Gibbs distribution is defined as any limiting measure, but this limit might not be unique. For example, for the Ising model on $\mathbb{Z}^2$ at sufficiently low temperatures, the probability of an interior vertex being assigned $+$ will be much higher if the boundary vertices were hard-wired to be $+$ than if they were hard-wired to be $-$, and this difference persists in the limit. The infinite volume Ising model was solved exactly by Onsager in 1944 [23], showing that there is a critical value $\beta_c = \ln(1 + \sqrt{2})/2$ such that for $\beta < \beta_c$ (i.e., high temperature), the limiting distribution is unique, and for $\beta > \beta_c$ (i.e., low temperature), spins on the boundary of the region persist and there are multiple limiting distributions. The all-plus and the all-minus boundary conditions are known to be extremal [1, 12]. and all other infinite-volume Gibbs measures are convex combinations of these extremal measures.

A related effect has been observed in the context of mixing times of local Markov chains for the Ising model on finite lattice regions with free boundaries (i.e., boundary vertices can take on either spin). The *mixing time* $\tau(\mathcal{M})$ of a chain $\mathcal{M}$, i.e., the number of steps required so that the distribution over configurations is close to its stationary distribution, also undergoes a phase change. When β is small, local dynamics are known to be efficient [18, 19, 15], while when β is large, local chains require exponential time to converge to equilibrium [31]. At low enough temperature, the Gibbs distribution strongly favors configurations that have predominantly one spin, and it will take exponential time to move from a mostly $+$ state to a mostly $-$ one using moves that only change $o(n^2)$ sites at a time [17].

Mixing times of Markov chains are known to be sensitive to boundary conditions. For example, local chains on Ising configurations are conjectured to converge in polynomial time at all temperatures for the "all $+$" boundary condition where all vertices on the boundary are hard-wired to have $+$ spins. While still open, Martinelli [16] showed mixing is indeed sub-exponential at all temperatures with all $+$ boundary conditions and subsequently Lubetsky et al. [15] showed that the chain converges in quasi-polynomial time. However, a standard Peierls argument can be used to show that when there are mixed boundary conditions with 4 connected components of like spins on the boundary, alternating "$+, -, +, -$", then the chain again will be slow at low temperatures. In particular, for *mixed boundary conditions* where we fix the boundary to be $+$ on all vertices on the vertical sides of the boundary and $-$ on the horizontal, then the chain provably requires time exponential in n at sufficiently low

temperature. For "p-shifted mixed boundary conditions" where we rotate the mixed boundary conditions clockwise p units, [6] for the Ising model establish bounds on the temperature below which convergence is slow, but they do not easily extend to other cases we consider.

Similar questions can be examined in the context of *spin glasses*, or magnetic alloys that are a natural generalization of the ferromagnetic and antiferromagnetic Ising models. We are given a graph $G = (V, E)$ and a set of couplings $J_{ij} \in \{-1, +1\}$ for each edge $(i, j) \in E$. The state space $\Omega = \{-1, +1\}^V$, where a configuration assigns a spin to each vertex in V. For a spin glass configuration $\sigma \in \Omega$, the Hamiltonian is defined as

$$H(\sigma) = - \sum_{(i,j) \in E} J_{ij}\sigma(i)\sigma(j)$$

and the Gibbs distribution is defined as for the Ising model as $\pi(\sigma) = e^{-\beta H(\sigma)}/Z(\beta)$.

When all the $J_{ij} = +1$, this model is precisely the ferromagnetic Ising model on G; when all the $J_{ij} = -1$, it is antiferromagnetic. In general, the behavior of a spin glass is much richer than simple models of magnetism because of the presence of *frustration*, or competition between local interactions. In the case of $G = \Lambda$, a square region in the lattice, a face of Λ is *frustrated* when $J_{ij} = -1$ for an odd number of edges around the face. No setting of the sites on the corners of such a face will satisfy all four edges, i.e., make each $J_{ij}\sigma(i)\sigma(j) = 1$. Even finding the ground states (or most likely configurations) reduces to solving an optimization problem that can be NP-hard (see, e.g., [2]. It will be convenient to refer to the dual lattice $\overline{\Lambda} = (\overline{V}, \overline{E})$ and refer to a frustrated face f of Λ by the frustrated vertex $v = \overline{f}$ in $\overline{V}$.

Here, we study spin glasses with exactly four well-separated frustrated vertices in order to understand the long-range interactions and their effects on mixing times. Notice that the Ising model with p-shifted mixed boundary conditions is a special case, where all four frustrated squares lie just inside the boundary. Models with well-separated defects are widely studied to understand long-range correlation; for example, in seminal work, Ciucu [4] studied the monomer-dimer model with a constant number of monomers to establish a connection with electrical networks and settle a nearly century old conjecture about long-range effects due to the separation of the monomers. It is natural to consider similar questions in the context of spin glasses with a few well-separated frustrated vertices.

We show that there is a universal temperature T below which the Markov chain $\mathcal{M}$ will be slow for any spin glass with exactly four frustrated vertices defining the corners of a (not necessarily axis-aligned) square in $\overline{\Lambda}$. We identify a bottleneck in the state space by looking at the how the *free energy* (i.e., $\ln Z/n^2$) changes as a parameter of the system is varied.

▶ **Theorem 1.1.** *Let Λ be the $kn \times kn$ lattice region, $k \geq 2$. Suppose that four distinguished faces $f_1, .., f_4$ are symmetric around the center of the lattice region under 90 degree rotations. There is a universal temperature $T = 0.360\ldots$ such that the Glauber dynamics $\mathcal{M}$ for the spin glass model on Λ with $f_1, ..., f_4$ the faces with frustration has mixing time $\tau(\mathcal{M}) \geq e^{cn}$, for some constant $c > 0$, whenever $t < T$.*

As a corollary, this gives a universal bound on the temperature for the Ising model with p-shifted mixed boundary conditions.

The proof of Theorem 1.1 requires several innovations. The standard argument to show slow mixing is based on the *conductance* of the Markov chain. The key is showing that the state space Ω can be partitioned into two sets, S and its complement S^C, such that getting from S to some subset S^C requires passing through a small cutset $\mathcal{C} \subset S^C$, and the stationary weights $\pi(S)$ and $\pi(S^C)$ are both exponentially larger than $\pi(\mathcal{C})$. This establishes that the chain has low conductance, which implies it takes exponential time to converge to equilibrium

[13]. The main ingredient is typically a *Peierls argument* [24], which introduces a map Ψ from $\mathcal{C}$ to $S \cup S^C$. Typically Ψ is chosen so that for all $\sigma \in \mathcal{C}$, we have $\pi(\Psi(\sigma)) \geq \pi(\sigma)e^{cn}$, mapping elements of $\mathcal{C}$ to configurations with exponentially larger weight. If we can show that Ψ is nearly injective (i.e., the cardinality of the inverse image of each configuration is bounded by a polynomial), then we can conclude that $\pi(\mathcal{C})$ is exponentially small.

In our setting, there is not always a natural candidate map that increases the probability of a configuration exponentially. In fact, the standard map gives no guaranteed increase to the stationary probability when each side of the boundary has close to an equal number of $+$ and $-$ spins (when $p = 0.5$ and the boundary changes spin at the center of the four sides of the boundary). In this case, we exploit the low *entropy* of $\mathcal{C}$ by defining an injective map from $\mathcal{C} \times 2^{cn} \to \Omega$, for some $c > 0$. The map never decreases the weight of a configuration, so we again can conclude that $\pi(\mathcal{C})$ is exponentially small. As we vary p, the *free energy* of $\mathcal{C}$ remains small compared to the two sides of the cut due to a derease in energy (when p is close to 0) or due to entropy (when p is close to 0.5); all other cases rely on both.

An important technical contribution in our proofs is in the construction of a new injective map. The *contour representation* of a spin glass configuration consists of edges in the dual lattice that cross edges $e = (i, j)$ where $J_{ij}\sigma(i)\sigma(j) = -1$; in this representation the frustrated vertices in the dual lattice have odd degree and all other vertices have even degree. Because of this property the contour representation can be decomposed into a even cycles (closed contours) and two long paths whose endpoints are the four frustrated vertices. In the standard case of the Ising model with alternating side boundary conditions, we can define an injective map that shifts the paths connecting the four frustrated vertices to paths with much shorter length, and therefore much larger probability. The new paths can be added along the boundary by shifting closed contours. In our case we cannot do this since we cannot always construct maps to configurations with larger probability. Therefore we define a map to a *set* of configurations of *at least equal* probability. To complete the proof we require a careful map that allows us to reconstruct the original path, the new path, and the closed contours that are intersected when the new path is added. Verifying that the map is injective now requires a very sensitive combinatorial encoding and decoding that is likely of independent interest.

2 Preliminaries

We review some standard background on Markov chains, convergence times, and the Ising model that are required for our results.

2.1 Markov chains and mixing times

Let $\mathcal{M}$ be an ergodic, reversible Markov chain with arbitrary finite state space $\mathcal{S}$, transition probability matrix P, and stationary distribution π. Let $P^t(x, y)$ be the t-step transition probability from x to y, and let $||\cdot, \cdot||$ denote total variation distance.

▶ **Definition 2.1.** For $\varepsilon > 0$, the *mixing time* is defined as

$$\tau(\epsilon) = \min\{t : \max_{x \in \mathcal{S}} \sum_{y \in \mathcal{S}} ||P^{t'}(x, y), \pi(y)|| \leq \epsilon, \text{ for all } t' \geq t\}.$$

A Markov chain is *rapidly* (or *polynomially) mixing* if the mixing time is bounded above by a polynomial in $\log \mathcal{S}$, the length of a description of a state in $\mathcal{S}$. A chain is *slowly mixing* if the mixing time is bounded below by an exponential function. The *conductance*, introduced by Jerrum and Sinclair [13], is useful to bound the mixing time [13].
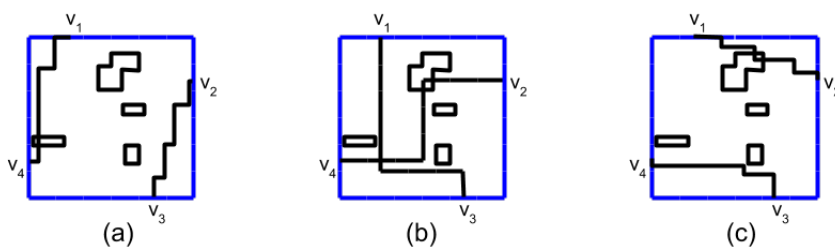
Figure 1 States with (a) positive orientation, (b) orientation 0, (c) negative orientation.

▶ **Definition 2.2.** For a Markov chain with stationary distribution π, the *conductance* Φ is

$$\Phi = \min_{S:0<\pi(S)\leq 1/2} \frac{\sum_{x\in S, y\notin S} \pi(x)P(x,y)}{\pi(S)}.$$

▶ **Theorem 2.3** (Jerrum and Sinclair [13])**.** *The mixing time of a Markov chain with conductance Φ satisfies:*

$$\tau(\epsilon) \geq \left(\frac{1-2\Phi}{2\Phi}\right)\ln\epsilon^{-1}.$$

To establish slow mixing, our strategy will be to define a set S along with sets $T \subset S^C$ and $\mathcal{C} \subset S^C \setminus T$ in the state space, such that $\pi(S) = \pi(T)$ and $\pi(\mathcal{C})/\pi(S) < e^{-cn}$ and such that getting from S to S^C in the Markov chain requires going through $\mathcal{C}$.

In this paper, we will focus on the simplest local Markov chain $\mathcal{M}$ for the Ising and spin glass models, known as *Glauber dynamics*, which connects pairs of configurations whose spins differ on at most one vertex. In a given step, the chain picks any vertex $v \in \Lambda$ at random and changes the spin with the appropriate transition probabilities so that the chain converges to the Gibbs distribution π. For our models, the transition probabilities of $\mathcal{M}$ are defined as

$$P(\sigma,\tau) = \frac{1}{2n^2}\min\left(1, \frac{\pi(\tau)}{\pi(\sigma)}\right),$$

if $|\{i : \sigma_i \neq \tau_i\}| = 1$, and with all remaining probability stay at the current configuration.

2.2 The Contour representation of the Ising and spin glass models

It will be convenient to view Ising and spin glass configurations in terms of *contours*. For every configuration $\sigma \in \Omega$, there is a contour representation $\Gamma(\sigma)$ in $\overline{\Lambda}$, the planar dual to Λ. We define $\overline{\Lambda} = (\overline{V}, \overline{E})$ by letting $\overline{V}$ correspond to the centers of unit squares in Λ and edges $\overline{E}$ connect any two vertices whose corresponding squares share an edge in Λ. An edge $e' \in \overline{E}$ that is dual to $e = (i,j) \in E$ is in $\Gamma(\sigma)$ if $J_{ij}\sigma(i)\sigma(j) = -1$ and we omit it if $J_{ij}\sigma(i)\sigma(j) = +1$. For the Ising model where all the $J_{ij} = +1$, the contour representation $\Gamma(\sigma)$ is precisely the set of edges separating $+$ and $-$ components in σ. Note that we can reconstruct the spin configuration σ from the contour representation (given a single spin) if we know the values of $\{J_{ij}\}$. The weight of a configuration σ is determined by $\Gamma(\sigma)$, and there is a weight-preserving bijection between the configurations of any two spin glasses with the same set of frustated vertices.

For the spin glass model considered here, all vertices of $\overline{V} \setminus \{v_1, ..., v_4\}$ have even degree in $\Gamma(\sigma)$ and the frustrated vertices $\{v_1, ..., v_4\}$ will have odd degree. It follows that $\Gamma(\sigma)$ must be the union of two paths terminating at the frustated vertices, along with even cycles.

(Note that these paths and cycles can intersect each other, and therefore are not necessarily unique.) In all that follows, it will be convenient to shift the primal lattice Λ by $(-1/2, -1/2)$ so that the vertices of $\overline{\Lambda}$ are integral. Now, recall that we assume that the four frustrated vertices lie on the boundary of a $2n \times 2n$ square S within $\overline{\Lambda}$ centered at (n, n), and they are the corners of a (not necessarily axis-aligned) square. Without loss of generality, we label these so that v_1 lies on the top side of S and is the ith vertex from the upper left corner for some $0 \le i \le n$. Setting $p = i/2n$, v_1 is at a distance of $2pn$ from the upper left corner, v_2 is on the right side of S a distance of $2pn$ from the upper right corner, v_3 is on the bottom of S a distance of $2pn$ from the lower right corner, and v_4 is on the left side of S a distance of $2pn$ from the lower left corner. The key to all of our arguments is how the two long paths in $\Gamma(\sigma)$ pair up these frustrated vertices. Let $\alpha(\sigma)$ be the length of the shortest path in $\overline{\Lambda}$ from the connected component of $\Gamma(\sigma)$ containing v_1 to the connected component containing v_4 (if v_1 a nd v_4 are connected, $\alpha(\sigma) = 0$). Likewise, let $\beta(\sigma)$ be the length of the shortest path between the component containing v_1 and the component containing v_2. Let $\gamma(\sigma) = \beta(\sigma) - \alpha(\sigma)$ be the *orientation* of the configuration σ. We partition the state space Ω into a disjoint union $\Omega = \cup_{i \in \mathbb{Z}} \Omega_i$, where $\sigma \in \Omega_i$ if $\gamma(\sigma) = i$.

The partition of Ω into $\cup_i \Omega_i$ allows us to define a cut in the state space in order to bound the conductance. In particular, we let $\Omega^- = \cup_{i<0} \Omega_i$ and $\Omega^+ = \cup_{i>0} \Omega_i$, and we observe that $\Omega = \Omega^- \cup \Omega_0 \cup \Omega_+$. We specify a subset of $\mathcal{C} \subset \Omega_0$ that will be critical to defining the cut as $\mathcal{C} = \{\sigma \in \Omega_0 : \alpha(\sigma) = \beta(\sigma) = 0\}$ (i.e., the configurations in which v_1 is connected to both v_2 and v_4). See Figure 1. Finally, we define $\mathcal{C}^* = \mathcal{C} \cup \Omega_{-1} \cup \Omega_1$ to be the configurations where the paths connecting the frustrated vertices are within distance 1 of each other. Following [25], for configurations in $\mathcal{C}$, we partition the cross into two paths, one from v_1 to v_3 and a one from v_2 to v_4; we do the same for configurations in Ω_{-1} and Ω_1, although it may be necessary to add a single "defect" that encodes where one or both of these paths incurs a jump by one unit. To move from a configuration in Ω^- to one in Ω^+ using Glauber dynamics, we must pass through a configuration in $\mathcal{C}^*$. We will show that the probability of $\mathcal{C}$ is exponentially small, and this will allow us to argue that the Glauber dynamics requires exponential time to converge to equilibrium.

3   Slow Mixing for the Ising model with Mixed Boundaries

We start with the standard approach used to show slow mixing when the boundary conditions alternate spins on the boundary of a $(2n + 1) \times (2n + 1)$ lattice region Λ. Here $\overline{\Lambda}$ is the $2n \times 2n$ lattice region centered in Λ. This will motivate the approach used in the general spin glass setting (when the frustrated vertices are not necessarily on the boundary of Λ) and will elucidate the difficulties in generalizing this simpler result.

Fix $0 \le p \le 1/2$ and let $q = 1 - p$. We define $v_1 = (2pn, 2n), v_2 = (2n, 2qn), v_3 = (2qn, 0)$ and $v_4 = (0, 2pn)$. Recall that all vertices on the boundary between v_1 and v_2 and between v_3 and v_4 are assigned $+$ and the others are assigned $-$. The vertices $v_1, ..., v_4$ define the endpoints of a pair of paths in each configuration. (There may be more than one choice of paths.) Using the strategy outlined in Section 2.2, we recall that $\mathcal{C}$ consists of those configurations where there are paths from v_1 to both v_2 and v_4 (and therefore also to v_3). Using the notion of "fault lines" introduced in [25], we note that this is the set of configurations that contain a *horizontal fault line*, i.e.,. a path from v_2 to v_4, and a *vertical fault line*, i.e., a path from v_1 to v_3. When both fault lines are present (and intersect) we call their union a *cross*. We define the cross so that it is a maximal component of the contour representation of the configuration.

Let C be a cross in $\overline{\Lambda}$. As we will show in Lemma 4.1, the minimum length of C is $L = 6n - 4np$. We write the length as $|C| = L + \ell$, for some $\ell \geq 0$. Let $\mathcal{C}_C$ be the set of configurations in $\mathcal{C}$ that have C as their cross.

We will write the weight of a configuration σ as $\lambda^{-H(\sigma)}$, $\lambda = e^\beta = e^{1/T}$, and note that the energy $H(\sigma)$ is the number of edges in the contour representation of σ.

▶ **Lemma 3.1.** *For any cross C, we have*

$$\pi(\mathcal{C}_C) \leq \lambda^{-(2n-4pn+\ell)}.$$

Proof. We define the injective map $\psi_C : \mathcal{C}_C \to \Omega$ so that $\pi(\psi_C(\sigma)) = \pi(\sigma)\lambda^{(L-4n+\ell)}$ for any fixed C. Given this map, we find

$$1 \;=\; \pi(\Omega) \geq \sum_{\sigma \in \mathcal{C}_C} \pi(\psi_C(\sigma)) \;=\; \sum_{\sigma \in \mathcal{C}_C} \pi(\sigma)\lambda^{(L-4n+\ell)} \;=\; \lambda^{(2n-4pn+\ell)}\pi(\mathcal{C}_C).$$

The map ψ_C is defined by removing C; then, along the upper-left boundary of Λ between v_1 and v_4 we add each edge not in σ and remove each edge in σ; then, along the lower-right boundary of Λ between v_3 and v_2 we add each edge not in σ and remove each edge in σ. ◄

▶ **Theorem 3.2.** *Let $\Lambda \subset \mathbb{Z}^2$ be an $(2n + 1) \times (2n + 1)$ lattice region and $0 \leq p \leq 1/2$ define a family of balanced mixed boundary conditions on Λ. Let Ω be the set of all Ising configurations and let $\mathcal{C}$ be the Ising configurations containing a cross. Then*

$$\pi(\mathcal{C}) \leq f(n)e^{-cn},$$

for some polynomial $f(n)$ and constant $c > 0$, whenever $\lambda^{(1-2p)} > 3^{(3-2p)}$.

Proof. By Lemma 3.1,

$$\pi(\mathcal{C}) \;\leq\; \sum_C \lambda^{-(2n-4pn+\ell)} \;\leq\; \sum_{\ell \geq 0} \lambda^{-(2n-4np+\ell)}3^{(6n-4np+\ell)} \;\leq\; 4n^2(3^{(3-2p)}\lambda^{-(1-2p)})^{2n},$$

which is exponentially small when $\lambda^{(1-2p)} > 3^{(3-2p)}$. The second inequality holds because there are at most $3^{(6n-4np+\ell)}$ ways to choose a cross of length $6n - 4np + \ell$. ◄

Thus, when $\lambda^{(1-2p)} > 3^{(3-2p)}$ we have that the size of the cut is exponentially small, and therefore the conductance of the graph is also exponentially small. By Theorem 2.3, this implies that the chain takes exponential time to mix.

▶ **Corollary 3.3.** *Glauber dynamics for the Ising model on Λ with balanced mixed boundary conditions takes time at least e^{cn} to mix, for some constant $c > 0$, when $\lambda^{(1-2p)} > 3^{(3-2p)}$.*

Notice that this gives $\lambda > 27$ when $p = 0$ and $\lambda > 3^{(2^{(k+1)}+1)}$ when $p = 1/2 - 1/2^k$ and when $p = 1/2$ this fails to give any useful bound.

4 Slow Mixing for Frustrated Spin Glasses Using Free Energy

We will now proceed to extend the result in Section 3 by establishing slow mixing below some temperature for spin glasses with four well-separated frustrated vertices.

In this setting we define Λ as the $kn \times kn$ lattice region, $k \geq 2$, centered at (n, n). Four distinguished faces are symmetric around the center of the lattice region under 90 degree rotations. The centers of these faces are four vertices $v_1, .., v_4$ in $\overline{\Lambda}$. As in Section 2.2 we define $\mathcal{C}$ to be the set of contour configurations in which v_1 is connected to both v_2 and
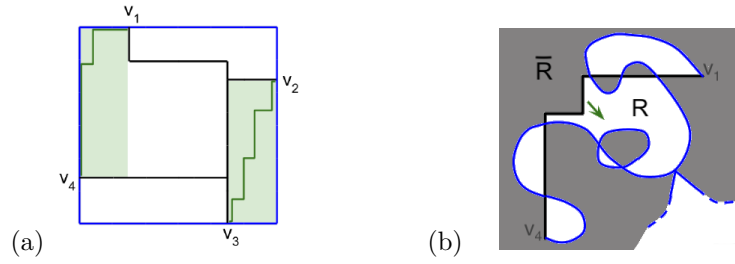
Figure 2 (a) A minimal cross is shown in black, with two possible monotone paths in green. Any monotone path in either shaded region is possible. (b) A staircase is shown in black, together with the part of a cross containing a path from v_1 to v_4. The green arrow shows the direction edges of σ are shifted in the region bounded by the middle section of the staircase and the cross.

v_4, and we define the cross C in such a configuration as the component containing v_1. The argument in Section 3 fails when $p = 1/2$, in particular when $\ell = o(n)$. The length of the cross C in that case is $4n + \ell$, and our injective map ψ_C removes C and replaces it with two paths of total length $4n$. The difference in energy, $H(\sigma) - H(\psi_C(\sigma)) = \ell$, is too small to show that σ has exponentially small probability.

The remedy comes from noticing that in exactly the case $\ell = o(n)$, C is nearly a minimal cross and there are many alternative choices of ψ_C. We will allow any monotone path that, in order to ensure loss of energy, does not intersect C. The set of possible paths is illustrated in Figure 2(a). We have the following lemma, whose proof appears in the Appendix.

▶ **Lemma 4.1.** *Let S_n be the $2n \times 2n$ axis-aligned square whose sides contain $v_1, .., v_4$. For some $\ell \geq 0$, $|C| = 6n - 4pn + \ell$. If $\ell < 2pn$ there are two $(2n - 2pn - \ell) \times (2pn - \ell)$ rectangular regions on opposite corners of the interior of S_n that contain no edges of C.*

Our new strategy is to use *all* possible choices of ψ_C, thereby defining an exponential family of images. We will define a function Ψ_C that involves mapping a configuration $\sigma \in \mathcal{C}_C$ to the union of possible $\psi_C(\sigma)$ defined by different pairs of monotone paths. Figure 2(a) also shows the tradeoff between energy and entropy for our method. As p decreases, the energy loss due to the map increases. As the width of each shaded area decreases, the number of possible paths, $\binom{2n}{2np}$, also decreases. This is what we mean by a decrease in entropy.

Just as we needed ψ_C to be injective in Section 3, we would like our new map to have the property that two different configurations map to disjoint sets of configurations. Instead, we define Ψ_C to pass a small amount of "side information," and with this definition we will get a disjointness property that serves our purpose. The side information is in the form of tokens placed on certain edges along each of the two paths that define the configuration σ is mapped to. Formally, for each path this information is encoded as a binary string of length $2n$: 0 for any plain edge, 1 for an edge with a token. The nice property that will make this side information small is that no two adjacent edges of a path are occupied by tokens.

Let $B(m)$ be the set of binary strings of length m with no consecutive 1's. Let $B = B_C = B(2n - \ell)$. Formally, we will define a function $\Psi_C : \mathcal{C}_C \to 2^{\Omega \times B \times B}$ that has the nice properties in the following lemma. To get our hands on the set of mapped configurations minus the tokens, we define the projection operator $\Pi : 2^{\Omega \times B \times B} \to 2^{\Omega}$, so that $\Pi(\{\sigma_i, b_i, b_i'\}) = \{\sigma_i\}$. Formally, $\Pi \circ \Psi_C$ is the map from one configuration to a set of configurations.

In the following lemmas, fix $0 \leq p \leq 1/2$ and let $L = 6n - 4np$.

▶ **Lemma 4.2.** *Let C be a maximal cross of length $|C| = L + \ell$. There exists a function $\Psi_C : \mathcal{C}_C \to 2^{\Omega \times B \times B}$ such that $\forall \sigma, \sigma' \in \mathcal{C}_C, \sigma'' \in \Pi \circ \Psi_C(\sigma)$,*

$$\Psi_C(\sigma) \cap \Psi_C(\sigma') = \emptyset,$$

$$|\Psi_C(\sigma)| = \binom{2n - 2\ell}{2pn - \ell}^2,$$

and $H(\sigma'') \leq H(\sigma) - (2n - 4np + \ell).$

We postpone constructing the function Ψ_C (and proving Lemma 4.2) until the next subsection. Theorem 4.5 is an analogue of Theorem 3.2 that gives an exponential bound for all p, $0 \leq p \leq 1/2$. As a corollary of Theorem 4.5, we will prove our main result, Theorem 1.1, asserting slow mixing for spin glasses with frustration.

We first bound the probability of the set of configurations containing a given cross C.

▶ **Lemma 4.3.** *For any maximal cross C of length $|C| = L + \ell$ we have*

$$\pi(\mathcal{C}_C) \leq \pi(\Pi \circ \Psi_C(\mathcal{C}_C)) \lambda^{-(2n-4np+\ell)} \phi^{4n-2\ell+1} \bigg/ \binom{2n - 2\ell}{2np - \ell}^2, \tag{1}$$

where $\phi = (1 + \sqrt{5})/2$.

Proof. It is well known that $|B(m)|$ is the mth Fibonacci number, which is within 1 of ϕ^m. Each $\sigma'' \in \Pi \circ \Psi_C(\sigma)$ appears in at most $|B|^2 \leq \phi^{4n-2\ell+1}$ elements of $\Psi_C(\sigma)$. The bound on $H(\sigma'')$ in Lemma 4.2, gives $\pi(\sigma'') \geq \pi(\sigma)\lambda^{-(2n-4np+\ell)}$ and the two equalities imply

$$\pi(\Pi \circ \Psi_C(\mathcal{C}_C)) \geq \sum_{\sigma \in \mathcal{C}_C} \pi(\sigma)\lambda^{(2n-4np+\ell)} \phi^{-(4n-2\ell+1)} \binom{2n - 2\ell}{2np - \ell}^2. \tag{2}$$

The inequality follows by replacing $\sum \pi(\sigma)$ with $\pi(\mathcal{C}_C)$.   ◀

Our main theorems establishing slow mixing of Glauber dynamics for spin glasses with well-separated frustrated vertices (Theorems 4.5 and 1.1) depend on the following technical lemma regarding the set $\mathcal{C}_\ell$ of configurations containing maximal crosses of fixed length $L + \ell$: $\mathcal{C}_\ell = \cup\{\mathcal{C}_C : |C| = L + \ell\}$. The idea of the lemma is to show that $\pi(\mathcal{C}_\ell)$ is exponentially small, where the constant in the exponent is independent of ℓ. This also means that the free energy $\ln \pi(\mathcal{C}_\ell)/n$ is less than some negative constant. Since there are polynomially many values of ℓ, it will follow that the whole set $\mathcal{C}$ is exponentially small.

▶ **Lemma 4.4.** *Let $\mathcal{C}_\ell$ be the spin glass configurations where $v_1, .., v_4$ are all connected by a maximal cross of length $L + \ell$. Then for $\lambda \geq 256$ we have*

$$\pi(\mathcal{C}_\ell) \leq 2^{-0.2n} \text{ poly}(n). \tag{3}$$

Proof. Let $s = 1/2 - p$ and $r = \ell/n$. We will actually prove that

$$\pi(\mathcal{C}_\ell) \leq \lambda^{-8sn} (3/\lambda)^{rn} 2^{n[(4-2r)\log_2 \phi + \mathcal{L}(r,s) + \mathcal{P}(r,s) - \mathcal{T}(r,s)]} \text{ poly}(n) , \tag{4}$$

where

$$\mathcal{L}(r,s) = (2 + 4s + r)h(\frac{r}{2 + 4s + r}),$$

$$\mathcal{P}(r,s) = (2 + 4s)h(\frac{2s}{1 + 2s}),$$

$$\mathcal{T}(r,s) = \max(0, 4 - 4r)h(\frac{1}{2} - \frac{s}{1 - r}),$$

and $h(x) = -x \log_2(x) - (1-x) \log_2(1-x)$. Then we will show that the right-hand side of Equation 4 is less than $2^{-0.2n}$.

First, we establish Equation 4. Each C consists of vertical path connecting v_1 to v_3 and a horizontal path connecting v_2 to v_4. The vertical path contains a minimal vertical path of $2n$ vertical edges and $2n - 4pn$ horizontal edges. There are $\binom{4n-4pn}{2n-4pn} = \binom{2n+4sn}{4sn}$ choices of minimal vertical path. There is one choice of minimal horizontal path, which contains only horizontal edges connecting v_2 and v_4 to the vertical path. Then there are $\binom{6n-4np+\ell}{\ell} = \binom{2n+4sn+rn}{rn}$ ways to choose the locations of the ℓ extra edges, and 3 possible directions for each extra edge. Applying Lemma 4.3 and Stirling's formula,

$$
\begin{aligned}
\pi(\mathcal{C}_\ell) &\leq \binom{6n-4np+\ell}{\ell}\binom{4n-4pn}{2n-4pn}3^\ell \max_{|C|=L+\ell} \pi(\mathcal{C}_C) \\
&\leq 2^{(2n+4sn+rn)h(r/(2+4s+r))}2^{(2n+4sn)h(2s/(1+2s))}3^{rn} \\
&\quad \cdot \lambda^{-(8sn+rn)}\phi^{4n-2rn+1}2^{-2(2n-2rn)h((1-r-2s)/(2-2r))}.
\end{aligned}
$$

Equation 4 follows immediately by collecting the terms in the exponents.

By taking logs and dividing by n it follows that $\log_2 \pi(\mathcal{C}_\ell)/n \leq \mathcal{F}(r,s)$, where

$$
\mathcal{F}(r,s) = (-r - 8s)\log_2 \lambda + r \log_2 3 + (4-2r)\log_2 \phi + \mathcal{L}(r,s) + \mathcal{P}(r,s) - \mathcal{T}(r,s)
$$

It remains to show that $\mathcal{F}(r,s) \leq -0.2$, for all s, r, $0 \leq s \leq 1/2$, $r > 0$, and large enough λ.

$\mathcal{L}(r,0)$ is concave as a function of r, $\mathcal{L}(r,s)$ and $\mathcal{P}(r,s)$ are concave as functions of s, and $-\mathcal{T}(r,s)$ is convex as a function of s. We numerically approximate the concave functions with a tangent line and the convex function with a secant, yielding these results:

$$
\begin{aligned}
\mathcal{L}(r,0) &\leq 0.5 + 2.9r; & \mathcal{P}(r,s) &\leq 0.5 + 12s; \\
\mathcal{L}(r,s) &\leq 0.5 + 2.9r + 2rs \;\leq\; 0.5 + 3.9r; & -\mathcal{T}(r,s) &\leq -4 + 4r + 8s.
\end{aligned}
$$

Also, $r\log_2 3 < 1.5r$ and $(4-2r)\log_2 \phi < (2.8 - 1.4)r$. Adding terms, for $\lambda \geq 256$, we get

$$
\mathcal{F}(r,s) \;\leq\; (-r - 8s)\log_2 \lambda + 8r + 20s - 0.2 \;\leq\; -0.2. \qquad \blacktriangleleft
$$

We now state the key theorem bounding the probability of the set $\mathcal{C}$ of configurations containing crosses.

▶ **Theorem 4.5.** *Let Ω be the set of all spin glass configurations in a $kn \times kn$ square lattice Λ centered at (n,n), $k \geq 2$. Suppose that four distinguished vertices $v_1,..,v_4$ lie on the boundary of an axis-aligned $2n \times 2n$ square S centered in $\overline{\Lambda}$, and these four vertices form the corners of a (not necessarily axis-aligned) square (i.e., they are shifted by $2p$ around the boundary of S). Let $\mathcal{C}$ be the set of configurations in which v_1 is connected to both v_2 and v_4. Then for $\lambda \geq 256$ we have*

$$
\pi(\mathcal{C}) \leq 2^{-0.2n}\mathrm{poly}(n). \tag{5}
$$

Proof. Since ℓ has at most $(cn)^2$ values, $\pi(\mathcal{C}) \leq (cn)^2 \max_\ell \pi(\mathcal{C}_\ell) \leq 2^{-0.2n}\mathrm{poly}(n)$. ◀

Proof of Theorem 1.1. Set $T = 2/\ln 256 = 0.360....$ Let $t < T$. The state space Ω contains the two disjoint subsets Ω_- and Ω_+, separated by a cut set $\mathcal{C}^*$ consisting of all configurations within two steps of $\mathcal{C}$. We have $\pi(\mathcal{C}^*) < \pi(\mathcal{C})\mathrm{poly}(n)$ and by symmetry $\pi(\Omega_-) = \pi(\Omega_+)$. The conductance Φ satisfies

$$
\Phi \;\leq\; \frac{\sum_{\sigma \in \Omega_-, \sigma' \in \Omega_c} \pi(\sigma)\Pr(\sigma,\sigma')}{\pi(\Omega_-)} \;\leq\; 4 \cdot \pi(\mathcal{C}^*) \;\leq\; 2^{-0.1n}, \quad \text{for large enough } n. \tag{6}
$$

Therefore the Markov chain mixes slowly. ◀
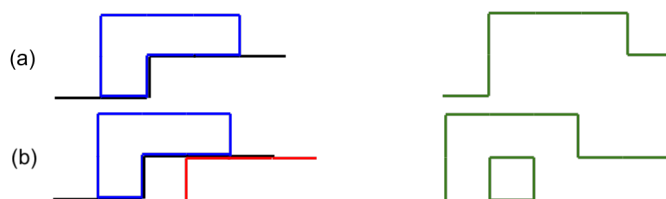
Figure 3 (a) A staircase and patch that share edges (left), and an encoding that loses information (right). (b) A staircase and patch with the default path (left), and an encoding that preserves information (right).

4.1 Construction of the Map

In this section we will construct the map Ψ_C using pairs of paths as shown in Figure 2(a). An *upper staircase* with respect to a cross C of length $L + \ell$ is a path of $\min(\ell, 2pn)$ west edges starting at v_1 followed by zero or more west and south edges, followed by $\min(\ell, 2pn)$ south edges ending at v_4. We refer to the section of west and south edges as the "middle $2n - 2\min(\ell, 2pn)$ edges." We define a *lower staircase* to be a path v_3 to v_2, which, when the configuration is rotated 180°, becomes an upper staircase. Note that the edges on a staircase need not be edges of a particular configuration. Given upper and lower staircases, we will map $\sigma \in \mathcal{C}_C$ to some $\sigma' \in \Omega$, marking certain edges with tokens. We will show that one can reconstruct σ from C, σ', and the marked edges, that no two marked edges are adjacent, and $H(\sigma') \leq H(\sigma) - |C| + 4n$, implying Lemma 4.2.

Our map is motivated by the map ψ_C in the proof of Lemma 3.1. In fact, the construction is the same along the first $\min(\ell, 2pn)$ edges and last $\min(\ell, 2pn)$ edges: we add each edge not in σ and remove each edge in σ. Along the middle section of the staircase that contains west and south edges, our map must encode the locations of the staircase edges in σ' without increasing $H(\sigma')$. The basic strategy is to remove C, shift edges in σ away from the staircase, toward the removed edges of C, then add the edges of the staircase.

Let S_U be an upper staircase and S_L be a lower staircase. The simple regions in the interior of $C \cup S_U \cup S_L$ may be two-colored gray and white, with the exterior, denoted $\overline{R}$, colored gray. Regions separated by an edge in $C \cap S_U$ or $C \cap S_L$ will have the same color. We assume in what follows that $\ell < 2pn$. In particular, S_U and S_L do not both contain edges in any one region boundary. When $\ell \geq 2pn$, S_U and S_L are contained in the boundary of the $2n \times 2n$ square S_n, and the proof of Lemma 3.1 applies.

By Lemma 4.1 there is one white simple region R whose boundary contains the middle $2n - 2\ell$ edges of S_U. The map will shift edges of σ in R southeast, and it will shift the corresponding region bounded by the middle $2n - 2\ell$ edges of S_L northwest. See Figure 2(b).

We may assign a + or − to each site in $R \cup \overline{R}$ so that the sites adjacent to C are + and the edges of σ restricted to $R \cup \overline{R}$ are exactly those edges between two neighboring sites of opposite sign. We define a *patch* to be a connected set of − sites in $R \cup \overline{R}$. The outer boundary of a patch is the unique cycle of edges in the configuration that, when traversed counterclockwise, has sites inside the left of each edge and sites outside to the right.

A naive map would remove C from the configuration and add the upper staircase and lower staircase to the configuration. The flaw in this approach is that σ cannot always be reconstructed when part of a staircase coincides with part of the boundary of a patch. Figure 3(a) shows an upper staircase in black that shares edges with a patch, shown in blue. Adding the staircase creates double edges. The natural recourse is removing double edges while preserving degrees, but shared edges are no longer recoverable from such a map.
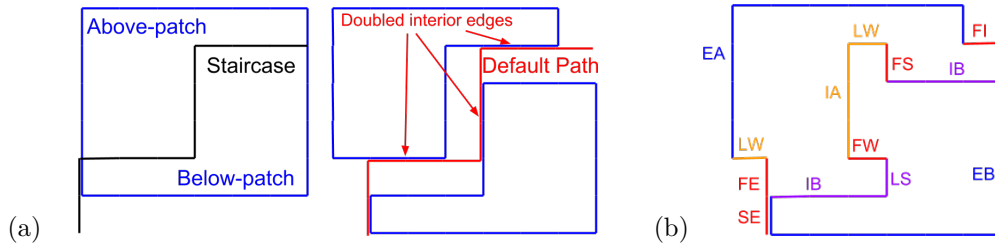
■ **Figure 4** (a) The components to encode. (b) The contour pieces defining the map.

Our map modifies the naive approach by shifting the staircases before adding them to the configuration, and shifting edges that are between the staircases toward the empty space left behind after the removal of C. Let S be the maximal contiguous section of S_U that forms part of the boundary of R and contains the middle $2n - 2\ell$ edges of S_U. We define the *default path* to be S shifted one step east. It consists of alternating west and south sections. The *first south* (northernmost) edge of each south section, and the *first west* edge of each west section, are each incident to S at just one vertex (with the exception of the first edge of S if it is a south edge preceded in S_U by a south edge). All other edges on the default path are on S or not incident to it. The *last south* and *last west* edges are defined accordingly.

Figure 3(b) shows the same staircase and patch, with the default path in red. σ is mapped to σ' by starting with the union of the patch and the default path, and removing double edges. The default path can be reconstructed from σ', because it contains the first-south and first-west edges of the default path. This is the information that was missing from the previous mapping. The mapping contains no more energy than the original.

A subtler problem of lost information arises when the staircase enters the interior of a patch. We define an *interior edge* of S to be one that bounds two $-$ sites. Each maximal contiguous segment of interior edges of S divides a patch into two patches, which we refer to as the *above-patch* (or *A-patch*) and the *below-patch* (or *B-patch*).

To solve the problem of interior segments, we triple each interior edge of the staircase, shifting the staircase and the B-patch one step east, and shifting the B-patch one step south. The drawing on the left of Figure 4(a) shows the staircase in black and the patch in blue before the two shifting steps, and the drawing on the right shows the default path in red and the two patches after the shifts. After the shifts, our mapping removes all double edges.

The doubled interior edges of the default path consist of all interior west edges of the A-patch except the *last-west edge of each west section*, and all interior south edges of the below patch except the *last-south edge of each south section*.

This mapping has the one final problem that it increases the energy of the configuration. This problem can be illustrated by labeling the edges as in Figure 4(b). *EA* and *EB* (blue) are exterior edges of the A-patch and B-patch, respectively. *IA* and *LW* (orange) are south and last-west interior edges of the A-patch, resp. *IB* and *LS* (purple) are west and last-south interior edges of the B-patch, resp. *FI*, *FW*, and *FS* (red) are the first interior edge and all first-west and first-south edges of the default path, resp. *FE* and *SE* (red) are the first and second "exterior" edges of the default path following this segment of interior edges. The first exterior edge will not be interior to any patch, but the second exterior edge may be interior to this or another patch.

The increase in energy is caused by the "detours" at *FS-LW* and *FW-LS*. The final mapping step is to flip the signs of sites bounded by corners of those two types and to place a token at each such site. The Appendix presents the map steps in detail.

4.2 Reconstruction

The default path (and hence σ restricted to R_L) can be reconstructed from σ', before token-placing, as it contains all of the first-west and first-south edges. Starting from the FI edge, the default path continues until it encounters the first FS or FW edge. Then it changes direction and the FS or FW edge inductively plays the role of the FI edge. The rest of the interior segment is reconstructed by induction on the number of south and west sections.

Reconstructing the default path in the presence of tokens is the same recursive process, except we look ahead one step. If the next edge has a token, we flip the adjacent site before proceeding. The adjacent site is unambiguous because it is between the A- and B-patches. The Appendix presents the reconstruction steps in detail.

4.3 Energy loss

Before token-placing and sign-flipping, σ' has more energy than $H(\sigma) - |C| + 4n$. The EA and EB naturally correspond 1-1 to the edges of the original patch. The IA and IB edges correspond 1-1 to the interior segment of the staircase. The excess energy consists of one pair of edges, FS-LW or FW-LS, for each corner of the interior segment, plus two more edges, the FI edge and one LW or LS edge incident to FE.

The mapping solves this problem by short-circuiting the corners. Each FS-LW pair occurs as part of a segment FS-LW-IA that form three sides of a site, and each FW-LS pair occurs as part of a segment FW-LS-IB that also form three sides. The mapping flips the sign of each such site, replacing three edges with one, and places a token at the flipped site.

Two such sites may be adjacent. This happens when an IA or IB section is one edge long. Then one of the two sites is bounded by an FS-LW-IA-FW segment or an FW-LS-IB-FS segment. In either case the mapping replaces four edges with zero. One sign-flip in the first traversal removes the excess energy of both sites, and one token is placed. It also flips one edge of the adjacent site, ensuring that no two tokens will be adjacent. (See Figure 5(a) steps (d)-(f).) Each sign-flip in the second traversal converts three edges to one, canceling excess energy due to this site. In this case, this site will not be adjacent to another token site.

The two remaining excess edges are the FI edge and one LW or LS edge. Suppose it is LW (the case of LS is similar). If FE and a LW form a double edge or SE and an EB form a double edge (the case pictured), the mapping removes the double edge, cancelling the excess energy. In the remaining case FE is an FS edge, SE is an FW edge, and the segment LW-FE-SE forms three sides of a site. The mapping flips the sign of that site and places a token. No two tokens are placed on adjacent sites. In the case considered in the previous paragraph, SE is not an interior edge of any patch, because the site is on the exterior side of FE. The first interior edge of a patch does not bound a site with a token.

—— **References** ——

1 M. Aizenman. Translation invariance and instability of phase coexistence in the two-dimensional Ising system. *Comm. Math. Phys.*, **73:** 83-94, 1980.

2 F. Barahona. On the computational complexity of Ising spin glass models. *Journal of Physics A: Math. and Gen.*, **15:** 3241–3253, 1982.

3 D. Chelkak and S. Smirnov. Universality in the 2D Ising model and conformal invariance of fermionic observables *Inventiones Mathematicae*, **189:** 1–66, 2009.

4 M. Ciucu. Dimer packings with gaps and electrostatics, *Proc. Natl. Acad. Sci.*, **105:** 2766–2772, 2008.

5 L. Coquille and Y. Velenik. A finite-volume version of Aizenman-Higuchi Theorem for the 2D Ising model. *Probab. Theory Relat. Fields,* **153:** 25–44, 2012.

6   R.L. Dobrushin, R. Kotecký and S. Shlosman. Wulff Construction: a Global Shape from Local Interaction. *AMS translations series,* **104**, Providence R.I.: AMS, 1992.

7   R. Fernandez, P.A. Ferrari, and N.L. Garcia. Loss network representation of Ising contours. *Annals of Probability* **29:** 902–937, 2001.

8   R. Fitzpatrick. Thermodynamics and Statistical Mechanics. Preprint at https://farside.ph.utexas.edu/teaching/sm1/Thermal.pdf.

9   S. Friedli and Y. Velenik. *Statistical Mechanics of Lattice Systems: a Concrete Mathematical Introduction.* Cambridge University Press, 2017, to appear (http://www.unige.ch/math/folks/velenik/smbook/).

10   H.O. Georgii. *Gibbs measures and phase transitions.* de Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, 1988.

11   S. Greenberg and D. Randall. Slow mixing of Glauber dynamics on perfect matchings of the square-octagon lattice. Preprint, 2006.

12   Y. Higuchi. On the absence of non-translation invariant Gibbs states for the two-dimensional Ising model. *Colloq. Math. Soc.*, J'anos Bolyai, 517–534, 1981.

13   M.R. Jerrum and A.J. Sinclair. Approximate counting, uniform generation and rapidly mixing Markov chains. *Information and Computation* **82**: 93–133, 1989.

14   G. Lawler. Scaling limits and the Schramm-Loewner evolution. *Probab. Surveys,* **8:** 442–495, 2011.

15   E. Lubetzky, F. Martinelli, A. Sly and F. Toninelli. Quasi-polynomial mixing of the 2D stochastic Ising model with "plus" boundary up to criticality. *J. Eur. Math. Soc.*, **15**: 339–386, 2013.

16   F. Martinelli. On the two-dimensional dynamical Ising model in the phase coexistence region. *Journal of Statistical Physics*, **76:** 1179–1246, 1994.

17   F. Martinelli. Lectures on Glauber dynamics for discrete spin models. *Lectures on Probability Theory and Statistics (Saint-Flour, 1997)*, Lecture notes in Mathematics **1717**: 93–191, Springer, Berlin, 1998.

18   F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case. *Comm. Math. Phys.*, **161:** 447–486, 1994.

19   F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case. *Comm. Math. Phys.*, **161:** 487–514, 1994.

20   F. Martinelli, A. Sinclair, D. Weitz, The Ising model on trees: Boundary conditions and mixing time. *Comm. Mathematical Physics* **250**: 301–334, 2004.

21   A. Messager and S. Miracle-Sole. Equilibrium states of the two-dimensional Ising model in the two-phase region. *Comm. Math. Phys.*, **40:**187–196, 1975.

22   E. Mossel and A. Sly. Exact thresholds for Ising – Gibbs samplers on general graphs. *Ann. Probab.*, **41:**294-328, 2013.

23   L. Onsager. Crystal statistics: a two-dimensional model with an order- disorder transition. *Phys. Rev.*, **65:** 117–149, 1944.

24   R. Peierls, On Ising's model of ferromagnetism. *Proc. Camb. Philos. Soc.*, **32**: 477–481, 1936.

25   D. Randall. Slow mixing of Glauber dynamics via topological obstructions. *17th Symposium on Discrete Algorihtms (SODA)*, 2006.

26   D. Randall and D.B. Wilson. Sampling Spin Configurations of an Ising System. *Proc. 10th ACM/SIAM Symposium on Discrete Algorithms*, S959–960, 1999.

27   O. Schramm. Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.*, **118**: 221–288, 2000.

28   S. Smirnov. Critical percolation in the plane. *Comptes Rendus de l'Académie des Sciences.* **333**: 239–244, 2001.

29   S. Smirnov and H. Duminil-Copin. Conformal invariance in lattice models, in Lecture notes of the 2010 Clay summer school, Buzios, 2010.

30   N. Sun. Conformally invariant scaling limits in planar critical percolation. *Probability Surveys*, **11:** 155–209, 2011.

31   L. Thomas. Bound on the mass gap for the finite volume stochastic Ising models at low temperature. *Comm. Mathematical Physics* **126:** 1–11, 1989.

A   Appendix

Proof of Lemma 4.1. The minimal cross contains a path from v_1 to v_3 and a path from v_2 to v_4. First let's assume that each of these is minimal. Then they each have length $4n - 4pn$ and the total length of the cross is $8n - 8pn - |o|$, where o is the length of the overlapping segments. Orient the edges along each path from v_1 to v_3 so that the edges all go right or down, and orient the path from v_2 to v_4 so that they go down or left. Then the overlapping segments are oriented the same way in both paths if the edge is vertical and in opposite directions if the edge is horizontal. But all horiztonal edges on the path from v_2 to v_4 after this shared edge are left of the edge, and those on the path from v_1 to v_3 are to the right; similarly, if they share a horizonal edge, all subsequent vertical edges must be to the left of the edge on one path and to the right on the other. Therefore, the overlapping segment must all be vertical or all horizontal. Furthermore, all the vertical edges that overlap have to lie between v_2 and v_4 and have length at most $2n - 4pn$; likewise if the horizontal edges that overlap since they lie between v_1 and v_3. It follows that when the two paths are minimal $|o| \leq 2n - 4pn$ and the length of the cross is at least $6n - 4pn$.

If either of the paths from v_1 to v_3 and v_2 to v_4 is not minimal, then the overlap can contain both horizontal and vertical edges. Notice that the overlapping segments must be contiguous along either path or the cross would contain a cycle, contradicting minimality. If this overlapping segment contains edges oriented both left and right (or down and up), then it can be shortened, again violating minimality. Therefore the overlapping segment must go down and left or down and right. If down and left, then the path from v_1 to v_3 has an extra edge to the right for each horizontal edge in the overlapping segment; if down and right then the path from v_2 to v_4 has an extra edge for each horiztonal edge in the overlap. Finally, if the number of vertical edges in the overlap exceeds the vertical distance between v_2 and v_4, then the path between them must contain at least that many additional vertical edges. Summing all of these up, we find that if there are $2n - 4pn + k$ edges in the overlap, then the sum of the lengths of the two paths must be at least $8n - 8pn + k$. Subtracting the length of the overlapping segment, we again find that the length of the cross is at least $6n - 4pn$.

If the cross is nearly minimal, with length $6n - 4pn + \ell$, the picture is similar. The paths from v_1 to v_3 and v_2 to v_4 must also be nearly minimal, each having length at most $4n - 4pn + \ell$ and the length of the overlapping segments must be at least $2n - 4pn - \ell$. It follows that the path from v_1 to v_3 lies in a $2n - 4pn + \ell \times 2n$ rectangle, the path from v_2 to v_4 lies in a $2n \times 2n - 4pn + \ell$ rectangle, and the overlapping segments lie in the center $2n - 4pn + \ell \times 2n - 4pn + \ell$ square. The overlapping segments do not have to be contiguous, but the distance between segments is at most ℓ. We find, by a similar argument to before, that all but ℓ edges on the overlap must have the same orientation, horiztonal or vertical. If the overlap is mostly vertical, then the $2pn - \ell \times 2n - 2pn - \ell$ rectangles adjacent to the upper-left and lower-right corners of the region cannot contain any edges from the cross. Similiarly, if the overlapping segments are mostly horizontal, then there cannot be any edges from the cross in the $2n - 2pn - \ell \times 2pn - \ell$ rectangles incident to the upper-right and bottom-left corners of the region.   ◄
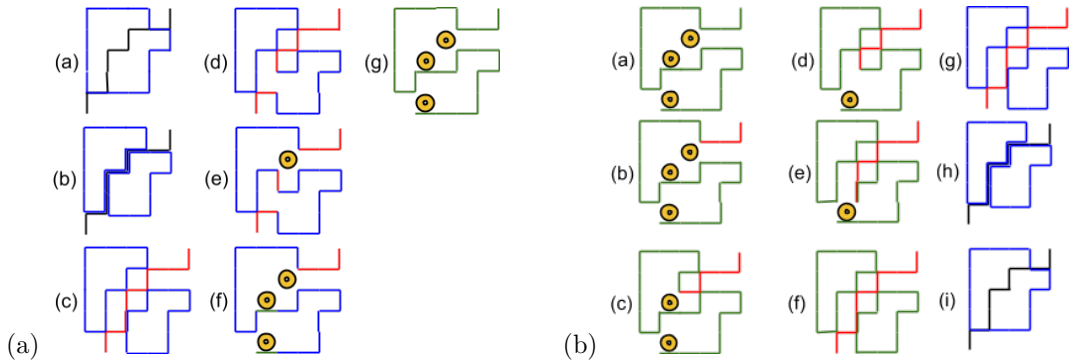
Figure 5 (a) The map: blue edges are the patch boundary, black edges are the staircase, red edges are the default path, and green edges are the final mapping. (b) Reconstruction steps: blue edges are the patch boundary, green edges are the mapping, black edges are the staircase, and red edges are the default path.

A.1 Map Steps

Given $\sigma \in \mathcal{C}_C$ pick an upper staircase S_U and a lower staircase S_L. Remove C from σ. Along the initial segment of ℓ edges and final segment of ℓ edges of S_U, add each edge not in σ and remove each edge in σ. Let S be the middle $2n - 2\ell$ edges of S_U.

1. Add S. If this doubles an edge, label one copy *on the staircase* and the other *above* (*below*) the staircase if it is on the boundary of an A-patch (B-patch).
2. Triple each interior edge of S. Label one copy *on the staircase*, the second *above the staircase*, and the third *below the staircase*. (Figure 5(a) step (b).)
3. Shift every edge on or below the staircase one step east.
4. Shift every edge below the staircase one step south. (Figure 5(a) step (c).)
5. Remove every double edge. (After the two shifts there are no triple edges.) (Figure 5(a) step (d).)
6. Traverse the default path twice from start to end (Figure 5(a) steps (e), (f)):
   a. First traversal: if the current edge and the next edge are interior FW or FS edges, then put a token on the site bounded by these two edges and flip its sign.
   b. Second traversal: if the current edge is either an interior FS or FW edge that is part of an FS-LW-IA or FW-LS-IB segment, or an SE edge that is FW or FS and is the third leg of an LW-FS-FW or LS-FW-FS segment, then flip the site bounded on three sides by the segment and place a token on it.

For S_L, rotate the configuration 180°, repeat steps 1-6, and rotate back.

A.2 Reconstruction steps

Given σ', the following steps reconstruct σ. For subpaths of the upper staircase that bound a white region to the left,

1. Infer and traverse the edges of the default path from start to end, but do not add them to the configuration. The first edge will be a west edge. Inductively, at a current edge, the next edge will be one of two possible edges that we'll call *straight*, for the edge that continues in the current direction, and *turning*, for the other edge.
   a. if there is a token by the next edge, flip the sign of the token site. (Figure 5(b) steps (c), (d), (f).)

    b. if the turning edge exists in the configuration (possibly after flipping), it is the next
    edge. (Figure 5(b) steps (b), (c), (d), (f).)

    c. otherwise the straight edge is the next edge; add it to the configuration if it doesn't
    exist. (Figure 5(b) steps (b), (e).)

2. Shift every edge in the white region to the left one step north.

3. Shift every edge in the white region to the left one step west.

For subpaths of the upper staircase that bound a white region to the right, reflect σ across the
line $y = x$, apply steps 1-3, and reflect back. For the lower staircase, rotate the configuration
180 degrees, repeat the process, and rotate back.

4. Remove all double edges.

5. Add C to the configuration.

On the Number of Variables in Special Classes of Random Lambda-Terms

Bernhard Gittenberger
Department of Discrete Mathematics and Geometry, Technische Universität Wien, Wiedner Hauptstraße 8-10/104, A-1040 Wien, Austria
gittenberger@dmg.tuwien.ac.at

Isabella Larcher
Department of Discrete Mathematics and Geometry, Technische Universität Wien, Wiedner Hauptstraße 8-10/104, A-1040 Wien, Austria
isabella.larcher@tuwien.ac.at

Abstract

We investigate the number of variables in two special subclasses of lambda-terms that are restricted by a bound of the number of abstractions between a variable and its binding lambda, and by a bound of the nesting levels of abstractions, respectively. These restrictions are on the one hand very natural from a practical point of view, and on the other hand they simplify the counting problem compared to that of unrestricted lambda-terms in such a way that the common methods of analytic combinatorics are applicable.

We will show that the total number of variables is asymptotically normally distributed for both subclasses of lambda-terms with mean and variance asymptotically equal to $C_1 n$ and $C_2 n$, respectively, where the constants C_1 and C_2 depend on the bound that has been imposed. So far we just derived closed formulas for the constants in case of the class of lambda-terms with a bounded number of abstractions between each variable and its binding lambda. However, for the other class of lambda-terms that we consider, namely lambda-terms with a bounded number of nesting levels of abstractions, we investigate the number of variables in the different abstraction levels and thereby exhibit very interesting results concerning the distribution of the variables within those lambda-terms.

1 Introduction

The lambda calculus was invented by Church and Kleene in the 30ies as a tool for the investigation of decision problems. Today it still plays an important role in computability theory and for automatic proof systems. Furthermore, it represents the basis for some programming languages, such as LISP. For a thorough introduction to the lambda calculus we refer to [1]. This paper does not require any preliminary knowledge of lambda calculus in order to follow the proofs. Instead we will study the basic objects of lambda calculus, namely lambda-terms, by considering them as combinatorial objects, or more precisely as a special class of directed acyclic graphs (DAGs).

▶ **Definition 1** (lambda-terms, [9, Definition 3]). Let $\mathcal{V}$ be a countable set of variables. The set Λ of lambda-terms is defined by the following grammar:
1. every variable in $\mathcal{V}$ is a lambda-term,
2. if T and S are lambda-terms then TS is a lambda-term, (application)
3. if T is a lambda-term and x is a variable then $\lambda x.T$ is a lambda-term. (abstraction)

The name application arises, since lambda-terms of the form TS can be regarded as functions $T(S)$, where the function T is applied to S, which in turn can be a function itself. An abstraction can be considered as a quantifier that binds the respective variable in the sub-lambda-term within its scope. Both application and repeated abstraction are not commutative, *i.e.*, in general the lambda-terms TS and ST, as well as $\lambda x.\lambda y.M$ and $\lambda y.\lambda x.M$, are different (with the exceptions of $T = S$ and none of the variables x or y occurring in M, respectively). Each λ binds exactly one variable (which may occur several times in the terms), and since we will just focus on a special subclass of closed lambda-terms, each variable is bound by exactly one λ.

We will consider lambda-terms modulo α-equivalence, which means that we identify two lambda-terms if they only differ by the names of their bound variables. For example $\lambda x.(\lambda y.(xy)) \equiv \lambda y.(\lambda z.(yz))$. There is a combinatorial interpretation of lambda-terms that considers them as DAGs and thereby naturally identifies two α-equivalent terms to be equal. Combinatorially, lambda-terms can be seen as rooted unary-binary trees containing special additional directed edges. Note that in general the resulting structures are not trees in the sense of graph theory, but due to their close relation to trees (see Definition 2) some authors call them lambda-trees or enriched trees. We will call them lambda-DAGs in order to emphasise that these structures are in fact DAGs, if we consider the undirected edges of the underlying tree to be directed away from its root.

▶ **Definition 2** (lambda-DAG, [9, Definition 5]). With every lambda-term T, the corresponding lambda-DAG $G(T)$ can be constructed in the following way:
1. If x is a variable then $G(x)$ is a single node labeled with x. Note that x is unbound.
2. $G(PQ)$ is a lambda-DAG with a binary node as root, having the two lambda-DAGs $G(P)$ (to the left) and $G(Q)$ (to the right) as subgraphs.
3. The DAG $G(\lambda x.P)$ is obtained from $G(P)$ in four steps:
   a. Add a unary node as new root.
   b. Connect the new root by an undirected edge with the root of G(P).
   c. Connect all leaves of $G(P)$ labelled with x by directed edges with the new root, where the root is start vertex of these edges.
   d. Remove all labels x from $G(P)$. Note that now x is bound.

Obviously, applications correspond to binary nodes and abstractions correspond to unary nodes of the underlying Motzkin-tree that is obtained by removing all directed edges. Of course in the lambda-DAG some of the vertices that were former unary nodes might have gained out-going edges, so they are no unary nodes in the lambda-DAG anymore. However, when we speak of unary nodes in the following, we mean the unary nodes of the underlying unary-binary tree, that forms the skeleton of the lambda-DAG.

Since the skeleton of a lambda-DAG is a tree, we sometimes call the variables leaves (*i.e.*, the nodes with out-degree zero), and the path connecting the root with a leaf (consisting of undirected edges) is called a branch. There are different approaches as to how one can define the size of a lambda-term ([4], [11]), but within this paper the size will be defined as the total number of nodes in the corresponding lambda-DAG.
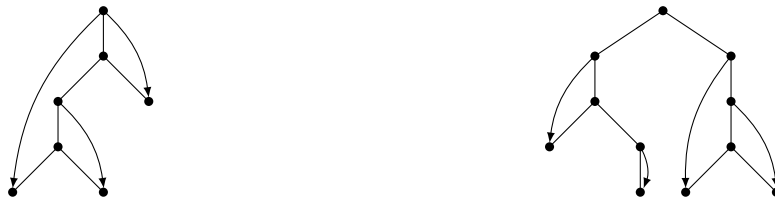
Recently rising interest in the number and structural properties of lambda-terms can be observed, due to the direct relationship between these random structures acting as computer programs and mathematical proofs ([7]). At first sight lambda-terms appear to be very simple structures, in the sense that their construction can easily be described, but so far no one has yet accomplished to derive their asymptotic number. However, the asymptotic equivalent of the logarithm of this number can be determined up to the second-order term (see [5]). The difficulty of counting unrestricted lambda-terms arises due to the fact that their number increases superexponentially with increasing size. Thus, if we translate the counting problem into generating functions, then the resulting generating function has a radius of convergence equal to zero, which makes the common methods of analytic combinatorics inapplicable. This fast growth of the number of lambda-terms can be explained by the numerous possible bindings of leaves by lambdas, *i.e.*, by unary nodes. Consequently, lately some simpler subclasses of lambda-terms, which reduce these multiple binding possibilities, have been studied, e.g. lambda-terms with prescribed number of unary nodes ([4]), or lambda-terms in which every lambda binds a prescribed ([5],[2],[9]) or a bounded ([6],[2],[9]) number of leaves. In this paper we will investigate structural properties of lambda-terms with a bounded number of abstractions between every variable and its binding lambda and lambda-terms with a bounded number of nesting levels of abstractions, which both have been introduced in [3] and [4]. From a practical point of view these restrictions appear to be very natural, since the number of abstractions in lambda-terms which are used for computer programming is in general assumed to be very low compared to their size.

Particular interest lies in the number and distribution of the variables within these special subclasses of lambda-terms. We will show within this paper that the total number of leaves in lambda-DAGs with bounded number of abstractions between the leaves and their respective binding lambdas as well as in lambda-terms with bounded number of nested abstractions is asymptotically normally distributed with mean and variance asymptotically Cn and $\tilde{C}n$, respectively, where the constants C and $\tilde{C}$ depend on the bound that has been imposed. For the latter class of lambda-terms we will also investigate the number of leaves on the different abstraction levels (so called unary levels, *cf.* Definition 11), which shows a very interesting behaviour. We will see that on the lower unary levels, *i.e.*, near the root of the lambda-DAG, there are very few leaves, while the majority of the leaves is located at the upper unary levels and these two domains will turn out to be strictly separated.

For lambda-terms that are locally restricted by a bounded number of abstractions located between the leaves and their binding lambdas the number of unary levels is not bounded and will tend to infinity for increasing size. The expected number of unary levels is unknown, which implies that the correct scaling cannot be determined. Thus, we have not been able to establish results concerning the leaves in the different unary levels for this class of lambda-terms so far. Nevertheless, further studies on this subject seem to be very interesting already for the simpler combinatorial class of Motzkin-trees.

Figure 2 The lambda-DAG of the term $\lambda x.((\lambda y.xy)(\lambda z.(z(\lambda t.tx))z))$, where left the unary length of all bindings, and right the unary height of the leaves is depicted at the respective leaves.

2   Main results

In this section we will introduce the basic definitions and summarize the main results that will be presented in this paper.

First, we will investigate the total number of leaves in lambda-DAGs with bounded unary length of their bindings, *i.e.*, with a bounded number of abstractions between each leaf and its binding lambda.

▶ **Definition 3** (unary length of a binding, [4, Definition 1]). Consider a lambda-term T and its associated lambda-DAG $G(T)$. The unary length of the binding of a leaf e by some abstraction v in T (directed edge from v to e in $G(T)$) is defined as the number of unary nodes on the path connecting v and e in the underlying Motzkin tree (*cf.* Figure 2, left).

Our first main result is the asymptotic distribution of the number of variables in random closed lambda-terms with bounded unary length of their bindings.

▶ **Theorem 4.** *Let X_n be the total number of leaves in lambda-DAGs of size n where the unary length of each binding is at most k. Then X_n is asymptotically normally distributed with*

$$\mathbb{E}X_n \sim \frac{k}{\sqrt{k}+2k}n, \quad \text{and} \quad \mathbb{V}X_n \sim \frac{k^2}{2\sqrt{k}(\sqrt{k}+2k)^2}n, \quad \text{as } n \longrightarrow \infty.$$

▶ Remark 5. Note that the number of leaves equals the number of binary nodes plus one. For $k = 1$ this implies that expections of the number of unary, the number of binary nodes, and the number of leaves are all asymptotically equal. Since the subtree attached to a unary node cannot contain further unary nodes, asymptotically almost all such trees are only a single leaf. So, almost all unary nodes are on the fringe of the tree.

On the other hand, as $k \to \infty$, we have $\mathbb{E}X_n \longrightarrow \frac{n}{2}$, and $\mathbb{V}X_n \longrightarrow 0$ for $k \longrightarrow \infty$. So, we can expect that a general lambda-term has $o(n)$ unary nodes and looks therefore like a slightly perturbed binary tree. So far, nothing isknown on the distribution of the locations of the unary nodes.

Next we turn to lambda-terms of bounded unary height.

▶ **Definition 6** (unary height, [4, Definition 1]). Consider a lambda-term T and its associated lambda-DAG $G(T)$. The unary height $h_u(v)$ of a vertex v of $G(T)$ is defined as the number of unary nodes on the path from the root to v in the underlying Motzkin tree.

The unary height of the lambda-term T is defined as the maximum number of unary nodes occurring in the separate branches of the underlying Motzkin tree (*cf.* Figure 2, right).

■ **Table 1** The coefficients occurring in the variance and the mean $k = 1, \ldots, 12$ and some larger values close to 135, the next value of some special sequence (*cf.* Definition 10), which is indicated by the lines in bold. The second column tells the number of (nested) radicands which must be considered for the determination of the dominant singularity.

bound k	$j+1$	$B''(1) + B'(1) - B'(1)^2$	$B'(1)$
1	**2**	**0**	**0**
2	2	0.0385234386	0.4381229337
3	2	0.0210625856	0.4414407371
4	2	0.0167136805	0.4463973717
5	2	0.0148700270	0.4504258849
6	2	0.0138224393	0.4536185043
7	2	0.0131157948	0.4561987871
8	**3**	**0.048**	**0.4**
9	3	0.0582322465	0.4566104777
10	3	0.0470481360	0.4560418340
11	3	0.0396601986	0.4560810348
12	3	0.0345090124	0.4564489368
$\vdots$	$\vdots$	$\vdots$	$\vdots$
133	3	0.0077469541	0.4821900098
134	3	0.0077234960	0.4822482745
135	**4**	**0.0108490182**	**0.4782608696**

▶ **Theorem 7.** *Let $\rho_k(u)$ be the root of smallest modulus of the function $z \mapsto R_{j+1,k}(z,u)$, where*

$$R_{j+1,k}(z,u) = 1 - 4(k-j)z^2u - 2z + 2z\sqrt{1 - 4(k-j+1)z^2u - 2z + \sqrt{\ldots + 2z\sqrt{1 - 4kz^2u}}},$$

and let us define $B(u) = \rho_k(u)/\rho_k(1)$.

If $B''(1) + B'(1) - B'(1)^2 \neq 0$, then the total number of leaves in lambda-DAGs with bounded unary height at most k is asymptotically normally distributed with asymptotic mean μn and asymptotic variance $\sigma^2 n$, where $\mu = B'(1)$ and $\sigma^2 = B''(1) + B'(1) - B'(1)^2$.

▶ **Remark 8.** The requirement $B''(1) + B'(1) - B'(1)^2 \neq 0$ obviously results from the fact that otherwise the variance would be equal to zero. However, this inequality seems to be very difficult to verify, since $B(u) = \frac{\rho_k(1)}{\rho_k(u)}$ and we do not know anything about the function $\rho_k(u)$, except for some crude bounds and its analyticity. In Table 1 we give *inter alia* the coefficients $B''(1) + B'(1) - B'(1)^2$ and $B'(1)$ for the variances and the mean values, respectively, for the first few values for k.

▶ **Remark 9.** No clear conclusion can be inferred from the numerical values given in Table 1. The mean seems to be slightly increasing, except for the special values belonging to the sequence given in Definition 10. But $k = 10$ is another exception in the interval $k = 9, \ldots, 134$ (not listed completely). The variance seems decreasing in any interval between two special values. If k belongs to the special sequence given in Definition 10 then we observe irregularities.

Lambda-terms of bounded unary height have been studied in [4], where a very unusual behaviour has been discovered. The asymptotic behaviour of the number of lambda-terms belonging to this subclass differs depending on whether the bound for the unary height is an element of a certain sequence $(N_i)_{i \geq 0}$, which will be given in Definition 10, or not (in Table 1
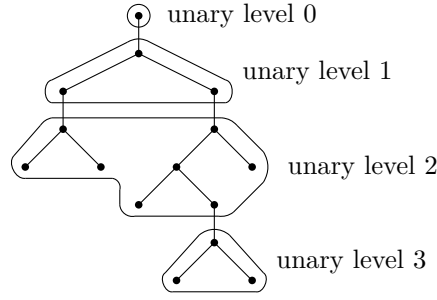
Figure 3 Underlying Motzkin tree of e.g. the lambda term $\lambda x.((\lambda y.yx)(\lambda z.(z(\lambda t.tx))z))$, where the different unary levels are encircled.

the rows belonging to elements of this sequence are therefore in bold). Though the behaviour of the counting sequences differs for these two cases, the result in Theorem 7 concerning lambda-terms of bounded unary height is the same after all. However, the method of proof is different in the two cases. For our subsequent results the distinction of cases will have an impact on the asymptotic behaviour of the investigated structures. Thus, we will have to distinguish between these two cases.

▶ **Definition 10** (auxiliary sequences $(u_i)_{i\geq 0}$ and $(N_i)_{i\geq 0}$, [4, Definition 6]). Let $(u_i)_{i\geq 0}$ and $(N_i)_{i\geq 0}$ be the integer sequence defined by $u_0 = 0$, $u_{i+1} = u_i^2 + i + 1$ for $i \geq 0$, and $N_i = u_i^2 - u_i + i$, for $i \geq 0$.

Finally, in the last section we investigate the number of leaves in lambda-DAGs with bounded unary height that are located in the different unary levels throughout the tree.

▶ **Definition 11** (unary level). A node is said to be in the i-th unary level, if there are exactly i unary nodes on the branch from the root to that node (the node itself is not counted). Thus, the i-th unary level contains all nodes with unary height i (*cf.* Figure 3).

The following theorem includes the results that we will present in Section 5, where we show that the number of leaves near the root of the lambda-DAG, *i.e.*, in the lower unary levels, is very low, while there are many leaves in the upper unary levels. Furthermore these two domains are strictly separated and the "separating level", *i.e.*, the first level with many leaves, depends on the bound of the unary height. We will show a very interesting behaviour, namely that, with growing bound of the unary height, the number of leaves within the unary level that is directly below the critical separating level increases, until the bound reaches a certain number, which makes this adjacent leaf-filled level become the new separating level.

▶ **Theorem 12.** *Let $\tilde{\rho}_{k,l}(u)$ be the root of smallest modulus of the function $z \mapsto \tilde{R}_{j+1,k}(z,u)$, where*

$$\tilde{R}_{i,k}(z,u) = 1 - 4(k-j)z^2 - 2z + \sqrt{\ldots + 2z\sqrt{1 - 4(k-l)z^2u - 2z + 2z\sqrt{\ldots + 2z\sqrt{1 - 4kz^2}}}},$$

i.e., the u is inserted only in the $(l+1)$-th radicand, and let us define $\tilde{B}_l(u) = \frac{\tilde{\rho}_{k,l}(u)}{\tilde{\rho}_{k,l}(1)}$.

1. *If $k \in (N_j, N_{j+1})$, then the average number of leaves in the first $k - j$ unary levels is $O(1)$, as $n \longrightarrow \infty$, while it is $\theta(n)$ for the last $j + 1$ unary levels.*
   In particular, if $\tilde{B}_l''(1) + \tilde{B}_l'(1) - \tilde{B}_l'(1)^2 \neq 0$, the number of leaves in each of the last $j + 1$ unary levels (i.e., $l = 0, \ldots, j$) is asymptotically normally distributed.
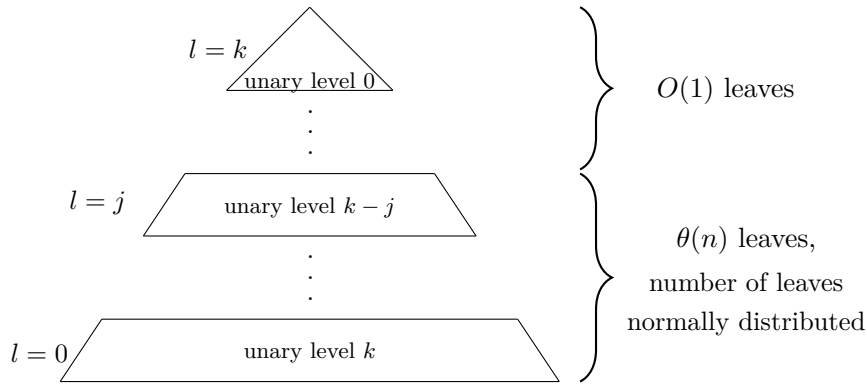
Figure 4 Summary of the mean values of the number of leaves in the different unary levels for the case $k \in (N_j, N_{j+1})$ in lambda-terms of unary height at most k.
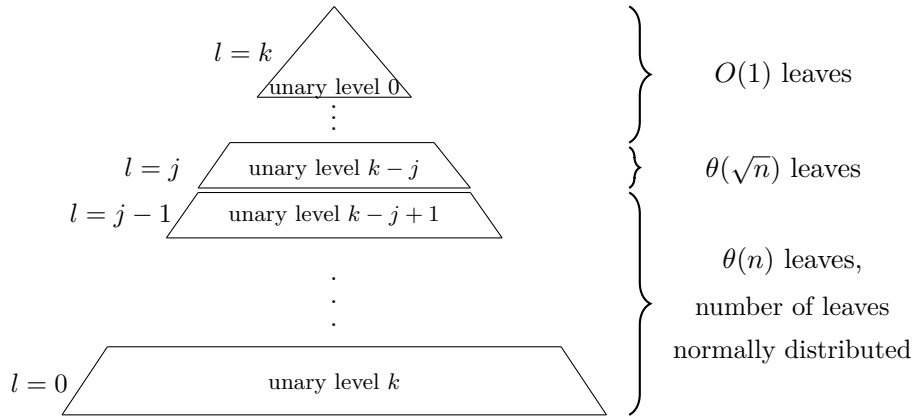
Figure 5 Summary of the mean values of the number of leaves in the different unary levels for the case $k = N_j$ in lambda-terms with unary height at most k.

2. *If $k = N_j$, then the average number of leaves in the first $k - j$ unary levels is $O(1)$, as $n \longrightarrow \infty$, while the average number of leaves in the j-th unary level is $\theta(\sqrt{n})$. The last j unary levels have asymptotically $\theta(n)$ leaves.*
   In particular, if $\tilde{B}_l''(1) + \tilde{B}_l'(1) - \tilde{B}_l'(1)^2 \neq 0$, the number of leaves in each of the last j unary levels (i.e., $l = 0, \ldots, j - 1$) is asymptotically normally distributed.

3 Total number of leaves in lambda-terms with bounded unary length of bindings

In this section we investigate the asymptotic number of all leaves in lambda-terms with bounded unary length of their bindings (*cf.* Definition 3). In order to get some quantitative results on this restricted class of lambda-terms we will use the well-known symbolic method (see [8]) and therefore we introduce certain combinatorial classes as it has been done in [4]: $\mathcal{Z}$ denotes the class of atoms, $\mathcal{A}$ the class of application nodes (*i.e.*, binary nodes), $\mathcal{U}$ the class of abstraction nodes (*i.e.*, unary nodes), and $\hat{\mathcal{P}}^{(i,k)}$ the class of unary-binary trees such that every leaf e can be labelled in $\min\{h_u(e) + i, k\}$ ways. The classes $\hat{\mathcal{P}}^{(i,k)}$ can be specified by

$$\hat{\mathcal{P}}^{(k,k)} = k\mathcal{Z} + (\mathcal{A} \times \hat{\mathcal{P}}^{(k,k)} \times \hat{\mathcal{P}}^{(k,k)}) + (\mathcal{U} \times \hat{\mathcal{P}}^{(k,k)}),$$

and

$$\hat{\mathcal{P}}^{(i,k)} = i\mathcal{Z} + (\mathcal{A} \times \hat{\mathcal{P}}^{(i,k)} \times \hat{\mathcal{P}}^{(i,k)}) + (\mathcal{U} \times \hat{\mathcal{P}}^{(i+1,k)}) \qquad \text{for } i < k.$$

Translating into generating functions with z marking the size and u marking the number of leaves, and solving for $\hat{P}^{(i,k)}(z,u)$ yields

$$\hat{P}^{(i,k)}(z,u) = \frac{1 - \mathbf{1}_{[i=k]}z - \sqrt{\hat{R}_{k-i+1}(z,u)}}{2z},$$

with $\hat{R}_{1,k}(z,u) = (1-z)^2 - 4kuz^2$, $\hat{R}_{2,k}(z,u) = 1 - 4(k-1)z^2u - 2z + 2z^2 + 2z\sqrt{\hat{R}_{1,k}(z,u)}$, and $\hat{R}_{i,k}(z,u) = 1 - 4(k-i+1)z^2u - 2z + 2z\sqrt{\hat{R}_{i-1,k}(z,u)}$,for $3 \leq i \leq k+1$.

Since the class $\hat{\mathcal{P}}^{(0,k)}$ is isomorphic to the class $\mathcal{G}_k$ of lambda-terms where all bindings have unary lengths not larger than k, we get for the corresponding bivariate generating function

$$G_k(z,u) = \hat{P}^{(0,k)}(z,u) = \frac{1 - \sqrt{\hat{R}_{k+1,k}(z,u)}}{2z}.$$

From [4] we know that the dominant singularity of $G_k(z,1)$ comes from the innermost radical and is of type $\frac{1}{2}$. Due to continuity arguments this implies that in a sufficiently small neighbourhood of $u = 1$ the dominant singularity $\hat{\rho}_k(u)$ of $G_k(z,u)$ comes also from the innermost radical and is also of type $\frac{1}{2}$. By calculating the smallest positive root of $\hat{R}_{1,k}(z,u)$ we get $\hat{\rho}_k(u) = \frac{1}{1+2\sqrt{ku}}$. Now we will determine the expansions of the radicands in a neighbourhood of the dominant singularity $\hat{\rho}_k(u)$.

▶ **Proposition 13.** *Let $\hat{\rho}_k(u)$ be the root of the innermost radicand $\hat{R}_{1,k}(z,u)$, i.e., $\hat{\rho}_k(u) = \frac{1}{1+2\sqrt{ku}}$. Then*

$$\hat{R}_{1,k}(\hat{\rho}_k(u)(1-\epsilon),u) = \left(2\hat{\rho}_k(u) - 2\hat{\rho}_k^2(u) + 8ku\hat{\rho}_k^2(u)\right)\epsilon + \mathcal{O}(\epsilon^2),$$

$$\hat{R}_{j,k}(\hat{\rho}_k(u)(1-\epsilon),u) = c_j\hat{\rho}_k^2(u) + \frac{4\hat{\rho}_k^2(u)((ku)^{\frac{1}{4}} + \sqrt{2ku})}{\prod_{l=2}^{j}\sqrt{c_l}}\sqrt{\epsilon} + \mathcal{O}(\epsilon^{\frac{3}{2}}),$$

for $2 \leq j \leq k+1$, where $c_1(u) = 1$ and $c_j(u) = 4(j-1)u - 1 + 2\sqrt{c_{j-1}(u)}$ for $2 \leq j \leq k+1$.

▶ **Theorem 14.** *Let for any fixed k, $G_k(z,u)$ denote the bivariate generating function of lambda-terms where all bindings have unary lengths not larger than k. Then*

$$[z^n]G_k(z,u) = \sqrt{\frac{\sqrt{ku} + 2ku}{4\pi\prod_{l=2}^{k+1}c_l(u)}}(1+2\sqrt{ku})^n n^{-\frac{3}{2}}\left(1 + O\left(\frac{1}{n}\right)\right), \qquad \text{for } n \longrightarrow \infty,$$

where $c_1(u) = 1$ and $c_j(u) = 4(j-1)u - 1 + 2\sqrt{c_{j-1}(u)}$, for $2 \leq j \leq k+1$.

From [4, Theorem 1] we know the following result.

$$[z^n]G_k(z,1) = \sqrt{\frac{\sqrt{k} + 2k}{4\pi\prod_{l=2}^{k+1}c_l(1)}}(1+2\sqrt{k})^n n^{-\frac{3}{2}}\left(1 + O\left(\frac{1}{n}\right)\right), \quad \text{as } n \longrightarrow \infty, \qquad (1)$$

with c_l defined as in Proposition 13.

Now we want to apply the well-known Quasi-Power Theorem.

▶ **Theorem 15** (Quasi-Power Theorem, [10])**.** *Let X_n be a sequence of random variables with the property that*

$$\mathbb{E}u^{X_n} = A(u)B(u)^{\lambda_n}\left(1 + \mathcal{O}\left(\frac{1}{\phi_n}\right)\right)$$

holds uniformly in a complex neighbourhood of $u = 1$, $\lambda_n \longrightarrow \infty$ and $\phi_n \longrightarrow \infty$, and $A(u)$ and $B(u)$ are analytic functions in a neighbourhood of $u = 1$ with $A(1) = B(1) = 1$. Set $\mu = B'(1)$ and $\sigma^2 = B''(1) + B'(1) - B'(1)^2$. If $\sigma^2 \neq 0$, then

$$\frac{X_n - \mathbb{E}X_n}{\sqrt{\mathbb{V}X_n}} \longrightarrow \mathcal{N}(0,1),$$

with $\mathbb{E}X_n = \mu\lambda_n + A'(1) + \mathcal{O}(1/\phi_n))$ and $\mathbb{V}X_n = \sigma^2\lambda_n + A''(1) + A'(1) - A'(1)^2 + \mathcal{O}(1/\phi_n))$.

Using Theorem 14 and (1), we get

$$\mathbb{E}u^{X_n} = \frac{[z^n]G_k(z,u)}{[z^n]G_k(z,1)} = \left(\frac{1 + 2\sqrt{ku}}{1 + 2\sqrt{k}}\right)^n \sqrt{\frac{\sqrt{ku} + 2ku}{2k + \sqrt{k}} \prod_{j=2}^{k+1} \frac{c_j(1)}{c_j(u)}}\left(1 + O\left(\frac{1}{n}\right)\right),$$

where $c_1(u) = 1$ and $c_j(u) = 4ju - 4u - 1 + 2\sqrt{c_{j-1}(u)}$.

Thus, all assumptions for the Quasi-Power Theorem are fulfilled, and we get that the number of leaves in lambda-DAGs with bounded unary length of their bindings is asymptotically normally distributed with

$$\mathbb{E}X_n \sim \frac{k}{\sqrt{k} + 2k}n, \quad \text{and} \quad \mathbb{V}X_n \sim \frac{k^2}{2\sqrt{k}(\sqrt{k} + 2k)^2}n, \quad \text{as } n \longrightarrow \infty,$$

and therefore Theorem 4 is shown.

4 Total number of leaves in lambda-terms with bounded unary height

This section is devoted to the enumeration of leaves in lambda-terms of bounded unary height (*cf.* Definition 6). As in [4] let us denote by $\mathcal{P}^{(i,k)}$ the class of unary-binary trees such that the unary height $h_u(e)$ of each leaf e is at most $k - i$ and every leaf can be colored with one out of $i + h_u(e)$ colors. These classes can be specified by

$$\mathcal{P}^{(k,k)} = k\mathcal{Z} + (\mathcal{A} \times \mathcal{P}^{(k,k)} \times \mathcal{P}^{(k,k)}),$$

and

$$\mathcal{P}^{(i,k)} = i\mathcal{Z} + (\mathcal{A} \times \mathcal{P}^{(i,k)} \times \mathcal{P}^{(i,k)}) + (\mathcal{U} \times \mathcal{P}^{(i+1,k)}) \quad \text{for } i < k.$$

Their bivariate generating functions can be derived analogously as the univariate ones in [4] and read as

$$P^{(i,k)}(z,u) = \frac{1 - \sqrt{R_{k-i+1,k}(z,u)}}{2z},$$

where $R_{1,k}(z,u) = 1 - 4kz^2u$, and $R_{i,k}(z,u) = 1 - 4(k - i + 1)z^2u - 2z + 2z\sqrt{R_{i-1,k}(z,u)}$, for $2 \leq i \leq k + 1$.

For the bivariate generating function of lambda-terms with bounded unary height this implies

$$H_k(z, u) = P^{(0,k)}(z, u) = \frac{1 - \sqrt{R_{k+1,k}(z, u)}}{2z}. \tag{2}$$

Thus, the generating function consists again of $k + 1$ nested radicals, but as stated in Section 2, the counting sequence of lambda-terms with bounded unary height has a very unusual behaviour, namely the location and the type of the dominant singularity changes with the bound k. More precisely, the following result has been shown in [4].

▶ **Theorem 16** ([4, Theorem 3]). *Let $(u_i)_{i \geq 0}$ and $(N_i)_{i \geq 0}$ be the integer sequences defined in Definition 10.*

(i) *If there exists $j \geq 0$ such that $N_j < k < N_{j+1}$, then there exists a constant h_k such that*

$$[z^n]H_k(z) \sim h_k n^{-3/2} \rho_k(1)^{-n}, \ as \ n \longrightarrow \infty.$$

(ii) *If there exists j such that $k = N_j$, then the following asymptotic relation holds:*

$$[z^n]H_k(z) \sim h_k n^{-5/4} \rho_k(1)^{-n} = h_k n^{-5/4} (2u_j)^n \ as \ n \longrightarrow \infty.$$

Thus, in order to investigate structural properties of this class of lambda-terms we perform a distinction of cases whether the bound k is an element of the sequence $(N_i)_{i \geq 0}$ or not.

4.1 The case $N_j < k < N_{j+1}$

From [4] we know that in this case the dominant singularity of the generating function $H_k(z, 1)$ comes from the $(j + 1)$-th radicand $R_{j+1,k}$ and is of type $\frac{1}{2}$. As in the previous section we can again use continuity arguments to guarantee that sufficiently close to $u = 1$ the dominant singularity $\rho_k(u)$ of $H_k(z, u)$ comes from the $(j + 1)$-th radicand $R_{j+1,k}(z, u)$ and is of type $\frac{1}{2}$. Now we will determine the expansions of the radicands in a neighbourhood of the dominant singularity.

▶ **Proposition 17.** *Let $\rho_k(u)$ be the dominant singularity of $H_k(z, u)$. Then*

(i) $\forall i < j + 1$ *(inner radicands)* : $R_{i,k}(\rho_k(u)(1 - \epsilon), u) = R_{i,k}(\rho_k(u), u) + \mathcal{O}(\epsilon)$

(ii) $R_{j+1,k}(\rho_k(u)(1 - \epsilon), u) = \rho_k(u)\gamma_{j+1}(u)\epsilon + \mathcal{O}(\epsilon^2)$, *with* $\gamma_{j+1}(u) = -\frac{\partial}{\partial z}R_{j+1,k}(\rho_k(u), u)$

(iii) $\forall i > j + 1$ *(outer radicands)* : $R_{i,k}(\rho_k(u)(1 - \epsilon), u) = a_i(u) + b_i(u)\sqrt{\epsilon} + \mathcal{O}(\epsilon^{\frac{3}{2}})$, *with*
$a_{i+1}(u) = 1 - 4(k-i)\rho_k^2(u)u - 2\rho_k(u) + 2\rho_k(u)\sqrt{a_i(u)}$, *and* $b_{i+1}(u) = \frac{b_i(u)\rho_k(u)}{\sqrt{a_i(u)}}$ *for* $j + 2 \leq i \leq k$, *with* $a_{j+2}(u) = 1 - 4(k-j)\rho_k^2(u)u - 2\rho_k(u)$ *and* $b_{j+2}(u) = 2\rho_k(u)\sqrt{\rho_k(u)\gamma_{j+1}(u)}$.

We know that for sufficiently large i the sequence u_i is given by $u_i = \lfloor \chi^{2^i} \rfloor$, with $\chi \approx 1.36660956\ldots$ (see [4, Lemma 18]). Therefore we have $N_j \sim u_j^2 \sim \chi^{2^j 2}$ and $N_j < k < N_{j+1} = O(N_j^2)$, which gives $j \asymp \log\log k$. This implies that $j + 1 < k + 1$, *i.e.*, that the dominant singularity $\rho_{j+1,k}(u)$ cannot come from the outermost radical.

▶ Remark 18. Obviously the same is true for the case $k = N_j$. Thus, the dominant singularity never comes from the outermost radical.

Using Proposition 17 and (2) we can prove

$$[z^n]H_k(z, u) = h_k(u)\rho_{j+1,k}(u)^{-n} \frac{n^{-\frac{3}{2}}}{\Gamma(-\frac{1}{2})} \left(1 + O\left(\frac{1}{n}\right)\right), \quad as \ n \longrightarrow \infty,$$

with $h_k(u) = -\frac{b_{k+1,k}(u)}{4\rho_{j+1,k}(u)\sqrt{a_{k+1}(u)}}$.

Taking a look at the recursive definitions of $a_i(u)$ and $b_i(u)$ (see Proposition 17), it can easily be seen that these functions are not equal to zero in a neighbourhood of $u = 1$, which implies that $h_k(u) \neq 0$ and thus we can apply the Quasi-Power Theorem. What is still left to show is, that $\sigma^2 = B''(1) + B'(1) - B'(1)^2 \neq 0$ with $B(u) = \frac{\rho_k(1)}{\rho_k(u)}$. Unfortunately, as stated in Section 2 this task appears to be quite difficult, since there is only very little known about the function $\rho_k(u)$. However, it seems very likely that this condition will be fulfilled for arbitrary $k \in (N_j, N_{j+1})$, so that the Quasi-Power Theorem can be applied and we get that the number of leaves in lambda-terms of bounded unary height is asymptotically normally distributed with asymptotic mean and variance μn and $\sigma^2 n$, respectively, where $\mu = B'(1)$ and $\sigma^2 = B''(1) + B'(1) - B'(1)^2$, with $B(u) = \frac{\rho_k(1)}{\rho_k(u)}$.

4.2 The case $k = N_j$

We know from [4] that in the case $k = N_j$ both radicands $R_{j,k}(z, 1)$ and $R_{j+1,k}(z, 1)$ vanish simultaneously and the dominant singularity is therefore of type $\frac{1}{4}$.

Now we will investigate how the radicands behave in a neighbourhood of the dominant singularity $\rho_k(u)$ for $u \neq 1$.

▶ **Lemma 19.** *Let $z = \rho_k(u)$ be the dominant singularity of the bivariate generating function $H_k(z, u)$. Then*

(i) $R_{j,k}\left(\rho_k(u)\left(1 + \frac{t}{n}\right), 1 + \frac{s}{n}\right) = \frac{1}{n}\left(c_{j,1} \cdot t + c_{j,2} \cdot s\right) + O\left(\frac{|t|^2 + |s|^2}{n^2},\right)$ *with* $c_{j,1} = 4\rho_k(1)^2 - 2\rho_k(1) - 8(k - j + 1)\rho_k(1)^2$, *and* $c_{j,2} = 4\rho_k(1)\rho_k'(1) - 4(k - j + 1)\rho_k(1)^2 - 8(k - j + 1)\rho_k(1)\rho_k'(1) - 2\rho_k'(1)$.

(ii) $R_{j+1,k}\left(\rho_k(u)\left(1 + \frac{t}{n}\right), 1 + \frac{s}{n}\right) = \frac{1}{n}\left(c_{j+1,1} \cdot t + c_{j+1,2} \cdot s\right) + 2\rho_k(1)\sqrt{R_{j,k}} + O\left(n^{-3/2}\right)$, *with* $c_{j+1,1} = -8(k - j)\rho_k(1)^2 - 2\rho_k(1)$, *and* $c_{j+1,2} = -2\rho_k'(1) - 4(k - j)\rho_k(1)^2 - 8\rho_k(1)\rho_k'(1)(k - j)$.

(iii) $R_{j+p,k}\left(\rho_k(u)\left(1 + \frac{t}{n}\right), 1 + \frac{s}{n}\right) = \hat{C}_{j+p} + \hat{D}_{j+p}\sqrt[4]{R_{j,k}} + O\left(\frac{1}{n}p_{j+p}(t, s)\right)$, *for* $2 \leq p \leq k - j + 1$, *where* $p_{j+p}(t, s)$ *is a polynomial that is linear in t and s, and $\hat{C}_{j+p}$ and $\hat{D}_{j+p}$ are constants.*

▶ **Proposition 20.** *Let $H_k(z, u)$ be the bivariate generating function of the class of lambda-terms with unary height at most k. Then the n-th coefficient of $H_k(z, u)$ is given by*

$$[z^n]H_k(z, u) = \tilde{h}_k(u)\rho_k(u)^{-n}n^{-\frac{5}{4}}\left(1 + O\left(n^{-\frac{3}{4}}\right)\right), \quad as \longrightarrow \infty,$$

with a constant $\tilde{h}_k(u) \neq 0$.

Thus, we apply the Quasi-Power Theorem and like in the previous case (where $k \in (N_j, N_{j+1})$) what is left to show is that the variance $\sigma^2 = B''(1) + B'(1) - B'(1)^2$ with $B(u) = \frac{\rho_k(1)}{\rho_k(u)}$ is positive. Assuming this requirement is valid we get that the total number of leaves in a lambda-term of bounded unary-height is asymptotically normally distributed for arbitrary bounds k.

5 Number of leaves in the unary levels in lambda-terms with bounded unary height

The aim of this section is the investigation of the distribution of the number of leaves in the different unary levels in lambda-terms with bounded unary height (*cf.* Definition 11). In order to do so, let us consider that each unary level in such a lambda-term corresponds to

one or more binary trees that contain different types of leaves, where the number of types depends on the respective level (*cf.* Figure 3). Let $\mathcal{C}$ be the class of binary trees. Using the notation from the previous sections we can specify this class by $\mathcal{C} = \mathcal{Z} + (\mathcal{A} \times \mathcal{C} \times \mathcal{C})$. Translating into generating functions and solving for $C(z, u)$, with z marking the size (*i.e.*, the total number of nodes) and u marking the number of leaves, yields $C(z, u) = \frac{1 - \sqrt{1-4uz^2}}{2z}$.

Let $_{k-l}H_k(z, u)$ be the generating function of lambda-terms with unary height at most k, where z marks the size and u marks the number of leaves on the $(k - l)$-th unary level $(0 \leq l \leq k)$. Then

$$_{k-l}H_k(z, u) = C(z, C(z, 1 + \ldots + C(z, (k - l) \cdot u + \ldots + C(z, (k - 1) + C(z, k))) \ldots) \ldots)),$$

which can be written as

$$_{k-l}H_k(z, u) = \frac{1 - \sqrt{\tilde{R}_{k+1}(z, u)}}{2z},$$

with $\tilde{R}_1(z, u) = 1 - 4z^2 k$, $\tilde{R}_i(z, u) = 1 - 4z^2(k - i + 1) - 2z + 2z\sqrt{\tilde{R}_{i-1}(z, u)}$, for $2 \leq i \leq k+1$, $i \neq l+1$, and $\tilde{R}_{l+1}(z, u) = 1 - 4z^2 u(k - l) - 2z + 2z\sqrt{\tilde{R}_{l-1}(z, u)}$.

▶ **Remark 21.** Note that the radicands $\tilde{R}_i$ that are introduced above are very similar to the radicands $R_{i,k}$ that were used in the previous section. The only difference is that now we have a u only in the $(l + 1)$-th radicand, while in the previous case u was occurring in all radicands. Thus, we will have further distinction of cases now depending on the relative position (w.r.t. l) of the radicand(s) where the dominant sigularity comes from.

We obtain the following result for the asymptotic mean values of the number of leaves in the different unary levels.

▶ **Proposition 22.** *Let X_n denote the number of leaves in the $(k - l)$-th unary level in a random lambda-term of unary height at most k with size n.*
1. *If $k \in (N_j, N_{j+1})$, then we get for the asymptotic mean*
   ▪ *in the case $l > j$:*

$$\mathbb{E}X_n = \frac{[z^n]\left(\frac{\partial}{\partial u} \, _{k-l}H_k(z, u)\right)|_{u=1}}{[z^n]_{k-l}H_k(z, 1)} = C_{k,l}\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right),$$

   ▪ *and in the case $l \leq j$:*

$$\mathbb{E}X_n = \frac{[z^n]\left(\frac{\partial}{\partial u} \, _{k-l}H_k(z, u)\right)|_{u=1}}{[z^n]_{k-l}H_k(z, 1)} = \tilde{C}_{k,l} \cdot n\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right),$$

   with constants $C_{k,l}$ and $\tilde{C}_{k,l}$ depending on l and k.
2. *If $k = N_j$, then the asymptotic mean reads as*
   ▪ *in the case $l > j$:*

$$\mathbb{E}X_n = \frac{[z^n]\left(\frac{\partial}{\partial u} \, _{k-l}H_k(z, u)\right)|_{u=1}}{[z^n]_{k-l}H_k(z, 1)} = D_{k,l}\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right),$$

   ▪ *in the case $l = j$:*

$$\mathbb{E}X_n = \frac{[z^n]\left(\frac{\partial}{\partial u} \, _{k-l}H_k(z, u)\right)|_{u=1}}{[z^n]_{k-l}H_k(z, 1)} = \hat{D}_{k,l} \cdot \sqrt{n}\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right),$$

- *and in the case $l < j$:*

$$\mathbb{E}X_n = \frac{[z^n]\left(\frac{\partial}{\partial u}\ _{k-l}H_k(z,u)\right)|_{u=1}}{[z^n]_{k-l}H_k(z,1)} = \tilde{D}_{k,l} \cdot n\left(1 + \mathcal{O}\left(n^{-\frac{1}{4}}\right)\right),$$

with constants $D_{k,l}$, $\hat{D}_{k,l}$ and $\tilde{D}_{k,l}$ depending on l and k.

Now that we derived the mean values for the number of leaves in the different unary levels, we are interested in their distribution. Therefore we perform the same distinction of cases as we did for the mean values. However, so far we only know the distribution of the leaves in the all those levels, which contain many leaves.

▶ **Proposition 23.** *Let $z = \tilde{\rho}_{k,l}(u)$ denote the dominant singularity of $_{k-l}H_k(z,u)$.*
1. *If $k \in (N_j, N_{j+1})$, then we get for $l \leq j$*

$$\frac{[z^n]\ _{k-l}H_k(z,u)}{[z^n]\ _{k-l}H_k(z,1)} = \frac{\tilde{h}_k(u)}{h_k}\left(\frac{\tilde{\rho}_{k,l}(1)}{\tilde{\rho}_{k,l}(u)}\right)^n\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right),$$

with constants $\tilde{h}_k(u)$ and h_k that are not equal to zero.
2. *If $k = N_j$, then it holds for $l < j$*

$$\frac{[z^n]\ _{k-l}H_k(z,u)}{[z^n]\ _{k-l}H_k(z,1)} = \frac{\hat{h}_k(u)}{h_k}\left(\frac{\tilde{\rho}_{k,l}(1)}{\tilde{\rho}_{k,l}(u)}\right)^n\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right),$$

with constants $\hat{h}_k(u)$ and h_k that are not equal to zero.

Finally, by using the Quasi-Power Theorem the proof of Theorem 12 is finished. Therefore we have to assume again that the variance is not equal to zero.

As stated before the generating function $_{k-j}H_k(z,u)$ consists of $k+1$ nested radicals, where a u is inserted in the $(l+1)$-th radicand counted from the innermost. In the case $k \in (N_j, N_{j+1})$ we know that the dominant singularity $\tilde{\rho}_k(u)$ comes from the $(j+1)$-th radicand. Thus, if $l > j$ then $\tilde{\rho}_k(u)$ is independent of u and we will not get a quasi-power. The same holds for the case $k = N_j$ and $l < j$, since we showed that in this case the dominant singularity comes from the j-th radicand. The $(j+1)$-th unary level for $k = N_j$ is a special case, because we do not know whether the dominant singularity comes from the j-th or the $(j+1)$-th radicand. However, it seems very unlikely that the number of leaves in this level will be asymptotically normally distributed, but further studies on this subject might be very interesting.

─── **References** ───

1. Henk P. Barendregt. *The lambda calculus, its syntax and semantics*, volume 40 of *Studies in Logic (London)*. College Publications, London, 2012. [Reprint of the 1984 revised edition, MR0774952], With addenda for the 6th imprinting, Mathematical Logic and Foundations.
2. O. Bodini, D. Gardy, and A. Jacquot. Asymptotics and random sampling for BCI and BCK lambda terms. *Theoret. Comput. Sci.*, 502:227–238, 2013. `doi:10.1016/j.tcs.2013.01.008`.
3. Olivier Bodini, Danièle Gardy, and Bernhard Gittenberger. Lambda terms of bounded unary height. In *ANALCO11—Workshop on Analytic Algorithmics and Combinatorics*, pages 23–32. SIAM, Philadelphia, PA, 2011.
4. Olivier Bodini, Danièle Gardy, Bernhard Gittenberger, and Zbigniew Gołębiewski. On the number of unary-binary tree-like structures with restrictions on the unary height. *Ann. Combin.*, 2018. To appear.

5   Olivier Bodini, Danièle Gardy, Bernhard Gittenberger, and Alice Jacquot. Enumeration of generalized BCI lambda-terms. *Electron. J. Combin.*, 20(4):Paper 30, 23, 2013.

6   Olivier Bodini and Bernhard Gittenberger. On the asymptotic number of $BCK(2)$-terms. In *ANALCO14—Meeting on Analytic Algorithmics and Combinatorics*, pages 25–39. SIAM, Philadelphia, PA, 2014. `doi:10.1137/1.9781611973204.3`.

7   Haskell Brooks Curry, Robert Feys, William Craig, J Roger Hindley, and Jonathan P Seldin. *Combinatory logic*, volume 1. North-Holland Amsterdam, 1958.

8   Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009. `doi:10.1017/CBO9780511801655`.

9   Katarzyna Grygiel, Paweł M. Idziak, and Marek Zaionc. How big is BCI fragment of BCK logic. *J. Logic Comput.*, 23(3):673–691, 2013. `doi:10.1093/logcom/exs017`.

10   H-K Hwang. On convergence rates in the central limit theorems for combinatorial structures. *European Journal of Combinatorics*, 19(3):329–343, 1998.

11   Pierre Lescanne. On counting untyped lambda terms. *Theoret. Comput. Sci.*, 474:80–97, 2013. `doi:10.1016/j.tcs.2012.11.019`.

Counting Ascents in Generalized Dyck Paths

Benjamin Hackl

Alpen-Adria-Universität Klagenfurt, Austria
benjamin.hackl@aau.at
🆔 https://orcid.org/0000-0003-2998-9599

Clemens Heuberger

Alpen-Adria-Universität Klagenfurt, Austria
clemens.heuberger@aau.at
🆔 https://orcid.org/0000-0003-0082-7334

Helmut Prodinger

Department of Mathematical Sciences, Stellenbosch University, South Africa
hproding@sun.ac.za
🆔 https://orcid.org/0000-0002-0009-8015

―― **Abstract** ――――――――――――――――――――――――――――――――

Non-negative Łukasiewicz paths are special two-dimensional lattice paths never passing below their starting altitude which have only one single special type of down step. They are well-known and -studied combinatorial objects, in particular due to their bijective relation to trees with given node degrees.

We study the asymptotic behavior of the number of ascents (i.e., the number of maximal sequences of consecutive up steps) of given length for classical subfamilies of general non-negative Łukasiewicz paths: those with arbitrary ending altitude, those ending on their starting altitude, and a variation thereof. Our results include precise asymptotic expansions for the expected number of such ascents as well as for the corresponding variance.

1 Introduction

Two-dimensional lattice paths can be defined as sequences of points in the plane $\mathbb{R}^2$ where for any point, the vector pointing to the succeeding point ("step") is from a predefined finite set, the *step set*. In general, lattice paths are very classical combinatorial objects with a variety of applications in, amongst others, Biology, Physics, and Chemistry.
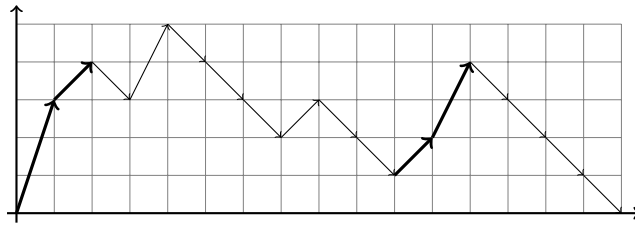
Figure 1 Simple Łukasiewicz excursion of length 16 with emphasized 2-ascents where $\mathcal{S} = \{-1, 1, 2, 3\}$

In this paper, our focus lies on a special class of two-dimensional lattice paths: *non-negative simple Łukasiewicz paths*. A lattice path is said to be *simple* if the horizontal coordinate is the same (e.g. is 1) for all possible steps. In case of a simple path family, we define the *step set* $\mathcal{S}$ as the set of allowed height differences, i.e., the respective y-coordinates between the points of the path. If, additionally, the step set $\mathcal{S} \subseteq \mathbb{Z}$ is integer-valued and contains -1 as the single negative value (meaning that all other values in $\mathcal{S}$ are non-negative), then the corresponding paths are called *simple Łukasiewicz paths*.

If a lattice path starts at the origin and never passes below the horizontal axis, then the path is said to be a *meander* (or non-negative path). And in case such a non-negative path ends on the horizontal axis, it is called an *excursion*.

Lattice path families of this type have been studied intensely, see [1] for a detailed survey on general simple lattice paths, and, for example, [2, 9] for investigations concerning Łukasiewicz paths.

We are interested in analyzing the number of *ascents* in these paths. An *ascent* is an inclusion-wise maximal sequence of up steps (i.e., steps in $\mathcal{S} \setminus \{-1\}$; this might also include the horizontal step corresponding to 0). For an integer $r \geq 1$, if an ascent consists of precisely r steps, then the ascent is said to be an r-ascent. As an example, Figure 1 depicts some non-negative Łukasiewicz excursion with emphasized 2-ascents.

In this paper, we give a precise analysis of the number of r-ascents for non-negative simple Łukasiewicz paths of given length, as well as of variants of this class of lattice paths. Our investigation is motivated by [6], where the number of 1-ascents in a special lattice path class related to the classic Dyck paths was analyzed explicitly by elementary methods.

Main Results

Within this paper, three special classes of non-negative Łukasiewicz paths are of interest:

- *excursions*, i.e., paths that end on the horizontal axis,
- *dispersed excursions*, i.e., excursions where horizontal steps are not allowed except on the horizontal axis,
- *meanders*, i.e., general non-negative Łukasiewicz paths without additional restrictions.

Formally, we conduct our analysis by investigating random variables $E_{n,r}$, $D_{n,r}$, $M_{n,r}$ which model the number of r-ascents in a random excursion, dispersed excursion, and meander of length n, respectively. The underlying probability models are based on equidistribution: within a family, all paths of length n are assumed to be equally likely.

Given $r \in \mathbb{N}$ and considering $n \in \mathbb{N}_0$ with $n \to \infty$, we prove that for excursions we have

$$\mathbb{E}E_{n,r} = \mu n + c_0 + O(n^{-1/2}) \qquad \text{and} \qquad \mathbb{V}E_{n,r} = \sigma^2 n + O(n^{1/2}),$$

for some constants μ, c_0, σ^2 depending on the chosen step set $\mathcal{S}$. The constants are given explicitly in Theorem 7. Additionally, if n is not a multiple of the so-called *period* of the step set, then the random variable degenerates and we have $E_{n,r} = 0$; see Theorem 7 for details.

For dispersed excursions, the corresponding computations get rather messy, which is why we restrict ourselves to the investigation of d_n, the number of dispersed excursions of length n, as well as the expected value $\mathbb{E}D_{n,r}$. In particular, for all step sets $\mathcal{S}$ (except for the special case of dispersed Dyck paths with $\mathcal{S} = \{-1, 1\}$), d_n satisfies

$$d_n = c_0 \kappa^n n^{-3/2} + O(\kappa^n n^{-5/2}),$$

with constants c_0 and κ depending on the chosen step set. For the expected number of ascents in this particular lattice path family, we find

$$\mathbb{E}D_{n,r} = \mu n + O(1)$$

for some constant μ depending on $\mathcal{S}$. Explicit values for these constants and more details are given in Theorem 10.

In the context of meanders we are able to show that for all step sets (with two special exceptions: Dyck meanders with $\mathcal{S} = \{-1, 1\}$, and Motzkin meanders with $\mathcal{S} = \{-1, 0, 1\}$) we have

$$\mathbb{E}M_{n,r} = \mu n + c_0 + O(n^{5/2} \kappa^n) \qquad \text{and} \qquad \mathbb{V}M_{n,r} = \sigma^2 n + O(1),$$

for constants μ, c_0, $\kappa \in (0, 1)$, σ^2 depending on $\mathcal{S}$. Also, the random variable $M_{n,r}$ is asymptotically normally distributed; see Theorem 12 for explicit formulas for the constants and more details.

In theory, our approach can be used to obtain arbitrarily precise asymptotic expansions for all the quantities above. For the sake of readability we have chosen to only give the main term as well as one additional term, wherever possible.

On a more technical note, in order to deal with general Łukasiewicz step sets in our setting, we make use of a generating function approach (see [3, Chapter I]). In particular, we heavily rely on the technique of *singular inversion* (see [3, Chapter VI.7], [8]), which deals with finding an asymptotic expansion for the growth of the coefficients of generating functions $y(z)$ satisfying a functional equation of the type

$$y = z \, \phi(y)$$

with a suitable function ϕ.

Notation and Special Cases

Throughout this paper, the step set is denoted as $\mathcal{S} = \{-1, b_1, \dots, b_{m-1}\}$ with integers $b_j \geq 0$ for all j and $m \geq 1$. The b_j are referred to as up steps – even if the step is a horizontal one.

The so-called characteristic polynomial of the lattice path class, i.e., the generating function corresponding to the set $\mathcal{S}$, is denoted by $S(u) := \sum_{s \in \mathcal{S}} u^s$. The strongly related generating function of the non-negative steps is denoted by $S_+(u) := \sum_{s \in \mathcal{S}, \, s \geq 0} u^s$.

In this context, observe that the particular step set $\mathcal{S} = \{-1, 0\}$ corresponds to a, in some sense, pathological family of Łukasiewicz paths. In this case, there is only precisely one non-negative Łukasiewicz path of any given length. The family of meanders and excursions coincides, and also the random variables degenerate in the sense that we have[1]

---

[1] We make use of the Iversonian notation popularized in [4, Chapter 2]: $[\![expr]\!]$ takes value 1 if *expr* is true, and 0 otherwise.

$E_{n,r} = M_{n,r} = [\![n = r]\!]$. Thus, further investigation of this case is not required – which is why we exclude the case $\mathcal{S} = \{-1, 0\}$ from now on.

While in the case of a general step set $\mathcal{S}$ we are forced to deal with implicitly given quantities, for special cases like $\mathcal{S} = \{-1, 1\}$ (Dyck paths), everything can be made completely explicit as we will demonstrate in the course of our investigations.

Finally, we make use of the following well-established notation: For a generating function $f(z) = \sum_{n \geq 0} f_n z^n$, the coefficient belonging to z^n is denoted as $f_n = [z^n]f(z)$.

Structure of this Abstract

In Section 2, we determine suitable generating functions required to analyze the number of ascents. The approach is based on the inherent relation between Łukasiewicz paths and plane trees with given vertex degrees. Formulas for the respective generating functions are given in Proposition 4. Note that in the full version of this extended abstract, [5], we demonstrate another approach (following the kernel method and the "Adding a new slice approach") to determine the suitable generating functions.

Section 3 contains the actual analysis of ascents for the different lattice path families mentioned above. In particular, in Section 3.1 we investigate excursions; the main result is stated in Theorem 7. Section 3.2 deals with the analysis of ascents in dispersed excursions. In this case, the expected number of r-ascents for all but one given step sets is analyzed within Theorem 10, and the analysis for the remaining one is conducted in Proposition 11. Finally, Section 3.3 contains our results for ascents in meanders. Similarly to the previous section, the analysis for most step sets is given in Theorem 12, and the remaining cases are investigated in Propositions 13 and 14.

Proofs and additional details can be found in the full version of this extended abstract, see [5]. Appendix A contains several important tools necessary for a detailed analysis of the inverse function in the center of this abstract. In particular, with Propositions 15 and 16 we prove useful extensions of [3, Theorem VI.6; Remark VI.17].

2   Generating Functions: A Combinatorial Approach

In this section we will introduce and discuss the preliminaries required in order to carry out the asymptotic analysis of ascents in the different path classes. We begin by taking a closer look at the structure of Łukasiewicz paths.

Of course, the number of excursions of given length n strongly depends on the structure of the step set $\mathcal{S}$. For example, in the case of Dyck paths, i.e., $\mathcal{S} = \{-1, 1\}$, there cannot be any excursions of odd length – Dyck paths are said to be periodic lattice paths.

▶ **Definition 1** (Periodicity of lattice paths)**.** Let $\mathcal{S}$ be a Łukasiewicz step set with corresponding characteristic polynomial $S(u) = \sum_{s \in \mathcal{S}} u^s$. Then the period of $\mathcal{S}$ (and the associated lattice path family) is the largest integer p for which a polynomial Q satisfying

$$u\,S(u) = Q(u^p)$$

exists. If $p = 1$, then $\mathcal{S}$ is said to be aperiodic, otherwise $\mathcal{S}$ is said to be p-periodic.

▶ Remark. Observe that if a step set $\mathcal{S}$ has period p, then there are only excursions of length n where $n \equiv 0 \pmod{p}$. This can be seen by considering the generating function enumerating unrestricted paths of length n with respect to their height, i.e., $S(u)^n$. Obviously, the number

of excursions of length n is at most the number of unrestricted paths ending at altitude 0, and the latter one can be written as

$$[u^0]\, S(u)^n = [u^n](u\, S(u))^n = [u^n]Q(u^p)^n.$$

Hence, if $n \not\equiv 0 \pmod{p}$, there are no unrestricted paths ending on the horizontal axis – and thus also no excursions.

The following proposition describes an integral relation which allows us to construct a suitable generating function later on.

▶ **Proposition 2.** *The excursions of Łukasiewicz paths of length n with respect to some step set $\mathcal{S}$ correspond to rooted plane trees with $n + 1$ nodes and node degrees contained in the set $1 + \mathcal{S}$.*

An r-ascent in a Łukasiewicz excursion with respect to the step set $\mathcal{S}$ corresponds to a rooted subtree such that the leftmost leaf in this subtree has height r, and additionally the root node of the subtree is not a leftmost child itself (in the original tree).

Proof. As pointed out in e.g. [1, Example 3], this bijection between rooted plane trees with given node degrees and Łukasiewicz excursions is well known. See [7, Section 11.3] for an approach using words. However, as this bijection and its consequences makes up an integral part of the argumentation within this paper, we present a short proof ourselves. Furthermore, proving the bijection allows us to find the substructure in the tree corresponding to an r-ascent.

Given a rooted plane tree T consisting of n nodes whose outdegrees are contained in $1 + \mathcal{S}$, we construct a lattice path as follows: when traversing the tree in preorder[2], if passing a node with outdegree d, take a step of height $d - 1$. The resulting lattice path thus consists of n steps, and always ends on altitude -1, which follows from

$$\sum_{v \in T}(\deg(v) - 1) = \sum_{v \in T}\deg(v) - n = (n - 1) - n = -1,$$

where $\deg(v)$ denotes the outdegree (i.e., the number of children) of a node v in the tree T. In particular, observe that by taking the first $n - 1$ steps of the lattice path, we actually end up with a Łukasiewicz excursion using the steps from $\mathcal{S}$. To see this, first observe that as the last node traversed in preorder certainly is a leaf, meaning that the nth step in the corresponding lattice path is a down step. As the path ends on altitude -1 after n steps, we have to arrive at the starting altitude after $n - 1$ steps.

Furthermore, as illustrated in Figure 2, adding one to the current height of the constructed lattice path gives the size of the stack remembering the children that still have to be visited while traversing the tree in preorder. Combining the two previous arguments proves that the first $n - 1$ steps in the constructed lattice path form a Łukasiewicz excursion.

Similarly, by simply reversing the lattice path construction, a rooted plane tree of size $n + 1$ with node degrees in $1 + \mathcal{S}$ can be constructed from any Łukasiewicz excursion of length n with respect to $\mathcal{S}$. This establishes the bijection between the two combinatorial families.

Finally, Figure 3 illustrates what r-ascents in Łukasiewicz paths are mapped to by means of the bijection above. ◀

---

[2] Traversing a tree in preorder corresponds to the order in which the nodes are visited when carrying out a depth-first search on it.
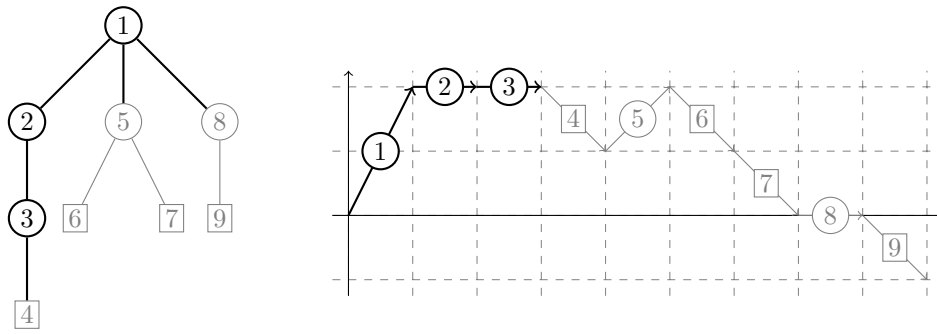
Figure 2 Bijection between Łukasiewicz paths and trees with given node degrees. The emphasized nodes and edges indicate the construction of the tree after the first three steps, which illustrates that the height of the Łukasiewicz path is one less than the number of available node positions in the tree.
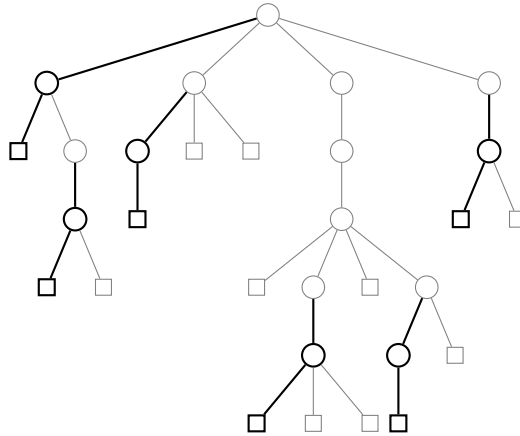
Figure 3 Plane tree with 30 nodes bijective to some Łukasiewicz excursion with respect to the step set $\mathcal{S} = \{-1, 0, 1, 2, 3\}$ whose number of 2-ascents is 6. The edges and nodes corresponding to the 2-ascents are emphasized.

In some sense, the bijection from Proposition 2 can be seen as a generalization of the well-known bijection between Dyck paths and binary trees where the tree is traversed in preorder, internal nodes correspond to up steps and leaves to down steps.

The fact that there is this bijection between Łukasiewicz excursions and these special trees with given node degrees allows us to draw an immediate conclusion regarding the corresponding generating functions.

▶ **Corollary 3.** *Let $V(z, t)$ be the generating function enumerating rooted plane trees with node degrees in $1 + \mathcal{S}$ where z marks the number of nodes and t marks the number of r-ascents in the corresponding Łukasiewicz excursion. Then $V(z, t)/z$ enumerates Łukasiewicz excursions with respect to $\mathcal{S}$ based on their length (marked by z) and the number of r-ascents (marked by t).*

Additionally, $V(z, t)$ satisfies the equations

$$V(0, t) = 0, \qquad V(z, t) = zL(z, t, V(z, t)), \tag{1}$$

$$\mathcal{V} \quad = \quad \sum_{s \in \mathcal{S}} \underbrace{\mathcal{V} \quad \mathcal{V} \quad \cdots \quad \mathcal{V}}_{1+s}$$
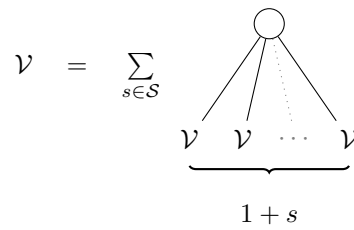
■ **Figure 4** Symbolic equation for the family of plane trees $\mathcal{V}$ with outdegrees in $1 + \mathcal{S}$. The generating function for $\mathcal{V}$ is $V(z)$, and the root node is enumerated by z.

where

$$L(z,t,v) = \frac{1}{1 - z\,S_+(v)} + (t-1)(z\,S_+(v))^r$$

enumerates sequences of up steps. The power series representation of $V(z,t)$ is given by

$$V(z,t) = \sum_{j \geq 0} g_j(t) z^{jp+1} \tag{2}$$

where p denotes the period of $\mathcal{S}$. In particular, $V(z) := V(z,1)$, the ordinary generating function enumerating plane trees with node degrees in $1+\mathcal{S}$ with respect to their size, satisfies

$$V(0) = 0, \qquad V(z) = zV(z)\,S(V(z)). \tag{3}$$

Proof. The first part of this statement is an immediate consequence of the bijection from Proposition 2. In order to prove (1), we observe that $\mathcal{V}$, the combinatorial class of plane trees with vertex outdegrees in $1 + \mathcal{S}$, can be constructed combinatorially by means of the symbolic equation

$$\mathcal{V} = \bigcirc \times \mathrm{SEQ}\Big(\bigcirc \times \sum_{\substack{s \in \mathcal{S} \\ s \geq 0}} \mathcal{V}^s\Big).$$

In a nutshell, this constructs trees in $\mathcal{V}$ by explicitly building the path to the leftmost leaf (the first factor in the equation above) in the tree as a sequence of nodes. Apart from a leftmost child, these nodes also have an additional $s \in \mathcal{S}$ branches, $s \geq 0$, where again a tree from $\mathcal{V}$ is attached. Considering that we obtain an r-ascent when using this construction with a sequence of length r, this is precisely what is enumerated by $L(z,t,V(z,t))$. Thus, the symbolic equation directly translates into the functional equation in (1). The condition $V(0,t) = 0$ is a consequence of the fact that there are no rooted trees without nodes.

The power series representation in (2) follows immediately from the considerations on periodicity at the beginning of this section.

Setting $t = 1$ in (1) leads to (3). We also want to give a combinatorial proof of (3):

The implicit equation follows from the observation that a tree with node degrees from $1 + \mathcal{S}$ can be seen as a root node (enumerated by z) where $1 + s$ for $s \in \mathcal{S}$ such trees are attached. Translating this into the language of generating functions via the symbolic equation illustrated in Figure 4, yields

$$V(z) = z \sum_{s \in \mathcal{S}} V(z)^{1+s} = zV(z)\,S(V(z)). \qquad \blacktriangleleft$$

The shape of the functional equation (3), which is an immediate consequence of the recursive structure of the underlying trees, is rather special. While it is tempting to cancel $V(z)$ on both sides of this equation, it is better to leave it in the present form: on the one hand, $S(u)$ starts with the summand $1/u$ – and on the other hand, we require (3) to be in this special form $y = z\,\phi(y)$ such that we can use singular inversion to obtain the asymptotic behavior of the coefficients of the generating function $V(z)$.

▶ **Proposition 4.** *Let $F(z,t,v)$ be the trivariate ordinary generating function counting non-negative Łukasiewicz paths with step set $\mathcal{S}$ starting at 0, where z marks the length of the path, t marks the number of r-ascents, and v marks the final altitude of the path. Then $F(z,t,v)$ can be expressed as*

$$F(z,t,v) = \frac{v - V(z,t)}{v - zL(z,t,v)} L(z,t,v), \tag{4}$$

where $V(z,t)$ and $L(z,t,v)$ are defined as in Corollary 3.

Proof. It is not hard to see that by considering a sequence of paths enumerated by $L(z,t,v)$ followed by a single down step (the corresponding generating function for this class is $\frac{1}{1-L(z,t,v)\,z/v}$), any unrestricted Łukasiewicz path with respect to $\mathcal{S}$ ending on a down step can be constructed.

We want to subtract all paths that pass below the starting altitude in order to obtain the trivariate generating function $\Phi(z,t,v)$ enumerating just the non-negative Łukasiewicz paths. The paths passing below the axis can be decomposed into an excursion enumerated by $V(z,t)/z$ (see Corollary 3), followed by an (illegal) down step enumerated by z/v, and ending with an unrestricted path again. Thus, the paths to be subtracted are enumerated by

$$\frac{V(z,t)}{z}\frac{z}{v}\frac{1}{1-L(z,t,v)\frac{z}{v}}.$$

Therefore, we find

$$\Phi(z,t,v) = \frac{v - V(z,t)}{v - zL(z,t,v)}.$$

Keeping in mind that $\Phi(z,t,v)$ only enumerates those non-negative Łukasiewicz paths ending on a down step $\searrow$, the generating function $F(z,t,v)$ enumerating all such paths can be obtained by appending another sequence of upsteps, i.e.,

$$F(z,t,v) = \Phi(z,t,v)L(z,t,v).$$

This proves the statement. ◀

Now, with an appropriate generating function at hand let us discuss our approach for the asymptotic analysis of the number of ascents in a nutshell.

Basically, we set $v = 0$ to obtain a bivariate generating function enumerating Łukasiewicz excursions, and we set $v = 1$ to obtain a generating function enumerating Łukasiewicz meanders. The appropriate generating functions for the factorial moments of $E_{n,r}$ and $M_{n,r}$ (from which expected value and variance can be computed) are then obtained by first differentiating the corresponding generating function with respect to t (possibly more often than once) and then setting $t = 1$ in this partial derivative. The growth of the coefficients of this function can then be extracted by means of singularity analysis.

In particular, this means that in order to compute the asymptotic expansions for the quantities we are interested in, we only need more information on $V(z,1)$.

▶ Notation. For the sake of simplicity, and because we will deal with these expressions throughout the entire paper, we omit the second argument in $V(z, t)$ in case $t = 1$, i.e., we set $V(z) := V(z, 1)$, $V_t(z) := V_t(z, 1) = \frac{\partial}{\partial t} V(z, t)|_{t=1}$, $V_z(z) := V_z(z, 1) = \frac{\partial}{\partial z} V(z, t)|_{t=1}$, and so on.

▶ **Example 5** (Explicit $F(z, t, v)$). In the case of $\mathcal{S} = \{-1, 1\}$ and $r = 1$ the generating function $F(z, t, v)$ can be computed explicitly and we find

$$F(z, t, v) =$$
$$\frac{(1 + (t-1)vz(1 - vz))((1 - 2vz)(1 - (t-1)z^2) - \sqrt{(1 - (t+3)z^2)(1 - (t-1)z^2)})}{2z(1 - (t-1)z^2)(z - v + v^2z + vz^2(t-1)(1-z))}. \tag{5}$$

Now, as we have derived a suitable generating function, we are interested in extracting information like, for example, asymptotic growth rates from $F(z, t, v)$. To this end, we need more information on the function $V(z, t)$.

▶ **Proposition 6.** *Let $V(z, t)$ be the bivariate generating function from Corollary 3. Let $\tau > 0$ be the uniquely determined positive constant satisfying $S'(\tau) = 0$. Then $V(z)$ has radius of convergence $\rho := 1/S(\tau)$ with a square-root singularity for $z \to \rho$. If $\mathcal{S}$ has period p, then the dominant singularities (i.e., singularities with modulus ρ) are located at $\zeta\rho$ with $\zeta \in G(p)$. The corresponding expansions are given by*

$$V(z) \stackrel{z \to \zeta\rho}{=} \zeta\tau - \zeta\sqrt{\frac{2S(\tau)}{S''(\tau)}} \left(1 - \frac{z}{\zeta\rho}\right)^{1/2} - \zeta \frac{S(\tau)S'''(\tau)}{3S''(\tau)^2} \left(1 - \frac{z}{\zeta\rho}\right) + O\left(\left(1 - \frac{z}{\zeta\rho}\right)^{3/2}\right). \tag{6}$$

3 Analysis of Ascents

3.1 Analysis of Excursions

In this section we focus on the analysis of *excursions*, i.e., paths that start and end on the horizontal axis. On the generating function level, this corresponds to setting $v = 0$ in $F(z, t, v)$ from (4). Also note that from this point on it is quite useful to replace $S_+(v) = S(v) - 1/v$ in $F(z, t, v)$.

Recall that $E_{n,r}$ is the random variable modeling the number of r-ascents in a random non-negative Łukasiewicz excursion of length n with respect to some given step set $\mathcal{S}$.

▶ **Theorem 7.** *Let $r \in \mathbb{N}$, $n \in \mathbb{N}_0$, and $p \geq 1$ be the period of the step set $\mathcal{S}$. Let τ be the structural constant, i.e., the unique positive solution of $S'(\tau) = 0$. Set $c := \tau S(\tau)$.*

Then, the expected number of r-ascents in Łukasiewicz paths of length n for $n \equiv 0$ (mod p) as well as the corresponding variance grow with $n \to \infty$ according to the asymptotic expansions

$$\mathbb{E}E_{n,r} = \frac{(c-1)^r}{c^{r+2}} n + \frac{(c-1)^{r-2}}{2\tau^2 c^{r+2} S''(\tau)^2} \Big(S''(\tau)^2\tau^2\big(4c^2 - (r+8)c + r + 4\big)$$
$$- S''(\tau)S(\tau)\big(6c^2 - 6(r+2)c + r^2 + 5r + 6\big)$$
$$- S'''(\tau)c(2c^2 - (r+4)c + r + 2)\Big)$$
$$+ O(n^{-1/2}) \tag{7}$$

and

$$\mathbb{V}E_{n,r} = \left(\frac{(c-1)^r}{c^{r+2}} + \frac{(2c-2r-3)(c-1)^{2r}}{c^{2r+4}} - \frac{(c-1)^{2r-2}(2c-r-2)^2}{c^{2r+3}\tau^3\,S''(\tau)} \right) n$$
$$+ O(n^{1/2}). \quad (8)$$

Additionally, for $n \not\equiv 0 \mod p$, we have $E_{n,r} = 0$. All O-constants depend implicitly on r.

By means of Theorem 7 we are immediately able to determine the asymptotic behavior of interesting special cases. We are particularly interested in the most basic setting: $\mathcal{S} = \{-1, 1\}$, i.e., Dyck paths.

▶ **Example 8** (*r*-Ascents in Dyck paths). In the case of Dyck paths, we have $u\,S(u) = 1 + u^2$. From there, it is easy to see that $\tau = 1$ and $\rho = 1/2$, and that the family of paths is 2-periodic. By the same approach as in the proof of Theorem 7, we can determine the expected number and variance of r-ascents in Dyck paths of length $2n$ with higher precision than stated in Theorem 7, namely as

$$\mathbb{E}D_{2n,r} = \frac{n}{2^{r+1}} - \frac{(r+1)(r-4)}{2^{r+3}} + \frac{(r^2 - 11r + 22)(r+1)r}{2^{r+6}}n^{-1} + O(n^{-2})$$

and

$$\mathbb{V}D_{2n,r} = \left(\frac{1}{2^{r+1}} - \frac{r^2 - 2r + 3}{2^{2r+3}} \right) n$$
$$- \left(\frac{r^2 - 3r - 4}{2^{r+3}} - \frac{3r^4 - 20r^3 + 29r^2 - 10r - 14}{2^{2r+5}} \right) + O(n^{-1/2}).$$

However, as we have a closed expression for $V(z)$, we can do even better. Because of

$$\frac{V(z)}{z} = \frac{1 - \sqrt{1 - 4z^2}}{2z^2},$$

we can also write down the generating function $V_t(z)/z$ for the expected number of r-ascents explicitly.

Ultimately, after extracting the corresponding coefficients of $V_t(z)/z$ we find

$$\mathbb{E}D_{2n,r} = \frac{1}{C_n} \binom{2n - r - 1}{n - 1}.$$

3.2   Analysis of Dispersed Excursions

Let $\mathcal{S}$ be a Łukasiewicz step set where $0 \notin \mathcal{S}$. In this setting, we define a *dispersed Łukasiewicz excursion* to be an $\mathcal{S}$-excursion where, additionally, horizontal steps can be taken whenever the path is on its starting altitude. Observe that, by our definition of r-ascents, these horizontal steps do not contribute towards ascents, as only the non-negative steps from $\mathcal{S}$ are relevant.

The motivation to study this specific family of Łukasiewicz paths originates from [6], where the authors investigate the total number of 1-ascents in dispersed Dyck paths using elementary methods. Our goal in this section is to find asymptotic expansions for the number of dispersed Łukasiewicz excursions of given length as well as for the expected number of r-ascents in these paths.

We begin our analysis by constructing a suitable bivariate generating function enumerating dispersed Łukasiewicz excursions with respect to their length and the number of r-ascents.

▶ **Proposition 9.** *Let $r \in \mathbb{N}$ and $V(z,t)$ as in Corollary 3. Then the generating function $D(z,t)$ enumerating dispersed $\mathcal{S}$-excursions where z marks the length of the excursion and t marks the number of r-ascents is given by*

$$D(z,t) = \frac{1}{z} \frac{V(z,t)}{1 - V(z,t)}. \tag{9}$$

Proof. Let $\mathcal{E}$ denote the combinatorial class of $\mathcal{S}$-excursions. The corresponding bivariate generating function is given by $V(z,t)/z$, as proved in Corollary 3.

By the symbolic method (see [3, Chapter I]), the combinatorial class $\mathcal{D}$ of dispersed excursions can be constructed as $\mathcal{D} = (\mathcal{E} \rightarrow)^* \mathcal{E}$, where $\rightarrow^*$ represents a (possibly empty) sequence of horizontal steps. Translating this combinatorial construction in the language of (bivariate) generating functions yields (9). ◀

In preparation for the analysis of the generating function $D(z,t)$, we have to investigate the structure of the dominant singularities. In particular, it can be shown that the radius of convergence of $D(z,1)$ (as well as for the corresponding partial derivatives with respect to t) is given by $\rho = 1/S(\tau)$ where $\tau > 0$ is the structural constant with respect to $\mathcal{S}$.

Thus, in the general case of $\tau \neq 1$, the singularities of $D(z,1)$ are of the same type as the singularities of $V(z)$. Therefore, the precise description of the singular structure of $V(z)$ given in Proposition 6 allows us to carry out the asymptotic analysis.

Recall that $D_{n,r}$ is the random variable modeling the number of r-ascents in a random dispersed Łukasiewicz excursion of length n with respect to some step set $\mathcal{S}$.

▶ **Theorem 10.** *Let $p \geq 1$ be the period of the step set $\mathcal{S}$. Assume additionally that for the structural constant τ we have $\tau \neq 1$.*

Then d_n, the number of dispersed Łukasiewicz excursions of length n, satisfies

$$d_n = \frac{1}{\sqrt{2\pi}} \frac{p\tau^k(\tau^p(p-k-1)+k+1)}{(1-\tau^p)^2} \sqrt{\frac{S(\tau)^3}{S''(\tau)}} S(\tau)^n n^{-3/2} + O(S(\tau)^n n^{-5/2}) \tag{10}$$

for $n \equiv k \mod p$ and $0 \leq k \leq p-1$. Furthermore, the expected number of r-ascents grows with $n \to \infty$ according to the asymptotic expansion

$$\mathbb{E}D_{n,r} = \frac{(\tau S(\tau) - 1)^r}{(\tau S(\tau))^{r+2}} n + O(1). \tag{11}$$

The O-constants depend implicitly on both r as well as on the residue class of n modulo p.

In a nutshell, the proof of this theorem involves a rigorous analysis of the generating functions $D(z,1)$ (for the overall number of dispersed excursions), as well as of $D_t(z,1) = \frac{1}{z}\frac{V_t(z)}{(1-V(z))^2}$ (for the expected number of ascents in these paths). Furthermore, while our results as stated in (10) and (11) only list the asymptotic main term, expansions with higher precision are available in the worksheet as well (they just become rather messy very quickly).

It can be shown that the only family of Łukasiewicz paths that is not covered by Theorem 10 is $\mathcal{S} = \{-1, 1\}$, the case of dispersed Dyck paths. However, as everything is explicitly given, the analysis is quite straightforward.

▶ **Proposition 11.** *Let d_n denote the total number of dispersed Dyck paths of length n, and let $D_{n,r}$ denote the random variable modeling the number of r-ascents in a random dispersed Dyck path of length n.*

Then, d_n is given by

$$d_n = \binom{n}{\lfloor n/2 \rfloor} = \sqrt{\frac{2}{\pi}} 2^n n^{-1/2} - \frac{2 - (-1)^n}{2\sqrt{2\pi}} 2^n n^{-3/2} + O(2^n n^{-5/2}), \tag{12}$$

and the expected number of r-ascents satisfies

$$\mathbb{E}D_{n,r} = \frac{n}{2^{r+2}} - \sqrt{\frac{\pi}{2}} \frac{r-2}{2^{r+2}} n^{1/2} + \frac{(r-1)(r-4)}{2^{r+3}}$$
$$- \sqrt{\frac{\pi}{2}} \frac{(r-2)(2-(-1)^n)}{2^{r+4}} n^{-1/2} + O(n^{-1}). \tag{13}$$

This completes our analysis of r-ascents in dispersed Łukasiewicz excursions.

3.3 Analysis of Meanders

In this section we study ascents in meanders, i.e., non-negative Łukasiewicz paths without further restriction. The corresponding generating function can be obtained from (4) by setting $v = 1$, which allows arbitrary ending altitude of the path.

In accordance to the results from [1, Theorem 4], the behavior of meanders depends on the sign of the drift (i.e., the quantity $S'(1)$). The following theorem handles the case of positive drift (which, in our setting, is equivalent to $\tau \neq 1$).

Recall that $M_{n,r}$ is the random variable modeling the number of r-ascents in a random non-negative Łukasiewicz path of length n with respect to some given step set $\mathcal{S}$.

▶ **Theorem 12.** *Let $\tau > 0$ be the structural constant, i.e., the unique positive solution of $S'(\tau) = 0$, and assume that $\tau \neq 1$.*

Then, with $\xi = 1/S(1)$, the expected number of r-ascents in Łukasiewicz meanders of length n as well as the corresponding variance grow with $n \to \infty$ according to the asymptotic expansions

$$\mathbb{E}M_{n,r} = \mu n + \frac{(S(1)-1)^r (2 S(1) - 1 - r)}{S(1)^{r+2}} + \frac{(S(1)-1)^r V_z(\xi)}{S(1)^{r+1}(1 - V(\xi))} - \frac{V_t(\xi)}{1 - V(\xi)}$$
$$+ O\left(n^{5/2} \left(\frac{S(\tau)}{S(1)}\right)^n\right), \tag{14}$$

and

$$\mathbb{V}M_{n,r} = \sigma^2 n + O(1), \tag{15}$$

where μ and σ^2 are given by

$$\mu = \frac{(S(1)-1)^r}{S(1)^{r+2}} \quad \text{and} \quad \sigma^2 = \frac{(S(1)-1)^r}{S(1)^{r+2}} + \frac{(S(1)-1)^{2r}(2 S(1) - 3 - 2r)}{S(1)^{2r+4}}.$$

Moreover, for $n \to \infty$, $M_{n,r}$ is asymptotically normally distributed. All O-constants depend implicitly on r.

It can be shown that Theorem 12 covers all step sets except for $\mathcal{S} = \{-1, 1\}$ and $\mathcal{S} = \{-1, 0, 1\}$. In these cases, we have a similar situation to what we had in Section 3.2: the square root singularity coming from $V(z)$ combines with the zero in the denominator.

The following propositions close this gap.

▶ **Proposition 13.** *The expected number of r-ascents in the Łukasiewicz meanders of length n associated to $\mathcal{S} = \{-1, 1\}$ as well as the corresponding variance grow with $n \to \infty$ according to the asymptotic expansions*

$$\mathbb{E}M_{n,r} = \frac{n}{2^{r+2}} + \frac{\sqrt{2\pi}(r-2)}{2^{r+3}}n^{1/2} - \frac{r^2 - r - 8}{2^{r+3}} + \frac{\sqrt{2\pi}((2 - (-1)^n)(r-2)}{2^{r+5}}n^{-1/2} + O(n^{-1}), \quad (16)$$

and

$$\mathbb{V}M_{n,r} = \frac{2^{r+3} - r^2(\pi - 2) + 4r(\pi - 3) - 4\pi + 10}{2^{2r+5}}n + O(n^{1/2}). \quad (17)$$

▶ **Proposition 14.** *The expected number of r-ascents in the Łukasiewicz meanders of length n associated to $\mathcal{S} = \{-1, 0, 1\}$ as well as the corresponding variance grow with $n \to \infty$ according to the asymptotic expansions*

$$\mathbb{E}M_{n,r} = \frac{2^r}{3^{r+2}}n + \frac{\sqrt{3\pi}(r-4)2^{r-2}}{3^{r+2}}n^{1/2} - (3r^2 - r - 96)\frac{2^{r-4}}{3^{r+2}} + \frac{\sqrt{3\pi}(r-4)2^{r-6}}{3^r}n^{-1/2} + O(n^{-1}) \quad (18)$$

and

$$\mathbb{V}M_{n,r} = \frac{3^{r+2}2^{r+4} - 2^{2r}(3r^2(\pi - 2) - 8r(3\pi - 10) + 48\pi - 144)}{16 \cdot 3^{2r+4}}n + O(n^{1/2}). \quad (19)$$

References

1   Cyril Banderier and Philippe Flajolet. Basic analytic combinatorics of directed lattice paths. *Theoret. Comput. Sci.*, 281(1–2):37–80, 2002. `doi:10.1016/S0304-3975(02)00007-5`.

2   David Bevan. Permutations avoiding 1324 and patterns in Łukasiewicz paths. *J. Lond. Math. Soc. (2)*, 92(1):105–122, 2015. `doi:10.1112/jlms/jdv020`.

3   Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009. `doi:10.1017/CBO9780511801655`.

4   Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete mathematics. A foundation for computer science*. Addison-Wesley, second edition, 1994.

5   Benjamin Hackl, Clemens Heuberger, and Helmut Prodinger. Ascents in non-negative lattice paths. arXiv:1801.02996 [math.CO], 2018. URL: `https://arxiv.org/abs/1801.02996`.

6   Kairi Kangro, Mozhgan Pourmoradnasseri, and Dirk Oliver Theis. Short note on the number of 1-ascents in dispersed Dyck paths. arXiv:1603.01422 [math.CO], 2016. URL: `https://arxiv.org/abs/1603.01422`.

7   M. Lothaire. *Combinatorics on words*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1997. With a foreword by Roger Lyndon and a preface by Dominique Perrin, Corrected reprint of the 1983 original, with a new preface by Perrin. `doi:10.1017/CBO9780511566097`.

8   Amram Meir and John W. Moon. On an asymptotic method in enumeration. *J. Combin. Theory Ser. A*, 51(1):77–89, 1989. `doi:10.1016/0097-3165(89)90078-2`.

9   Helmut Prodinger. Returns, hills, and t-ary trees. *J. Integer Seq.*, 19(7):Article 16.7.2, 8, 2016. URL: `http://emis.ams.org/journals/JIS/VOL19/Prodinger/prod41.html`.

10   The SageMath Developers. *SageMath Mathematics Software (Version 8.1)*, 2017. URL: `http://www.sagemath.org`.

A Singularity Analysis of Inverse Functions

The aim of this appendix is, on the one hand, to state and prove an extension of [3, Remark VI.17]. In fact, we simply confirm what is announced in the footnote in [3, p. 405] and give more details. Then, we use these results in order to derive relevant information on the generating function $V(z, t)$ from before.

For the following two propositions, we borrow the notation used in [3, Chapter VI.7].

▶ **Proposition 15.** *Let $\phi(u)$ be analytic with radius of convergence $0 < R \leq \infty$, $\phi(0) \neq 0$, $[u^n]\,\phi(u) \geq 0$ for all $n \geq 0$ and $\phi(u)$ not affine linear. Assume that there is a positive $\tau \in (0, R)$ such that $\tau\,\phi'(\tau) = \phi(\tau)$. Finally assume that $\phi(u)$ is a p-periodic power series for some maximal p. Denote the set of all pth roots of unity by $G(p)$.*

Then there is a unique function $y(z)$ satisfying $y(z) = z\,\phi(y(z))$ which is analytic in a neighborhood of 0 with $y(0) = 0$. It has radius of convergence $\rho = \tau/\phi(\tau)$ around the origin. For $|z| \leq \rho$, it has exactly singularities at $z = \rho\zeta$ for $\zeta \in G(p)$. For $z \to \rho$, we have the singular expansion

$$y(z) \overset{z \to \rho}{\sim} \sum_{j \geq 0} (-1)^j d_j \left(1 - \frac{z}{\rho}\right)^{j/2}$$

for some computable constants d_j, $j \geq 0$. We have $d_0 = \tau$ and $d_1 = \sqrt{2\,\phi(\tau)/\phi''(\tau)}$. Additionally, we have $[z^n]y(z) = 0$ for $n \not\equiv 1 \pmod{p}$.

Proof. Existence, uniqueness, radius of convergence as well as singular expansion around $z \to \rho$ of $y(z)$ are shown in [3, Theorem VI.6].

As ϕ is a p-periodic power series and $\phi(0) \neq 0$, there exists an aperiodic function χ such that $\phi(u) = \chi(u^p)$. From the non-negativity of the coefficients of $\phi(u)$, it is clear that $\chi(u)$ has non-negative coefficients and is analytic for $|u| < R^p$. We consider $\psi(u) := \chi(u)^p$. Then ψ is again analytic for $|u| < R^p$, it has clearly non-negative coefficients, $\psi(0) \neq 0$ and $\psi(u)$ is not an affine linear function. If $[u^m]\chi(u) > 0$ and $[u^n]\chi(u) > 0$ for some $m < n$, then $[u^{pm}]\psi(u) > 0$ as well as $[u^{pm+(n-m)}]\psi(u) > 0$, which implies that ψ is aperiodic.

Finally, we have

$$\tau^p\psi'(\tau^p) = p\tau^p\chi(\tau^p)^{p-1}\chi'(\tau^p) = \tau\,\phi(\tau)^{p-1}\,\phi'(\tau) = \phi(\tau)^p = \chi(\tau^p)^p = \psi(\tau^p).$$

Considering the functional equation $Y(Z) = Z\psi(Y(Z))$, we see that all assumptions of [3, Theorem VI.6] are satisfied; thus it has a unique solution $Y(Z)$ with $Y(0) = 0$ which is analytic around the origin. By the same result, $Y(Z)$ has radius of convergence

$$\frac{\tau^p}{\psi(\tau^p)} = \frac{\tau^p}{\chi(\tau^p)^p} = \left(\frac{\tau}{\phi(\tau)}\right)^p = \rho^p$$

and, as ψ is aperiodic, the only singularity of $Y(Z)$ with $|Z| \leq \rho^p$ is $Z = \rho^p$.

We consider the function $\widetilde{y}(z) := z\chi(Y(z^p))$. By definition, it is analytic for $|z| < \rho$ and its only singularities with $|z| \leq \rho$ are those z with $z^p = \rho^p$, i.e., $z = \rho\zeta$ for $\zeta \in G(p)$. It is also clear by definition that $[z^n]\widetilde{y}(z) = 0$ for $n \not\equiv 1 \pmod{p}$. We have $\widetilde{y}(0) = 0$ and

$$z\,\phi(\widetilde{y}(z)) = z\chi((\widetilde{y}(z))^p) = z\chi(z^p\chi(Y(z^p))^p) = z\chi(z^p\psi(Y(z^p))) = z\chi(Y(z^p)) = \widetilde{y}(z)$$

because $z^p\psi(Y(z^p)) = Y(z^p)$ by definition of Y. This implies that $y = \widetilde{y}$. ◀

While the following proposition is particularly useful in the context of the previous one, it also holds in a slightly more general setting. It gives a detailed description of the singular expansions for p-periodic power series like above.

▶ **Proposition 16.** *Let p be a positive integer and let y be analytic with radius of convergence $0 < \rho \leq \infty$, where $[z^n]y(z) = 0$ for $n \not\equiv 1 \pmod{p}$. Assume that $y(z)$ has p dominant singularites located at $\zeta\rho$ for $\zeta \in G(p)$, and that for some $L \geq 0$ and $z \to \rho$, we have the singular expansion*

$$y(z) \overset{z \to \rho}{=} \sum_{j=0}^{L-1} d_j \left(1 - \frac{z}{\rho}\right)^{-\alpha_j} + O\left(\left(1 - \frac{z}{\rho}\right)^{-\alpha_L}\right),$$

where α_0, α_1, ..., α_L are complex numbers such that $\mathrm{Re}(\alpha_j) \geq \mathrm{Re}(\alpha_{j+1})$ for all $0 \leq j < L$. Then, for $\zeta \in G(p)$, the singular expansion of $y(z)$ for $z \to \zeta\rho$ is given by

$$y(z) \overset{z \to \zeta\rho}{=} \sum_{j=0}^{L-1} \zeta d_j \left(1 - \frac{z}{\zeta\rho}\right)^{-\alpha_j} + O\left(\left(1 - \frac{z}{\zeta\rho}\right)^{-\alpha_L}\right),$$

i.e., the expansion for $z \to \zeta\rho$ can be obtained by multiplying the expansion for $z \to \rho$ with ζ and substituting $z \mapsto \zeta/\rho$. Finally, for the coefficients of $y(z)$ we find

$$[z^n]y(z) = [\![p \mid 1 - n]\!][z^n]\left(p \sum_{j=0}^{L-1} d_j \left(1 - \frac{z}{\rho}\right)^{-\alpha_j} + O\left(\left(1 - \frac{z}{\rho}\right)^{-\alpha_L}\right)\right), \tag{20}$$

which can be made explicit easily by means of singularity analysis (cf. [3, Chapter VI.4]). In particular,

$$[z^n]y(z) = [\![p \mid 1 - n]\!]\left(\sum_{j=0}^{L-1} \frac{pd_j}{\Gamma(\alpha_j)} n^{\alpha_j - 1}\rho^{-n} + O\left(n^{\mathrm{Re}(\alpha_0)-2}\rho^{-n}\right) + O\left(n^{\mathrm{Re}(\alpha_L)-1}\rho^{-n}\right)\right).$$

Proof. As $[z^n]y(z) = 0$ for $n \not\equiv 1 \pmod{p}$ there is a function χ, analytic around the origin, such that $y(z) = z\chi(z^p)$. Thus, for every $\zeta \in G(p)$, we have

$$y(\zeta z) = \zeta z \chi((\zeta z)^p) = \zeta z \chi(z^p) = \zeta y(z)$$

or, equivalently,

$$y(z) = \zeta y\left(\frac{z}{\zeta}\right).$$

Thus the singular expansion for $z \to \zeta\rho$ follows from that for $z \to \rho$ by replacing z with z/ζ and multiplication by ζ.

With the singular expansions at all the dominant singularities located at $\zeta\rho$ for $\zeta \in G(p)$ at hand, we are able to extract the overall growth of the coefficients of $y(z)$ by first applying singularity analysis to every expansion separately, and then summing up all these contributions. When doing so, we use the well-known property of roots of unity that

$$\sum_{\zeta \in G(p)} \zeta^m = p[\![p \mid m]\!] \tag{21}$$

for $m \in \mathbb{Z}$ in order to rewrite the occurring sums as $\sum_{\zeta \in G(p)} \zeta^{1-n} = p[\![p \mid 1 - n]\!]$. Comparing the resulting asymptotic expansion with (20) proves the statement. ◀

Analysis of Summatory Functions of Regular Sequences: Transducer and Pascal's Rhombus

Clemens Heuberger

Institut für Mathematik, Alpen-Adria-Universität Klagenfurt, Universitätsstraße 65–67, 9020 Klagenfurt am Wörthersee, Austria
clemens.heuberger@aau.at
https://orcid.org/0000-0003-0082-7334

Daniel Krenn

Institut für Mathematik, Alpen-Adria-Universität Klagenfurt, Universitätsstraße 65–67, 9020 Klagenfurt am Wörthersee, Austria
math@danielkrenn.at
https://orcid.org/0000-0001-8076-8535

Helmut Prodinger

Department of Mathematical Sciences, Stellenbosch University, 7602 Stellenbosch, South Africa
hproding@sun.ac.za
https://orcid.org/0000-0002-0009-8015

— **Abstract** —

The summatory function of a q-regular sequence in the sense of Allouche and Shallit is analysed asymptotically. The result is a sum of periodic fluctuations multiplied by a scaling factor. Each summand corresponds to an eigenvalues of absolute value larger than the joint spectral radius of the matrices of a linear representation of the sequence. The Fourier coefficients of the fluctuations are expressed in terms of residues of the corresponding Dirichlet generating function. A known pseudo Tauberian argument is extended in order to overcome convergence problems in Mellin–Perron summation.

Two examples are discussed in more detail: The case of sequences defined as the sum of outputs written by a transducer when reading a q-ary expansion of the input and the number of odd entries in the rows of Pascal's rhombus.

29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018).
Editors: James Allen Fill and Mark Daniel Ward; Article No. 27; pp. 27:1–27:18
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1   Introduction

In this paper, we study the asymptotic behaviour of the summatory function of q-regular sequences.[1] Regular sequences have been introduced by Allouche and Shallit [2] (see also [3, Chapter 16]); these are sequences which are intimately related to the q-ary expansion of their arguments. Many special cases have been investigated in the literature; this is also due to their relation to divide-and-conquer algorithms. Our goal is to provide a single result decomposing the summatory function into periodic fluctuations multiplied by some scaling functions and to provide the Fourier coefficients of these periodic fluctuations.

Note that it is well-known that the summatory function of a q-regular sequence is itself q-regular. (This is an immediate consequence of [2, Theorem 3.1].) Similarly, the sequence of differences of a q-regular sequence is q-regular. Therefore, we might also start to analyse a regular sequence by considering it to be the summatory function of its sequence of differences. However, when modelling a quantity by a regular sequences, its asymptotic behaviour is often not smooth, but the asymptotic behaviour of its summatory functions is. Moreover, we will see throughout this work that from a technical perspective, considering partial sums is appropriate. Therefore, we adopt this point of view of summatory functions of q-regular sequences in this paper. This also enlightens us about the expectation of a random element of the sequence (with respect to uniform distribution on the non-negative integers smaller than a certain N).

In the remaining paper, we first recall the definition of q-regular sequences in Section 1.1, then formulate a somewhat simplified version of our main result in Section 1.2. In Section 1.3, we give a heuristic non-rigorous argument to explain why the result is expected. We outline the relation to previous work in Section 1.4. We give two examples in Sections 2 and 3. In principle, these examples are straight-forward applications of the results, but still, we have to reformulate the relevant questions in terms of a q-regular sequence and will then provide shortcuts for the computation of the Fourier series. The first example is generic and deals with sequences defined as the sum of outputs of transducer automata; the second example—which motivated us to conduct this study at this point—is a concrete problem counting the number of odd entries in Pascal's rhombus.

The full formulation of our results is given in the appendix; their proofs are given in the appendix of the arXiv version [17] of this extended abstract.

1.1 q-Regular Sequences

We start by giving a definition of q-regular sequences, see Allouche and Shallit [2]. Let $q \geq 2$ be a fixed integer and $(x(n))_{n \geq 0}$ be a sequence.

Then $(x(n))_{n \geq 0}$ is said to be $(\mathbb{C}, q)$-*regular* (briefly: q-*regular* or simply *regular*) if the $\mathbb{C}$-vector space generated by its q-*kernel*

$$\left\{ \left(x(q^j n + r) \right)_{n \geq 0} : \text{integers } j \geq 0, \, 0 \leq r < q^j \right\}$$

has finite dimension. In other words, $(x(n))_{n \geq 0}$ is q-regular if there is an integer D and sequences $(x_1(n))_{n \geq 0}, \ldots, (x_D(n))_{n \geq 0}$ such that for every $j \geq 0$ and $0 \leq r < q^j$ there exist integers $c_1, \ldots, c_D$ such that

$$x(q^j n + r) = c_1 x_1(n) + \cdots + c_D x_D(n) \qquad \text{for all } n \geq 0.$$

---

[1]   In the standard literature [2, 3] these sequences are called k-regular sequences (instead of q-regular sequences).

By Allouche and Shallit [2, Theorem 2.2], $(x(n))_{n\geq 0}$ is q-regular if and only if there exists a vector valued sequence $(v(n))_{n\geq 0}$ whose first component coincides with $(x(n))_{n\geq 0}$ and there exist square matrices $A_0, \ldots, A_{q-1} \in \mathbb{C}^{d\times d}$ such that

$$v(qn + r) = A_r v(n) \qquad \text{for } 0 \leq r < q,\, n \geq 0. \tag{1.1}$$

This is called a *q-linear representation* of $x(n)$.

The best-known example for a 2-regular function is the binary sum-of-digits function.

▶ **Example 1.** For $n \geq 0$, let $x(n) = s(n)$ be the binary sum-of-digits function. We clearly have

$$\begin{aligned} x(2n) &= x(n), \\ x(2n + 1) &= x(n) + 1 \end{aligned} \tag{1.2}$$

for $n \geq 0$.

Indeed, we have

$$x(2^j n + r) = x(n) + x(r) \cdot 1$$

for integers $j \geq 0$, $0 \leq r < 2^j$ and $n \geq 0$; i.e., the complex vector space generated by the 2-kernel is generated by $(x(n))_{n\geq 0}$ and the constant sequence $(1)_{n\geq 0}$.

Alternatively, we set $v(n) = (x(n), 1)^\top$ and have

$$v(2n) = \begin{pmatrix} x(n) \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} v(n),$$

$$v(2n + 1) = \begin{pmatrix} x(n) + 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} v(n)$$

for $n \geq 0$. Thus (1.1) holds with

$$A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

We defer the discussion of other examples, both generic such as sequences defined by transducer automata as well as a specific example involving the number of odd entries in Pascal's rhombus to Sections 2 and 3.

At this point, we note that a linear representation (1.1) immediately leads to an explicit expression for $x(n)$ by induction.

▶ Remark. Let $r_{\ell-1} \ldots r_0$ be the q-ary digit expansion[2] of n. Then

$$x(n) = e_1 A_{r_0} \cdots A_{r_{\ell-1}} v(0)$$

where $e_1 = \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix}$.

---

[2] Whenever we write that $r_{\ell-1} \ldots r_0$ is the q-ary digit expansion of n, we mean that $r_j \in \{0, \ldots, q-1\}$ for $0 \leq j < \ell$, $r_{\ell-1} \neq 0$ and $n = \sum_{j=0}^{\ell-1} r_j q^j$. In particular, the q-ary expansion of zero is the empty word.

1.2   Main Result

We are interested in the asymptotic behaviour of the summary function $X(N) = \sum_{0 \le n < N} x(n)$.

At this point, we give a simplified version of our results. We choose any vector norm $\|\cdot\|$ on $\mathbb{C}^d$ and its induced matrix norm. We set $C := \sum_{r=0}^{q-1} A_r$. We choose $R > 0$ such that $\|A_{r_1} \cdots A_{r_\ell}\| = O(R^\ell)$ holds for all $\ell \ge 0$ and $0 \le r_1, \dots, r_\ell < q$. In other words, R is an upper bound for the joint spectral radius of $A_1, \dots, A_{q-1}$. The spectrum of C, i.e., the set of eigenvalues of C, is denoted by $\sigma(C)$. For $\lambda \in \mathbb{C}$, let $m(\lambda)$ denote the size of the largest Jordan block of C associated with λ; in particular, $m(\lambda) = 0$ if $\lambda \notin \sigma(C)$. Finally, we consider the Dirichlet series[3]

$$\mathcal{X}(s) = \sum_{n \ge 1} n^{-s} x(n), \qquad \mathcal{V}(s) = \sum_{n \ge 1} n^{-s} v(n)$$

where $v(n)$ is the vector valued sequence defined in (1.1). Of course, $\mathcal{X}(s)$ is the first component of $\mathcal{V}(s)$. The principal value of the complex logarithm is denoted by $\log$. The fractional part of a real number z is denoted by $\{z\} := z - \lfloor z \rfloor$.

▶ **Theorem 2.** *With the notations above, we have*

$$X(N) = \sum_{\substack{\lambda \in \sigma(C) \\ |\lambda| > R}} N^{\log_q \lambda} \sum_{0 \le k < m(\lambda)} (\log_q N)^k \Phi_{\lambda k}(\{\log_q N\})$$

$$+ O\big(N^{\log_q R}(\log N)^{\max\{m(\lambda) \colon |\lambda| = R\}}\big) \quad (1.3)$$

for suitable 1-periodic continuous functions $\Phi_{\lambda k}$. If there are no eigenvalues $\lambda \in \sigma(C)$ with $|\lambda| \le R$, the O-term can be omitted.

For $|\lambda| > R$ and $0 \le k < m(\lambda)$, the function $\Phi_{\lambda k}$ is Hölder continuous with any exponent smaller than $\log_q(|\lambda|/R)$.

The Dirichlet series $\mathcal{V}(s)$ converges absolutely and uniformly on compact subsets of the half plane $\Re s > \log_q R + 1$ and can be continued to a meromorphic function on the half plane $\Re s > \log_q R$. It satisfies the functional equation

$$(I - q^{-s}C)\mathcal{V}(s) = \sum_{n=1}^{q-1} n^{-s} v(n) + q^{-s} \sum_{r=0}^{q-1} A_r \sum_{k \ge 1} \binom{-s}{k} \left(\frac{r}{q}\right)^k \mathcal{V}(s+k) \quad (1.4)$$

for $\Re s > \log_q R$. The right side converges absolutely and uniformly on compact subsets of $\Re s > \log_q R$. In particular, $\mathcal{V}(s)$ can only have poles where $q^s \in \sigma(C)$.

For $\lambda \in \sigma(C)$ with $|\lambda| > \max\{R, 1/q\}$, the Fourier series

$$\Phi_{\lambda k}(u) = \sum_{\ell \in \mathbb{Z}} \varphi_{\lambda k \ell} \exp(2\ell \pi i u)$$

converges pointwise for $u \in \mathbb{R}$ where

$$\varphi_{\lambda k \ell} = \frac{(\log q)^k}{k!} \operatorname{Res}\left(\frac{\big(x(0) + \mathcal{X}(s)\big)\big(s - \log_q \lambda - \frac{2\ell \pi i}{\log q}\big)^k}{s}, s = \log_q \lambda + \frac{2\ell \pi i}{\log q}\right) \quad (1.5)$$

for $\ell \in \mathbb{Z}$, $0 \le k < m(\lambda)$.

---

[3]   Note that the summary function $X(N)$ contains the summand $x(0)$ but the Dirichlet series cannot. This is because the choice of including $x(0)$ into $X(N)$ will lead to more consistent results.

This theorem is proved in the arXiv version [17, Appendix G] of this extended abstract. Note that we write $\Phi_{\lambda k}(\{\log_q N\})$ to optically emphasise the 1-periodicity; technically, we have $\Phi_{\lambda k}(\{\log_q N\}) = \Phi_{\lambda k}(\log_q N)$. Note that the arguments in the proof could be used to meromophically continue the Dirichlet series to the complex plane, but we do not need this result for our purposes. See [1] for the corresponding argument for automatic sequences.

We come back to the binary sum of digits.

▶ **Example 3** (Continuation of Example 1). We have $C = A_0 + A_1 = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right)$. As A_0 is the identity matrix, any product $A_{r_1} \cdots A_{r_\ell}$ has the shape $A_1^k = \left(\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right)$ where k is the number of factors A_1 in the product. This implies that R with $\|A_{r_1} \cdots A_{r_\ell}\| = O(R^\ell)$ may be chosen to be any number greater than 1. As C is a Jordan block itself, we simply read off that the only eigenvalue of C is $\lambda = 2$ with $m(2) = 2$.

Thus Theorem 2 yields

$$X(N) = N(\log_2 N)\,\Phi_{21}(\{\log_2 N\}) + N\,\Phi_{20}(\{\log_2 N\})$$

for suitable 1-periodic continuous functions Φ_{21} and Φ_{20}.

In principle, we can now use the functional equation (1.4). Due to the fact that one component of v is the constant sequence where everything is known, it is more efficient to use an ad-hoc calculation for $\mathcal{X}$ by splitting the sum according to the parity of the index and using the recurrence relation (1.2) for $x(n)$. We obtain

$$\begin{aligned}\mathcal{X}(s) &= \sum_{n \geq 1} \frac{x(2n)}{(2n)^s} + \sum_{n \geq 0} \frac{x(2n+1)}{(2n+1)^s} \\ &= 2^{-s} \sum_{n \geq 1} \frac{x(n)}{n^s} + \sum_{n \geq 0} \frac{x(n)}{(2n+1)^s} + \sum_{n \geq 0} \frac{1}{(2n+1)^s} \\ &= 2^{-s}\mathcal{X}(s) + \frac{x(0)}{1^s} + \sum_{n \geq 1} \frac{x(n)}{(2n)^s} + \sum_{n \geq 1} x(n)\Big(\frac{1}{(2n+1)^s} - \frac{1}{(2n)^s}\Big) \\ &\quad + 2^{-s}\sum_{n \geq 0} \frac{1}{\left(n+\frac{1}{2}\right)^s} \\ &= 2^{1-s}\mathcal{X}(s) + 2^{-s}\,\zeta\left(s, \tfrac{1}{2}\right) + \sum_{n \geq 1} x(n)\Big(\frac{1}{(2n+1)^s} - \frac{1}{(2n)^s}\Big)\end{aligned}$$

where the Hurwitz zeta function $\zeta(s, \alpha) := \sum_{n+\alpha>0}(n+\alpha)^{-s}$ has been used. We get

$$(1 - 2^{1-s})\mathcal{X}(s) = 2^{-s}\,\zeta\left(s, \tfrac{1}{2}\right) + \sum_{n \geq 1} x(n)\Big(\frac{1}{(2n+1)^s} - \frac{1}{(2n)^s}\Big). \tag{1.6}$$

As the sum of digits is bounded by the length of the expansion, we have $x(n) = O(\log n)$. By combining this estimate with

$$(2n+1)^{-s} - (2n)^{-s} = \int_{2n}^{2n+1} \Big(\frac{\mathrm{d}}{\mathrm{d}t} t^{-s}\Big)\,\mathrm{d}t = \int_{2n}^{2n+1} (-s)t^{-s-1}\,\mathrm{d}t = O(|s|n^{-\Re s-1}),$$

we see that the sum in (1.6) converges absolutely for $\Re s > 0$ and is therefore analytic for $\Re s > 0$.

Therefore, the right side of (1.6) is a meromorphic function for $\Re s > 0$ whose only pole is simple and at $s = 1$ which originates from $\zeta\left(s, \tfrac{1}{2}\right)$. Therefore, $\mathcal{X}(s)$ is a meromorphic function for $\Re s > 0$ with a double pole at $s = 1$ and simple poles at $1 + \frac{2\ell\pi i}{\log 2}$ for $\ell \in \mathbb{Z} \setminus \{0\}$.

Thus

$$
\begin{aligned}
\Phi_{21}(u) = \varphi_{210} &= (\log 2) \operatorname{Res}\Big(\frac{\mathcal{X}(s)(s-1)}{s}, s = 1\Big) \\
&= (\log 2) \operatorname{Res}\Big(\frac{2^{-s}(s-1)}{1 - 2^{1-s}}\,\zeta\big(s, \tfrac{1}{2}\big), s = 1\Big) = \frac{1}{2}
\end{aligned}
\tag{1.7}
$$

by (1.5) and (1.6).

We conclude that

$$
X(N) = \frac{1}{2} N \log_2 N + N\,\Phi_{20}(\{\log_2 N\}).
$$

We refrain from computing the Fourier coefficients of $\Phi_{20}(u)$ explicitly at this point; numerically, they could be computed from (1.6). However, an explicit expression can be obtained by rewriting the residues of $\mathcal{X}(s)$ in terms of shifted residues of $\sum_{n\geq 1}\big(x(n) - x(n-1)\big)n^{-s}$ and computing the latter explicitly; see [18, Proof of Corollary 2.5]. This yields the well-known result by Delange [6].

It will also turn out that (1.7) being a constant function is an immediate consequence of the fact that $\begin{pmatrix} 0 & 1 \end{pmatrix}$ is a left eigenvector of both A_0 and A_1 associated with the eigenvalue 1.

1.3   Heuristic Approach: Mellin–Perron Summation

The purpose of this section is to explain why the formula (1.5) for the Fourier coefficients is expected. The approach here is heuristic and non-rigorous because we do not have the required growth estimates. See also [7].

By the Mellin–Perron summation formula of order 0 (see, for example, [12, Theorem 2.1]), we have

$$
\sum_{1 \leq n < N} x(n) + \frac{x(N)}{2} = \frac{1}{2\pi i} \int_{\max\{\log_q R+2,1\}-i\infty}^{\max\{\log_q R+2,1\}+i\infty} \mathcal{X}(s)\,\frac{N^s \, \mathrm{d}s}{s}.
$$

By Remark 1.1 and the definition of R, we have $x(N) = O(R^{\log_q N}) = O(N^{\log_q R})$. Adding the summand $x(0)$ to match our definition of $X(N)$ amounts to adding $O(1)$. Shifting the line of integration to the left—we have *no analytic justification* that this is allowed—and using the location of the poles of $\mathcal{X}(s)$ claimed in Theorem 2 yield

$$
X(N) = \sum_{\substack{\lambda \in \sigma(C) \\ |\lambda| > R}} \sum_{\ell \in \mathbb{Z}} \operatorname{Res}\Big(\frac{\mathcal{X}(s)N^s}{s}, s = \log_q \lambda + \frac{2\ell\pi i}{\log q}\Big)
$$

$$
+ \frac{1}{2\pi i} \int_{\log_q R+\varepsilon-i\infty}^{\log_q R+\varepsilon+i\infty} \mathcal{X}(s)\,\frac{N^s \, \mathrm{d}s}{s} + O(N^{\log_q R} + 1)
$$

for some $\varepsilon > 0$. Expanding N^s as

$$
N^s = \sum_{k \geq 0} \frac{(\log N)^k}{k!} N^{\log_q \lambda + \frac{2\ell\pi i}{\log q}} \Big(s - \log_q \lambda - \frac{2\ell\pi i}{\log q}\Big)^k
$$

and assuming that the remainder integral converges absolutely yields

$$
X(N) = \sum_{\substack{\lambda \in \sigma(C) \\ |\lambda| > R}} N^{\log_q \lambda} \sum_{0 \leq k < m_{\lambda\ell}} (\log_q N)^k \sum_{\ell \in \mathbb{Z}} \varphi_{\lambda k \ell} \exp\big(2\ell\pi i \log_q N\big)
$$

$$
+ O(N^{\log_q R+\varepsilon} + 1)
$$

where $m_{\lambda\ell}$ denotes the order of the pole of $\mathcal{X}(s)/s$ at $\log_q \lambda + \frac{2\ell\pi i}{\log q}$ and $\varphi_{\lambda k\ell}$ is as in (1.5).

Summarising, this heuristic approach explains most of the formulæ in Theorem 2. Some details (exact error term and order of the poles) are not explained by this approach. A result "repairing" the zeroth order Mellin–Perron formula is known as Landau's theorem, see [4, § 9]. It is not applicable to our situation due to multiple poles along vertical lines which then yield the periodic fluctuations. Instead, we prove a theorem which provides the required justification (not by estimating the relevant quantities, but by reducing the problem to higher order Mellin–Perron summation). The essential assumption is that the summatory function can be decomposed into fluctuations multiplied by some growth factors such as in (1.3).

1.4 Relation to Previous Work

Sequences defined as the output sum of transducer automata in the sense of [18] are a special case of regular sequences; these are a generalisation of many previously studied concepts. In that case, much more is known (variance, limiting distribution, higher dimensional input). See [18] for references and results. A more detailed comparison can be found in Section 2. Divide and Conquer recurrences (see [19] and [8]) can also be seen as special cases of regular sequences.

The asymptotics of the summatory function of specific examples of regular sequences has been studied in [14], [15], [11].

Dumas [9, 10] finally proved the first part of Theorem 2. We re-prove it here in a self-contained way because we need more explicit results than obtained by Dumas (e.g., we need explicit expressions for the fluctuations) for proving Hölder continuity and to explicitly get the precise structure depending on the eigenspaces. Notice that Dumas' paper introduces linear representations as we do in (1.1), but then the order of factors is reversed in his equivalent of Remark 1.1, which means that some transpositions have to be silently introduced.

The first version of our pseudo-Tauberian argument was provided in [12]: there, no logarithmic factors were allowed and the growth conditions on the Dirichlet series were stronger.

2 Sequences Defined by Transducer Automata

Let $q \geq 2$ be a positive integer. We consider a complete deterministic subsequential transducer $\mathcal{T}$ with input alphabet $\{0, \ldots, q-1\}$ and output alphabet $\mathbb{C}$, see [5, Chapter 1] and [18]. Recall that a transducer is said to be *deterministic* and *complete* if for every state and every digit of the input alphabet, there is exactly one transition starting in this state with this input label. A *subsequential* transducer has a final output label for every state.

For a non-negative integer n, let $\mathcal{T}(n)$ be the sum of the output labels (including the final output label) encountered when the transducer reads the q-ary expansion of n. This concept has been thoroughly studied in [18]: there, $\mathcal{T}(n)$ is considered as a random variable defined on the probability space $\{0, \ldots, N-1\}$ equipped with uniform distribution. The expectation in this model corresponds (up to a factor of N) to our summatory function $\sum_{0 \leq n < N} \mathcal{T}(n)$. We remark that in [18], the variance and limiting distribution of the random variable $\mathcal{T}(n)$ have also been investigated. Most of the results there are also valid for higher dimensional input.

The purpose of this section is to show that $\mathcal{T}(n)$ is a q-regular sequence and to see that our results here coincide with the corresponding results in [18]. We note that the binary sum of digits considered in Example 1 is the special case of $q = 2$ and the transducer consisting of a single state which implements the identity map. For additional special cases of this

concept, see [18]. Note that our result here for the summatory function contains (fluctuating) terms for all eigenvalues λ of the adjacency matrix of the underlying digraph with $1 < |\lambda|$ whereas in [18] only contributions of those eigenvalues λ with $|\lambda| = q$ are available, all other contributions are absorbed by the error term there.

By a *component* of a digraph we always mean a strongly connected component. We call a component *final* if there are no arcs leaving the component. The *period* of a component is the greatest common divisor of its cycle lengths. The *final period* of a digraph is the least common multiple of the periods of its final components.

We consider the states of $\mathcal{T}$ to be numbered by $\{1, \ldots, d\}$ for some positive integer $d \geq 1$ such that the initial state is state 1. We set $\mathcal{T}_j(n)$ to be the sum of the output labels (including the final output label) encountered when the transducer reads the q-ary expansion of n when starting in state j. By construction, we have $\mathcal{T}(n) = \mathcal{T}_1(n)$ and $\mathcal{T}_j(0)$ is the final output label of state j. We set $y(n) = (\mathcal{T}_1(n), \ldots, \mathcal{T}_d(n))$. For $0 \leq r < q$, we define the $(d \times d)$-$\{0,1\}$-matrix P_r in such a way that there is a one in row j, column k if and only if there is a transition from state j to state k with input label r. The vector o_r is defined by setting its jth coordinate to be the output label of the transition from state j with input label r.

For $n_0 \geq 1$, we set

$$\mathcal{X}(s) = \sum_{n \geq 1} n^{-s} \mathcal{T}(n), \qquad \mathcal{Y}_{n_0}(s) = \sum_{n \geq n_0} n^{-s} y(n), \qquad \zeta_{n_0}(s, \alpha) = \sum_{n \geq n_0} (n + \alpha)^{-s}.$$

The last Dirichlet series is a truncated version of the Hurwitz zeta function.

▶ **Corollary 4.** *Let $\mathcal{T}$ be a transducer as described at the beginning of this section. Let M and p be the adjacency matrix and the final period of the underlying digraph, respectively. For $\lambda \in \mathbb{C}$ let $m(\lambda)$ be the size of the largest Jordan block associated with the eigenvalue λ of M.*

Then $(\mathcal{T}(n))_{n \geq 0}$ is a q-regular sequence and

$$\sum_{0 \leq n < N} \mathcal{T}(n) = e_{\mathcal{T}} N \log_q N + N \Phi(\log_q N)$$

$$+ \sum_{\substack{\lambda \in \sigma(M) \\ 1 < |\lambda| < q}} N^{\log_q \lambda} \sum_{0 \leq k < m(\lambda)} (\log_q N)^k \Phi_{\lambda k}(\log_q N) \qquad (2.1)$$

$$+ O\big((\log N)^{\max\{m(\lambda) : \, |\lambda| = 1\}}\big)$$

for some continuous p-periodic function Φ, some continuous 1-periodic functions $\Phi_{\lambda k}$ for $\lambda \in \sigma(M)$ with $1 < |\lambda| < q$ and $0 \leq k < m(\lambda)$ and some constant $e_{\mathcal{T}}$.

Furthermore,

$$\Phi(u) = \sum_{\ell \in \mathbb{Z}} \varphi_\ell \exp\Big(\frac{2\ell \pi i}{p} u\Big)$$

with

$$\varphi_\ell = \operatorname{Res}\Big(\frac{\mathcal{X}(s)}{s}, s = 1 + \frac{2\ell \pi i}{p \log q}\Big)$$

for $\ell \in \mathbb{Z}$. The Fourier series expansion of $\Phi_{\lambda k}$ for $\lambda \in \sigma(M)$ with $1 < |\lambda| < q$ is given in Theorem 2.

The Dirichlet series $\mathcal{Y}_{n_0}(s)$ satisfies the functional equation

$$(I - q^{-s}M)\mathcal{Y}_{n_0}(s) = \sum_{n_0 \le n < qn_0} n^{-s}y(n) + q^{-s}\sum_{0 \le r < q}\zeta_{n_0}\left(s, \tfrac{r}{q}\right)o_r$$
$$+ q^{-s}\sum_{0 \le r < q}P_r\sum_{k \ge 1}\binom{-s}{k}\left(\tfrac{r}{q}\right)^k\mathcal{Y}_{n_0}(s+k).$$

Sketch of the Proof. The proof is split into several steps.

Recursive Description. We set $v(n) = \big(\mathcal{T}_1(n), \ldots, \mathcal{T}_d(n), 1\big)^{\top}$. For $1 \le j \le d$ and $0 \le r < q$, we define $t(j, r)$ and $o(j, r)$ to be the target state and output label of the unique transition from state j with input label r, respectively. Therefore,

$$\mathcal{T}_j(qn + r) = \mathcal{T}_{t(j,r)}(n) + o(j, r) \tag{2.2}$$

for $1 \le j \le d$, $n \ge 0$, $0 \le r < q$ with $qn + r > 0$.

For $0 \le r < q$, define $A_r = (a_{rjk})_{1 \le j, k \le d+1}$ by

$$a_{rjk} = \begin{cases} [t(j, r) = k] & \text{if } j, k \le d, \\ o(j, r) & \text{if } j \le d, k = d+1, \\ [k = d+1] & \text{if } j = d+1 \end{cases}$$

where we use Iverson's convention $[expr] = 1$ if $expr$ is true and 0 otherwise; see Graham, Knuth, and Patashnik [16]. Then (2.2) is equivalent to

$$v(qn + r) = A_r v(n)$$

for $n \ge 0$, $0 \le r < q$ with $qn + r > 0$.

q-Regular Sequence. If we insist on a proper formulation as a regular sequence, we rewrite (2.2) to

$$\mathcal{T}_j(qn + r) = \mathcal{T}_{t(j,r)}(n) + o(j, r) + [r = 0][n = 0](\mathcal{T}_j(0) - \mathcal{T}_{t(j,0)}(0) - o(j, 0)) \tag{2.3}$$

for $1 \le j \le d$, $n \ge 0$, $0 \le r < q$. Setting $\widetilde{v}(n) = (\mathcal{T}_1(n), \ldots, \mathcal{T}_d(n), 1, [n = 0])$ and $\widetilde{A}_r = (\widetilde{a}_{rjk})_{1 \le j, k \le d+2}$ with

$$\widetilde{a}_{rjk} = \begin{cases} [t(j, r) = k] & \text{if } j, k \le d, \\ o(j, r) & \text{if } j \le d, k = d+1, \\ [r = 0](\mathcal{T}_j(0) - \mathcal{T}_{t(j,0)}(0) - o(j, 0)) & \text{if } j \le d, k = d+2, \\ [k = d+1] & \text{if } j = d+1, \\ [k = d+2][r = 0] & \text{if } j = d+2, \end{cases}$$

the system (2.3) is equivalent to

$$\widetilde{v}(qn + r) = \widetilde{A}_r\widetilde{v}(n)$$

for $n \ge 0$, $0 \le r < q$.

The rest of the proof (relating the eigenvalues of M with those of C) can be found in the arXiv version [17, Appendix H] of this extended abstract.   ◀
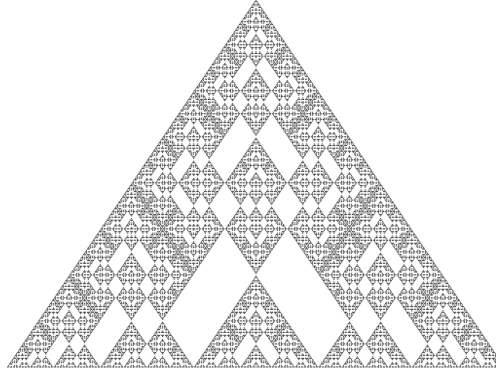
■ **Figure 3.1** Pascal's rhombus modulo 2.

3   Pascal's Rhombus

We consider Pascal's rhombus $\mathfrak{R}$ which is, for integers $i \geq 0$ and j, the array with entries $r_{i,j}$, where

- $r_{0,j} = 0$ all j,
- $r_{1,0} = 1$ and $r_{1,j} = 0$ for all $j \neq 0$,
- and

$$r_{i,j} = r_{i-1,j-1} + r_{i-1,j} + r_{i-1,j+1} + r_{i-2,j}$$

for $i \geq 1$.

Let $\mathfrak{X}$ be equal to $\mathfrak{R}$ but with entries takes modulo 2; see also Figure 3.1. We partition $\mathfrak{X}$ into the four sub-arrays

- $\mathfrak{E}$ consisting only of the rows and columns of $\mathfrak{X}$ with even indices, i.e., the entries $r_{2i,2j}$,
- $\mathfrak{Y}$ consisting only of the rows with odd indices and columns with even indices, i.e., the entries $r_{2i-1,2j}$,
- $\mathfrak{Z}$ consisting only of the rows with even indices and columns with odd indices, i.e., the entries $r_{2i,2j-1}$, and
- $\mathfrak{N}$ consisting only of the rows and columns with odd indices, i.e., the entries $r_{2i-1,2j-1}$.

Note that $\mathfrak{E} = \mathfrak{X}$ and $\mathfrak{N} = 0$; see [13].

3.1   Recurrence Relations and 2-Regular Sequences

Let $X(N)$, $Y(N)$ and $Z(N)$ be the number of ones in the first N rows (starting with row index 1) of $\mathfrak{X}$, $\mathfrak{Y}$ and $\mathfrak{Z}$ respectively.

Using results by Goldwasser, Klostermeyer, Mays and Trapp [13] leads to recurrence relations for the backward differences $x(n) = X(n) - X(n-1)$, $y(n) = Y(n) - Y(n-1)$ and $z(n) = Z(n) - Z(n-1)$, namely

$$
\begin{array}{lll}
x(2n) = x(n) + z(n), & x(2n+1) = y(n+1), & \text{(3.1a)} \\
y(2n) = x(n-1) + z(n), & y(2n+1) = x(n+1) + z(n), & \text{(3.1b)} \\
z(2n) = 2x(n), & z(2n+1) = 2y(n+1) & \text{(3.1c)}
\end{array}
$$

for $n \geq 1$, and $x(0) = y(0) = z(0) = 0$, $x(1) = 1$, $y(1) = 1$ and $z(1) = 2$. (See the arXiv version [17, Appendix I.1] of this extended abstract for details.) These $x(n)$, $y(n)$ and $z(n)$ are the number of ones in the nth row of $\mathfrak{X}$, $\mathfrak{Y}$ and $\mathfrak{Z}$ respectively.
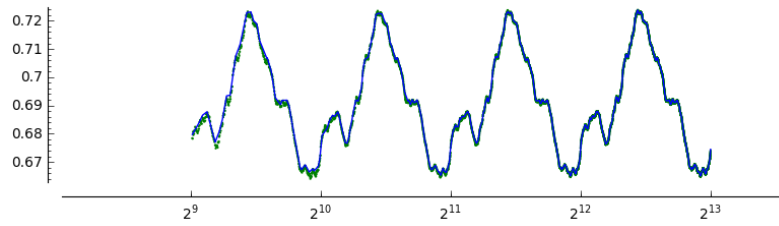
■ **Figure 3.2** Fluctuation in the main term of the asymptotic expansion of $X(N)$. The figure shows $\Phi(\log_2 N)$ (blue) approximated by its trigonometric polynomial of degree 99 as well as $X(N)/N^\kappa$ (green).

Let use write our coefficients as the vector

$$v(n) = \big(x(n), x(n+1), y(n+1), z(n), z(n+1)\big)^\top. \tag{3.2}$$

It turns out that the components included into $v(n)$ are sufficient for a self-contained linear representation of $v(n)$. In particular, it is not necessary to include $y(n)$. By using the recurrences (3.1), we find that

$$v(2n) = A_0 v(n) \qquad \text{and} \qquad v(2n+1) = A_1 v(n)$$

for all[4] $n \geq 0$ with the matrices

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad A_1 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \end{pmatrix},$$

and with $v(0) = (0, 1, 1, 0, 2)^\top$. Therefore, the sequences $x(n)$, $y(n)$ and $z(n)$ are 2-regular.

3.2 Asymptotics

▶ **Corollary 5.** *We have*

$$X(N) = \sum_{1 \leq n \leq N} x(n) = N^\kappa \, \Phi(\{\log_2 N\}) + O(N \log_2 N) \tag{3.3}$$

with $\kappa = \log_2\big(3 + \sqrt{17}\,\big) - 1 = 1.83250638358045\ldots$ and a 1-periodic function Φ which is Hölder continuous with any exponent smaller than $\kappa - 1$.

Moreover, we can effectively compute the Fourier coefficients of Φ.

We get analogous results for the sequences $Y(N)$ and $Z(N)$ (each with its own periodic function Φ, but the same exponent κ). The fluctuation Φ of $X(N)$ is visualized in Figure 3.2 and its first few Fourier coefficients are shown in Table 3.1.

At this point, we only prove (3.3) of Corollary 5. We deal with the Fourier coefficients in the arXiv version [17, Appendix I.2] of this extended abstract. As in the introductory example of the binary sum-of-digits functions (Example 1), we could get Fourier coefficients by

---

[4] Note that $v(0) = A_0 v(0)$ and $v(1) = A_1 v(0)$ are indeed true.

Table 3.1 Fourier coefficients of Φ (Corollary 5). All stated digits are correct.

ℓ	α_ℓ
0	$0.69116151112341912755021246$
1	$-0.010792163112404078729505110 - 0.0023421761940286789685827i$
2	$0.0027937863735049517211671200 - 0.0006673612865972891134775600i$
3	$-0.0002007825832364584252264000 - 0.0031973663977645462669373i$
4	$0.0002494467892174674728133800 - 0.0005912995467076061497650i$
5	$-0.0003886698612765803447578000 + 0.0000672386631993014856843100i$
6	$-0.0006223575988893574655258000 + 0.0004321722061493985978154200i$
7	$0.0002303431736418138313047600 - 0.0005866316877285609142768800i$
8	$0.0005339060804798716172593000 - 0.0002119380802590974909465100i$
9	$0.0000678983897701759285290000 - 0.0003830782328548623528018500i$
10	$-0.0001998174599735525506199100 - 0.0003139456906014279980817500i$

Theorem 2 and the 2-linear representation of Section 3.1 directly. However, the information in the vector $v(n)$ (see (3.2)) is redundant with respect to the asymptotic main term as it contains $x(n)$ and $z(n)$ as well as $x(n+1)$ and $z(n+1)$; both pairs are asymptotically equal in the sense of (3.3). Therefore, we head for an only 3-dimensional functional system of equations for our Dirichlet series of $x(n)$, $y(n)$ and $z(n)$ (instead of a 5-dimensional system).

Proof of (3.3). We use Theorem 2.

Joint Spectral Radius.   First we compute the joint spectral radius ρ of A_0 and A_1. Both matrices have a maximum absolute row sum equal to 2, thus $\rho \leq 2$, and both matrices have 2 as an eigenvalue. Therefore we obtain $\rho = 2$. Moreover, the finiteness property of the linear representation is satisfied by considering only products with exactly one matrix factor A_0 or A_1.

Thus, we have $R = \rho = 2$.

Eigenvalues.   Next, we compute the spectrum $\sigma(C)$ of $C = A_0 + A_1$. The matrix C has the eigenvalues $\lambda_1 = \left(3 + \sqrt{17}\right)/2 = 3.5615528128088\ldots$, $\lambda_2 = 2$, $\lambda_3 = -2$, $\lambda_4 = -1$ and $\lambda_5 = \left(3 - \sqrt{17}\right)/2 = -0.5615528128088\ldots$ (each with multiplicity one). (Note that λ_1 and λ_5 are the zeros of the polynomial $U^2 - 3U - U$.)

Asymptotic Formula.   By using Theorem 2, we obtain an asymptotic formula for $X(N-1)$. Shifting from $N-1$ to N does not change this asymptotic formula, as this shift is absorbed by the error term $O(N \log_2 N)$.   ◀

3.3   Dirichlet Series and Meromorphic Continuation

Let $n_0 \geq 2$ be an integer and define

$$\mathcal{X}_{n_0}(s) = \sum_{n \geq n_0} \frac{x(n)}{n^s}, \qquad \mathcal{Y}_{n_0}(s) = \sum_{n \geq n_0} \frac{y(n)}{n^s}, \qquad \mathcal{Z}_{n_0}(s) = \sum_{n \geq n_0} \frac{z(n)}{n^s}.$$

▶ **Lemma 6.** *Set*

$$
C = I - \begin{pmatrix} 2^{-s} & 2^{-s} & 2^{-s} \\ 2^{1-s} & 0 & 2^{1-s} \\ 2^{1-s} & 2^{1-s} & 0 \end{pmatrix}.
$$

Then

$$
C \begin{pmatrix} \mathcal{X}_{n_0}(s) \\ \mathcal{Y}_{n_0}(s) \\ \mathcal{Z}_{n_0}(s) \end{pmatrix} = \begin{pmatrix} \mathcal{J}_{n_0}(s) \\ \mathcal{K}_{n_0}(s) \\ \mathcal{L}_{n_0}(s) \end{pmatrix}, \tag{3.4}
$$

where

$$
\mathcal{J}_{n_0}(s) = 2^{-s}\, \Sigma(s, -\tfrac{1}{2}, \mathcal{Y}_{n_0}) + \mathcal{I}_{\mathcal{J}_{n_0}}(s),
$$

$$
\mathcal{I}_{\mathcal{J}_{n_0}}(s) = -\frac{y(n_0)}{(2n_0-1)^s} + \sum_{n_0 \le n < 2n_0} \frac{x(n)}{n^s},
$$

$$
\mathcal{K}_{n_0}(s) = 2^{-s}\, \Sigma(s, 1, \mathcal{X}_{n_0}) + 2^{-s}\, \Sigma(s, -\tfrac{1}{2}, \mathcal{X}_{n_0}) + 2^{-s}\, \Sigma(s, \tfrac{1}{2}, \mathcal{Z}_{n_0}) + \mathcal{I}_{\mathcal{K}_{n_0}}(s),
$$

$$
\mathcal{I}_{\mathcal{K}_{n_0}}(s) = \frac{x(n_0-1)}{(2n_0)^s} - \frac{x(n_0)}{(2n_0-1)^s} + \sum_{n_0 \le n < 2n_0} \frac{y(n)}{n^s},
$$

$$
\mathcal{L}_{n_0}(s) = 2^{1-s}\, \Sigma(s, -\tfrac{1}{2}, \mathcal{Y}_{n_0}) + \mathcal{I}_{\mathcal{L}_{n_0}}(s),
$$

$$
\mathcal{I}_{\mathcal{L}_{n_0}}(s) = -\frac{2y(n_0)}{(2n_0-1)^s} + \sum_{n_0 \le n < 2n_0} \frac{z(n)}{n^s},
$$

with

$$
\Sigma(s, \beta, \mathcal{D}) = \sum_{k \ge 1} \binom{-s}{k} \beta^k \mathcal{D}(s+k)
$$

provides meromorphic continuations of the Dirichlet series $\mathcal{X}_{n_0}(s)$, $\mathcal{Y}_{n_0}(s)$, and $\mathcal{Z}_{n_0}(s)$ for $\Re s > \kappa_0 = 1$ with the only possible poles at $\kappa + \chi_\ell$ for $\ell \in \mathbb{Z}$, all of which are simple poles.

The proof of Lemma 6 can be found in the arXiv version [17, Appendix I] of this extended abstract. The idea is to rewrite the Dirichlet series corresponding to (3.1a), (3.1b) and (3.1c) to obtain the functional equation. The poles in the meromorphic continuation come from

$$
\Delta(s) = \det C = 2^{-3s}(2^{2s} - 3 \cdot 2^s - 2)(2^s + 2).
$$

The Fourier coefficients (rest of Corollary 5) can then be computed by applying Theorem 2.

───── **References** ─────

1   Jean-Paul Allouche, Michel Mendès France, and Jacques Peyrière. Automatic Dirichlet series. *J. Number Theory*, 81(2):359–373, 2000. `doi:10.1006/jnth.1999.2487`.
2   Jean-Paul Allouche and Jeffrey Shallit. The ring of k-regular sequences. *Theoret. Comput. Sci.*, 98(2):163–197, 1992. `doi:10.1016/0304-3975(92)90001-V`.
3   Jean-Paul Allouche and Jeffrey Shallit. *Automatic sequences: Theory, applications, generalizations.* Cambridge University Press, Cambridge, 2003. `doi:10.1017/CBO9780511546563`.
4   Valérie Berthé, Loïck Lhote, and Brigitte Vallée. Probabilistic analyses of the plain multiple gcd algorithm. *J. Symbolic Comput.*, 74:425–474, 2016. `doi:10.1016/j.jsc.2015.08.007`.

5 Valérie Berthé and Michel Rigo, editors. *Combinatorics, automata and number theory*, volume 135 of *Encyclopedia Math. Appl.* Cambridge University Press, Cambridge, 2010. `doi:10.1017/CBO9780511777653`.

6 Hubert Delange. Sur la fonction sommatoire de la fonction "somme des chiffres". *Enseignement Math. (2)*, 21:31–47, 1975.

7 Michael Drmota and Peter J. Grabner. Analysis of digital functions and applications. In Valérie Berthé and Michel Rigo, editors, *Combinatorics, automata and number theory*, volume 135 of *Encyclopedia Math. Appl.*, pages 452–504. Cambridge University Press, Cambridge, 2010. `doi:10.1017/CBO9780511777653.010`.

8 Michael Drmota and Wojciech Szpankowski. A master theorem for discrete divide and conquer recurrences. *J. ACM*, 60(3):Art. 16, 49 pp., 2013. `doi:10.1145/2487241.2487242`.

9 Philippe Dumas. Joint spectral radius, dilation equations, and asymptotic behavior of radix-rational sequences. *Linear Algebra Appl.*, 438(5):2107–2126, 2013. `doi:10.1016/j.laa.2012.10.013`.

10 Philippe Dumas. Asymptotic expansions for linear homogeneous divide-and-conquer recurrences: Algebraic and analytic approaches collated. *Theoret. Comput. Sci.*, 548:25–53, 2014. `doi:10.1016/j.tcs.2014.06.036`.

11 Philippe Dumas, Helger Lipmaa, and Johan Wallén. Asymptotic behaviour of a non-commutative rational series with a nonnegative linear representation. *Discrete Math. Theor. Comput. Sci.*, 9(1):247–272, 2007. URL: `https://dmtcs.episciences.org/399`.

12 Philippe Flajolet, Peter Grabner, Peter Kirschenhofer, Helmut Prodinger, and Robert F. Tichy. Mellin transforms and asymptotics: digital sums. *Theoret. Comput. Sci.*, 123:291–314, 1994. `doi:10.1016/0304-3975(92)00065-Y`.

13 John Goldwasser, William Klostermeyer, Michael Mays, and George Trapp. The density of ones in Pascal's rhombus. *Discrete Math.*, 204(1–3):231–236, 1999. `doi:10.1016/S0012-365X(98)00373-2`.

14 Peter J. Grabner and Clemens Heuberger. On the number of optimal base 2 representations of integers. *Des. Codes Cryptogr.*, 40(1):25–39, 2006. `doi:10.1007/s10623-005-6158-y`.

15 Peter J. Grabner, Clemens Heuberger, and Helmut Prodinger. Counting optimal joint digit expansions. *Integers*, 5(3):A9, 2005. URL: `http://www.integers-ejcnt.org/vol5-3.html`.

16 Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete mathematics. A foundation for computer science.* Addison-Wesley, second edition, 1994.

17 Clemens Heuberger, Daniel Krenn, and Helmut Prodinger. Analysis of summatory functions of regular sequences: Transducer and Pascal's rhombus. arXiv:1802.03266 [math.CO], 2018. URL: `https://arxiv.org/abs/1802.03266`.

18 Clemens Heuberger, Sara Kropf, and Helmut Prodinger. Output sum of transducers: Limiting distribution and periodic fluctuation. *Electron. J. Combin.*, 22(2):1–53, 2015. URL: `http://www.combinatorics.org/ojs/index.php/eljc/article/view/v22i2p19`.

19 Hsien-Kuei Hwang, Svante Janson, and Tsung-Hsi Tsai. Exact and asymptotic solutions of a divide-and-conquer recurrence dividing at half: Theory and applications. *ACM Trans. Algorithms*, 13(4):Art. 47, 43 pp., 2017. `doi:10.1145/3127585`.

A   Results

We formulate the full version of our results here in Appendix A. Formulating them will need quite a number of definitions provided in Appendix A.2. The proofs are given in the appendix of the arXiv version [17] of this extended abstract.

As announced in the introduction, we study matrix products instead of regular sequences.

A.1 Problem Statement

Let $q \geq 2$, $d \geq 1$ be fixed integers and $A_0, \ldots, A_{q-1} \in \mathbb{C}^{d \times d}$. We investigate the sequence $(f(n))_{n \geq 0}$ of $d \times d$ matrices such that

$$f(qn + r) = A_r f(n) \quad \text{for } 0 \leq r < q, \ 0 \leq n \text{ with } qn + r \neq 0 \tag{A.1}$$

and $f(0) = I$.

Let n be an integer with q-ary expansion $r_{\ell-1} \ldots r_0$. Then it is easily seen that (A.1) implies that

$$f(n) = A_{r_0} \ldots A_{r_{\ell-1}}. \tag{A.2}$$

We are interested in the asymptotic behaviour of $F(N) := \sum_{0 \leq n < N} f(n)$.

A.2 Definitions and Notations

In this section, we give all definitions and notations which are required in order to state the results.

The following notations are essential:

- Let $\|\cdot\|$ denote a fixed norm on $\mathbb{C}^d$ and its induced matrix norm on $\mathbb{C}^{d \times d}$.
- We set $B_r := \sum_{0 \leq r' < r} A_{r'}$ for $0 \leq r < q$ and $C := \sum_{0 \leq r < q} A_r$.
- The joint spectral radius of $A_0, \ldots, A_{q-1}$ is denoted by

$$\rho := \inf_{\ell} \sup \{ \|A_{r_1} \ldots A_{r_\ell}\|^{1/\ell} : r_1, \ldots, r_\ell \in \{0, \ldots, q-1\} \}.$$

  If the set of matrices $A_0, \ldots, A_{q-1}$ has the finiteness property, i.e., there is an $\ell > 0$ such that

$$\rho = \sup \{ \|A_{r_1} \ldots A_{r_\ell}\|^{1/\ell} : r_1, \ldots, r_\ell \in \{0, \ldots, q-1\} \},$$

  then we set $R = \rho$. Otherwise, we choose $R > \rho$ in such a way that there is no eigenvalue λ of C with $\rho < |\lambda| \leq R$.
- The spectrum of C, i.e., the set of eigenvalues of C, is denoted by $\sigma(C)$.
- For a positive integer n_0, set

$$\mathcal{F}_{n_0}(s) := \sum_{n \geq n_0} n^{-s} f(n)$$

  for a complex variable s.
- Set $\chi_k := \frac{2\pi i k}{\log q}$ for $k \in \mathbb{Z}$.

In the formulation of Theorem 7 and Corollary 8, the following constants are needed additionally:

- Choose a regular matrix T such that $TCT^{-1} =: J$ is in Jordan form.
- Let D be the diagonal matrix whose jth diagonal element is 1 if the jth diagonal element of J is not equal to 1; otherwise the jth diagonal element of D is 0.
- Set $C' := T^{-1}DJT$.
- Set $K := T^{-1}DT(I - C')^{-1}(I - A_0)$.
- For a $\lambda \in \mathbb{C}$, let $m(\lambda)$ be the size of the largest Jordan block associated with λ. In particular, $m(\lambda) = 0$ if $\lambda \notin \sigma(C)$.

- For $m \geq 0$, set

$$\vartheta_m := \frac{1}{m!} T^{-1}(I - D)T(C - I)^{m-1}(I - A_0);$$

  here, ϑ_0 remains undefined if $1 \in \sigma(C)$.[5]
- Define $\vartheta := \vartheta_{m(1)}$.

All implicit O-constants depend on q, d, the matrices A_0, ..., A_{q-1} (and therefore on ρ) as well as on R.

A.3   Decomposition into Periodic Fluctuations

Instead of considering $F(N)$, it is certainly enough to consider $wF(N)$ for all generalised left eigenvectors w of C, e.g., the rows of T. The result for $F(N)$ then follows by taking appropriate linear combinations.

▶ **Theorem 7.** *Let w be a generalised left eigenvector of rank m of C corresponding to the eigenvalue λ.*
1. *If $|\lambda| < R$, then*

$$wF(N) = wK + (\log_q N)^m w\vartheta_m + O(N^{\log_q R}).$$

2. *If $|\lambda| = R$, then*

$$wF(N) = wK + (\log_q N)^m w\vartheta_m + O(N^{\log_q R}(\log N)^m).$$

3. *If $|\lambda| > R$, then there are 1-periodic continuous functions $\Phi_k \colon \mathbb{R} \to \mathbb{C}^d$, $0 \leq k < m$, such that*

$$wF(N) = wK + (\log_q N)^m w\vartheta_m + N^{\log_q \lambda} \sum_{0 \leq k < m} (\log_q N)^k \Phi_k(\{\log_q N\})$$

   for $N \geq q^{m-1}$. The function Φ_k is Hölder-continuous with any exponent smaller than $\log_q|\lambda|/R$.
   If, additionally, the left eigenvector $w(C - \lambda I)^{m-1}$ of C happens to be a left eigenvector to each matrix A_0, ..., A_{q-1} associated with the eigenvalue 1, then

$$\Phi_{m-1}(u) = \frac{1}{q^{m-1}(m-1)!} w(C - qI)^{m-1}$$

   is constant.
Here, $wK = 0$ for $\lambda = 1$ and $w\vartheta_m = 0$ for $\lambda \neq 1$.

Note that in general, the three summands in the theorem have different growths: a constant, a logarithmic term and a term whose growth depends essentially on the joint spectral radius and the eigenvalues larger than the joint spectral radius, respectively. The vector w is not directly visible in front of the third summand; instead, the vectors of its Jordan chain are part of the function Φ_k.

Expressing the identity matrix as linear combinations of generalised left eigenvalues and summing up the contributions of Theorem 7 essentially yields the following corollary.

---

[5] If $1 \in \sigma(C)$, then the matrix $C - I$ is singular. In that case, ϑ_0 will never be used.

▶ **Corollary 8.** *With the notations above, we have*

$$F(N) = \sum_{\substack{\lambda \in \sigma(C) \\ |\lambda| > \rho}} N^{\log_q \lambda} \sum_{0 \le k < m(\lambda)} (\log_q N)^k \Phi_{\lambda k}(\{\log_q N\}) + (\log_q N)^{m(1)} \vartheta + K$$

$$+ O\big(N^{\log_q R}(\log N)^{\max\{m(\lambda) : \, |\lambda| = R\}}\big)$$

for suitable 1-periodic continuous functions $\Phi_{\lambda k}$. If 1 is not an eigenvalue of C, then $\vartheta = 0$. If there are no eigenvalues $\lambda \in \sigma(C)$ with $|\lambda| \le \rho$, then the O-term can be omitted.

For $|\lambda| > R$, the function $\Phi_{\lambda k}$ is Hölder continuous with any exponent smaller than $\log_q(|\lambda|/R)$.

A.4 Dirichlet Series

This section gives the required result on the Dirichlet series $\mathcal{F}_{n_0}$. For theoretical purposes, it is enough to study $\mathcal{F} := \mathcal{F}_1$; for numerical purposes, however, convergence improves for larger values of n_0.

▶ **Theorem 9.** *Let n_0 be a positive integer. Then the Dirichlet series $\mathcal{F}_{n_0}(s)$ converges absolutely and uniformly on compact subsets of the half plane $\Re s > \log_q \rho + 1$, thus is analytic there.*

We have

$$(I - q^{-s}C)\mathcal{F}_{n_0}(s) = \mathcal{G}_{n_0}(s) \tag{A.3}$$

for $\Re s > \log_q \rho + 1$ with

$$\mathcal{G}_{n_0}(s) = \sum_{n=n_0}^{qn_0 - 1} n^{-s} f(n) + q^{-s} \sum_{r=0}^{q-1} A_r \sum_{k \ge 1} \binom{-s}{k}\left(\frac{r}{q}\right)^k \mathcal{F}_{n_0}(s+k). \tag{A.4}$$

The series in (A.4) converge absolutely and uniformly on compact sets for $\Re s > \log_q \rho$. Thus (A.3) gives a meromorphic continuation of $\mathcal{F}_{n_0}$ to the half plane $\Re s > \log_q \rho$ with possible poles at $s = \log_q \lambda + \chi_\ell$ for each $\lambda \in \sigma(C)$ with $|\lambda| > \rho$ and $\ell \in \mathbb{Z}$ whose pole order is at most $m(\lambda)$.

Let $\delta > 0$. For real z, we set

$$\mu_\delta(z) = \max\{1 - (z - \log_q \rho - \delta), 0\},$$

i.e., the linear function on the interval $[\log_q \rho + \delta, \log_q \rho + \delta + 1]$ with $\mu_\delta(\log_q \rho + \delta) = 1$ and $\mu_\delta(\log_q \rho + \delta + 1) = 0$. Then

$$\mathcal{F}_{n_0}(s) = O\big(|\Im s|^{\mu_\delta(\Re s)}\big) \tag{A.5}$$

holds uniformly for $\log_q \rho + \delta \le \Re s$ and $|q^s - \lambda| \ge \delta$ for all eigenvalues $\lambda \in \sigma(C)$. Here, the implicit O-constant also depends on δ.

▶ **Remark.** By the identity theorem for analytic functions, the meromorphic continuation of $\mathcal{F}_{n_0}$ is unique on the domain given in the theorem. Therefore, the bound (A.5) does not depend on the particular expression for the meromorphic continuation given in (A.3) and (A.4).

A.5 Fourier Coefficients

As discussed in Section 1.3, we would like to apply the zeroth order Mellin–Perron summation formula but need analytic justification. In the following theorem we prove that whenever it is known that the result is a periodic fluctuation, the use of zeroth order Mellin–Perron summation can be justified. In contrast to the remaining paper, this theorem does *not* assume that $f(n)$ is a matrix product.

▶ **Theorem 10.** *Let $f(n)$ be a sequence, let $\kappa_0 \in \mathbb{R} \setminus \{0\}$ and $\kappa \in \mathbb{C}$ with $\Re\kappa > \kappa_0 > -1$, $\delta > 0$, $q > 1$ be real numbers with $\delta \leq \pi/(\log q)$ and $\delta < \Re\kappa - \kappa_0$, and let m be a positive integer. Moreover, let Φ_k be Hölder-continuous (with exponent α with $\Re\kappa - \kappa_0 < \alpha \leq 1$) 1-periodic functions for $0 \leq k < m$ such that*

$$F(N) := \sum_{1 \leq n < N} f(n) = \sum_{0 \leq k < m} N^\kappa (\log_q N)^k \Phi_k(\{\log_q N\}) + O(N^{\kappa_0}) \tag{A.6}$$

for integers $N \to \infty$.

For the Dirichlet series $\mathcal{F}(s) := \sum_{n \geq 1} n^{-s} f(n)$ assume that

- *there is some real number $\sigma_a \geq \Re\kappa$ such that $\mathcal{F}(s)$ converges absolutely for $\Re s > \sigma_a$;*
- *the Dirichlet series $\mathcal{F}(s)$ can be continued to a meromorphic function for $\Re s > \kappa_0 - \delta$ such that poles can only occur at $\kappa + \chi_\ell$ for $\ell \in \mathbb{Z}$ and such that these poles have order at most m;*
- *there is some real number $\eta > 0$ such that for $\kappa_0 \leq \Re s \leq \sigma_a$ and $|s - \kappa - \chi_\ell| \geq \delta$ for all $\ell \in \mathbb{Z}$, we have*

$$\mathcal{F}(s) = O\big(|\Im s|^\eta\big) \tag{A.7}$$

for $|\Im s| \to \infty$.

All implicit O-constants may depend on f, q, m, κ, κ_0, α, δ, σ_a and η.

Then

$$\Phi_k(u) = \sum_{\ell \in \mathbb{Z}} \varphi_{k\ell} \exp(2\ell\pi i u)$$

for $u \in \mathbb{R}$ where

$$\varphi_{k\ell} = \frac{(\log q)^k}{k!} \operatorname{Res}\left(\frac{\mathcal{F}(s)(s - \kappa - \chi_\ell)^k}{s}, s = \kappa + \chi_\ell\right) \tag{A.8}$$

for $\ell \in \mathbb{Z}$ and $0 \leq k < m$.

If $-1 < \kappa_0 < 0$ and $\kappa \notin \frac{2\pi i}{\log q} \mathbb{Z}$, then $\mathcal{F}(0) = 0$.

The theorem is more general than necessary for q-regular sequences because Theorem 9 shows that we could use some $0 < \eta < 1$. However, it might be applicable in other cases, so we prefer to state it in this more general form.

Distribution of the Number of Corners in Tree-like and Permutation Tableaux

Paweł Hitczenko[1]

Department of Mathematics, Drexel University, Philadelphia, PA 19104, USA
phitczenko@math.drexel.edu

Aleksandr Yaroslavskiy[2]

Department of Mathematics, Drexel University, Philadelphia, PA 19104, USA
amy46@drexel.edu

──── **Abstract** ────

In this abstract, we study tree-like tableaux and some of their probabilistic properties. Tree-like tableaux are in bijection with other combinatorial structures, including permutation tableaux, and have a connection to the partially asymmetric simple exclusion process (PASEP), an important model of interacting particles system. In particular, in the context of tree-like tableaux, a corner corresponds to a node occupied by a particle that could jump to the right while inner corners indicate a particle with an empty node to its left. Thus, the total number of corners represents the number of nodes at which PASEP can move, i.e., the total current activity of the system. As the number of inner corners and regular corners is connected, we limit our discussion to just regular corners and show that, asymptotically, the number of corners in a tableaux of length n is normally distributed. Furthermore, since the number of corners in tree-like tableaux are closely related to the number of corners in permutation tableaux, we will discuss the corners in the context of the latter tableaux.

1 Introduction

In this report, we study tree-like tableaux, a combinatorial object introduced in [1]. They are in bijection with permutation tableaux and alternative tableaux but are interesting in their own right as they exhibit a natural tree structure. Aside from being in bijection with permutations and permutation tableaux, they can be used to study the partially asymmetric simple exclusion process (PASEP). The PASEP (see e. g. [6, 9] and references therein) is a model in which n nodes on a 1-dimensional lattice each either contains a particle or not. At each time interval, a particle can either move left or right to an empty adjacent node with fixed probabilities and the probability of a move left is q times the probability of jumping to the right. New particle may also enter from the left with probability α (if the first node

---

is unoccupied) and a particle on the nth node may leave the lattice with probability β. A state of the PASEP is a configuration of occupied and unoccupied nodes and it naturally corresponds to border edges of tree-like tableaux. In this association, corners in tree-like tableaux, correspond to sites at which a particle can move (we will give more details below, see also [13] for an explanation). In physics literature this is known as (total) current activity [7, 8] and was studied for the TASEP (a special case of the PASEP with $q = 0$) in [15].

It was conjectured (see [13, Conjecture 4.1]) that the expected number of corners in a randomly chosen tree-like tableaux of size n is $(n+4)/6$. This conjecture (and its companion for symmetric tree-like tableaux) was proved in [12, Theorem 4] and subsequently also in [10, Theorem 4.1]). However, not much beyond that has been known (even the asymptotic value of the variance). In the present paper we take the next step in the analysis of tree-like tableaux. First, since permutation tableaux are in bijection with tree-like tableaux and the number of corners in both are related, we shift our discussion to be solely concerned with permutation tableaux and derive our results in that context. In particular, we obtain the variance of the number of corners. Furthermore, we also show that the number of corners in random permutation, and therefore tree-like tableau of size n is asymptotically normal as n goes to infinity.

The rest of the paper is organized as follows. In the next section we introduce the necessary definitions and notation. We also explain the relation between the tree-like tableaux and the PASEP. In Section 3 we present a recursive relation for the generating function involving the corners in a similar combinatorial object, namely permutation tableaux. This recursion will be used in Section 4 to obtain a recursion for the moment generation function of the number of corners in permutation tableaux and in Section 5 to establish its asymptotic normality. As mentioned earlier, this will imply the same result for the number of corners in tree-like tableaux.

2   Preliminaries

2.1   Tree-like Tableaux and Permutation Tableaux

We endeavor to introduce the background for studying tree-like tableaux. We start by recalling the necessary notions and properties.

▶ **Definition 1.** A Ferrers diagram is an up and left justified arrangement of cells with weakly decreasing number of cells in rows. Depending on the situation, some rows may or may not be empty. The length of a Ferrers diagram is the number of columns plus the number of rows.

Let us recall the following definition introduced in [1].

▶ **Definition 2.** A tree-like tableau of size n is a Ferrers diagram of length $n + 1$ with no empty rows and with some cells (called pointed cells) filled with a point according to the following rules:
1. The cell in the first column and first row is always pointed (this point is known as the root point).
2. Every row and every column contains at least one pointed cell.
3. For every non-root pointed cell, either all the cells above are empty or all the cells to the left are empty (but not both).
We denote the set of all tree-like tableaux of size n by $\mathcal{T}_n$.

We will also need a notion of permutation tableaux originally introduced in [14].
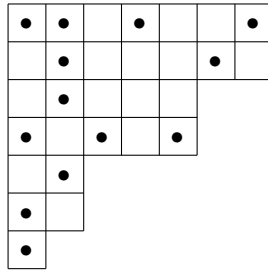
Figure 1 A tree-like tableau of size 13.
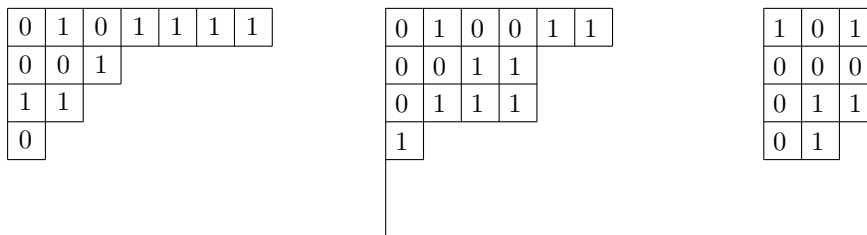
Figure 2 Examples of permutation tableaux. The tableau in the middle has two empty rows.

▶ **Definition 3.** A permutation tableau of size n is a Ferrers diagram of length n whose non-empty rows are filled with 0's and 1's according to the following rules:
1. Each column has at least one 1.
2. Any 0 cannot have a 1 both above it and to the left of it simultaneously.
We denote the set of all permutation tableaux of size n by $\mathcal{P}_n$.

In a tree-like or a permutation tableau, the edges outlining the southeast border are often called border edges. We also refer to those edges as steps. Each step is either a south step or a west step if we move along border edges from northeast to southwest or a north step or an east step if we move in the opposite direction.

▶ **Definition 4.** A corner in a tableau is a south step followed immediately by a west step as we traverse the border edges starting from the northeast and going to the southwest end. We denote by $c(T)$ the number of corners of the tableau T. If $\mathcal{T}$ is a set of tableaux we let

$$c(\mathcal{T}) = \sum_{T \in \mathcal{T}} c(T)$$

denote the total number of corners of tableaux in $\mathcal{T}$.

Tree-like tableaux correspond to the states of the PASEP as follows: traverse the border edges of a tree-like tableau beginning at the southwest end. Ignoring the first and the last step, a north step corresponds to an unoccupied node and an east step corresponds to an occupied node. Thus, for example, the tree-like tableau depicted in Figure 1 corresponds to the following state of the PASEP on 12 nodes: (In this state of the PASEP a particle could enter from the left, the particle in the second node could jump in either direction, the particle in the fifth or the tenth node could jump to the left and a particle in the seventh or the eleventh node could jump to the right.)

Figure 3 The state of the PASEP corresponding to the tree-like tableau in Figure 1.

With this association, the corners in tree-like tableau correspond to occupied sites, in which the particle could jump to the right (or enter from the left, or leave to the right) and any inner corner (north step followed by the east step) corresponds to an occupied node with a particle that can jump to the left. Since the number of inner corners is one less than the number of corners, the total number of possible moves for the PASEP in a state corresponding to $T \in \mathcal{T}_n$ is $2c(T) - 1$. For example the tableau in Figure 1 has four corners and thus the PASEP in the state depicted in Figure 3 has seven possible moves as described above. As we mentioned earlier, in physics literature the number of nodes at which a particle can move is called the current activity of the system, see e. g. [7, 8, 15].

It is known (see [1, Proposition 3.1]) that tree-like tableaux of length $n+1$ are in bijection with permutation tableaux of length n (and both are in bijection with permutations of $[n]$, see e. g. [3, 5, 14, 1]). The corners need not be preserved, but a difference between their number in a tableau and its image under that bijection is at most one (see [12, Section 3]). Therefore, in order to study corners in tree-like tableaux it will be enough to study corners in permutation tableaux, and this is what we are going to do. We need a few more notions associated with permutation tableaux.

▶ **Definition 5.** We say a zero in a permutation tableau is restricted if it has a one above it. Otherwise, the zero is unrestricted. We say a row is restricted if it contains a restricted zero, otherwise it is unrestricted. We denote by $u(T)$ the number of unrestricted rows of T.

In the first example in Figure 2, the top and the third row are unrestricted, but the other two rows are restricted. Note that the top row of a permutation tableau is necessarily unrestricted.

An important feature of permutation tableaux is that they can be constructed recursively. Given a permutation tableau, we can increase its length incrementally and fill in the new columns as they come.

▶ **Definition 6.** We say a tableau $T' \in \mathcal{P}_{n+1}$ is an extension of a tableau $T \in \mathcal{P}_n$ if T' is obtained either by adding a south step to the southwest corner of T or by adding a west step and filling the new column according to the rules.

Notice that there is only one way to extend a tableau by adding a south step, but multiple ways by adding a west step. When a west step is added, a new column is formed which must be filled. In a cell that is part of a restricted row, it must have a zero. The cells that are part of the unrestricted rows leave us options. It is not difficult to count the number of extensions (see e. g. [4, 11]) and we have:

▶ **Proposition 7.** The number of extensions of $T \in \mathcal{P}_n$ into $T' \in \mathcal{P}_{n+1}$ is $2^{u(T)}$.

This, however, tells us nothing of the number of unrestricted rows of the extended tableau, which is often of relevance. But the evolution of the number of unrestricted rows can be traced down (see [4] or [11]) and is given by:

▶ **Proposition 8.** *Let $T \in \mathcal{P}_n$ be a permutation tableau of length n, and let $u(T)$ be the number of unrestricted rows of T. The number of ways to extend T so that the extension has exactly k unrestricted rows, $1 \le k \le u(T)$, is:*

$$\sum_{j=1}^{k} \binom{u(T) - j}{k - j} = \binom{u(T)}{k - 1}.$$

In the following sections we prefer to use probabilistic language and thus, instead of talking about the number of corners in tableaux, we let $\mathbb{P}_n$ be the uniform probability measure on $\mathcal{X}_n$ (where $\mathcal{X}_n$ is either $\mathcal{T}_n$ or $\mathcal{P}_n$) and consider a random variable C_n on the probability space $(\mathcal{X}_n, \mathbb{P}_n)$ defined by $C_n(T) = c(T)$, the number of corners of $T \in \mathcal{X}_n$. A tableau chosen from $\mathcal{X}_n$ according to the probability measure $\mathbb{P}_n$ is usually referred to as a random tableau of size n and C_n is referred to as the number of corners in a random tableau of size n. We let $\mathbb{E}_n$ denote the expected value with respect to the measure $\mathbb{P}_n$. Then, of course, we have:

$$\mathbb{E}_n C_n = \frac{c(\mathcal{X}_n)}{|\mathcal{X}_n|}.$$

As we will see below, the variance of the number of corners, $\mathbb{Var}(C_n)$, grows to infinity as $n \to \infty$ (in fact, $\mathbb{Var}(C_n) \sim 11n/180$). Furthermore if $\phi_n : \mathcal{T}_n \to \mathcal{P}_n$ is the bijection described in [1, 12] then for $T \in \mathcal{T}_n$, $c(T) = c(\phi_n(T)) + I$, where I is 0 or 1 depending on the shape of T. Therefore, for every $x \in \mathbb{R}$

$$\mathbb{P}_n \left(T \in \mathcal{T}_n : \ \frac{C_n(T) - \mathbb{E}C_n}{\sqrt{\mathbb{Var}(C_n)}} \le x \right) = \mathbb{P}_n \left(T \in \mathcal{P}_n : \ \frac{C_n(T) - \mathbb{E}C_n + O(1)}{\sqrt{\mathbb{Var}(C_n)}} \le x \right).$$

Thus, the limiting distribution of the number of corners in a random tree-like tableau is the same as that of the number of corners in a random permutation tableau, so we will focus on the latter.

3 Generating Function and the First Two Moments

We wish to construct a generating function for the number of corners in permutation tableaux of length n. We can do it recursively by using the extension procedure for permutation tableaux mentioned earlier. In order to do this we need to keep track of the number of unrestricted rows, and we use it as a 'catalytic' variable. Proposition 8 allows us to follow the evolution of the number of unrestricted rows under the extension and with its help we can derive a recurrence for the bivariate generating function. Because of the space limitation, the presentation of our proof is deferred to the full version of the paper.

▶ **Proposition 9.** *Let for $n \ge 0$*

$$C_n(x, z) = \sum_{T \in \mathcal{P}_n} x^{c(T)} z^{u(T)}$$

be the bivariate generation function of permutation tableaux of length n, where x marks the number of corners and z marks the number of unrestricted rows. Then we have the following recurrence for $C_n(x, z)$:

$$C_n(x, z) = z C_{n-1}(x, z+1) + (x - 1)\Big(z(z+1) C_{n-2}(x, z+1) - z^2 C_{n-2}(x, z) \Big) \qquad (1)$$

with

$$C_0(x, z) = 1, \quad C_1(x, z) = z.$$

3.1 Expectation

The above proposition allows us to recover the expected value of the number of corners, a result conjectured in [13], first proved in [12], and then also in [10]. To do this, note that it is clear from (1) that

$$C_n(1, z) = zC_{n-1}(1, z+1) = \cdots = z^{\overline{n}},$$

where

$$z^{\overline{n}} = z(z+1) \cdots \cdots (z+n-1),$$

is the rising factorial. We can treat

$$\frac{C_n(x, z)}{C_n(1, z)} = \frac{C_n(x, z)}{z^{\overline{n}}}$$

as the probability generating function of a random variable that depends on a parameter z and, in fact, is defined on a probability space that depends on z. Ultimately, we will be interested in $z = 1$ but it is convenient to proceed with more generality.

When we write $C_n(x, z)$ in the form

$$C_n(x, z) = \sum_{m=0}^{\lfloor n/2 \rfloor} c_{n,m}(z)(x-1)^m,$$

then the expected value of such random variable is $c_{n,1}(z)/z^{\overline{n}}$. Note that (1) yields

$$c_{n,m}(z) = zc_{n-1,m}(z+1) + z(z+1)c_{n-2,m-1}(z+1) - z^2 c_{n-2,m-1}(z),$$

with the initial conditions $c_{n,0} = z^{\overline{n}}$, $n \geq 0$. Iteration gives

$$
\begin{aligned}
c_{n,m}(z) &= z(z+1)c_{n-2,m}(z+2) \\
&\quad + z(z+1)\Big((z+2)c_{n-3,m-1}(z+2) - (z+1)c_{n-3,m-1}(z+1)\Big) \\
&\quad + z\Big((z+1)c_{n-2,m-1}(z+1) - zc_{n-2,m-1}(z)\Big) \\
&= z^{\overline{k}} c_{n-k,m}(z+k) \\
&\quad + \sum_{j=1}^{k} z^{\overline{j}}\Big((z+j)c_{n-j-1,m-1}(z+j) - (z+j-1)c_{n-j-1,m-1}(z+j-1)\Big) \\
&= z^{\overline{n-2m}} c_{2m,m}(z+n-2m) \qquad\qquad\qquad\qquad\qquad\qquad (2) \\
&\quad + \sum_{j=1}^{n-2m} z^{\overline{j}}\Big((z+j)c_{n-j-1,m-1}(z+j) - (z+j-1)c_{n-j-1,m-1}(z+j-1)\Big).
\end{aligned}
$$

When $m = 1$ this becomes

$$
\begin{aligned}
c_{n,1}(z) &= z^{\overline{n-2}} c_{2,1}(z + n - 2) + \\
&\quad \sum_{j=1}^{n-2} z^{\overline{j}} \Big((z + j) c_{n-j-1,0}(z + j) - (z + j - 1) c_{n-j-1,0}(z + j - 1) \Big) \\
&= z^{\overline{n-2}} (z + n - 2) + \\
&\quad \sum_{j=1}^{n-2} z^{\overline{j}} \Big((z + j)(z + j)^{\overline{n-j-1}} - (z + j - 1)(z + j - 1)^{\overline{n-j-1}} \Big) \\
&= z^{\overline{n-1}} + z^{\overline{n-2}} \sum_{j=1}^{n-2} \Big((z + j)(z + n - 2) - (z + j - 1)^2 \Big) \\
&= z^{\overline{n-1}} + z^{\overline{n-2}}(n - 2)(z + n - 2) + \\
&\quad z^{\overline{n-2}} \sum_{j=1}^{n-2} \Big((z + j - 1)(z + n - 2) - (z + j - 1)^2 \Big) \\
&= z^{\overline{n-2}} \Big((n - 1)(z + n - 2) + \sum_{j=1}^{n-2} (z + j - 1)(n - j - 1) \Big) \\
&= z^{\overline{n-2}}(n - 1) \Big((z + n - 2) + \frac{(n - 2)(n + 3z - 3)}{6} \Big) \\
&= z^{\overline{n-2}}(n - 1) \frac{n^2 + 3zn + n - 6}{6}.
\end{aligned}
$$

Therefore,

$$
\frac{c_{n,1}(z)}{z^{\overline{n}}} = \frac{(n - 1)(n^2 + 3zn + n - 6)}{6(z + n - 1)_2}
$$

where $(w)_k = w(w-1)\dots(w-(k-1))$ is the falling factorial. When $z = 1$ the above formula gives

$$
\mathbb{E} C_n = \frac{n^2 + 4n - 6}{6n} = \frac{n + 4}{6} - \frac{1}{n}
$$

which agrees with [12, Theorem 2].

3.2   Variance

Calculation of the expected value can be pushed further and we can obtain the variance of the number of corners, which has not been known before.

▶ **Proposition 10.** *For $n \geq 4$ we have*

$$
\mathbb{V}ar(C_n) = \frac{11n^4 - 191n^2 + 360n + 180}{180n^2(n - 1)} \sim \frac{11}{180} n
$$

as $n \to \infty$. In addition

$$
\mathbb{V}ar(C_1) = 0, \quad \mathbb{V}ar(C_2) = \frac{1}{4}, \quad \mathbb{V}ar(C_3) = \frac{5}{36}.
$$

Proof. Consider $n \geq 4$ (the other three cases can be calculated directly). Our first goal is to extract $c_{n,2}(z)$. From (2) used with $m = 2$ we have

$$
\begin{aligned}
c_{n,2}(z) &= z^{\overline{n-4}}c_{4,2}(z + n - 4) \\
&\quad + \sum_{j=1}^{n-4} z^{\overline{j}}\Big((z + j)c_{n-j-1,1}(z + j) - (z + j - 1)c_{n-j-1,1}(z + j - 1)\Big).
\end{aligned}
$$

Since

$$
c_{4,2}(z) = z(z + 1)c_{2,1}(z + 1) - z^2 c_{2,1}(z) = z(z + 1)^2 - z^3 = z(2z + 1),
$$

we see that

$$
z^{\overline{n-4}}c_{4,2}(z + n - 4) = z^{\overline{n-3}}(2(z + n) - 7). \tag{3}
$$

Furthermore,

$$
\begin{aligned}
&z^{\overline{j}}(z + j)c_{n-j-1,1}(z + j) \\
&\quad = z^{\overline{n-3}}(z + j)(n - j - 2)\frac{(n - j - 1)^2 + 3(z + j)(n - j - 1) + n - j - 7}{6}
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
&z^{\overline{j}}(z + j - 1)c_{n-j-1,1}(z + j - 1) \\
&\quad = z^{\overline{n-4}}(z + j - 1)^2(n - j - 2)\frac{(n - j - 1)^2 + 3(z + j - 1)(n - j - 1) + n - j - 7}{6}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\sum_{j=1}^{n-4} z^{\overline{j}}\Big((z + j)c_{n-j-1,1}(z + j) - (z + j - 1)c_{n-j-1,1}(z + j - 1)\Big) \\
&\quad = \frac{z^{\overline{n-4}}}{6}\sum_{j=1}^{n-4}(n - j - 2)\Big\{(z + j)(z + n - 4)\Big((n - j - 1)^2 + 3(z + j)(n - j - 1) + n - j - 7\Big) \\
&\qquad -(z + j - 1)^2\Big((n - j - 1)^2 + 3(z + j - 1)(n - j - 1) + n - j - 7\Big)\Big\}.
\end{aligned}
$$

When $z = 1$ this equals

$$
\frac{(n - 4)!}{360}(n - 3)(n - 4)(5n^4 + 26n^3 - 38n^2 - 83n - 150).
$$

Combining with (3) we get

$$
\begin{aligned}
c_{n,2}(1) &= (n - 3)!(2n - 5) + \frac{(n - 4)!}{360}(n - 3)(n - 4)(5n^4 + 26n^3 - 38n^2 - 83n - 150) \\
&= \frac{(n - 2)!}{360}(5n^4 + 16n^3 - 110n^2 - 151n + 600).
\end{aligned}
$$

The second factorial moment for the number of corners is thus given by:

$$
\mathbb{E}(C_n)_2 = \mathbb{E}C_n(C_n - 1) = \frac{2!}{n!}c_{n,2}(1) = \frac{5n^4 + 16n^3 - 110n^2 - 151n + 600}{180n(n - 1)}
$$

and therefore,

$$
\begin{aligned}
\mathbb{V}\mathrm{ar}(C_n) &= \mathbb{E}(C_n)_2 - (\mathbb{E}C_n)^2 + \mathbb{E}C_n \\
&= \frac{5n^4 + 16n^3 - 110n^2 - 151n + 600}{180n(n-1)} - \left(\frac{n+4}{6} - \frac{1}{n}\right)^2 + \left(\frac{n+4}{6} - \frac{1}{n}\right) \\
&= \frac{11n^4 - 191n^2 + 360n + 180}{180n^2(n-1)}
\end{aligned}
$$

as claimed. ◄

It is, however, increasingly difficult to find $c_{n,m}$ for higher m. Instead, we will use (1) to derive a recurrence for the moment generating function and rely on method of moments (see e. g. [2, Theorem 30.2]) to establish the asymptotic normality of suitably normalized (C_n).

4   Moment Generating Function

Consider

$$
P_n(t,z) := e^{-\mu_n(z)t} \frac{C_n(e^t, z)}{z^{\overline{n}}}, \quad P_0(t,z) = P_1(t,z) = 1
$$

where

$$
\mu_0(z) = \mu_1(z) = 0; \quad \mu_n(z) = \frac{(n-1)(n^2 + 3zn + n - 6)}{6(z+n-1)_2}, \quad n \geq 2.
$$

(Notice that $\mu_n(1)$ is the expected value of C_n, the number of corners in permutation tableaux of size n.) Then, recurrence (1) translates into

$$
\begin{aligned}
P_n(t,z) &= e^{\alpha_n(z)t} P_{n-1}(t, z+1) \\
&+ \frac{e^t - 1}{(z+n-1)_2} \left((z+1)(z+n-2) e^{\beta_n(z)t} P_{n-2}(t, z+1) - z^2 e^{\delta_n(z)t} P_{n-2}(t, z) \right),
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha_n(z) &= \mu_{n-1}(z+1) - \mu_n(z) = -\frac{n + zn - z - 2}{(z+n-1)_2}, \\
\beta_n(z) &= \mu_{n-2}(z+1) - \mu_n(z), \\
\delta_n(z) &= \mu_{n-2}(z) - \mu_n(z).
\end{aligned}
$$

This gives a linear recurrence of the first order for $P_n^{(m)}(0,z)$; first

$$
\begin{aligned}
P_n^{(m)}(t,z) &= e^{\alpha_n(z)t} P_{n-1}^{(m)}(t, z+1) + \sum_{k=0}^{m-1} \binom{m}{k} \alpha_n^{m-k}(z) e^{\alpha_n(z)t} P_{n-1}^{(k)}(t, z+1) \\
&+ \frac{e^t}{(z+n-1)_2} \sum_{k=0}^{m-1} \binom{m}{k} \Bigg[(z+1)(z+n-2) \sum_{i=0}^{k} \binom{k}{i} \beta_n^{k-i}(z) e^{\beta_n(z)t} P_{n-2}^{(i)}(t, z+1) \\
&\qquad\qquad - z^2 \sum_{i=0}^{k} \binom{k}{i} \delta_n^{k-i}(z) e^{\delta_n(z)t} P_{n-2}^{(i)}(t, z) \Bigg] \\
&+ \frac{e^t - 1}{(z+n-1)_2} \left((z+1)(z+n-2) e^{\beta_n(z)t} P_{n-2}^{(i)}(t, z+1) - z^2 e^{\delta_n(z)t} P_{n-2}^{(i)}(t, z) \right)^{(m)}.
\end{aligned}
$$

At $t = 0$ the last term vanishes and letting $P_n^{(m)}(z) := P_n^{(m)}(0, z)$ we get

$$
P_n^{(m)}(z) = P_{n-1}^{(m)}(z+1) + \sum_{k=0}^{m-1} \binom{m}{k} \alpha_n^{m-k}(z) P_{n-1}^{(k)}(z+1) +
$$

$$
+ \frac{1}{(z+n-1)_2} \sum_{k=0}^{m-1} \left\{ (z+1)(z+n-2) P_{n-2}^{(k)}(z+1) \sum_{i=k}^{m-1} \binom{m}{i}\binom{i}{k} \beta_n^{i-k}(z) \right. \tag{4}
$$

$$
\left. - z^2 P_{n-2}^{(k)}(z) \sum_{i=k}^{m-1} \binom{m}{i}\binom{i}{k} \delta_n^{i-k}(z) \right\}.
$$

This recurrence is the starting point for establishing asymptotic normality for the number of corners in permutation tableaux. We outline the argument in the forthcoming section.

5 Main Result

We can now state our main result.

▶ **Theorem 11.** *Let $\{C_n\}$ be a sequence of random variables where C_n is the number of corners in a random permutation tableau of length n. Let:*

$$
\mu_n = \frac{n+4}{6} - \frac{1}{n} \sim \frac{n}{6}
$$

and

$$
\sigma_n^2 = \mathbb{V}ar(C_n) \sim \frac{11}{180} n.
$$

Then

$$
\frac{C_n - \mu_n}{\sigma_n} \xrightarrow{d} \mathcal{N}(0,1) \quad or \quad \frac{C_n - \frac{n}{6}}{\sqrt{\frac{11}{180} n}} \xrightarrow{d} \mathcal{N}(0,1),
$$

where $\xrightarrow{d}$ is convergence in distribution and $\mathcal{N}(0,1)$ is the standard normal random variable.

Proof. *(Sketch)* The proof relies on the method of moments (see e. g. [2, Theorem 30.2] and on the analysis of recurrence (4) for the moments which will allow us to establish that:

$$
\frac{P_n^{(m)}(1)}{\left(\frac{11}{180} n\right)^{\frac{m}{2}}} \to \begin{cases} 0, & m \text{ odd} \\ \frac{m!}{2^{\frac{m}{2}} \cdot (m/2)!}, & m \text{ even.} \end{cases} \tag{5}
$$

The complete proof will be presented in the full version of the paper; here we indicate the main steps in the argument. First, we retain only the two highest degree terms in recurrence (4) (the remaining terms are of lower order and thus do not contribute significantly). Then (4) simplifies to

$$
P_n^{(m)}(z) = P_{n-1}^{(m)}(z+1) + \binom{m}{m-1} \alpha_n(z) P_{n-1}^{(m-1)}(z+1) + \binom{m}{m-2} \alpha_n^2(z) P_{n-1}^{(m-2)}(z+1)
$$

$$
+ \frac{1}{(z+n-1)_2} \left\{ (z+1)(z+n-2) P_{n-2}^{(m-1)}(z+1) \binom{m}{m-1} - z^2 P_{n-2}^{(m-1)}(z) \binom{m}{m-1} \right.
$$

$$
+ (z+1)(z+n-2) P_{n-2}^{(m-2)}(z+1) \left[\binom{m}{m-2} + \binom{m}{m-1}\binom{m-1}{m-2} \beta_n(z) \right]
$$

$$
\left. - z^2 P_{n-2}^{(m-2)}(0,z) \left[\binom{m}{m-2} + \binom{m}{m-1}\binom{m-1}{m-2} \delta_n(z) \right] \right\}.
$$

The next simplification is based on the observation (which, again, will be justified in the full version of the paper) that $P_{n-2}^{(k)}(z+1) \sim P_{n-1}^{(k)}(z+1) \sim P_{n-2}^{(k)}(z)$. Thus, replacing all the kth derivatives in the curly brackets by $P_{n-1}^{(k)}(z+1)$, $k = m-1, m-2$ the above is

$$P_n^{(m)}(z) = P_{n-1}^{(m)}(z+1) + m P_{n-1}^{(m-1)}(z+1) \left[\alpha_n(z) + \frac{(z+1)(z+n-2) - z^2}{(z+n-1)_2} \right]$$
$$+ \binom{m}{2} P_{n-1}^{(m-2)}(z+1) \left\{ \alpha_n^2(z) + \frac{(z+1)(z+n-2)(1+2\beta_n(z)) - z^2(1+2\delta_n(z))}{(z+n-1)_2} \right\}.$$

By the definition of $\alpha_n(z)$, the term in the square brackets is zero. Denote the term in the curly brackets by $T_{n-1}(z+1)$. Then, our recurrence becomes

$$P_n^{(m)}(z) = P_{n-1}^{(m)}(z+1) + \binom{m}{2} T_{n-1}(z+1) P_{n-1}^{(m-2)}(z+1).$$

Iterating and using $P_1^{(m)}(z) = 0$ for all z and $m \geq 1$, we get

$$P_n^{(m)}(z) = \binom{m}{2} \sum_{j=1}^{n-1} T_{n-j}(z+j) P_{n-j}^{(m-2)}(z+j). \tag{6}$$

Let now $m = 2r$ be even. Iterating (6) yields

$$\begin{aligned}
P_n^{(2r)}(z) &= \binom{2r}{2} \sum_{j=1}^{n-1} T_{n-j}(z+j) P_{n-j}^{(2r-2)}(z+j) \\
&= \binom{2r}{2} \binom{2r-2}{2} \sum_{j_1=1}^{n-1} \sum_{j_2=1}^{n-j_1-1} T_{n-j_1}(z+j_1) T_{n-j_1-j_2}(z+j_1+j_2) P_{n-j_1-j_2}^{(2r-4)}(z+j_1+j_2) \\
&= \binom{2r}{2} \binom{2r-2}{2} \cdots \binom{4}{2} \sum_{j_1=1}^{n-1} \sum_{j_2=1}^{n-j_1-1} \cdots \sum_{j_r=1}^{n-1-\sum_{i=1}^{r-1} j_i} \prod_{i=1}^{r} T_{n-\sum_{l=1}^{i} j_l}\left(z + \sum_{l=1}^{i} j_l\right) \\
&= \frac{(2r)!}{2^r} \sum_{1 \leq k_1 < k_2 < \cdots < k_r < n} \prod_{i=1}^{r} T_{n-k_i}(z+k_i) \\
&= \frac{(2r)!}{2^r} \frac{1}{r!} \sum_{\substack{1 \leq k_1, \ldots, k_r < n \\ \text{distinct}}} \prod_{i=1}^{r} T_{n-k_i}(z+k_i) \\
&= \frac{(2r)!}{2^r} \frac{1}{r!} \left(\sum_{\substack{1 \leq k_1, \ldots, k_r < n \\ \text{all}}} \prod_{i=1}^{r} T_{n-k_i}(z+k_i) - \sum_{\substack{1 \leq k_1, \ldots, k_r < n \\ \text{not all distinct}}} \prod_{i=1}^{r} T_{n-k_i}(z+k_i) \right).
\end{aligned}$$

Set $z = 1$. The first sum is

$$\left(\sum_{k=1}^{n-1} T_{n-k}(k+1) \right)^r = \left(P_n^{(2)}(1) \right)^r \sim \left(\frac{11}{180} n \right)^r$$

where the first equality follows from (6) used with $m = 2$ and $P_n^{(0)}(z) = 1$. (This can also be verified by evaluating the sum of the $T_{n-k}(1+k)$ directly; for the purpose of asymptotic evaluation it suffices to use the highest order term approximating $T_n(z)$, i.e.,

$$T_n(z) \sim \frac{nz(2n^4 + 6n^3 z + n^2 z^2 + 3z^4)}{3(n+z)^6}$$

which is obtained by using the highest order term approximations, for example, $\mu_n(z) \sim n^2(n + 3z)/(6(z + n - 1)_2)$ and similarly for $\alpha_n(z)$, $\beta_n(z)$, and $\delta_n(z)$, but all of these approximations require justifications.)

For the second summation, observe that $|T_{n-j}(1 + j)| \leq C$ for all $1 \leq j < n$ and a universal constant C. Thus:

$$\left| \sum_{\substack{1 \leq k_1, \ldots, k_r < n \\ \text{not all distinct}}} \prod_{i=1}^{r} T_{n-k_i}(1 + k_i) \right| \leq \sum_{\substack{1 \leq k_1, \ldots, k_r < n \\ \text{not all distinct}}} \prod_{i=1}^{r} |T_{n-k_i}(1 + k_i)| \leq C^r \cdot O(n^{r-1}),$$

which is of lower order than the first sum. This proves (5) for m even.

Let now $m = 2r + 1$ be odd. We wish to show that $P_n^{(2r+1)}(1) = O(n^r)$ as $n \to \infty$. In fact, we proceed to prove by induction that for every $r \geq 0$, $P_{n-k}^{(2r+1)}(k + 1) = O(n^r)$ uniformly in $1 \leq k < n$ as $n \to \infty$. Since $P_n^{(1)}(z) = 0$ (and approximation errors are bounded) this is true for $r = 0$. Now assume $P_{n-k}^{(2r-1)}(k + 1) = O(n^{r-1})$ uniformly in $1 \leq k < n$. Then from (6) we have that:

$$
\begin{aligned}
|P_{n-k}^{(2r+1)}(k + 1)| &\leq \binom{2r + 1}{2} \max_{1 \leq k \leq n-1} |P_{n-k}^{(2r-1)}(k + 1)| \sum_{j=1}^{n-1} |T_{n-j}(z + j)| \\
&\leq \binom{2r + 1}{2} O(n^{r-1}) \cdot Cn = O(n^r)
\end{aligned}
$$

as desired.   ◀

References

1 J.-C. Aval, A. Boussicault, and P. Nadeau. Tree-like tableaux. *Electron. J. Combin.*, 20(4):Paper 34, 24, 2013.

2 P. Billingsley. *Probability and measure.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

3 A. Burstein. On some properties of permutation tableaux. *Ann. Comb.*, 11(3-4):355–368, 2007. `doi:10.1007/s00026-007-0323-0`.

4 S. Corteel and P. Hitczenko. Expected values of statistics on permutation tableaux. In *2007 Conference on Analysis of Algorithms, AofA 07*, Discrete Math. Theor. Comput. Sci. Proc., AH, pages 325–339. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2007.

5 S. Corteel and P. Nadeau. Bijections for permutation tableaux. *European J. Combin.*, 30(1):295–310, 2009. `doi:10.1016/j.ejc.2007.12.007`.

6 S. Corteel and L. K. Williams. Tableaux combinatorics for the asymmetric exclusion process. *Adv. in Appl. Math.*, 39(3):293–310, 2007. `doi:10.1016/j.aam.2006.08.002`.

7 M. Depken and R. B. Stinchcombe. Exact joint density–current probability function for the asymmetric exclusion process. *Phys. Rev. Lett.*, 93(2):040602, 2004. `doi:10.1016/j.jcta.2006.04.001`.

8 M. Depken and R. B. Stinchcombe. Exact probability function for bulk density and current in the asymmetric exclusion process. *Phys. Rev.*, 71(3):036120, 2005. `doi:10.1016/j.jcta.2006.04.001`.

9 B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier. Exact solution of a 1D asymmetric exclusion model using a matrix formulation. *J. Phys. A*, 26(7):1493–1517, 1993. URL: `http://stacks.iop.org/0305-4470/26/1493`.

10 A. L. L. Gao, E. X. L. Gao, P. Laborde-Zubieta, and B. Y. Sun. Enumeration of corners in tree-like tableaux. *Discrete Math. Theor. Comput. Sci.*, 18(3):Paper No. 17, 26, 2016.

11 P. Hitczenko and S. Janson. Asymptotic normality of statistics on permutation tableaux. In *Algorithmic probability and combinatorics*, volume 520 of *Contemp. Math.*, pages 83–104. Amer. Math. Soc., Providence, RI, 2010. `doi:10.1090/conm/520/10255`.

12 P. Hitczenko and A. Lohss. Corners in tree-like tableaux. *Electron. J. Combin.*, 23(4):Paper 4.26, 18, 2016.

13 P. Laborde-Zubieta. Occupied corners in tree-like tableaux. *Sém. Lothar. Combin.*, 74:Art. B74b, 14, [2015-2017].

14 E. Steingrímsson and L. K. Williams. Permutation tableaux and permutation patterns. *J. Combin. Theory Ser. A*, 114(2):211–234, 2007. `doi:10.1016/j.jcta.2006.04.001`.

15 R. B. Stinchcombe and S. L. A. de Quieros. Statistics of current activity fluctuations in asymmetric flow with exclusion. *Phys. Rev. E*, 85(4):041111, 2012. `doi:10.1016/j.jcta.2006.04.001`.

Asymptotic Expansions for Sub-Critical Lagrangean Forms

Hsien-Kuei Hwang[1]

Institute of Statistical Science, Academia Sinica, Taiwan
hkhwang@stat.sinica.edu.tw
https://orcid.org/0000-0002-9410-6476

Mihyun Kang[2]

Institute of Discrete Mathematics, TU Graz, Austria
kang@math.tugraz.at
https://orcid.org/0000-0001-8729-2779

Guan-Huei Duh

Institute of Statistical Science, Academia Sinica, Taiwan
arthurduh1@gmail.com

Abstract

Asymptotic expansions for the Taylor coefficients of the Lagrangean form $\phi(z) = zf(\phi(z))$ are examined with a focus on the calculations of the asymptotic coefficients. The expansions are simple and useful, and we discuss their use in some enumerating sequences in trees, lattice paths and planar maps.

1 Introduction

Singularity analysis and saddle-point method represent the two major standard approaches used in analytic combinatorics to compute the asymptotics of, say the Taylor coefficients $[z^n]\phi(z)$ for large n; see [9, Chap. VI & VII]. The choice of which method to use depends crucially on the growth order of the functions in question near the dominant singularity or the saddle-point. The general principle is to apply the saddle-point method when the growth order of f near the saddle-point is large (e.g., $\log \phi(z) \gg (\log|1-z|)^{1+\varepsilon}$) and to apply the singularity analysis otherwise. In most cases, only one of the two works if one is interested in more precise asymptotic approximations. The Lagrangean form (frequently encountered in diverse areas; see [9, §A.6])

$$\phi(z) = zf(\phi(z)) \tag{1}$$

---

with both ϕ and f analytic functions, is one of the few situations in which both methods apply, and the key tool bridging the two different approaches is the Lagrange Inversion Formula [9, §A.6]

$$[z^n]\phi(z) = n^{-1}[t^{n-1}]f(t)^n \qquad (n \geqslant 1). \tag{2}$$

This form of large powers shows generally that saddle-point method is a good candidate for deriving the corresponding asymptotics, while the functional form (1) favors the use of singularity analysis (coupling with the implicit function theorem).

For the purpose of more precise asymptotics, we assume the following conditions.

C1 (nonnegativity and aperiodicity) $a_j := [t^j]f(t) \geqslant 0$ for every $j \geqslant 0$ and $\gcd\{j : a_j > 0\} = 1$;

C2 (analyticity) f is analytic in $|z| < R$ for $0 < R \leqslant \infty$;

C3 (sub-criticality) there exists an $r \in (0, R)$ such that $rf'(r) = f(r)$.

Note that the conditions C1 and C3 imply that

$$a_0 = f(0) = \sum_{j \geqslant 2}(j-1)a_j r^j > 0.$$

Note further that the condition C3 fails when f is linear, namely, $f(z) = a_0 + a_1 z$, which gives rise to

$$\phi(z) = \frac{a_0 z}{1 - a_1 z} \implies [z^n]\phi(z) = a_0 a_1^{n-1}.$$

Under the conditions C1–C3, it is known, by singularity analysis or saddle-point method, that (see [9, §IV.7] or [13, 14])

$$[z^n]\phi(z) \sim C\, n^{-\frac{3}{2}}\, \rho^{-n}, \quad \text{where } \rho := \frac{r}{f(r)} \text{ and } C := \sqrt{\frac{f(r)}{2\pi f''(r)}}.$$

The aim of this extended abstract is to examine the asymptotic expansions of the Lagrangean form (1). In particular, we will prove the following theorem, which can be regarded as an alternative version of Theorem VI.6 in [9, §VI.7] with the coefficients not just "computable" but by a more precise formula. Also we prefer the use of binomial coefficients to negative powers of n.

▶ **Theorem 1.** *Assume that ϕ and f satisfy* (1). *Then, under the conditions C1–C3, we have*

$$[z^n]\phi(z) \sim \rho^{-n} \sum_{k \geqslant 0} c_{2k+1} \binom{n-k-\frac{3}{2}}{n}, \tag{3}$$

where the coefficients c_j's are expressible in yet another Lagrangean form

$$c_k = k^{-1}[v^{k-1}]F(v)^k, \quad \text{with} \quad F(v) := -\left(\frac{1 - \frac{(r+v)f(r)}{rf(r+v)}}{v^2}\right)^{-\frac{1}{2}} \qquad (k \geqslant 1). \tag{4}$$

This succinct expression for c_k shows that the Lagrangean form (1) is not only useful for computing the Taylor expansion of ϕ at $z = 0$ (as is most commonly used), but also at the dominant singularity in subcritical situations (the latter is little known). The singular Lagrangean form (4) can further be used to derive the asymptotic behavior of c_k (although in

most cases the sub-criticality condition C3 fails), which in turn will be helpful in determining the number of terms used in order to reduce the numerical errors; see Section 4–6 for the discussion of some examples.

Let $\tau := \sqrt{\frac{2f(r)}{f''(r)}}$. Then we have (with $f_j := f^{(j)}(r)$)

$$\frac{c_1}{\tau} = -1 \quad \text{and} \quad \frac{c_3}{\tau^3} = -\frac{1}{8r^2} - \frac{f_2}{4f_0} + \frac{f_3}{4rf_2} - \frac{5f_3^2}{72f_2^2} + \frac{f_4}{24f_2}.$$

While the expressions of c_k are becoming messy as k increases, the neat expression (4) is not commonly available in most asymptotic expansions and reflects certain intrinsic properties of the Lagrangean form (1).

The asymptotic expansion (3) is to be compared with the usual one (see [9, Theorem VI.6]):

$$[z^n]\phi(z) \sim \rho^{-n} \sum_{k \geq 0} d_k n^{-k-\frac{3}{2}}, \tag{5}$$

where the coefficients d_k can be computed recursively but no simple expression such as (4) is available; see for example the next section for the usual constructive procedures to compute d_k. Alternatively, we can convert (3) to (5) by the following argument. Recall first Euler's reflection formula

$$\binom{n-k-\frac{3}{2}}{n} = \frac{\Gamma\left(n-k-\frac{1}{2}\right)}{n!\,\Gamma\left(-k-\frac{1}{2}\right)} = \frac{(-1)^{k+1}\Gamma\left(k+\frac{3}{2}\right)}{\pi} \cdot \frac{\Gamma\left(n-k-\frac{1}{2}\right)}{\Gamma(n+1)}. \tag{6}$$

Then we need the following asymptotic expansion.

▶ **Lemma 2** ([17]). *For $\alpha \in \mathbb{C}$*

$$\frac{\Gamma(z+\alpha)}{\Gamma(z)} \sim \sum_{j \geq 0} \lambda_j(\alpha) z^{\alpha-j} \qquad (|z| \to \infty),$$

uniformly for $|\arg(z)| \leq \pi - \varepsilon$, $\varepsilon > 0$. Here $\lambda_0(\alpha) = 1$ and

$$\lambda_j(\alpha) = \frac{1}{j} \sum_{0 \leq l < j} \binom{\alpha-l}{j+1-l} \lambda_l(\alpha) \qquad (j \geq 1).$$

This expression of $\lambda_j(\alpha)$ is simpler than that given in [8, Proposition 1]. Applying Lemma 2 to (6), we obtain

$$\binom{n-k-\frac{3}{2}}{n} \sim \frac{(-1)^{k+1}\Gamma\left(k+\frac{3}{2}\right)}{\pi} \sum_{j \geq 0} \lambda_j\left(-k-\frac{1}{2}\right) n^{-k-\frac{3}{2}-j},$$

from which we deduce the relation between d_k and c_k, which in turn results in the effective version (5) of [9, Theorem VI.6].

▶ **Theorem 3.** *Assume that ϕ and f satisfy* (1). *Then, under the conditions C1–C3, the expansion* (5) *holds with*

$$d_k = \frac{1}{\pi} \sum_{0 \leq j \leq k} (-1)^j c_{2j+1} \Gamma\left(j+\frac{3}{2}\right) \lambda_{k-j}\left(-j-\frac{1}{2}\right) \qquad (k \geq 0). \tag{7}$$

In view of the computational complexity of the coefficients, the expansion (3) is preferable and recommended for most numerical purposes because the binomial coefficients can be easily computed in most softwares.

On the other hand, the expansion (3) can be extended to a more general context of the form (or Lagrange-Bürmann formula)

$$[z^n]G(\phi(z)) = n^{-1}[t^{n-1}]G'(t)f(t)^n \qquad (n \geqslant 1).$$

▶ **Theorem 4.** *Let G be an analytic function in $|z| \leqslant r$. Under the conditions C1–C3, we have*

$$[z^n]G(\phi(z)) \sim \rho^{-n} \sum_{k \geqslant 0} e_{2k+1} \binom{n - k - \frac{3}{2}}{n}, \tag{8}$$

where $e_k = k^{-1}[v^{k-1}]G'(r + v)F(v)^k$ for $k \geqslant 1$.

Since $e_1 = -G'(r)\tau$, we see that $e_1 = 0$ when $G'(r) = 0$ (very common in the context of planar maps [1]), and we then get the $n^{-\frac{5}{2}}$-asymptotics

$$[z^n]G(\phi(z)) \sim \rho^{-n} \sum_{k \geqslant 0} e_{2k+3} \binom{n - k - \frac{5}{2}}{n},$$

where in particular (with $G_j = G^{(j)}(r)$) $\frac{e_3}{\tau^3} = -\frac{G_3}{6} - \frac{G_2}{2r} + \frac{f_3 G_2}{6 f_2}$. The usefulness of the two expansions (3) and (8) will be demonstrated through a few examples of trees and planar maps.

In the next section, we give a procedure to compute the coefficients d_k in (5). Then we prove (3) and (8) in Section 3. Some applications are discussed in the remaining sections.

2   An asymptotic expansion by saddle-point method

For comparison and for more methodological interest, we derive (5) in this section by a direct saddle-point method. Since the analysis is standard (see [9, 14]), we focus on the computation of the asymptotic coefficients d_k as follows.

1. Compute first the expansion $f(re^v) = \sum_{k \geqslant 0} d_k^{[1]} v^k$, where ($S(k, j)$ being Stirling numbers of the second kind)

$$d_k^{[1]} = \frac{1}{k!} \sum_{0 \leqslant j \leqslant k} S(k, j)f^{(j)}(r)r^j \qquad (k \geqslant 0).$$

2. Expand $\log f(re^v) = \sum_{k \geqslant 0} d_k^{[2]} v^k$, where $d_0^{[2]} = \log f(r)$ and

$$d_k^{[2]} = \frac{d_k^{[1]}}{f(r)} - \frac{1}{kf(r)} \sum_{1 \leqslant j < k} j d_j^{[2]} d_{k-j}^{[1]} \qquad (k \geqslant 2).$$

By $rf'(r) = r$, we see that $d_1^{[2]} = 1$ and $d_2^{[2]} = \frac{r^2 f''(r)}{2f(r)}$.

3. Now expand

$$\exp\left(vx + \frac{1}{2d_2^{[2]}} \sum_{j \geqslant 1} d_{j+2}^{[2]} v^{j+3} x^j\right) = \sum_{k \geqslant 0} d_k^{[3]}(v)x^k,$$

where $d_0^{[3]} = 1$ and

$$d_k^{[3]}(v) = \frac{v}{k} d_{k-1}^{[3]}(v) + \frac{1}{2d_2^{[2]}k} \sum_{1 \leqslant j \leqslant k} j v^{j+2} d_{j+2}^{[2]} d_{k-j}^{[3]}(v) \qquad (k \geqslant 1).$$

Note that $d_k^{[3]}(v)$ contains only powers of v with the same parity as k of degree $3k$.

4. Then, with $\sigma := r\sqrt{\frac{f''(r)}{f(r)}}$,

$$[z^{n-1}]f(z)^n \sim \frac{r^{1-n}f(r)^n}{\sqrt{2\pi n}\,\sigma}\left(1 + \sum_{k \geqslant 1}\frac{d_k^{[4]}}{\sigma^{2k}n^k}\right),$$

where

$$d_k^{[4]} := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2}t^2}d_{2k}^{[3]}(it)\,\mathrm{d}t = \sum_{0 \leqslant j \leqslant 3k}(-1)^j\frac{(2j)!}{j!2^j}[v^{2j}]d_{2k}^{[3]}(v).$$

Thus, by comparing with (5), we have (with $\tau := \sqrt{\frac{2f(r)}{f''(r)}}$)

$$d_k = \frac{r}{\sqrt{2\pi}\,\sigma^{2k+1}}d_k^{[4]} = \frac{\tau^{2k+1}}{2^{k+\frac{1}{2}}r^{2k}}d_k^{[4]}. \tag{9}$$

A non-recursive procedure is possible via Bell polynomials but not simpler; see [7] and the references therein. In particular (with $f_j := f^{(j)}(r)$), $d_1 = \frac{1}{8} - \frac{3r^2f_2}{8f_0} + \frac{rf_3}{4f_2} - \frac{r^2f_3^2}{24f_2^2} + \frac{r^2f_4}{8f_2}$. For larger k the expressions of d_n become very messy.

3 An asymptotic expansion by singularity analysis

We prove (3) and (8) in this section by singularity analysis. As in the previous section, we focus on the computations of c_k, the analytic justification being done as in [9, Theorem VI.6]. Following the exposition there, the idea starts from the equation (writing $w = \phi(z)$)

$$\rho - z = \frac{r}{f(r)} - \frac{w}{f(w)}. \tag{10}$$

Then invert (10) by expanding w in terms of $\rho - z$. For convenience, we find that it is simpler to use the expansion

$$1 - \frac{wf(r)}{rf(w)} = \sum_{j \geqslant 2}b_j(w - r)^j. \tag{11}$$

In particular, we have (with $f_j = f^{(j)}(r)$)

$$b_2 = \frac{f_2}{2f_0}, \quad b_3 = -\frac{b_2}{r} + \frac{f_3}{6f_0}, \quad b_4 = -\frac{b_3}{r} - \frac{f_2^2}{4f_0^2} + \frac{f_4}{24f_0}.$$

Now write $w = r + t$, and rearrange the expansion (11) as

$$f(r+t) - \frac{f(r)}{r}(r+t) = f(r+t)\sum_{j \geqslant 2}b_jt^j,$$

which then leads to the recurrence

$$b_m = \frac{f_m}{m!f_0} - \sum_{2 \leqslant j \leqslant m-2}b_{m-j}\frac{f_j}{f_0j!} - \frac{b_{m-1}}{r} \qquad (m \geqslant 3).$$

These coefficients can be computed in linear time (in m) once the derivatives of f at r are available.

Let $\Delta := \sqrt{1-z}$. We now examine the local behavior of Δ for $z \sim 1$ by first inverting the relation

$$\Delta^2 = 1 - \frac{(r+t)f(r)}{rf(r+t)} = \sum_{j \geqslant 2} b_j t^j,$$

or

$$\Delta^2 = t^2 \sum_{j \geqslant 0} b_{j+2} t^j \quad \Longrightarrow \quad t = \Delta F(t), \tag{12}$$

where

$$F(t) := -\left(\sum_{j \geqslant 0} b_{j+2} t^j\right)^{-\frac{1}{2}} = -\left(\frac{1 - \frac{(r+t)f(r)}{rf(r+t)}}{t^2}\right)^{-\frac{1}{2}}.$$

Interestingly, this is again of a Lagrangean form, and we see that

$$t = \sum_{k \geqslant 1} c_k \Delta^k,$$

where c_k is given in (4). Then we are led to the singular expansion

$$t = \phi(\rho z) - r = \sum_{k \geqslant 1} c_k (1-z)^{\frac{k}{2}},$$

which is convergent in a neighborhood of unity excluding the branch-cut $[1, \infty)$ (the exact range depending on the zeros or singularities of F). Then, by singularity analysis, we obtain (3).

The proof for (8) is similar, because

$$G(\phi(\rho z)) = G(r+t) = G(r) + \sum_{k \geqslant 1} e_k \Delta^k.$$

4   Applications I: $[z^n]\phi(z)$ and the $n^{-\frac{3}{2}}$-asymptotics

We discuss in this section the use of our asymptotic expansions in some popular counting sequences in combinatorics.

The following simple observation is useful for justifying sub-criticality of the Lagrangean form (1); see also [9, Proposition IV.5] for a slightly more general version.

▶ **Lemma 5** (Sub-criticality). *Let the radius of convergence of the series $f(z) = \sum_{j \geqslant 0} a_j z^j$ be $R > 0$ with $a_0 > 0$ and $a_j \geqslant 0$ for $j \geqslant 1$. If f is not linear and $\lim_{z \to R} f(z) \to \infty$, then the condition C3 is satisfied, namely, there exists an $r \in (0, R)$ such that $rf'(r) = f(r)$.*

Proof. Consider the function $g(z) := \frac{z}{f(z)}$, which is well-defined at least in $[0, R)$. Since $g(0) = 0$ (because $a_0 > 0$) and $\lim_{z \to R} g(z) \to 0$, by Rolle's Theorem, there exists an $r \in (0, R)$ such that $g'(r) = 0$. But $g'(r) = 0$ is equivalent to $rf'(r) - f(r) = 0$. ◀

In particular, if f is a rational function of z with nonnegative Taylor coefficients, then the Lagrangean form (1) is always sub-critical. For example, $f(z) = \frac{1}{1-z}$ gives the Catalan numbers $\phi(z) = \frac{1-\sqrt{1-4z}}{2}$ (A000108 in Sloane's OEIS). We see that no further treatment is needed because the singular expansion contains only one term.

Motzkin numbers (A001006).

Consider now the Motzkin numbers with $f(z) = 1 + z + z^2$ and

$$\phi(z) = \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z}. \tag{13}$$

Note that $zf'(z) - f(z) = z^2 - 1$, implying that $r = 1$ and $\rho = \frac{1}{3}$. Thus

$$[z^n]\phi(z) = \sum_{0 \leqslant j \leqslant \lfloor \frac{n-1}{2} \rfloor} \frac{(n-1)!}{j!(j+1)!(n-1-2j)!} \sim 3^n \sum_{k \geqslant 0} c_{2k+1} \binom{n - k - \frac{3}{2}}{n}. \tag{14}$$

For finite k, the coefficients c_k can be readily computed by (4) with $F(v) = -\sqrt{3 + 3v + v^2}$. We observe that while the asymptotics of the left-hand side of (14) remains less visible even for the exponential order, that of the right-hand side is transparent if we regard the binomial factor as decreasing powers in n. Furthermore, the right-hand side is a direct consequence of Theorem 1, and thus even without any explicit formula for $[z^n]\phi(z)$, which is often the case, we can still apply the expansion (3) and obtain very effective approximations.

We now look at the large k-asymptotics of c_k. Note that $vF'(v) - F(v) = \frac{3(2+v)}{3\sqrt{3+3v+v^2}}$, which equals zero when $v = -2$, and has a pair of conjugate singularities at $-\frac{3}{2} \pm \frac{\sqrt{3}}{2}i$ with modulus $\sqrt{3} < 2$, so the Lagrangean form (12) is not sub-critical (and thus the saddle-point at -2 is not dominant). Indeed, by the closed-form expression (13) of ϕ, we have

$$t = \Delta F(t) \quad \implies \quad t = \frac{3\Delta^2 - \Delta\sqrt{3(4 - \Delta^2)}}{2(1 - \Delta^2)}.$$

This implies that

$$c_{2k+1} = -\sqrt{3} \left(1 - 2 \sum_{1 \leqslant j \leqslant k} \binom{2j - 2}{j - 1} \frac{16^{-j}}{j} \right) = -\frac{3}{2} + O\left(k^{-\frac{3}{2}} 16^{-k} \right).$$

Thus they can be replaced by $-\frac{3}{2}$ for moderate values of k (depending on the desired degree of precision).

Schröder numbers (A001003).

In this case, we have $f(z) = \frac{1-z}{1-2z}$ and

$$\phi(z) = \frac{1 + z - \sqrt{1 - 6z + z^2}}{4}, \tag{15}$$

implying that $r = 1 - \frac{1}{\sqrt{2}} < \frac{1}{2}$, $\rho = \frac{r}{f(r)} = 3 - 2\sqrt{2}$, and ϕ has the dominant singularity at $3 - 2\sqrt{2}$. Furthermore, $F(v) = -r\sqrt{\frac{1}{\sqrt{2}} - v}$, and therefore the equation $vF'(v) = F(v)$ has a solution at $v = \sqrt{2} > \frac{1}{\sqrt{2}}$ lying outside the circle where F is analytic. So the Lagrangean form $t = \Delta F(t)$ is not sub-critical. On the other hand, t can be solved in terms of Δ as

$$t = \Delta F(t) \quad \implies \quad t = -\frac{(\sqrt{2} - 1)^2\Delta^2 + (\sqrt{2} - 1)\Delta\sqrt{(\sqrt{2} - 1)^2\Delta^2 + 4\sqrt{2}}}{4}.$$

Thus

$$c_{2k+1} = [\Delta^{2k+1}]t = \frac{(-1)^{k+1}}{2^{\frac{3}{4}}} \binom{k - \frac{3}{2}}{k} 2^{-\frac{5}{2}k} (\sqrt{2} - 1)^{2k+1}.$$

We get the same $k^{-\frac{3}{2}}$-asymptotics for the coefficients c_k even though $t = \Delta F(t)$ is not sub-critical. Note that c_{2k+1} are asymptotic to $c' k^{-\frac{3}{2}} \rho_c^{-k}$ for large k, where $\rho_c \approx 32.97$, meaning that they converge exponentially to zero. By the closed-form expression for Schröder numbers, we have the asymptotic expansion

$$[z^n]\phi(z) = \frac{1}{n} \sum_{0 \leqslant j \leqslant n-2} \binom{n-2}{j}\binom{n}{j+1} 2^{n-2-j} \sim (3+2\sqrt{2})^n \sum_{k \geqslant 0} c_{2k+1}\binom{n-k-\frac{3}{2}}{n}.$$

Again, the right-hand side is preferable for large-n numerics and the numerical fits are very good even for small values of n. For example, for $n = 10$, $\left| \frac{[z^n]\phi(z)}{(3+2\sqrt{2})^n} - c_1\binom{n-\frac{3}{2}}{n}\right| \leqslant 6.2 \times 10^{-6}$.

The same approach applies to many other sequences with f a polynomial or a rational form. Indeed, several hundred of sequences were found in OEIS whose generating functions satisfy the Lagrangean form (1) with polynomial or rational f. Some of these will be compiled and discussed in the journal version.

5   Applications II: $[z^n]G(\phi(z))$ and the $n^{-\frac{5}{2}}$-asymptotics

A map is an *embedding of a connected planar multigraph on the sphere*, up to orientation preserving homeomorphism. Asymptotic enumeration of planar maps often features a universal $n^{-\frac{5}{2}}$-behavior, in contrast to $n^{-\frac{3}{2}}$ for that of trees; see [1, 2] for more information and references. Given a class $\mathscr{M}$ of maps, let m_n denote the number of maps in $\mathscr{M}$ with n edges. Let $M(z) := \sum_{n \geqslant 0} m_n z^n$ be the generating function of m_n, which is specified by the Lagrangean form

$$M(z) = G(\phi(z)), \qquad \phi(z) = z f(\phi(z)). \tag{16}$$

As the number of different types of maps is huge (see, e.g., [2, 4, 10, 11, 12, 18]), we content ourselves in this extended abstract only with the discussion of Table 2 in [1] (a total of 14 examples reformatted below with a correction for $\mathscr{M}_2$) for representative asymptotic patterns, focusing on the calculations of the asymptotic coefficients e_k, a missing facet in most previous publications. See [1] for precise definitions of the diverse terms used here (such as non-separable, bridgeless, singular, irreducible, etc.).

Incremental maps

We first categorize the 14 examples into two groups according to the availability for the counting function $m_n = r_n m_{n-1}$ for some rational function r_n—such a counting formula entails a Markovian property and in turn an incremental construction procedure. This group includes (see the following tables) the six examples $\mathscr{M}_1, \mathscr{M}_2, \mathscr{M}_3, \mathscr{B}_1, \mathscr{T}_1$ and $\mathscr{T}_2$. Note that $\mathscr{M}_2 = \mathscr{T}_2$; see [3]. Since the application of our expansions is straightforward, we omit the details of the expansions.

type of maps	$\mathscr{M}_1$ general	$\mathscr{M}_2$ bridgeless $= \mathscr{T}_2$ triangulations	$\mathscr{M}_3$ non-separable	$\mathscr{B}_1$ bipartite	$\mathscr{T}_1$ singular triangulations
$f(z)$	$3(1+z)^2$	$(1+z)^4$	$(1+z)^3$	$2(1+z)^2$	$2(1+z^3)$
$[z^n]\phi(z)$ or $\phi(z)$	$\frac{3^n}{n+1}\binom{2n}{n}$	$\frac{1}{3n+1}\binom{4n}{n}$	$\frac{1}{2n+1}\binom{3n}{n}$	$\frac{2^n}{n+1}\binom{2n}{n}$	$\frac{2^n}{2n+1}\binom{3n}{n}$
OEIS(ϕ)	A005159	A002293	A001764	A151374	A153231

type of maps	$\mathcal{M}_1$ general	$\mathcal{M}_2$ bridgeless $=\mathcal{T}_2$ triangulations	$\mathcal{M}_3$ non-separable	$\mathcal{B}_1$ bipartite	$\mathcal{T}_1$ singular triangulations
$G(z)$	$\frac{z(2-z)}{3}$	$(1-z)(1+z)^2$	$\frac{z(2+z-z^2)}{(1+z)^3}$	$\frac{z(2-z)}{4}$	$\frac{z(1-z)}{2}$
$[z^n]G(\phi(z))$	$\frac{2\cdot 3^n(2n)!}{n!(n+2)!}$	$\frac{2\cdot(4n+1)!}{(n+1)!(3n+2)!}$	$\frac{2(3n)!}{(n+1)!(2n+1)!}$	$\frac{3\cdot 2^{n-1}(2n)!}{n!(n+2)!}$	$\frac{2^n(3n)!}{(n+1)!(2n+1)!}$
OEIS$(G(\phi))$	A000168	A000260	A000139	A000257	A000309
r	1	$\frac{1}{3}$	$\frac{1}{2}$	1	$\frac{1}{2}$
ρ	$\frac{1}{12}$	$\frac{27}{256}$	$\frac{4}{27}$	8	$\frac{2}{27}$
$F(t)$	$-(2+t)$	$-\frac{\sqrt{3}(4+3t)^2}{9\sqrt{32+16t+3t^2}}$	$-\frac{(3+2t)^{\frac{3}{2}}}{2\sqrt{9+2t}}$	$-(2+t)$	$-\frac{(3+2t)^{\frac{3}{2}}}{2\sqrt{9+2t}}$
$G'(r+v)$	$-\frac{2}{3}v$	$-v(4+3v)$	$\frac{-32v}{(3+2v)^3}$	$-\frac{1}{2}v$	$-v$

Non-incremental maps

The remaining eight maps are further divided into two sub-groups:

(1) $[z^n]f(z)\geqslant 0$: $\mathcal{M}_4$, $\mathcal{M}_6$, and $\mathcal{T}_3$;

(2) $[z^n]f(z)$ contains negative coefficients: $\mathcal{M}_5, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4$, and $\mathcal{B}_5$.

Our theorems in the introduction are directly applicable to the first sub-group, and can be readily modified for the second for which the condition C1 (nonnegativity of coefficients) fails.

Non-incremental maps with $[z^n]f(z)\geqslant 0$

Similar to the incremental maps given above, we summarize the major properties of the three non-incremental maps in the following table.

type of maps	$\mathcal{M}_4$ simple	$\mathcal{M}_6$ 3-connected	$\mathcal{T}_3$ irreducible triangulations
$f(z)$	$\frac{(3+z)^2}{3-z}$	$\frac{1}{1-z}$	$\frac{1}{(1-z)^2}$
$[z^n]\phi(z)$ or $\phi(z)$	$\frac{3-\sqrt{1-8z}}{2(1+z)}$	$\frac{1}{n}\binom{2n-2}{n-1}$	$\frac{(3n+1)!}{(n+1)!(2n+1)!}$
OEIS(ϕ)	A062062	A000108	A006013
$G(z)$	$\frac{z(9-3z-z^2)}{27}$	$\frac{z^5(1-z-z^2)}{(1+z)^3(1+z-z^2)}$	$\frac{z(1-z-z^2)}{(1-z)(1+z)^2}$
OEIS$(G(\phi))$	A022558	A000287	A000256
r	1	$\frac{1}{2}$	$\frac{1}{3}$
ρ	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{4}{27}$
$F(t)$	$-\frac{4+t}{3}$	$-\frac{1}{2}$	$\frac{-2}{3\sqrt{3(1-t)}}$
$G'(r+v)$	$\frac{-v(4+v)}{9}$	$-\frac{2v(1+2v)^4\left(167+176v-24v^2-64v^3-16v^4\right)}{(3+2v)^4(5-4v^2)^2}$	$\frac{-729v}{(2-3v)^2(4+3v)^3}$

We consider only the simple maps $\mathcal{M}_4$ with $f(z) = \frac{(3+z)^2}{3-z}$ and $G(z) = \frac{z(9-3z-z^2)}{27}$. In this case, ϕ is given by

$$\phi(z) = \frac{3}{2} \cdot \frac{1 - 2z - \sqrt{1-8z}}{1+z},$$

implying that

$$G(\phi(z)) = -1 + \frac{4}{1+z} + \frac{18}{(1+z)^2} - \frac{27}{2(1+z)^3} + \frac{(1-8z)^{\frac{3}{2}}}{2(1+z)^3}, \tag{17}$$

which then gives

$$[z^n]G(\phi(z)) = (-1)^n \left(\frac{1}{2} - \frac{9}{4}n - \frac{27}{4}n^2 \right) + \frac{1}{2} \sum_{0 \leqslant j \leqslant n} \binom{j+2}{2}(-1)^j \binom{n-j-\frac{5}{2}}{n-j} 8^{n-j},$$

for $n \geqslant 1$. On the other hand, we also have, by Lagrange inversion formula,

$$[z^n]G(\phi(z)) = 2 \sum_{0 \leqslant j < n} \frac{(2n+1)!(2n-j-2)!(n-2j+1)}{n!j!(n-1-j)!(2n-j+2)!} \qquad (n \geqslant 1).$$

The main difference is that the former expands at $z = \frac{1}{8}$, while the latter at the origin.

On the other hand, without relying on the explicit forms, we also have (with $\Delta = \sqrt{1-z}$)

$$F(t) = -\frac{4+t}{3} \quad \Longrightarrow \quad t = -\frac{4\Delta}{3+\Delta} \Longrightarrow G(\phi(\tfrac{z}{8})) = -1 + \frac{32}{3+\Delta^2} - \frac{64}{(3+\Delta)^3}.$$

We then obtain the same singular expansion as above, which is convergent in the region with $|1-8z| < 9$. It follows that

$$[z^n]G(\phi(z)) \sim 8^n \sum_{k \geqslant 0} e_{2k+3} \binom{n-k-\frac{5}{2}}{n}, \quad \text{where } e_{2k+3} = \frac{1}{2} \left(\frac{8}{9} \right)^3 \binom{k+2}{2} 9^{-k}.$$

Non-incremental maps with $[z^n]f(z) \lessgtr 0$

The remaining five cases are listed below.

type of maps	$\mathcal{M}_5$ non-separable simple	$\mathcal{B}_2$ bipartite simple	$\mathcal{B}_3$ bipartite bridgeless	$\mathcal{B}_4$ bipartite non-separable	$\mathcal{B}_5$ bipartite non-separable simple
$f(z)$	$\frac{(1+z)^6}{(1+2z)^2}$	$\frac{8(1+z)^2}{4+2z-z^2}$	$\frac{(2+z)^6}{32(1+z)^2}$	$\frac{32(1+z)^2}{(4+2z-z^2)^2}$	$\frac{128(1+z)^2}{(4+2z-z^2)^3}$
$G(z)$	$\frac{z(1+z-z^2)}{(1+z)^3}$	$\frac{z(2-z)}{4}$	$\frac{z^2(8+4z-4z^2-z^3)}{32(1+z)^2}$	$\frac{z(2-z)}{4}$	$\frac{z(2-z)}{4}$
OEIS($G(\phi)$)	$-$	$-$	$-$	A069728	A298358
r	$\frac{1}{2}$	1	1	1	1
ρ	$\frac{128}{729}$	$\frac{5}{32}$	$\frac{128}{729}$	$\frac{25}{128}$	$\frac{25}{512}$
$F(t)$	$(\star)_1$	$-\frac{\sqrt{5}(2+t)}{\sqrt{9+4t}}$	$(\star)_2$	$-t$	$(\star)_3$
$G'(r+v)$	$\frac{-64v(1+v)}{9}$	$-\frac{1}{2}v$	$-\frac{v(1+v)(v+3)^2(7+3v)}{32(2+v)^3}$	$-\frac{1}{2}v$	$-\frac{1}{2}v$

Here $(\star)_1 := \frac{-(3+2t)^3}{\sqrt{1215+2862t+2160t^2+576t^3+64t^4}}$, $(\star)_2 := \frac{-2(3+t)^3}{\sqrt{1215+1431t+540t^2+72t^3+4t^4}}$, and $(\star)_3 :=$ $\frac{-5\sqrt{5}(2+t)}{\sqrt{425+300t-60t^2-60t^3+4t^4+4t^5}}$.

We now show how to extend the same analysis to the cases when $[z^n]f(z)$ contains negative coefficients. From the viewpoint of the saddle-point method, a sufficient condition replacing the condition C1 is as follows (see also [6]):

C1' (Concentration of $|f(z)|$)

$$f(r) > 0 \text{ for } 0 < r < R \text{ and } |f(re^{i\theta})| < f(r) \text{ for } 0 < |\theta| < \pi. \tag{18}$$

Briefly, this condition implies, by the saddle-point method, that the major contribution to the integral representation of $[t^{n-1}]G'(r+t)f(t)^n$ comes from a small neighborhood of $t = r$, and in turn that the asymptotic expansion (8) holds.

$\mathscr{B}_2$: simple bipartite maps

In this case, f and G are given by $f(z) = \frac{8(1+z)^2}{4+2z-z^2}$ and $G(z) := \frac{1}{4}z(2-z)$; thus $r = 1$, $\rho = \frac{5}{32}$, and

$$F(t) = -\frac{\sqrt{5}(2+t)}{\sqrt{9+4t}}.$$

We can check the condition C1' by elementary calculus and then derive the expansion (8); in particular, we have

$$\frac{1}{\sqrt{5}}\left(\frac{5}{32}\right)^n [z^n]G(\phi(z)) \sim \frac{50}{243}\binom{n-\frac{5}{2}}{n} + \frac{1100}{59049}\binom{n-\frac{7}{2}}{n} + \cdots.$$

Whether the left-hand side is easy to compute or not is irrelevant here. Also we can compute e_{2k+3} by the following closed-form expression

$$e_{2k+3} = \frac{2 \cdot 5^{k+\frac{3}{2}}}{(2k+3)3^{2k+3}} \sum_{0 \leqslant j \leqslant 2k+1} \binom{k+j+\frac{1}{2}}{j}\binom{2k+3}{j+2}\left(-\frac{8}{9}\right)^j \qquad (k \geqslant 0).$$

The same technique applies to $\mathscr{B}_4$ and $\mathscr{B}_5$, but not to $\mathscr{M}_5$ and $\mathscr{B}_3$ because the condition C1' fails when z is near $-\frac{1}{2}$ (for $\mathscr{M}_5$) and near -1 (for $\mathscr{B}_3$) because of polar singularities there. However, we can modify suitably the integration contour so as to avoid the polar singularities and prove that the contribution comes principally from $z \sim r$ in the corresponding Cauchy integral. So we still get the expansions (8) for both cases; we omit the details here.

6 Applications III: Other examples

Due to a space constraint, we mention in this section only two interesting examples for which our expansions apply.

- The number of rooted 3-connected bicubic maps of $2n$ vertices (see [15, 16] and A298358) is given by

$$m_n = \frac{3}{n-1}[z^{n-1}]g(z)f(z)^n,$$

where $f(z) = \frac{(1+2z)^2}{(1+z-z^2)^3}$ and $g(z) = \frac{z^3(1-2z)}{(1+2z)(1+z-z^2)}$. By modifying our expansions (the condition C1' holds), we have

$$m_n \sim \frac{3n}{n-1}\left(\frac{512}{125}\right)^n \sum_{k\geqslant 0} e_{2k+3}\binom{n-k-\frac{5}{2}}{n},$$

where

$$e_k := \frac{1}{k}[t^{k-1}]g\left(\tfrac{1}{2}+t\right)\left(\frac{-5\sqrt{5}(1+t)}{\sqrt{425-240t^2+64t^4+600t-480t^3+128t^5}}\right)^k.$$

- The number of labeled rooted trees of subsets of an n-set (see [5] or A005172) is given by $[z^n]G(\phi(z))$, where $\phi = zf(\phi)$ with $f(z) = \frac{z}{1-e^z+\log(2e^z-1)}$ and $G(z) = e^z - 1$. Interestingly, all coefficients $[z^n]f(z)$ are positive for $1 \leqslant n \leqslant 47$, but negative coefficients appear from $n = 48$ on. By checking the conditions C1', C2 and C3, we then obtain (8) with

$$F(t) = -\frac{\sqrt{2\log 2 - 1}\, t}{\sqrt{2\log 2 + 3(e^t - 1) - 2\log(3e^t - 1)}}.$$

Then we deduce that (with $r = \log\frac{3}{2}$)

$$G(r+t) = \tfrac{2}{3}\left(T\left(e^{-1-(\log 2 - \frac{1}{2})\Delta^2}\right) - 1\right),$$

where $T(z) = \sum_{n\geqslant 1}\frac{n^{n-1}}{n!}z^n$ denotes the generating function for Cayley trees.
- All examples treated in [16] can be dealt with by our expansions. Consider, as in [16], the asymptotics of the Stirling numbers of the second kind:

$$\begin{Bmatrix}2n\\n\end{Bmatrix} = \frac{(2n)!}{n!}[z^{2n}](e^z - 1)^n = \frac{(2n)!}{(n-1)!}\cdot\frac{1}{n}[z^{n-1}]z^{-1}f(z)^n,$$

where $f(z) := \frac{e^z-1}{z}$. Although $G = \log z$ is not analytic at $z = 0$, we can still apply the same expansion (8) and obtain

$$\frac{(n-1)!}{(2n)!}\begin{Bmatrix}2n\\n\end{Bmatrix} \sim \rho^{-n}\sum_{j\geqslant 0} e_{2j+1}\binom{n-j-\frac{3}{2}}{n},$$

where $r = 2 - T(2e^{-2})$, $\rho = \frac{r}{f(r)}$, and $e_k = k^{-1}[t^{k-1}](r+t)^{-1}F(t)^k$ for $k \geqslant 1$. The Stirling numbers of the first kind with $f = -z^{-1}\log(1-z)$ is similar.
Other examples in [16] include the relations

$$\frac{1}{n\cdot(n-1)!^2}\sum_{0\leqslant k<n}\begin{Bmatrix}n-1\\k\end{Bmatrix}(n-1+k)! = \frac{1}{n}[z^{n-1}](2-e^z)^{-n}$$

$$\frac{1}{n}\sum_{0\leqslant k<n}\binom{n}{k}\binom{2n-2-k}{n-1}(-1)^k 2^{n-k} = \frac{1}{n}[z^{n-1}]\frac{(1+z)^n(1-2z)^n}{(1+z)^2},$$

and the following table for the form $[z^n]g(z)f(z)^n$:

g	f	g	f
1	$(1-\frac{3}{8}z+\frac{1}{24}z^2)^{-1}$	1	$(1-\sqrt{3}z+z^2)^{-1}$
1	$(1-z-z^2)^{-1}$	1	$\frac{1}{2}(1+e^z)$
$z(1-z)^{-1}$	$e^{\alpha z}$	$z(1-z)^{-\alpha-1}$	e^z
$(1-\alpha z)^{-1}$	e^z	$(1-z)^{-1}$	$(1+z)^\alpha$
$-\log(1-3z)$	$(1+z)^4$	$-z(1-z)^{-1}\log(1-z)$	e^z

─────── **References** ───────

1   C. Banderier, P. Flajolet, G. Schaeffer, and M. Soria. Random maps, coalescing saddles, singularity analysis, and airy phenomena. *Random Struct. Algorithms*, 19(3-4):194–246, 2001. `doi:10.1002/rsa.10021`.

2   E. A Bender and L. B. Richmond. A survey of the asymptotic behaviour of maps. *Journal of Combinatorial Theory, Series B*, 40(3):297–329, 1986. `doi:10.1016/0095-8956(86) 90086-9`.

3   O. Bernardi and E. Fusy. A bijection for triangulations, quadrangulations, pentagulations, etc. *Journal of Combinatorial Theory, Series A*, 119(1):218–244, 2012. `doi:10.1016/j. jcta.2011.08.006`.

4   M. Bousquet-Mélou and G. Schaeffer. Enumeration of planar constellations. *Advances in Applied Mathematics*, 24(4):337–368, 2000. `doi:10.1006/aama.1999.0673`.

5   F. Chapoton, F. Hivert, and J.-C. Novelli. A set-operad of formal fractions and dendriform-like sub-operads. *Journal of Algebra*, 465:322–355, 2016. `doi:10.1016/j.jalgebra.2016. 07.001`.

6   V. De Angelis. Asymptotic expansions and positivity of coefficients for large powers of analytic functions. *Int. J. Math. Math. Sci.*, 2003(16):1003–1025, 2003. `doi:10.1155/ S0161171203205056`.

7   M. Drmota. A Bivariate Asymptotic Expansion of Coefficients of Powers of Generating Functions. *Europ. J. Combinatorics*, 15:139–152, 1994.

8   P. Flajolet and A. M. Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990. `doi:10.1137/0403019`.

9   P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, New York, NY, USA, 1 edition, 2009.

10   Zh. Gao and N. C. Wormald. Enumeration of rooted cubic planar maps. *Annals of Combinatorics*, 6(3):313–325, Dec 2002. `doi:10.1007/s000260200006`.

11   V. A. Liskovets and T. R. Walsh. Counting unrooted maps on the plane. *Adv. in Appl. Math.*, 36(4):364–387, 2006. `doi:10.1016/j.aam.2005.03.006`.

12   Y. Liu. *Enumerative theory of maps*, volume 468 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht; Science Press Beijing, Beijing, 1999.

13   A. Meir and J. W Moon. On the altitude of nodes in random trees. *Canadian Journal of Mathematics*, 30(1978):997–1015, 1978. `doi:10.4153/CJM-1978-085-0`.

14   A. Meir and J. W. Moon. The asymptotic behaviour of coefficients of powers of certain generating functions. *European J. Combin.*, 11(6):581–587, 1990. `doi:10.1016/ S0195-6698(13)80043-1`.

15   R. C. Mullin, L. B. Richmond, and R. G. Stanton. An asymptotic relation for bicubic maps. In *Proceedings of the Third Manitoba Conference on Numerical Mathematics (Winnipeg, Man., 1973)*, pages 345–355. Utilitas Math., Winnipeg, Man., 1974.

16   R. Sprugnoli and M. C. Verri. Asymptotics for lagrange inversion. *Pure Mathematics and Applications*, 5(1):79–104, 1994. URL: `http://EconPapers.repec.org/RePEc:cmt: pumath:puma1994v005pp0079-0104`.

17   F. G. Tricomi and A. Erdélyi. The asymptotic expansion of a ratio of gamma functions. *Pacific J. Math.*, 1:133–142, 1951. URL: `http://projecteuclid.org/euclid.pjm/ 1102613160`.

18   W. T. Tutte. On the enumeration of planar maps. *Bull. Amer. Math. Soc.*, 74:64–74, 1968. `doi:10.1090/S0002-9904-1968-11877-4`.

Periods of Iterations of Mappings over Finite Fields with Restricted Preimage Sizes

Rodrigo S. V. Martins
Universidade Tecnológica Federal do Paraná, Apucarana, PR 86812-460, Brazil
rodrigomartins@utfpr.edu.br

Daniel Panario[1]
Carleton University, Colonel By Drive, Ottawa, ON K1S 5B6, Canada
daniel@math.carleton.ca

Claudio Qureshi[2]
Universidade Estadual de Campinas, Campinas, SP 13083-859, Brazil
cqureshi@ime.unicamp.br

Eric Schmutz
Drexel University, Philadelphia, Pa 19104, USA
Eric.Jonathan.Schmutz@drexel.edu

─── **Abstract** ───────────────────────────────

Let f be a uniformly random element of the set of all mappings from $[n] = \{1, \ldots, n\}$ to itself. Let $\mathbf{T}(f)$ and $\mathbf{B}(f)$ denote, respectively, the least common multiple and the product of the lengths of the cycles of f. Harris proved in 1973 that $\log \mathbf{T}$ converges in distribution to a standard normal distribution and, in 2011, Schmutz obtained an asymptotic estimate on the logarithm of the expectation of $\mathbf{T}$ and $\mathbf{B}$ over all mappings on n nodes. We obtain analogous results for uniform random mappings on $n = kr$ nodes with preimage sizes restricted to a set of the form $\{0, k\}$, where $k = k(r) \geq 2$. This is motivated by the use of these classes of mappings as heuristic models for the statistics of polynomials of the form $x^k + a$ over the integers modulo p, where k divides $p - 1$. We exhibit and discuss our numerical results on this heuristic.

1 Introduction

Let $f : [n] \to [n]$ be a mapping from a finite set to itself. The iterations of mappings has attracted interest in recent years due to applications in areas such as physics, biology, coding theory and cryptography. Every polynomial f over a finite field $\mathbb{F}_p$ is a particular case of a mapping, and there are a number of applications where one considers the iterations of polynomials over finite fields. We highlight Pollard's classical factorization method for integers, which is based on iterations of quadratic polynomials; it allowed Brent and Pollard to obtain the previously unknown factorization of the eighth Fermat number. The adaptation

---

of Pollard's method to the discrete logarithm problem also relies on iterations of mappings; it is considered by some authors the best attack on the elliptic curve version of this problem [20].

Let $f = f^{(0)}$ be a mapping on n elements and consider the sequence of functional compositions $f^{(m)} = f \circ f^{(m-1)}$, $m \geq 1$. The least integer $\mathbf{T} = \mathbf{T}(f)$ such that $f^{(m+T)} = f^{(m)}$ for all $m \geq n$ equals the order of the permutation obtained by restricting the mapping f to its cyclic vertices. Erdös and Turán proved in [8] that the logarithm of the corresponding random variable defined over the symmetric group S_n converges in distribution to a standard normal distribution, when properly centered and normalized. By adapting Erdős and Turán's "statistical group theory approach"[8], Harris was able to prove an analogous result for the space of mappings with uniform distribution [12]. The logarithm of the expected value of $\mathbf{T}$ was estimated in [18].

The parameter $\mathbf{T}$ can be proven to be the least common multiple of the cycle lengths of the components of the functional graph of f. If $\mathbf{B}(f)$ is the product of all cycle lengths of f including multiplicities, then it is clear that $\mathbf{B}(f)$ represents an upper bound for $\mathbf{T}(f)$; moreover, one might consider $\mathbf{B}$ as an approximation for $\mathbf{T}$. For instance, Proposition 1.2 of [18] implies that, for any $\delta > 0$, the sequence of nonnegative random variables $X_n = (\log \mathbf{B} - \log \mathbf{T})/\log^{1+\delta} n$, $n \geq 1$, converges in probability to zero. However, it is proved in [18] that the expectation of $\mathbf{B}$ deviates significantly from the expectation of $\mathbf{T}$.

In this paper we derive similar results for the classes of $\{0, k\}$-mappings, $k \geq 2$, defined as mappings $f : [n] \rightarrow [n]$ such that $|f^{-1}(y)| \in \{0, k\}$ for all $y \in [n]$. In [1, 14] the authors consider the case where k is a fixed integer. Although this case is arguably of the most interest due to connections with polynomials over finite fields, we derive our results in a more general context, as explained at the end of this section. This might be desirable, for example, when modeling polynomials whose degree depends on the size of the prime p; see [6].

By now there is a rather large literature on the asymptotic distribution of random variables defined on mappings, both with and without indegree restrictions. One motivation is methodological. Random mappings are important examples that serve as benchmarks for both probabilistic and analytic methods. On the analytic side, combinatorial methods can be used to identify generating functions whose coefficients are the quantities of interest. In many cases it is possible to estimate the coefficients asymptotically using complex analysis. A standard reference is [10], which includes several applications to random mappings; see also [7, 9, 13]. In another direction, random mappings correspond to a large class of random graphs G_f for which the joint distribution of components sizes can be realized as independent random variables, conditioned on the number of vertices that the graph has. Stein's method and couplings have been used to prove strong and general results [2, 3]. One application of this theory is a generalization of the theorem of Harris [12] that was mentioned above. However the proofs in our paper are elementary, and do not directly use any of these probabilistic techniques (except indirectly by citing a theorem from [4]).

The research on random mappings with such restrictions is also motivated by the Brent-Pollard heuristic, where one uses these objects as a model for the statistics of polynomials. It was introduced by Pollard in the analysis of his factorization method: he conjectured that quadratic polynomials modulo large primes behave like random mappings with respect to their average rho length [15]. However, the indegree distribution of a class of mappings impacts the asymptotic distribution of a number of parameters [1, 11]. Since it is known that the functional graph of a quadratic polynomial over $\mathbb{F}_p$, p odd, has just one node with indegree 1 and the remaining nodes are split in half between indegrees 0 or 2, $\{0, 2\}$-mappings could provide a better heuristic model for quadratic polynomials; see [14] for a discussion of alternative models for the Brent-Pollard heuristic. Furthermore, the class of $\{0, k\}$-mappings

provides a good heuristic model for polynomials of the form $x^k + a \in \mathbb{F}_p[x]$ with $p \equiv 1$ (mod k). This heuristic model was used in [5] to predict that Pollard's method is sped up in some cases if these polynomials are used, eventually leading to the factorization of the eighth Fermat number.

It is discussed in [14] that unrestricted mappings and $\{0,2\}$-mappings represent equally accurate models for the expected rho length of quadratic polynomials. This is the case because both classes of mappings present the same asymptotic average coalescence, defined as the variance of its distribution of indegrees under uniform distribution; see [1, 14]. For example, the coalescence λ of a $\{0,k\}$-mapping f on $n = kr$ nodes satisfies

$$\lambda = \sum_{y \in [n]} \frac{|f^{(-1)}(y)|^2}{n} - 1 = r\frac{k^2}{n} - 1 = k - 1.$$

It is curious that the knowledge of the indegree distribution of these polynomials does not represent an improvement on the heuristic. Thus asymptotic estimates for a different parameter, such as **B** or **T**, represents an interesting problem: it could provide a significant deviation between polynomials over finite fields and their heuristic models, or reinforce the similarities between these classes. We exhibit our numerical results on the behavior of **T** and **B** over different classes of polynomials over finite fields and investigate different classes of mappings as heuristic models for the behavior of **T** and **B** over these classes of polynomials.

Preliminaries and notation. For f a mapping, let $\mathcal{Z} = \mathcal{Z}(f)$ be the set of cyclic nodes of f and let $\mathbf{Z} = |\mathcal{Z}|$. To avoid confusion, we index probabilities and expected values by the set of allowed indegrees of the class of mappings in question: $\mathbb{N}$ in the unrestricted case [18] or $\{0,k\}$ in our case. For example, the expected value of **T** over all mappings on n nodes is denoted by $\mathbb{E}_n^{\mathbb{N}}(\mathbf{T})$, whereas $\mathbb{E}_n^{\{0,k\}}(\mathbf{T})$ denotes the expectation of **T** over $\{0,k\}$-mappings on n nodes. In this work we consider $\{0,k\}$-mappings on $n = kr$ elements, where r denotes the size of their range and $k = k(r)$ is a sequence of integers satisfying $k \geq 2$ for all $r \geq 1$. Although $n(r)$ and $k(r)$ are functions of r, we omit this dependence on our notation. We emphasize that *all asymptotic calculations and results in this work are taken as r approaches infinity*, unless said otherwise. We assume throughout the paper that, for some $0 < \alpha < 1$, $k = o(n^{1-\alpha})$ as r approaches infinity, or equivalently, $\log n = O(\log(\frac{n}{\lambda}))$ where $\lambda = k - 1$.

▶ Remark. Due to the lack of space all proofs are given in `https://arxiv.org/abs/1701.09148`

2   Expected Value of T and B

In this section we obtain asymptotic estimates for $\mathbb{E}_n^{\{0,k\}}(\mathbf{T})$ and $\mathbb{E}_n^{\{0,k\}}(\mathbf{B})$ following a similar strategy as in [18] with some differences that we describe next. It is known that the restriction of a random uniform mapping to its cyclic nodes represents a random uniform permutation. Therefore, if we let M_m be the expected order of a uniform random permutation of S_m, then the expected value of **T** over all $\{0,k\}$-mappings can be written as

$$\mathbb{E}_n^{\{0,k\}}(\mathbf{T}) = \sum_{m=1}^{n} \mathbb{P}_n^{\{0,k\}}(\mathbf{Z} = m)M_m. \tag{1}$$

The author in [18] combines an exact result for $\mathbb{P}_n^{\mathbb{N}}(\mathbf{Z} = m)$ with Lemma 2 below to estimate the expected value of **T** asymptotically in the case of unrestricted mappings. In our case we use Lemma 1 for the distribution of **Z** over $\{0,k\}$-mappings.

▶ **Lemma 1** (Equation (3.17) of [17]). *Let $n = kr$, $\lambda = k - 1 \geq 1$ and $1 \leq m \leq r$. A random uniform $\{0, k\}$-mapping on n nodes has exactly m cyclic nodes with probability*

$$\mathbb{P}_n^{\{0,k\}}(\mathbf{Z} = m) = \lambda k^{m-1} \binom{r-1}{m-1} \binom{n-1}{m}^{-1} = \frac{\lambda m k^{m-1} \Gamma(r) \Gamma(n - m)}{\Gamma(r - m + 1) \Gamma(n)}.$$

▶ **Lemma 2** ([19]). *Let M_m be the expected order of a random permutation of S_m. Let $\beta_0 = \sqrt{8I}$, where*

$$I = \int_0^\infty \log\log\left(\frac{e}{1 - e^{-t}}\right) dt. \tag{2}$$

Then, as m approaches infinity,

$$\log M_m = \beta_0 \sqrt{\frac{m}{\log m}} + O\left(\frac{\sqrt{m}\log\log m}{\log m}\right).$$

It is clear from Equation (1) that, if m_* is the integer that maximizes $\mathbb{P}_n^{\{0,k\}}(\mathbf{Z} = m)M_m$ for $1 \leq m \leq n$ and m_0 is an integer in $(1, n)$, then

$$\mathbb{P}_n^{\{0,k\}}(\mathbf{Z} = m_0)M_{m_0} \leq \mathbb{E}_n^{\{0,k\}}(\mathbf{T}) \leq n\mathbb{P}_n^{\{0,k\}}(\mathbf{Z} = m_*)M_{m_*}. \tag{3}$$

Let $n \geq 1$ and $\varepsilon \in (-1, 1)$. Let $\beta_\varepsilon = \beta_0 + \varepsilon$. We define the following real function that provides a tight upper or lower bound for the summand in Equation (1), according to the value of ε:

$$\phi_{n,\varepsilon}(x) = \lambda x k^{x-1} \frac{\Gamma(r)}{\Gamma(r - x + 1)} \frac{\Gamma(n - x)}{\Gamma(n)} \exp\left(\beta_\varepsilon \sqrt{\frac{x}{\log x}}\right). \tag{4}$$

▶ **Proposition 3.** *Let $n = kr$, $\lambda = k - 1 \geq 1$ and $\varepsilon \in (-1, 1)$. If, for some $0 < \alpha < 1$, $k = o(n^{1-\alpha})$ as r approaches infinity, then there exists a constant $c > 0$ such that, for sufficiently large n, the function $x \longmapsto \phi_{n,\varepsilon}(x)$ assumes a unique maximum x_* for $x \in (c, r)$. Moreover, if $k_\varepsilon = \sqrt[3]{3^5 \beta_\varepsilon^4}/8$, then*

$$\log \phi_{n,\varepsilon}(x_*) = k_\varepsilon \frac{\left(\frac{n}{\lambda}\right)^{1/3}}{\log^{2/3}\left(\frac{n}{\lambda}\right)} (1 + o(1)).$$

The calculation of the maximum value that $\phi_{n,\varepsilon}(x)$ assumes for $x \in (1, n)$ is a main ingredient in the proof of the asymptotic estimate on $\mathbb{E}_n^{\{0,k\}}(\mathbf{T})$. It allows us to obtain an upper bound for the rightmost term in Equation (3). The maximum x_* also allows us to select an integer m_0 that provides a lower bound in Equation (3) that is good enough for our purposes.

▶ **Theorem 4.** *Let $k = k(r)$ and $n = n(r)$ be sequences such that $n = kr$ and, for some $0 < \alpha < 1$, $k = o(n^{1-\alpha})$ as r approaches infinity. Let $\mathbb{E}_n^{\{0,k\}}(\mathbf{T})$ be the expected value of $\mathbf{T}$ over the class of mappings on n nodes with indegrees restricted to the set $\{0, k\}$. Then,*

$$\log \mathbb{E}_n^{\{0,k\}}(\mathbf{T}) = k_0 \frac{\left(\frac{n}{\lambda}\right)^{1/3}}{\log^{2/3}\left(\frac{n}{\lambda}\right)} (1 + o(1)),$$

as r approaches infinity, where $\lambda = k - 1$, $k_0 = \frac{3}{2}(3I)^{2/3}$ and I is given in Equation (2). In particular, the estimate above holds if $k \geq 2$ is a fixed integer.

We obtain asymptotic estimates for the expectation of **B** over $\{0, k\}$-mappings using the same arguments.

▶ **Theorem 5.** *Let $k = k(r)$ and $n = n(r)$ be sequences such that $n = kr$ and, for some $0 < \alpha < 1$, $k = o(n^{1-\alpha})$ as r approaches infinity. For $r \geq 1$, let $\mathbb{E}_n^{\{0,k\}}(\mathbf{B})$ be the expected value of $\mathbf{B}$ over the class of mappings on n nodes with indegrees restricted to the set $\{0, k\}$. Then, as r approaches infinity,*

$$\log \mathbb{E}_n^{\{0,k\}}(\mathbf{B}) = \frac{3}{2} \left(\frac{n}{\lambda}\right)^{1/3} (1 + o(1)),$$

where $\lambda = k - 1$. In particular, the estimate above holds if $k \geq 2$ is a fixed integer.

3 Lognormality

Let

$$\mu_n^* = \frac{1}{2} \log^2(\sqrt{n}), \quad \sigma_n^* = \frac{1}{\sqrt{3}} \log^{3/2}(\sqrt{n})$$

and

$$\mu_n = \frac{1}{2} \log^2\left(\sqrt{n/\lambda}\right), \quad \sigma_n = \frac{1}{\sqrt{3}} \log^{3/2}\left(\sqrt{n/\lambda}\right).$$

Harris proved that the sequence of random variables defined over the space of random mappings on n nodes as $X_n = (\log \mathbf{T} - \mu_n^*)/\sigma_n^*$, $n \geq 1$, converges weakly to a standard normal distribution [12]. In this section we prove an analogue of this result for $\{0, k\}$-mappings:

$$\lim_{n \to \infty} \mathbb{P}_n^{\{0,k\}} \left(\frac{\log \mathbf{T} - \mu_n}{\sigma_n} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt. \tag{5}$$

The analogous result for the parameter **B** is proved from Equation (5) by showing that the random variable $\chi_n = \log \mathbf{B} - \log \mathbf{T}$, when properly normalized, converges in probability to zero.

We write the probability in Equation (5) using the law of total probability, where we partition the space of $\{0, k\}$-mappings as follows. It is possible to prove that, for $k \geq 2$, $r \geq 1$ fixed integers and $n = kr$, there exists a positive real number $m_\#$ such that the sequence $z_m = \mathbb{P}_n^{\{0,k\}}(\mathbf{Z} = m)$, $m \geq 1$, is increasing for $m < m_\#$ and decreasing for $m > m_\#$. Furthermore, $m_\# = \sqrt{n/\lambda} + O(1)$. Let $\varepsilon_n = \log^{-3/4}(\sqrt{n/\lambda})$, $\xi_1 = m_\#^{1-\varepsilon_n}$ and $\xi_2 = m_\#^{1+\varepsilon_n}$. We partition the interval $[1, r]$ into three subintervals:

- $I_1 = \{m : 1 \leq m < \xi_1\}$,
- $I_2 = \{m : \xi_1 \leq m \leq \xi_2\}$,
- $I_3 = \{m : \xi_2 < m \leq r\}$.

For $k \geq 2$ fixed, it is proved in [1] that $\mathbb{E}_n^{\{0,k\}}(\mathbf{Z}) \sim \sqrt{\pi n/2\lambda}$, hence the mode $m_\#$ has the same order of growth as the expectation of **Z**.

▶ **Lemma 6.** *Let $\varepsilon_n = \log^{-3/4}(\sqrt{n/\lambda})$. If $\xi_1 = m_\#^{1-\varepsilon_n}$ and $\xi_2 = m_\#^{1+\varepsilon_n}$, then $\mathbb{P}_n^{\{0,k\}}(\mathbf{Z} < \xi_1) = o(1)$, $\mathbb{P}_n^{\{0,k\}}(\mathbf{Z} > \xi_2) = o(1)$ and $\mathbb{P}_n^{\{0,k\}}(\xi_1 \leq \mathbf{Z} \leq \xi_2) \sim 1$.*

It follows from the law of total probability that

$$\mathbb{P}_n^{\{0,k\}} (\log \mathbf{T} \leq \mu_n + x\sigma_n) = \zeta_1 + \zeta_2 + \zeta_3, \tag{6}$$

where

$$\zeta_j = \sum_{m \in I_j} \mathbb{P}_n^{\{0,k\}}(\mathbf{Z} = m)\mathbb{P}_n^{\{0,k\}}(\log \mathbf{T} \leq \mu_n + x\sigma_n|\mathbf{Z} = m). \tag{7}$$

Lemma 6 clearly implies that $\zeta_1 = o(1)$ and $\zeta_3 = o(1)$. We prove next that ζ_2 provides the asymptotic main term in (6). We use the special case $\theta = 1$ of Theorem 1.2 of [4], that represents a stronger version of Erdös and Turán's famous result [8]. We denote by $\mathbf{Q}_m$ the uniform probability measure on the symmetric group S_m and by $\phi(x) = \frac{1}{2\pi}\int_{-\infty}^{x} e^{-t^2/2}dt$ the standard normal distribution.

▶ **Theorem 7** ([4]). *Let $\alpha_m = \frac{1}{2}\log^2 m + \log m \log \log m$ and $\beta_m = \frac{1}{\sqrt{3}}\log^{3/2} m$. Then there exists a constant $K > 0$ such that, for all real numbers x and all integers $m > 1$,*

$$\left|\mathbf{Q}_m\left(\log \mathbf{T} \leq \alpha_m + x\beta_m\right) - \phi(x)\right| \leq \frac{K}{\sqrt{\log m}}.$$

▶ **Lemma 8.** *For $n = kr$ and $m \in I_2$, let*

$$\delta_x(m,n) = \mathbb{P}_n^{\{0,k\}}\left(\log \mathbf{T} \leq \mu_n + x\sigma_n|\mathbf{Z} = m\right) - \phi(x),$$

and let $\Delta_x(n) = \max\{|\delta_x(m,n)|, m \in I_2\}$. Then, for any fixed $x \in \mathbb{R}$, $\Delta_x(n) = o(1)$ as r approaches infinity. Moreover, if $|x| \leq c\sqrt{\log n}$, for some positive constant c, then $\Delta_x(n) \leq K_4 \log^{-1/4}(\sqrt{n/\lambda})$, for some $K_4 > 0$.

Sketch. Let α_m and β_m be as in Theorem 7 and define $y = y(n,m,x)$ to be the real number for which $\mu_n + x\sigma_n = \alpha_m + y\beta_m$. Then, for any $m \in I_2$,

$$|\delta_x(m,n)| \leq \left|\mathbf{Q}_m\left(\frac{\log \mathbf{T} - \alpha_m}{\beta_m} \leq y\right) - \phi(y)\right| + |\phi(y) - \phi(x)|. \tag{8}$$

We note that Theorem 7 implies that, for some constant $K_1 > 0$,

$$\left|\mathbf{Q}_m\left(\frac{\log \mathbf{T} - \alpha_m}{\beta_m} \leq y\right) - \phi(y)\right| \leq \frac{K_1}{\sqrt{\log m}}. \tag{9}$$

Using Equations (8) and (9) and $|\phi(y) - \phi(x)| \leq |y - x|$ we obtain

$$|\delta_x(m,n)| \leq \frac{K_1}{\sqrt{\log m}} + |y - x|. \tag{10}$$

We note that the definition of y implies

$$y - x = \frac{(\mu_n - \alpha_m) + x(\sigma_n - \beta_m)}{\beta_m},$$

where $\sigma_n - \beta_m = O(\beta_m \varepsilon_n)$ and $\alpha_m - \mu_n = O\left(\beta_m \log^{-1/4}\left(\sqrt{n/\lambda}\right)\right)$. Hence,

$$y - x = O\left(\log^{-1/4}\left(\sqrt{n/\lambda}\right)\right) + O(|x|\varepsilon_n). \tag{11}$$

The result follows from Equations (10) and (11) and $m > \xi_1 = O\left(\log\sqrt{\frac{n}{\lambda}}\right)$.   ◀

With Lemma 8 in hand, it is straight-forward to deduce the following result.

▶ **Theorem 9.** *Let $k = k(r)$ and $n = n(r)$ be sequences such that $n = kr$ and, for some $0 < \alpha < 1$, $k = o(n^{1-\alpha})$ as r approaches infinity. Let $\mu_n = \frac{1}{2}\log^2(\sqrt{\frac{n}{\lambda}})$, $\sigma_n^2 = \frac{1}{3}\log^3(\sqrt{\frac{n}{\lambda}})$. Let $\mathbf{T}(f)$ denote the least common multiple of the length of the cycles of a mapping f. Then, for any real number x, as r approaches infinity,*

$$\mathbb{P}_n^{\{0,k\}}\left(\log \mathbf{T} \leq \mu_n + x\sigma_n\right) = \phi(x) + o_x(1),$$

where $o_x(\cdot)$ indicates that the error term depends on x. Moreover, if c is a positive constant, then the convergence is uniform for $|x| \leq c\sqrt{\log n}$.

Comment. We observe that $\mathbf{Z}$ and $\log \mathbf{T}$ are concentrated in the interval $[\xi_1, \xi_2]$. However this interval does *not* contain the terms that contribute most to the expected value of $\mathbf{T}$. Most of the contribution for the sum in (1) is from mappings with $\Theta\left(\frac{(n/\lambda)^{2/3}}{\log^{1/3}(n/\lambda)}\right)$ cyclic vertices.

In order to prove asymptotic lognormality for the parameter $\mathbf{B}$, we use Theorem 10 below, where it is proved that the normalized difference between $\log \mathbf{B}$ and $\log \mathbf{T}$ converges in probability to zero. We consider this result of independent interest. Lognormality for the parameter $\mathbf{B}$ follows at once from Slutsky's Theorem; see Theorem 15 in Section 6.2 of [16].

▶ **Theorem 10.** *Let $k = k(r)$ and $n = n(r)$ be sequences such that $n = kr$ and, for some $0 < \alpha < 1$, $k = o(n^{1-\alpha})$ as r approaches infinity. For $r \geq 1$, let χ_n be the random variable defined over $\{0, k\}$-mappings on n nodes as $\chi_n = (\log \mathbf{B} - \log \mathbf{T})/\sigma_n$, where $\sigma_n = \frac{1}{\sqrt{3}} \log^{3/2}(\sqrt{\frac{n}{\lambda}})$. Then the sequence defined by χ_n converges in probability to zero. In other words, for all $\varepsilon > 0$ we have*

$$\mathbb{P}_n^{\{0,k\}}(\chi_n > \varepsilon) = o(1),$$

as r approaches infinity.

▶ **Theorem 11.** *Let $k = k(r)$ and $n = n(r)$ be sequences such that $n = kr$ and, for some $0 < \alpha < 1$, $k = o(n^{1-\alpha})$ as r approaches infinity. Let $\mu_n = \frac{1}{2} \log^2(\sqrt{\frac{n}{\lambda}})$, $\sigma_n^2 = \frac{1}{3} \log^3(\sqrt{\frac{n}{\lambda}})$. Let $\mathbf{B}(f)$ denote the product of the length of the cycles of a mapping f. Then, for any real number x,*

$$\mathbb{P}_n^{\{0,k\}}(\log \mathbf{B} \leq \mu_n + x\sigma_n) = \phi(x) + o_x(1),$$

as r approaches infinity. Moreover, if c is a positive constant, then the convergence is uniform for $|x| \leq c\sqrt{\log n}$.

4 Heuristics

In the analysis of his factorization method [15], Pollard conjectured that quadratic polynomials modulo large primes behave like random mappings with respect to their average rho length. However, it should be noted that the indegree distribution of a class of mappings impacts the asymptotic distribution of a number of parameters [1]; the indegree distribution of a mapping f on n nodes is defined as the sequence $n_j = \#\{y \in [n] : |f^{-1}(y)| = j\}$, $j \geq 0$. Since a quadratic polynomial modulo an odd prime p has a very particular indegree distribution, namely $(n_0, n_1, n_2) = (\frac{p-1}{2}, 1, \frac{p-1}{2})$, one might wonder if $\{0, 2\}$-mappings do not represent a better heuristic model. Furthermore, there are classes of polynomials from which one might not expect the typical random mapping behavior, and it is possible to use different classes of mappings as heuristic models. This is the case for the polynomials of the form $f(x) = x^d + a \in \mathbb{F}_p[x]$, where, as usual, $\mathbb{F}_p$ denotes the finite field on p elements. Their indegree distribution satisfies

$$n_0 = \left(1 - \frac{1}{k}\right)(p-1), \quad n_1 = 1, \quad n_k = \frac{1}{k}(p-1). \tag{12}$$

where $k = \gcd(p-1, d)$. We refer to the polynomials with indegree distribution (12) as $\{0, k\}$-*polynomials*. As a particular case, we note that a polynomial of the form $x^k + a \in \mathbb{F}_p[x]$, $p \equiv 1 \pmod{k}$, is a $\{0, k\}$-polynomial.

Algorithm 1: GENERATING A RANDOM UNIFORM $\{0, k\}$-MAPPING.

Input: Integers $r \geq 1$ and $k \geq 2$.

Output: $\{0, k\}$-mapping f on $n = kr$ nodes.

1 Pick a permutation $\sigma = \sigma_1 \cdots \sigma_n \in S_n$ uniformly at random.

2 Pick a permutation $\tau = \tau_1 \cdots \tau_n \in S_n$ uniformly at random.

3 **for** $i = 0, \ldots, r - 1$ **do**

4   **for** $j = 1, \ldots, k$ **do**

5        $f(\tau[ik + j]) = \sigma[i + 1]$ `// ` $\tau[\ell]$ ` denotes ` τ_ℓ `, same for ` $\sigma[\ell]$ `.`

6   **end**

7 **end**

8 **return** f.

In this section we consider classes of $\{0, k\}$-mappings, treated in the previous sections, as heuristic models for $\{0, k\}$-polynomials. Our focus lies on polynomials of a certain degree modulo large prime numbers, hence from this point on we restrict our attention to $\{0, k\}$-mappings with $k \geq 2$ fixed, even though the results of the previous sections hold in a more general setting. The asymptotic results in this section are taken as n approaches infinity.

The interest in the heuristic approximation mentioned above can be attributed at least in part to the wealth of asymptotic results on the statistics of mappings with indegree restrictions, when compared to the literature on the number theoretical setting; see for example [1, 7]. The main term of several asymptotic results on the statistics of a class $\mathcal{F}$ of mappings with restrictions on the indegrees depends on its asymptotic average coalescence $\lambda = \lambda(\mathcal{F})$, defined as in Section 1. This is the case for the rho length of a random node, a parameter involved in the analysis of Pollard factorization algorithm. Since $\lambda = 1$ for unrestricted mappings and $\{0, 2\}$-mappings, these two classes represent equally accurate models for the average rho length of quadratic polynomials [14]. It is curious that the knowledge of the indegree distribution of these polynomials does not represent an improvement on the heuristic in this case. It is worth noting that our asymptotic results on different classes of $\{0, k\}$-mappings are determined by their coalescence λ as well; compare Theorems 4 and 5 with Theorems 1.3 and 1.4 of [18]. Compare μ_n and μ_m^* with σ_n and σ_m^* as well, under the light of the fact that the expected number of cyclic nodes over all unrestricted or $\{0, k\}$-mappings are asymptotically equivalent to $\sqrt{\pi n / 2}$ and $\sqrt{\pi n / 2\lambda}$, respectively. We note that if $\log k = o(\log n)$, then $\mu_n \sim \mu_n^*$ and $\sigma_n \sim \sigma_n^*$ as r approaches infinity.

4.1 Sampling $\{0, k\}$-Mappings

In our experiments, for each prime number $p \equiv 1 \pmod{k}$ considered, we select p $\{0, k\}$-mappings on $n = p - 1$ nodes uniformly at random according to the following algorithm. We determine the range of each random mapping f by selecting the first $r = n/k$ elements of a random uniform permutation $\sigma = \sigma_1 \cdots \sigma_n \in S_n$. The image $f(x)$ of every element $x \in [n]$ is defined by dividing a random uniform permutation $\tau = \tau_1 \cdots \tau_n \in S_n$ in blocks of k elements. We make this process precise in Algorithm 1. Assuming that σ, τ are random uniform permutations, it is possible to prove that Algorithm 1 returns a random uniform $\{0, k\}$-mapping on n nodes.

It should be stressed that our experimental results are based on sampling, so a number of problems can occur in the numerical estimate of the expectation of a random variable. Typically one must sample a very large number of $\{0, k\}$-mappings until a mapping with

a substantial value of T or B is revealed (Theorems 12 and 13). To simplify notation, let $\vec{S} = f_1, f_2, f_3, \ldots$ denote a sequence of independent random samples chosen uniformly at random from the class of $\{0, k\}$-mappings on n nodes.

▶ **Theorem 12.** *Let* $\xi = \xi(n) = \left(\mathbb{E}_n^{\{0,k\}}(\mathbf{T}) \right)^a$, *where* $a = a(n)$, *and define* $\mathbf{N} = \min\{t : \mathbf{T}(f_t) \geq \xi\}$. *If* $a^{-1} \log^{-1/3} n = o(1)$, *then for sufficiently large n, we have*

$$\mathbb{E}_n^{\{0,k\}}(\mathbf{N}) > \exp\left(\frac{\lambda n^{1/3}}{3 \log^6 n} \right),$$

and

$$\mathbb{P}_n^{\{0,k\}} \left(\mathbf{N} \leq \exp\left(\frac{\lambda n^{1/3}}{4 \log^6 n} \right) \right) \leq \exp\left(-\frac{\lambda n^{1/3}}{12 \log^6 n} \right).$$

▶ **Theorem 13.** *Let* $\widetilde{\xi} = \left(\mathbb{E}_n^{\{0,k\}}(\mathbf{B}) \right)^b$, *where* $b = b(n)$, *and let* $\widetilde{\mathbf{N}} = \widetilde{\mathbf{N}}(n, \vec{S}, b) = \min\{t : \mathbf{B}(f_t) \geq \widetilde{\xi}\}$. *If* $b^{-1} \log^{-2} n = o(1)$, *then there exist positive constants c_1, c_2 such that, for sufficiently large n,*

$$\mathbb{E}_n^{\{0,k\}}(\widetilde{\mathbf{N}}) > \exp\left(\frac{c_1 (\frac{n}{\lambda})^{1/3}}{\log^3(\frac{n}{\lambda})} \right),$$

and

$$\mathbb{P}_n^{\{0,k\}} \left(\widetilde{\mathbf{N}} \leq \exp\left(\frac{c_2 (\frac{n}{\lambda})^{1/3}}{\log^3(\frac{n}{\lambda})} \right) \right) \leq \exp\left(\frac{-c_2 (\frac{n}{\lambda})^{1/3}}{\log^3(\frac{n}{\lambda})} \right).$$

The proof of Theorem 12 relies on tail estimates for the number $\mathbf{Z}$ of cyclic vertices, and bounds on the maximum order that a permutation can have. The proof of Theorem 13 is based on tail estimates for the number $\mathbf{C}$ of cycles, together with the inequality $\mathbf{B} \leq (\frac{\mathbf{Z}}{\mathbf{C}})^{\mathbf{C}}$.

4.2 Numerical Results

We exhibit in Table 1 our numerical results on the behavior of $\mathbf{T}$ and $\mathbf{B}$ over different classes of polynomials over finite fields and different classes of mappings. For each value of k, we consider the first 100 primes greater than 10^3 of the form indicated in Table 1. For each such prime, we select, according to Algorithm 1, p mappings on $n = p - 1$ nodes; we also consider all p polynomials of the form indicated in Table 1. We compute the exact value of $\mathbf{T}$ for each function and compute the corresponding average values $\overline{\mathbf{T}}(p)$. We compute the ratio $R_{\mathbf{T}}(p)$ between $\log \overline{\mathbf{T}}(p)$ and the quantity in Theorem 4. In Table 1 we exhibit the average value $\overline{R_{\mathbf{T}}}$ of $R_{\mathbf{T}}(p)$ over all primes considered; we stress the dependence of this calculation on the coalescence λ of the corresponding class by adopting the notation $\overline{R_{\mathbf{T}}}(\lambda)$. The same is done for the parameter $\mathbf{B}$.

It is not surprising to have the ratio $\overline{R_{\mathbf{T}}}$ distant from 1 even in the case of $\{0, k\}$-mappings, where we have an asymptotic result proved on the logarithm of the expectation of $\mathbf{T}$. It is proved in Theorem 12 that most of the contribution to $\mathbb{E}_n^{\{0,k\}}(\mathbf{T})$ comes from a relatively small set of exceptional maps. Unless the number of samples is enormous, as stated in the first part of the theorem, none of these exceptional maps is likely to be sampled, so our empirical estimate for $\mathbb{E}_n^{\{0,k\}}(\mathbf{T})$ is likely to be poor. The ratios $\overline{R_{\mathbf{T}}}$ appear to decrease as λ grows large, but this agrees, in a way, with the fact that the upper bound in Theorem 12 decreases as k grows large.

■ **Table 1** Experimental results on mappings and polynomials according to their coalescence.

Class of functions	Asymp. Coalescence	$\overline{R_{\mathbf{T}}}(\lambda)$	$\overline{R_{\mathbf{B}}}(\lambda)$
Unrestricted mappings	1	0.8090	0.7247
$\{0,2\}$-mappings	1	0.7929	0.7097
$x^2 + a \in \mathbb{F}_p[x]$	1	0.8031	2.4183
$x^4 + a \in \mathbb{F}_p[x]$, $p \equiv 3 \pmod 4$	1	0.8033	3.9237
$\{0,3\}$-mappings	2	0.7700	0.7043
$x^3 + a \in \mathbb{F}_p[x]$, $p \equiv 1 \pmod 3$	2	0.7631	2.5067
$\{0,4\}$-mappings	3	0.7436	0.7007
$x^4 + a \in \mathbb{F}_p[x]$, $p \equiv 1 \pmod 4$	3	0.7391	2.6055
$\{0,5\}$-mappings	4	0.7465	0.7041
$x^5 + a \in \mathbb{F}_p[x]$, $p \equiv 1 \pmod 5$	4	0.7435	3.3597
$\{0,6\}$-mappings	5	0.6986	0.6789
$x^6 + a \in \mathbb{F}_p[x]$, $p \equiv 1 \pmod 6$	5	0.6989	1.3522

Regardless of the sampling problem explained in Section 4.1, it is remarkable that the ratio between any two entries in the table above for $\overline{R_{\mathbf{T}}}$ with the same value of λ lies in the interval $(0.97, 1.03)$. This suggests that the behavior of a typical $\{0, k\}$-polynomial can be approximated by the behavior of a typical $\{0, k\}$-mapping. However, one must be careful when using the asymptotic estimate in Theorem 4 as a reference, due to the results in Theorem 12. The numerical results for the parameter **B**, on the other hand, represent a different scenario, where the ratio between numerical results for classes with the same value of asymptotic coalescence were found to be as high as 4.8835. It is interesting but not clear why the heuristic performs so poorly in the approximation of the statistics of polynomials by mappings in the case of the parameter **B**.

─── **References** ───────────

1   James Arney and Edward A. and Bender. Random mappings with constraints on coalescence and number of origins. *Pacific J. Math.*, 103(2):269–294, 1982.

2   R. Arratia, A. D. Barbour, and S. Tavaré. Limits of logarithmic combinatorial structures. *Ann. Probab.*, 28(4):1620–1644, 2000. `doi:10.1214/aop/1019160500`.

3   Richard Arratia, A. D. Barbour, and Simon Tavaré. *Logarithmic combinatorial structures: a probabilistic approach.* EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2003. `doi:10.4171/000`.

4   A. D. Barbour and Simon Tavaré. A rate for the Erdős-Turán law. *Combin. Probab. Comput.*, 3(2):167–176, 1994. `doi:10.1017/S0963548300001097`.

5   Richard P. Brent and John M. Pollard. Factorization of the eighth Fermat number. *Math. Comp.*, 36(154):627–630, 1981. `doi:10.2307/2007666`.

6   Charles Burnette and Eric Schmutz. Periods of iterated rational functions. *Int. J. Number Theory*, 13(5):1301–1315, 2017. `doi:10.1142/S1793042117500713`.

7   Michael Drmota and Michèle Soria. Images and preimages in random mappings. *SIAM J. Discrete Math.*, 10(2):246–269, 1997. `doi:10.1137/S0895480194268421`.

8   P. Erdős and P. Turán. On some problems of a statistical group-theory. III. *Acta Math. Acad. Sci. Hungar.*, 18:309–320, 1967. `doi:10.1007/BF02280290`.

9 Philippe Flajolet and Andrew M. Odlyzko. Random mapping statistics. In *Advances in cryptology—EUROCRYPT '89 (Houthalen, 1989)*, volume 434 of *Lecture Notes in Comput. Sci.*, pages 329–354. Springer, Berlin, 1990. `doi:10.1007/3-540-46885-4_34`.

10 Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009. `doi:10.1017/CBO9780511801655`.

11 Jennie C. Hansen and Jerzy Jaworski. Random mappings with exchangeable in-degrees. *Random Structures Algorithms*, 33(1):105–126, 2008. `doi:10.1002/rsa.20187`.

12 Bernard Harris. The asymptotic distribution of the order of elements in symmetric semigroups. *J. Combinatorial Theory Ser. A*, 15:66–74, 1973. `doi:10.1016/0097-3165(73)90036-8`.

13 Valentin F. Kolchin. *Random mappings*. Translation Series in Mathematics and Engineering. Optimization Software, Inc., Publications Division, New York, 1986. Translated from the Russian, With a foreword by S. R. S. Varadhan.

14 Rodrigo S. V. Martins and Daniel Panario. On the heuristic of approximating polynomials over finite fields by random mappings. *Int. J. Number Theory*, 12(7):1987–2016, 2016. `doi:10.1142/S1793042116501219`.

15 J. M. Pollard. A Monte Carlo method for factorization. *Nordisk Tidskr. Informationsbehandling (BIT)*, 15(3):331–334, 1975. `doi:10.1007/BF01933667`.

16 Vijay K. Rohatgi and A. K. Md. Ehsanes Saleh. *An introduction to probability and statistics*. Wiley Series in Probability and Statistics: Texts and References Section. Wiley-Interscience, New York, second edition, 2001. `doi:10.1002/9781118165676`.

17 H Rubin and R Sitgreaves. Probability distributions related to random transformations of a finite set. Technical Report SOL ONR 19A, Stanford, Stanford, 1954.

18 Eric Schmutz. Period lengths for iterated functions. *Combin. Probab. Comput.*, 20(2):289–298, 2011. `doi:10.1017/S0963548310000337`.

19 Richard Stong. The average order of a permutation. *Electron. J. Combin.*, 5:Research Paper 41, 6, 1998.

20 Michael J. Wiener and Robert J. Zuccherato. Faster attacks on elliptic curve cryptosystems. In *Selected areas in cryptography (Kingston, ON, 1998)*, volume 1556 of *Lecture Notes in Comput. Sci.*, pages 190–200. Springer, Berlin, 1999. `doi:10.1007/3-540-48892-8_15`.

Modularity of Erdős-Rényi Random Graphs

Colin McDiarmid

Department of Statistics, University of Oxford, Oxford, UK
cmcd@stats.ox.ac.uk

Fiona Skerman[1]

Department of Mathematics, Uppsala University, Uppsala, Sweden
fiona.skerman@math.uu.se
  https://orcid.org/0000-0003-4141-7059

⎯⎯ Abstract ⎯⎯

For a given graph G, modularity gives a score to each vertex partition, with higher values taken to indicate that the partition better captures community structure in G. The modularity $q^*(G)$ (where $0 \leq q^*(G) \leq 1$) of the graph G is defined to be the maximum over all vertex partitions of the modularity value. Given the prominence of modularity in community detection, it is an important graph parameter to understand mathematically.

For the Erdős-Rényi random graph $G_{n,p}$ with n vertices and edge-probability p, the likely modularity has three distinct phases. For $np \leq 1 + o(1)$ the modularity is $1 + o(1)$ with high probability (whp), and for $np \to \infty$ the modularity is $o(1)$ whp. Between these regions the modularity is non-trivial: for constants $1 < c_0 \leq c_1$ there exists $\delta > 0$ such that when $c_0 \leq np \leq c_1$ we have $\delta < q^*(G) < 1 - \delta$ whp. For this critical region, we show that whp $q^*(G_{n,p})$ has order $(np)^{-1/2}$, in accord with a conjecture by Reichardt and Bornholdt in 2006 (and disproving another conjecture from the physics literature).

1 Introduction

We start this section with some background and definitions, and then present our results on the modularity of the random graph $G_{n,p}$. After that, we sketch previous work on modularity, and then give a plan of the rest of the paper, which essentially consists of the proofs of the three phases. The remaining proofs will be given in the extended version of this paper.

1.1 Definitions

The large and increasing quantities of network data available in many fields has led to great interest in techniques to discover network structure. We want to be able to identify if a network can be decomposed into communities or highly clustered components.

Modularity was introduced by Newman and Girvan in 2004 [27]. It gives a measure of how well a graph can be divided into communities, and now forms the backbone of the most popular algorithms used to cluster real data [18]. Here a 'community' is a collection of nodes which are more densely interconnected than one would expect – see the discussion following the definition of modularity below. There are many applications, including protein discovery, identifying connections between websites, and mapping the onset of schizophrenia on neuron

---

clusters in the brain [2]. Its widespread use and empirical success in finding communities in networks makes modularity an important function to understand mathematically. See [11] and [28] for surveys on the use of modularity for community detection in networks.

Given a graph G, modularity gives a score to each vertex partition: the modularity $q^*(G)$ (sometimes called the 'maximum modularity') of G is defined to be the maximum of these scores over all vertex partitions. For a set A of vertices, let $e(A)$ be the number of edges within A, and let the *volume* $\mathrm{vol}(A)$ be the sum over the vertices v in A of the degree d_v.

▶ **Definition 1.1** (Newman & Girvan [27], see also Newman [26])**.** Let G be a graph with $m \geq 1$ edges. For a vertex partition $\mathcal{A}$ of G, the modularity of $\mathcal{A}$ on G is

$$q_{\mathcal{A}}(G) = \frac{1}{2m} \sum_{A \in \mathcal{A}} \sum_{u,v \in A} \left(\mathbf{1}_{uv \in E} - \frac{d_u d_v}{2m} \right) = \frac{1}{m} \sum_{A \in \mathcal{A}} e(A) - \frac{1}{4m^2} \sum_{A \in \mathcal{A}} \mathrm{vol}(A)^2;$$

and the modularity of G is $q^*(G) = \max_{\mathcal{A}}(G)$, where the maximum is over all partitions $\mathcal{A}$ of the vertices of G.

Isolated vertices are irrelevant. We need to give empty graphs (graphs with no edges) some modularity value. Conventionally we set $q^*(G) = 1$ for each empty graph G [5] (though the value will not be important). The second equation for $q_{\mathcal{A}}(G)$ expresses modularity as the difference of two terms, the *edge contribution* or coverage $q_{\mathcal{A}}^E(G) = \frac{1}{m} \sum_A e(A)$, and the *degree tax* $q_{\mathcal{A}}^D(G) = \frac{1}{4m^2} \sum_A \mathrm{vol}(A)^2$. Since $q_{\mathcal{A}}^E(G) \leq 1$ and $q_{\mathcal{A}}^D(G) > 0$, we have $q_{\mathcal{A}}(G) < 1$ for any non-empty graph G. Also, the trivial partition $\mathcal{A}_0$ with all vertices in one part has $q_{\mathcal{A}_0}^E(G) = q_{\mathcal{A}_0}^D(G) = 1$, so $q_{\mathcal{A}_0}(G) = 0$. Thus we have

$$0 \leq q^*(G) \leq 1.$$

Suppose that we pick uniformly at random a multigraph with degree sequence $(d_1, \ldots, d_n)$ where $\sum_v d_v = 2m$. Then the expected number of edges between vertices u and v is $d_u d_v / (2m - 1)$. This is the original rationale for the definition: whilst rewarding the partition for capturing edges within the parts, we should penalise by (approximately) the expected number of edges.

A differentiation between graphs which are truly modular and those which are not can … only be made if we gain an understanding of the intrinsic modularity of random graphs. – Reichardt and Bornholdt [30]. In this paper we investigate the likely value of the modularity of an Erdős-Rényi random graph. Let n be a positive integer. Given $0 \leq p \leq 1$, the random graph $G_{n,p}$ has vertex set $[n] := \{1, \ldots, n\}$ and the $\binom{n}{2}$ possible edges appear independently with probability p. Given an integer m with $0 \leq m \leq \binom{n}{2}$, the random graph $G_{n,m}$ is sampled uniformly from the m-edge graphs on vertex set $[n]$. These two random graphs are closely related when $m \approx \binom{n}{2}p$: we shall investigate only $q^*(G_{n,p})$ here, but in the extended version of the paper we shall also deduce corresponding results for $q^*(G_{n,m})$.

For a sequence of events A_n we say that A_n holds *with high probability (whp)* if $\mathbb{P}(A_n) \to 1$ as $n \to \infty$. For a sequence of random variables X_n and a real number a, we write $X_n \xrightarrow{p} a$ if X_n converges in probability to a as $n \to \infty$ (that is, if for each $\varepsilon > 0$ we have $|X_n - a| < \varepsilon$ whp).

1.2 Results on the modularity of the random graph $G_{n,p}$

Our first theorem, the Three Phases Theorem, gives the big picture. The three phases correspond to when (a) the expected vertex degree (essentially np) is at most about 1, (b) bigger than 1 but bounded, or (c) tending to infinity.

▶ **Theorem 1.2.** *Let $p = p(n)$ satisfy $0 \leq p \leq 1$.*

(a) *If $n^2 p \to \infty$ and $np \leq 1 + o(1)$ then $q^*(G_{n,p}) \xrightarrow{p} 1$.*

(b) *Given constants $1 < c_0 \leq c_1$, there exists $\delta = \delta(c_0, c_1) > 0$ such that if $c_0 \leq np \leq c_1$ for n sufficiently large, then whp $\delta < q^*(G_{n,p}) < 1 - \delta$.*

(c) *If $np \to \infty$ then $q^*(G_{n,p}) \xrightarrow{p} 0$.*

We are able to confirm the $(np)^{-1/2}$ growth rate conjectured to hold for the critical region in [30]. The edge probabilities p correspond to parts (b) or (c) of Theorem 1.2.

▶ **Theorem 1.3.** *There exists b such that for all $0 < p = p(n) \leq 1$ we have $q^*(G_{n,p}) < \frac{b}{\sqrt{np}}$ whp. Also, given $0 < \varepsilon < 1$, there exists $a = a(\varepsilon) > 0$ such that, if $p = p(n)$ satisfies $np \geq 1$ and $p \leq 1 - \varepsilon$ for n sufficiently large, then $q^*(G_{n,p}) > \frac{a}{\sqrt{np}}$ whp.*

Observe that the upper bound here on $q^*(G_{n,p})$ implies part (c) of Theorem 1.2. As an immediate corollary of Theorem 1.3 we have:

▶ **Corollary 1.4.** *There exists $0 < a < b$ such that, if $1/n \leq p = p(n) \leq 0.99$ then*

$$\frac{a}{\sqrt{np}} < q^*(G_{n,p}) < \frac{b}{\sqrt{np}} \quad whp.$$

This result confirms the $\Theta((np)^{-1/2})$ growth rate predicted to hold in this range by Reichardt and Bornholdt [30]: further details of their prediction are given in Section 1.3.

In this extended abstract we give a full proof of the Three Phases Theorem, Theorem 1.2. For Theorem 1.3 we give a proof of the upper bound. We also give a sketch proof of the lower bound, based on an algorithm we call *Swap*, which whp outputs a bipartition achieving the required modularity.

A higher modularity score is taken to indicate a better community division. Thus to determine whether a clustering $\mathcal{A}$ in a graph G shows significant community structure we should compare $q_{\mathcal{A}}(G)$ to the likely (maximum) modularity for an appropriate null model, that is, to the likely value of $q^*(\tilde{G})$ for null model $\tilde{G}$. It is an interesting question which null model may be most appropriate in a given situation. For example, real networks have been shown to exhibit power law degree behaviour and so null models which can mimic this have been suggested; for example the Chung-Lu model [1] and random hyperbolic graphs [17]. However, a natural minimum requirement is not to consider a community division of a real network as statistically significant unless it has higher modularity than the Erdős-Rényi random graph of the same edge density.

1.3 Previous work on Modularity

The vast majority of papers referencing modularity are papers in which real data, clustered using modularity based algorithms, are analysed. Alongside its use in community detection, many interesting properties of modularity have been documented. A basic observation is that, given a graph G without isolated vertices, in each optimal partition, for each part the corresponding induced subgraph of G must be connected.

Properties and modularity of graph classes

The idea of a *resolution limit* was introduced by Fortunato and Barthélemy [12] in 2007: in particular, if a connected component C in an m-edge graph has strictly fewer than $\sqrt{2m}$ edges, then every optimal partition will cluster the vertices of C together. This is so even if the connected component C consists of two large cliques joined by a single edge. This

property highlights the sensitivity of modularity to noise in the network: if that edge between the cliques, perhaps a mistake in the data, had not been there, then the cliques would be in separate parts in every optimal partition.

The complexity is known. Brandes et al. showed in 2007 that finding the (maximum) modularity of a graph is NP-hard [4]. The reduction required some properties of optimal partitions; for example it was shown that a vertex of degree 1 will be placed in the same part as its neighbour in every optimal partition. Indeed, every part in every optimal partition has size at least 2 or is an isolated vertex, see Lemma 1.6.5 in [31]. The paper [4] also began the rigorous study of the modularity of classes of graphs, in particular of cycles and complete graphs. Later Bagrow [3] and Montgolfier et al. [9] proved that some classes of trees have high modularity, and this was extended in [21] to all trees with maximum degree $o(n)$, and indeed to all graphs where the product of treewidth and maximum degree grows more slowly than the number of edges. There is a growing literature concerning the modularity behaviour of different classes of graphs, see for example [3, 9, 20, 21, 29, 32].

Franke and Wolfe in [13] look at a very different topic, namely the distribution of the modularity of a random partition of a graph or random graph, rather than the modularity of the graph, which is the maximum modularity of a partition. The paper covers some random weighted models where the probability of an edge is proportional to the product of the weights of the end-vertices, including the case of the Erdős-Rényi random graph $G_{n,p}$ for $np \to \infty$. They show that the modularity of a random partition is asymptotically normally distributed. Their results do not imply anything about the (maximum) modularity $q^*(G_{n,p})$; see also the discussion in the conclusion of [21].

Statistical Physics predictions

In 2004 Guimera et al. [15] observed through simulations that the modularity of random graphs can be surprisingly high. In [15] they conjectured that, for each (large) constant $c > 1$, if $p = c/n$ then whp $q^*(G_{n,p}) \approx c^{-2/3}$. In 2006 Reichardt and Bornholdt [30] made a different conjecture for the modularity in this range. They assumed that an optimal partition will have parts of equal size, then approximated the number of edges between parts using results from [16], where the authors give spin glass predictions for the minimum number of crossing edges in an equipartition of a random graph. For $p = c/n$ their prediction was $q^*(G_{n,c/n}) \approx 0.97\, c^{-1/2}(1 + o(1))$ whp and we confirm this growth rate. Indeed they predicted $q^*(G_{n,p}) \approx 0.97\sqrt{(1-p)/np}$ which is $\Theta((np)^{-1/2})$ for $1/n \le p \le 0.99$. Hence Corollary 1.4 proves that for a large range of p the prediction of Reichardt and Bornholdt [30] is correct up to constant factors (and refutes that of of Guimera et al.).

1.4   Plan of the paper

The three phases theorem Theorem 1.2 gave an overview of the behaviour of the modularity $q^*(G_{n,p})$, with the three parts (a), (b) and (c) corresponding to increasing edge-probability p, starting with the sparse case. Theorem 1.3, gave more detailed results for the critical region and confirmed the $(np)^{-1/2}$ growth rate conjectured in the physics community.

Our proofs are organised by starting with the sparse case. In Section 2 we prove Theorem 1.2 part (a), by showing that the partition $\mathcal{C}$ into connected components satisfies $q_{\mathcal{C}}(G_{n,p}) \xrightarrow{p} 1$ in the sparse case. We prove Theorem 1.2 part (b) in Section 3: the lower bound follows quickly from counting isolated edges in $G_{n,p}$, and to prove the upper bound we use expansion properties of the giant component. Section 4 concerns the $a(np)^{-1/2}$ lower bound on $q^*(G_{n,p})$ in Theorem 1.3, and we a give a sketch of the proof. In Section 5, we use a robustness result and spectral methods to prove the upper bound $b(np)^{-1/2}$ on $q^*(G_{n,p})$ in Theorem 1.3.

2 The sparse phase: proof of Theorem 1.2 (a)

We can prove that sufficiently sparse random graphs whp have modularity near 1 without developing any extra theory, and we do so here. Lemma 2.2 gives part (a) of the three phases result Theorem 1.2. It is convenient to record first one standard preliminary result on degree tax.

▶ **Lemma 2.1.** *Let the graph G have $m \geq 1$ edges, and let $\mathcal{A}$ be a k-part vertex partition. If $\mathcal{A}$ has k parts then $q_{\mathcal{A}}^D(G) \geq 1/k$; and if x, y are respectively the largest, second largest volume of a part, then $q_{\mathcal{A}}^D(G) \leq x/2m$ and $q_{\mathcal{A}}^D(G) \leq (x/2m)^2 + y/2m$.*

Proof. All the bounds follow from the convexity of $f(t) = t^2$. Let x_i be the volume of the ith part in $\mathcal{A}$. For the lower bound, observe that $x_1, \ldots, x_k \geq 0$ and $\sum_{i=1}^k x_i = 2m$ together imply that $\sum_{i=1}^k x_i^2 \geq k (2m/k)^2 = 4m^2/k$; and thus $q_{\mathcal{A}}^D(G) = \sum_i x_i^2/(2m)^2 \geq 1/k$.

For the upper bounds, observe that $0 \leq x_1, \ldots, x_k \leq x$ and $\sum_{i=1}^k x_i = 2m$ together imply that $\sum_{i=1}^k x_i^2 \leq (2m/x) x^2 = 2mx$; and so $q_{\mathcal{A}}^D(G) \leq x/2m$. Similarly, supposing that $x_k = x$ and $x_i \leq y$ for $i = 1, \ldots, k-1$, we have $\sum_{i=1}^{k-1} x_i^2 \leq (2m-x)y \leq 2my$; and so $q_{\mathcal{A}}^D(G) \leq (x^2 + 2my)/(2m)^2 = (x/2m)^2 + y/2m$. ◀

▶ **Lemma 2.2.** *Let $0 < \varepsilon \leq 1/4$, and let $p = p(n)$ satisfy $n^2 p \to \infty$ and $np \leq 1 + \varepsilon$ for n sufficiently large. Then $q^*(G_{n,p}) \geq q_{\mathcal{C}}(G_{n,p}) > 1 - (4\varepsilon)^2$ whp.*

Proof. Let $m = e(G_{n,p})$, and let X be the maximum number of edges in a connected component of $G_{n,p}$. Note that for the connected components partition $\mathcal{C}$, the edge contribution is 1, and so by the first upper bound on the degree tax in Lemma 2.1, we have $q_{\mathcal{C}}(G_{n,p}) \geq 1 - \frac{X}{m}$. We shall see that when $np \leq 1$ we have $X/m = o(1)$ whp, and so $q_{\mathcal{C}}(G_{n,p}) = 1 - o(1)$ whp. To prove this we break into three ranges of p. The final range, when $1 < np \leq 1 + \varepsilon$ will require a little more care. Observe that since $n^2 p \to \infty$ we have $m \sim n^2 p/2$ whp.

Range 1: $n^2 p \to \infty$ and $np \leq n^{-3/4}$. Whp $G_{n,p}$ consists of disjoint edges. This follows by the first moment method, since the expected number of paths on three vertices is $\Theta(n^3 p^2)$. Hence whp $X/m = 1/m = o(1)$.

Range 2: $n^{-3/4} \leq np \leq 1/2$. Whp all components are trees or unicyclic and have $O(\log n)$ vertices. Hence whp $X = O(\log n)$ and whp $X/m = O\left(\log n / n^2 p\right) = o(1)$.

Range 3: $1/2 \leq np \leq 1$. Since $np \leq 1$, whp the maximum number of edges in any component is $o(n)$ (see the next range). But whp $m = \Theta(n)$, and so whp $X/m = o(1)$.

Range 4: $1 < np \leq 1 + \varepsilon/4$. Let $c = 1 + \varepsilon$. Let $x = x(c)$ be the unique root in $(0,1)$ of $xe^{-x} = ce^{-c}$. Then, for $G_{n,c/n}$, whp $X = (1 + o(1))(1 - x^2/c^2)cn/2$ and each component other than the giant has $O(\log n)$ edges (see for example Theorem 2.14 of [14]). We claim that

$$(1 - x^2/c^2)c < 4\varepsilon/(1+\varepsilon). \tag{1}$$

To see this, let $f(t) = (1+t)e^{-(1+t)} - (1-t)e^{-(1-t)}$ for $t \geq 0$. Then $f(0) = 0$; and for $t > 0$,

$$f'(t) = e^{-(1+t)}(-(1+t) + 1) - e^{-(1-t)}((1-t) - 1) = te^{-1}(e^t - e^{-t}) > 0;$$

and so $f(t) > 0$ for all $t > 0$. Then $f(\varepsilon) > 0$, that is $(1 - \varepsilon)e^{-(1-\varepsilon)} < (1 + \varepsilon)e^{-(1+\varepsilon)}$, and it follows that $1 - x < \varepsilon$. Hence, $1 - x^2/c^2 < 1 - (1 - \varepsilon)^2/(1 + \varepsilon)^2$. But now

$$(1 - x^2/c^2)\,c \;<\; (1 - (1-\varepsilon)^2/(1+\varepsilon)^2)(1+\varepsilon) = \big((1+\varepsilon)^2 - (1-\varepsilon)^2\big)/(1+\varepsilon) = 4\varepsilon/(1+\varepsilon),$$

and we have proved (1). Hence, for $G_{n,c/n}$, whp $X \leq \frac{4\varepsilon}{1+\varepsilon}\frac{n}{2}$; and so by monotonocity this holds also for $G_{n,p}$ (with $p \leq cn$ as here). Also, $e(G_{n,1/n}) \geq \frac{1+\varepsilon/2}{1+\varepsilon}\frac{n}{2}$ whp, and so by monotonocity this holds also for $G_{n,p}$. Now by the last part of Lemma 2.1, whp

$$q_{\mathcal{C}}(G_{n,p}) \geq 1 - (X/m)^2 - O((\log n)/n) \geq 1 - \left(\frac{4\varepsilon}{1+\varepsilon/2}\right)^2 - O((\log n)/n) > 1 - (4\varepsilon)^2.$$

This completes the proof of the lemma.   ◀

3   The middle phase: proof of Theorem 1.2 (b)

It is straightforward to use known results to prove Theorem 1.2 part (b). First we show that the connected components partition $\mathcal{C}$ yields the lower bound. The lower bound will follow also from the lower bound in Theorem 1.3 part (b), but that has quite a long and involved proof, whereas the proof below is only a few lines. As we noted earlier, the upper bound in Theorem 1.3 part (b) will give the upper bound in Theorem 1.2 part (b) for large np, but not when np is small.

3.1   Proof of lower bound

There is a simple reason why the modularity $q^*(G_{n,p})$ is bounded away from 0 whp when the average degree is bounded, namely that whp there is a linear number of isolated edges. First, here is a deterministic lemma.

▶ **Lemma 3.1.** *Let the graph G have $m \geq 2$ edges, and $i \geq \eta m$ isolated edges, where $0 < \eta \leq \frac{1}{2}$. Then $q_{\mathcal{C}}(G) \geq \eta$.*

Proof. Note first that if $i = m$ then $q_{\mathcal{C}}(G) = 1 - 1/m \geq \eta$. Thus we may assume that $i < m$, and so $i \leq m - 2$. Since there are in total $m - i$ edges in the components which are not isolated edges,

$$q_{\mathcal{C}}(G) \geq 1 - \frac{(m-i)^2}{m^2} - \frac{i}{m^2}.$$

Treating i as a continuous variable and differentiating, we see that the bound is an increasing function of i for $i \leq m - 1$; and so, setting $i = \eta m$,

$$q_{\mathcal{C}}(G) \geq 1 - (1 - \eta)^2 - \eta/m = \eta + \eta(1 - \eta - 1/m) \geq \eta,$$

as required.   ◀

Assume that $1 \leq np \leq c_1$. Let X be the number of isolated edges in $G_{n,p}$. Then

$$\mathbb{E}[X] = \binom{n}{2}p(1-p)^{2n-4} = n \cdot (\tfrac{1}{2} + o(1))np\, e^{-2np} \geq n \cdot (\tfrac{1}{2} + o(1))c_1 e^{-2c_1},$$

since $f(x) = xe^{-2x}$ is decreasing for $x > \frac{1}{2}$. A simple calculation shows that the variance of X is $o((\mathbb{E}[X])^2)$: thus by Chebyshev's inequality, whp $X \geq n \cdot \frac{1}{3}c_1 e^{-2c_1}$. Similarly, whp $m = e(G_{n,p}) \leq \frac{2}{3}c_1 n$; and so whp $X/m \geq \frac{1}{2}e^{-2c_1}$. Finally, Lemma 3.1 shows that whp $q_{\mathcal{C}}(G_{n,p}) \geq \eta = \frac{1}{2}e^{-2c_1}$. This completes the proof of the lower bound.

3.2 Proof of upper bound

It is convenient to spell out the upper bound in Theorem 1.2(b) as the following lemma.

▶ **Lemma 3.2.** *Given constants $1 < c_0 < c_1$, there exists $\varepsilon = \varepsilon(c_0, c_1) > 0$ such that, if $c_0 \le np \le c_1$ for n sufficiently large, then whp $q^*(G_{n,p}) < 1 - \varepsilon$.*

For the proof of this lemma we use a result from [19] concerning edge expansion in the giant component. Define a (δ, η)-cut of $G = (V, E)$ to be a bipartition of V into V_1, V_2 such that both sets have at least $\delta|V|$ vertices and $e(V_1, V_2) < \eta|V|$. We need only the case $\delta = 1/3$.

Proof of Lemma 3.2. We employ double exposure. Let $G' \sim \mathcal{G}_{n,c_0/n}$. For each non-edge of G' resample with probability $p' = (p - c_0/n)/(1 - c_0/n)$ to obtain G, so $G \sim G_{n,p}$. Let $\mathcal{A}$ be an optimal partition of G. Observe that whp $m = e(G) < c_1 n$, and then

$$1 - q^*(G) = \frac{1}{2m} \sum_{A \in \mathcal{A}} \left(e_G(A, \bar{A}) + \frac{\mathrm{vol}_G(A)^2}{2m} \right) > \frac{1}{2c_1 n} \sum_{A \in \mathcal{A}} \left(e_{G'}(A, \bar{A}) + \frac{\mathrm{vol}_{G'}(A)^2}{2c_1 n} \right).$$

Thus it suffices to show that whp, for each vertex partition $\mathcal{A}$,

$$\sum_{A \in \mathcal{A}} \left(e_{G'}(A, \bar{A}) + \frac{\mathrm{vol}_{G'}(A)^2}{2c_1 n} \right) \ge 2\varepsilon c_1 n. \tag{2}$$

We will now work solely with G', so we shall drop the subscripts. Whp G' has a unique giant component H, such that H does not admit a $(1/3, \eta)$-cut for a constant $\eta = \eta(c_0) > 0$ by [19] [Lemma 2], and such that $|V(H)| \sim (1 - t_0/c_0)n$ where $t_0 < 1$ satisfies $t_0 e^{-t_0} = c_0 e^{-c_0}$ [10]. Let F be the event that G' has a unique giant component H, such that H does not admit a $(1/3, \eta)$-cut, and $|V(H)| \ge \frac{1}{2}(1 - t_0/c_0)n + 3$. Then the event F holds whp. Let W be a set of vertices such that $|W| \ge \frac{1}{2}(1 - t_0/c_0)n + 3$, and let F_W be the event that F holds and $V(H) = W$. To prove the lemma, it suffices to show that, conditioning on F_W holding, the inequality (2) holds with

$$\varepsilon = \min\{(1 - t_0/c_0)^2/36c_1^2, \eta(1 - t_0/c_0)/2c_1\}.$$

Let $\mathcal{A}$ be any vertex partition which minimises the left side of (2), and let $\mathcal{H}$ be the partition of the giant component H induced by $\mathcal{A}$, that is, $\mathcal{H}$ consists of the parts $A \in \mathcal{A}$ with $A \cap W$ non-empty (since the induced subgraph on A is connected). Relabel $\mathcal{H}$ as $\{W_1, \ldots, W_h\}$ where $h \ge 1$ and $|W_1| \ge \ldots \ge |W_h|$. We will restrict our attention to $\mathcal{H}$. There are two cases to consider.

Case 1. Suppose $|W_1| \ge |W|/3$. As the subgraph induced by W_1 is connected,

$$\mathrm{vol}(W_1) \ge 2(|W_1| - 1) \ge (1 - t_0/c_0)n/3;$$

and so

$$\sum_{A \in \mathcal{A}} \frac{\mathrm{vol}(A)^2}{2c_1 n} \ge \frac{\mathrm{vol}(W_1)^2}{2c_1 n} \ge \frac{(1 - t_0/c_0)^2 n^2}{18c_1 n} \ge 2\varepsilon c_1 n,$$

which yields (2).

Case 2. Now suppose that $|W_i| < |W|/3$ for all $W_i \in \mathcal{H}$. We group the parts to make a bipartition $W = B_1 \cup B_2$ with B_1 and B_2 of similar size. We may for example start with B_1 and B_2 empty, consider the W_i in turn, and each time add W_i to a smaller of B_1 and B_2. This clearly gives $||B_1| - |B_2|| < |W|/3$. Since there is no $(1/3, \eta)$-cut of H in G', we have $e(B_1, B_2) \geq \eta|W|$. But each edge between B_1 and B_2 lies between the parts of $\mathcal{A}$, and so

$$\sum_{A \in \mathcal{A}} e(A, \bar{A}) \geq 2e(B_1, B_2) \geq 2\eta|W| > \eta(1 - t_0/c_0)n \geq 2\varepsilon c_1 n,$$

which again yields (2), and completes the proof.   ◀

4   The $a(np)^{-1/2}$ lower bound on the modularity $q^*(G_{n,p})$

We consider a simple algorithm *Swap* which, given a graph G, runs in linear time (in time $O(n + m)$ if G has n vertices and m edges), and yields a balanced bipartition $\mathcal{A}$ of the vertices. The theorem below shows that $q_{\mathcal{A}}(G_{n,p})$ yields a good lower bound for $q^*(G_{n,p})$.

▶ **Theorem 4.1.** *There are constants c_0 and $a > 0$ such that (a) if $p = p(n)$ satisfies $c_0 \leq np \leq n - c_0$ for n sufficiently large, then whp $q_{\mathcal{A}}(G_{n,p}) \geq \frac{1}{5}\sqrt{\frac{1-p}{np}}$; and (b) if $p = p(n)$ satisfies $1 \leq np \leq n - c_0$ for n sufficiently large, then whp $q_{\mathcal{A}}(G_{n,p}) \geq a\sqrt{\frac{1-p}{np}}$.*

The idea of the proof of Theorem 4.1 is as follows. The algorithm Swap starts with a balanced bipartition of the vertex set into $A \cup B$, which has modularity very near 0 whp. By swapping some pairs (a_i, b_i) between A and B, whp we can increase the edge contribution significantly, without changing the distribution of the degree tax (and without introducing dependencies which would be hard to analyse). We give a sketch proof below but defer the full proof to the extended paper.

Proof of Theorem 4.1 (sketch of the main ideas). Let $n \geq 6$, and let $V = [n]$. We start with the initial bipartition $\mathcal{A}$ of V into $A = \{j \in V : j \text{ is odd}\}$ and $B = \{j \in V : j \text{ is even}\}$. Let $k = k(n) = \lfloor n/6 \rfloor$. Let $V_0 = [4k]$, let $V_1 = \{4k + 1, \ldots, 6k\}$ and let $V_2 = \{6k + 1, \ldots, n\}$. Note that $0 \leq |V_2| \leq 5$: we shall essentially ignore any vertices in V_2. Let $A_i = A \cap V_i$ and $B_i = B \cap V_i$ for $i = 0, 1, 2$. The six sets A_i, B_i are pairwise disjoint with union V. Currently V_0 is partitioned into $A_0 \cup B_0$: the algorithm Swap 'improves' this partition, keeping the other 4 sets fixed. For $i = 1, \ldots, 2k$ let $a_i = 2i - 1$ and $b_i = 2i$, so $A_0 = \{a_1, \ldots, a_{2k}\}$ and $B_0 = \{b_1, \ldots, b_{2k}\}$. The way that we improve the partition $V_0 = A_0 \cup B_0$ is by swapping a_i and b_i for certain values i.

Consider the initial bipartition $\mathcal{A}$. Write G for $G_{n,p}$. It is not hard to show that whp $q_{\mathcal{A}}(G)$ is very near 0. For each $i \in [2k]$ let

$$T_i = e(a_i, B_1) - e(a_i, A_1) + e(b_i, A_1) - e(b_i, B_1),$$

and note that the random variables $T_1, \ldots, T_{2k}$ are iid. Observe that if $T_i > 0$ and we swap a_i and b_i between A_0 and B_0 (that is, replace A_0 by $(A_0 \setminus \{a_i\}) \cup \{b_i\}$ and similarly for B_0) then $e(A, B)$ decreases by T_i, so the edge contribution of the partition increases. The algorithm Swap makes all such swaps (looking only at possible edges between V_0 and V_1). For each $i \in [2k]$, let $(a_i', b_i') = (b_i, a_i)$ if we perform a swap, and let $(a_i', b_i') = (a_i, b_i)$ if not; and let $A_0' = \{a_1', \ldots, a_{2k}'\}$ and $B_0' = \{b_1', \ldots, b_{2k}'\}$. Let us call the resulting balanced bipartition $\mathcal{A}' = (A', B')$, where $A' = A_0' \cup A_1 \cup A_2$ and $B' = B_0' \cup B_1 \cup B_2$. We shall see that $q_{\mathcal{A}'}(G)$ is as required.
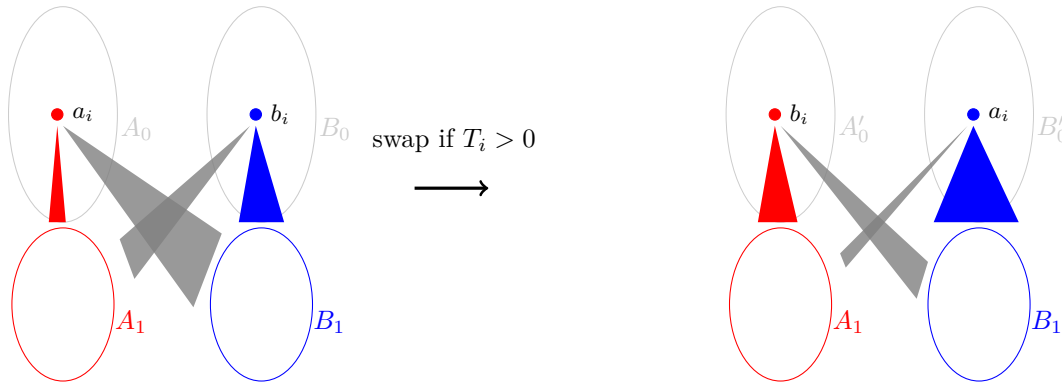
Figure 1 An illustration of the constructed partition in the proof of Theorem 4.1.

Let $T^* = \sum_{i \in [2k]} |T_i|$. Observe that

$$e(A_0', A_1) + e(B_0', B_1) - (e(A_0', B_1) + e(A_1, B_0')) = T^*,$$

so

$$e(A_0', B_1) + e(A_1, B_0') = \tfrac{1}{2}e(V_0, V_1) - \tfrac{1}{2}T^*. \tag{3}$$

This is where $\mathcal{A}'$ will gain over $\mathcal{A}$. The main technical part of the proof is to show that whp T^* is large and we leave this to the full version of the paper. We will also show that the degree tax for $\mathcal{A}'$ has exactly the same distribution as for the initial bipartition $\mathcal{A}$, and it will follow that it is very close to $1/2$ whp.   ◀

5   Upper bounds on modularity

In this section we prove the upper bound on $q^*(G_{n,p})$ in Theorem 1.3, which establishes both part (c) of Theorem 1.2, and the upper bound in part (b) of Theorem 1.2. In Section 5.1 we give bounds on the modularity of a graph G in terms of the eigenvalues of its normalised Laplacian $\mathcal{L}(G)$. In Section 5.2, these results are used, together with spectral bounds from [7] and [8], and a 'robustness' result on modularity, to complete the proof.

5.1   Spectral upper bound on modularity

The main task of this subsection is prove that the modularity of a graph is bounded above by the spectral gap of the normalised Laplacian. We begin with a definition. For an n-vertex graph G with adjacency matrix A_G and no isolated vertices define the *degrees matrix* D to be the diagonal matrix $\operatorname{diag}(d_1, \ldots, d_n)$ and the *normalised Laplacian* to be $\mathcal{L} = I - D^{-1/2}A_G D^{-1/2}$. Here $D^{-1/2}$ is $\operatorname{diag}(d_1^{-1/2}, \ldots, d_n^{-1/2})$. Denote the eigenvalues of $\mathcal{L}$ by $0 = \lambda_0 \leq \ldots \leq \lambda_{n-1}(\leq 2)$, see [6]. We call

$$\max_{i \neq 0} |1 - \lambda_i| = \max\{|1 - \lambda_1|, |\lambda_{n-1} - 1|\}$$

the *spectral gap* of G, and denote it be $\bar{\lambda}(G)$. (In terms of the eigenvalues $\tilde{\lambda}_0 \geq \cdots \geq \tilde{\lambda}_{n-1}$ of $D^{-1/2}A_G D^{-1/2}$, we have $\tilde{\lambda}_i = 1 - \lambda_i$ and so $\bar{\lambda}(G) = \max_{i \neq 0} |\tilde{\lambda}_i| = \max\{|\tilde{\lambda}_1|, |\tilde{\lambda}_{n-1}|\}$.)

▶ **Lemma 5.1.** *Let G be a graph with at least one edge and no isolated vertices. Then*

$$q_{\mathcal{A}}(G) \leq \bar{\lambda}(G)\,(1 - 1/k) \leq \bar{\lambda}(G)$$

for each k-part vertex partition $\mathcal{A}$, and so $q^*(G) \leq \bar{\lambda}(G)$.

The proof of Lemma 5.1 relies on a corollary of the Discrepancy Inequality, Theorem 5.4 of [6], which is an extension of the Expander-Mixing Lemma to non-regular graphs. Write $\bar{S} = V \backslash S$ where $V = V(G)$.

▶ **Lemma 5.2** (Corollary 5.5 of [6]). *Let G be a graph with at least one edge and no isolated vertices. Then for each $S \subseteq V$*

$$e(S, \bar{S}) \geq (1 - \bar{\lambda}(G))\,\mathrm{vol}(S)\mathrm{vol}(\bar{S})/\mathrm{vol}(G).$$

Proof of Lemma 5.1. Let G have $m \geq 1$ edges. Let $\mathcal{A} = \{A_1, \ldots, A_k\}$ be a vertex partition of G. Lemma 5.2 guarantees many edges between the parts of $\mathcal{A}$. The edge contribution satisfies

$$1 - q_{\mathcal{A}}^{E}(G) = \frac{1}{2m} \sum_i e(A_i, \bar{A}_i) \geq (1 - \bar{\lambda}) \frac{1}{4m^2} \sum_i \mathrm{vol}(A_i)\mathrm{vol}(\bar{A}_i);$$

and

$$\frac{1}{4m^2} \sum_i \mathrm{vol}(A_i)\mathrm{vol}(\bar{A}_i) = \frac{1}{4m^2} \sum_i \mathrm{vol}(A_i)(2m - \mathrm{vol}(A_i)) = 1 - q_{\mathcal{A}}^{D}(G).$$

Hence

$$1 - q_{\mathcal{A}}^{E}(G) \geq (1 - \bar{\lambda})(1 - q_{\mathcal{A}}^{D}(G)),$$

and so

$$q_{\mathcal{A}}(G) = q_{\mathcal{A}}^{E}(G) - q_{\mathcal{A}}^{D}(G) \leq \bar{\lambda}(1 - q_{\mathcal{A}}^{D}(G)) \leq \bar{\lambda}(1 - \tfrac{1}{k})$$

(since $q_{\mathcal{A}}^{D}(G) \geq 1/k$ by Lemma 2.1). This completes the proof. ◀

5.2 The $b(np)^{-1/2}$ upper bound on the modularity $q^*(G_{n,p})$.

We are now ready to prove the spectral upper bound for $q^*(G_{n,p})$. Let us restate the upper bound in Theorem 1.3 as a lemma.

▶ **Lemma 5.3.** *There is a constant b such that for $0 < p = p(n) \leq 1$*

$$q^*(G_{n,p}) \leq \frac{b}{\sqrt{np}} \quad whp.$$

Proof. Notice first that it suffices to show that there exist c_0 and b such that for $np \geq c_0$ whp $q^*(G_{n,p}) \leq b/\sqrt{np}$, and then replace b by $\max\{\sqrt{c_0}, b\}$.

For $p \gg \log^2 n/n$, the result follows directly from Lemma 5.1, and Theorem 3.6 of Chung, Vu and Lu [7] (see also (1.2) in [8]), which shows that

$$\bar{\lambda}(G_{n,p}) \leq 4(np)^{-1/2}(1 + o(1)) \quad whp.$$

For the remainder of the proof we assume that $c_0/n \leq p \leq 0.99$ for some large constant $c_0 \geq 1$. We will use the spectral bound in Lemma 5.1 on a subgraph H which is obtained from the random graph $G = G_{n,p}$ by deleting a small subset of the vertices (and the incident edges).

Following the construction in [8], let H be the induced subgraph of G obtained as follows.

- Initially set $H = G \setminus \{v \in V(G) \; : \; d_v < (n-1)p/2\}$.
- While there is a vertex $v \in V(H)$ with at least 100 neighbours in $V(G) \setminus V(H)$, remove v from H.

Let V' be the set of deleted vertices, and let E' be the set of deleted edges (the edges incident with vertices in V'). Then by Theorem 1.2 of Coja-Oghlan [8], assuming that c_0 is sufficiently large, there are positive constants c_1 and c_2 such that whp $|V'| \leq ne^{-np/c_2}$ and $\bar{\lambda}(H) \leq c_1(np)^{-1/2}$.

We want a bound on $|E'|$, not $|V'|$. By the proof of Corollary 2.3 in [8], whp in $G_{n,p}$ we have $\mathrm{vol}(S) \leq 2np|S| + ne^{-np/1500}$ simultaneously for each set S of vertices. (The result is stated with $\mathrm{vol}(S)$ replaced by $|N_G(S)|$, the number of neighbours of S outside S, but the proof actually shows the result for $\mathrm{vol}(S)$.) Hence, noting also that $np \geq 1$ and setting $c_3 = \max\{c_2, 1500\}$, whp

$$|E'| \leq \mathrm{vol}(V') \leq 2n^2 p \, e^{-np/c_2} + ne^{-np/1500} \leq 3n^2 p \, e^{-np/c_3} \leq e(G) \cdot 9e^{-np/c_3},$$

where the last inequality follows since whp $e(G) \geq n^2 p/3$. By making c_0 larger if necessary, we can ensure that $9e^{-np/c_3} \leq (1/3)(np)^{-1/2}$, and so whp $|E'|/e(G) \leq (1/3)(np)^{-1/2}$. Now, by Lemma 5.1, whp

$$q^*(G \setminus E') = q^*(H) \leq \bar{\lambda}(H) \leq c_1(np)^{-1/2}.$$

One of the 'robustness' results in the full paper says that, if H is a graph and E' is a proper subset of the edges, then $|q^*(H) - q^*(H \setminus E')| \leq 3|E'|/e(H)$. Using this result, whp

$$q^*(G) \leq q^*(G \setminus E') + 3|E'|/e(G) \leq (c_1 + 1)(np)^{-1/2},$$

and the proof is complete. ◄

6 Concluding remarks

We have presented results on $q^*(G_{n,p})$, focussing on the three phases as the average degree moves past 1 and then grows to ∞, and on the $\Theta((np)^{-1/2})$ result. The full paper [24] (as mentioned earlier) also contains corresponding results for $q^*(G_{n,m})$; and it contains some other results, including concentration for both $q^*(G_{n,p})$ and $q^*(G_{n,m})$.

There is further related work in progress: concerning the modularity of very dense graphs and random graphs, see [25]; concerning modularity and edge-sampling (it may be expensive to test if an edge is present), see [23]; and concerning extreme values of modularity (to set random results in context), see [22].

References

1   William Aiello, Fan Chung, and Linyuan Lu. A random graph model for power law graphs. *Experimental Mathematics*, 10(1):53–66, 2001. `doi:10.1080/10586458.2001.10504428`.

2   Aaron F. Alexander-Bloch, Nitin Gogtay, David Meunier, Rasmus Birn, Liv Clasen, Francois Lalonde, Rhoshel Lenroot, Jay Giedd, and Edward T. Bullmore. Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Frontiers in Systems Neuroscience*, 4, 2010. `doi:10.3389/fnsys.2010.00147`.

3   James P. Bagrow. Communities and bottlenecks: Trees and treelike networks have high modularity. *Physical Review E*, 85(6):066118, 2012. `doi:10.1103/PhysRevE.85.066118`.

4   Ulrik Brandes, Daniel Delling, Marco Gaertler, Robert Görke, Martin Hoefer, Zoran Nikoloski, and Dorothea Wagner. On finding graph clusterings with maximum modularity. In *Graph-Theoretic Concepts in Computer Science*, pages 121–132. Springer, 2007. `doi:10.1007/978-3-540-74839-7_12`.

5   Ulrik Brandes, Daniel Delling, Marco Gaertler, Robert Gorke, Martin Hoefer, Zoran Niko-loski, and Dorothea Wagner. On modularity clustering. *Knowledge and Data Engineering, IEEE Transactions on*, 20(2):172–188, 2008. `doi:10.1109/TKDE.2007.190689`.

6   Fan Chung. *Spectral graph theory*, volume 92. American Mathematical Soc. Providence, RI, 1997. `doi:10.1090/cbms/092`.

7   Fan Chung, Linyuan Lu, and Van Vu. The spectra of random graphs with given expected degrees. *Internet Mathematics*, 1(3):257–275, 2003. `doi:10.1073/pnas.0937490100`.

8   Amin Coja-Oghlan. On the Laplacian Eigenvalues of $G_{n,p}$. *Combinatorics, Probability and Computing*, 16:923–946, 2007. `doi:10.1017/S0963548307008693`.

9   Fabien De Montgolfier, Mauricio Soto, and Laurent Viennot. Asymptotic modularity of some graph classes. In *Algorithms and Computation*, pages 435–444. Springer, 2011. `doi:10.1007/978-3-642-25591-5_45`.

10   P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.

11   Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3):75–174, 2010. `doi:10.1016/j.physrep.2009.11.002`.

12   Santo Fortunato and Marc Barthélemy. Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, 104(1):36–41, 2007. `doi:10.1073/pnas.0605965104`.

13   Beate Franke and Patrick J. Wolfe. Network modularity in the presence of covariates. *preprint arXiv:1603.01214*, 2016.

14   Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2015. `doi:10.1017/cbo9781316339831`.

15   Roger Guimerà, Marta Sales-Pardo, and Luís A. Nunes Amaral. Modularity from fluctuations in random graphs and complex networks. *Physical Review E*, 70:025101, 2004. `doi:10.1103/physreve.70.025101`.

16   Ido Kanter and Haim Sompolinsky. Graph optimisation problems and the Potts glass. *Journal of Physics A*, 20(11):L673, 1987. `doi:10.1088/0305-4470/20/11/001`.

17   Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marián Boguná. Hyperbolic geometry of complex networks. *Physical Review E*, 82(3):036106, 2010. `doi:10.1103/physreve.82.036106`.

18   Andrea Lancichinetti and Santo Fortunato. Limits of modularity maximization in community detection. *Physical Review E*, 84(6):066122, 2011. `doi:10.1103/physreve.84.066122`.

19   Malwina J. Luczak and Colin McDiarmid. Bisecting sparse random graphs. *Random Structures & Algorithms*, 18(1):31–38, 2001. `doi:10.1002/1098-2418(200101)18:1<31::aid-rsa3>3.3.co;2-t`.

20   Colin McDiarmid and Fiona Skerman. Modularity in random regular graphs and lattices. *Electronic Notes in Discrete Mathematics*, 43:431–437, 2013. `doi:10.1016/j.endm.2013.07.063`.

21   Colin McDiarmid and Fiona Skerman. Modularity of regular and treelike graphs. *Journal of Complex Networks*, 5, 2017. `doi:10.1093/comnet/cnx046`.

22   Colin McDiarmid and Fiona Skerman. Extreme values of modularity. In preparation, 2018.

23   Colin McDiarmid and Fiona Skerman. Modularity and edge-sampling. In preparation, 2018.

24   Colin McDiarmid and Fiona Skerman. Modularity of Erdős-Rényi random graphs. Manuscript, 2018.

25   Colin McDiarmid and Fiona Skerman. Modularity of very dense graphs. In preparation, 2018.

26 Mark E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010. `doi:10.1093/acprof:oso/9780199206650.001.0001`.

27 Mark E. J. Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69(2):026113, 2004. `doi:10.1103/physreve.69.026113`.

28 Mason A. Porter, Jukka-Pekka Onnela, and Peter J. Mucha. Communities in networks. *Notices of the AMS*, 56(9):1082–1097, 2009.

29 Liudmila Ostroumova Prokhorenkova, Paweł Prałat, and Andrei Raigorodskii. Modularity in several random graph models. *Electronic Notes in Discrete Mathematics*, 61:947–953, 2017. `doi:10.1016/j.endm.2017.07.058`.

30 Jörg Reichardt and Stefan Bornholdt. When are networks truly modular? *Physica D: Nonlinear Phenomena*, 224(1):20–26, 2006. `doi:10.1016/j.physd.2006.09.009`.

31 Fiona Skerman. *Modularity of Networks*. PhD thesis, University of Oxford, 2016.

32 Stojan Trajanovski, Huijuan Wang, and Piet Van Mieghem. Maximum modular graphs. *The European Physical Journal B-Condensed Matter and Complex Systems*, 85(7):1–14, 2012. `doi:10.1140/epjb/e2012-20898-3`.

Counting Planar Tanglegrams

Dimbinaina Ralaivaosaona

Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, South Africa
naina@sun.ac.za

Jean Bernoulli Ravelomanana

Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, South Africa
jeanr@sun.ac.za

Stephan Wagner[1]

Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, South Africa
swagner@sun.ac.za

──── **Abstract** ────

Tanglegrams are structures consisting of two binary rooted trees with the same number of leaves and a perfect matching between the leaves of the two trees. We say that a tanglegram is planar if it can be drawn in the plane without crossings. Using a blend of combinatorial and analytic techniques, we determine an asymptotic formula for the number of planar tanglegrams with n leaves on each side.

1 Introduction and statement of results

A *tanglegram* is a structure consisting of two (unordered, non-plane) binary trees with the same number of leaves and a perfect matching between the leaf sets. Tanglegrams occur naturally in the study of cospeciation and coevolution (see [10, 12]), where the two trees are phylogenetic trees, and also in computer science in the analysis of software projects and clustering problems [2]. Formally, we can define a tanglegram as a triplet (T, ϕ, S), where T and S are two rooted binary trees with the same number of leaves n, and ϕ is a bijection between the leaf sets. The size of a tanglegram is the number of leaves in each tree. We draw a tanglegram (T, ϕ, S) with one tree on top and the other on the bottom; the corresponding bijection ϕ is represented by inter-tree edges (see Figure 1 for an example of a tanglegram: edges between leaves are dashed). In such a representation, tree edges are not allowed to cross, while inter-tree edges may have crossings.

Note that a tanglegram can usually be drawn in many different ways. Figure 2 shows two representations of a tanglegram (the same as in Figure 1), where corresponding leaves are indicated by identical colours and labels. Formally, we consider two tanglegrams (T, ϕ, S) and (T', ϕ', S') isomorphic if there are (rooted tree) isomorphisms g from T to T' and h from S to S' such that $\phi' = g \circ \phi \circ h^{-1}$.
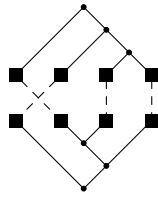
Figure 1 A tanglegram of size 4.
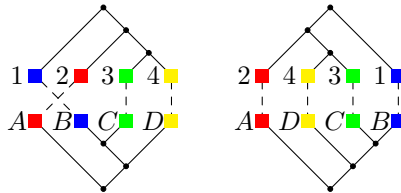
Figure 2 Two different representations of a tanglegram.

An equivalence class of tanglegrams under this definition of isomorphism formally corresponds to a double coset of the symmetric group, see [1] for details. We point out that a tanglegram isomorphism cannot interchange the top and the bottom tree.

It is desirable, both for aesthetic and practical purposes, to represent a tanglegram with a minimum number of crossings between inter-tree edges. For instance, the left representation in Figure 2 has one crossing, the right representation is crossing-free. The problem of determining the minimum number of crossings for a given tanglegram is known as the *Tanglegram Layout* (TL) problem [2]. This problem is, just like the crossing number problem for graphs, NP-hard in general [5].

In this paper, we consider a related enumerative question. In [1], Billey, Konvalinka and Matsen established a formula for the number t_n of tanglegrams of size n (up to isomorphism). The first few terms of the sequence t_n are given by $1, 1, 2, 13, 114, 1509, 25595, 535753, 13305590, 382728552, \ldots$, see also [11, A258620]. They also obtained an asymptotic formula for t_n: for $n \to \infty$, we have

$$t_n \sim \frac{2^{2n-\frac{3}{2}} \cdot n^{n-\frac{5}{2}}}{\sqrt{\pi} \cdot e^{n-\frac{1}{8}}}.$$

Based on the results of Billey, Konvalinka and Matsen, properties of random tanglegrams were investigated in [9].

Here, we ask a similar question: how many tanglegrams of size n (up to isomorphism) are there that can be drawn without crossing? In analogy to planar graphs, we will call them *planar tanglegrams*. For example, the tanglegram in Figure 1 has a crossing-free representation (as Figure 2 shows) and is thus planar. All tanglegrams of size $1, 2$, or 3 are easily seen to be planar. Among the thirteen tanglegrams of size 4, only two are not planar, see Figure 3. Indeed, it can be shown that a tanglegram is planar if and only if all of the subtanglegrams induced by four leaf pairs are planar, in analogy to Kuratowski's celebrated characterisation of planar graphs (see [3]). Here, an induced subtanglegram is a tanglegram obtained in the following way: pick some leaf pairs, then take the smallest subtree on each side that contains the respective leaves, and suppress internal vertices that are no longer needed.

Our approach to the enumeration of planar tanglegrams is based on generating functions. Our first main result characterises the generating function $T(x)$ of planar tanglegrams by means of a functional equation.
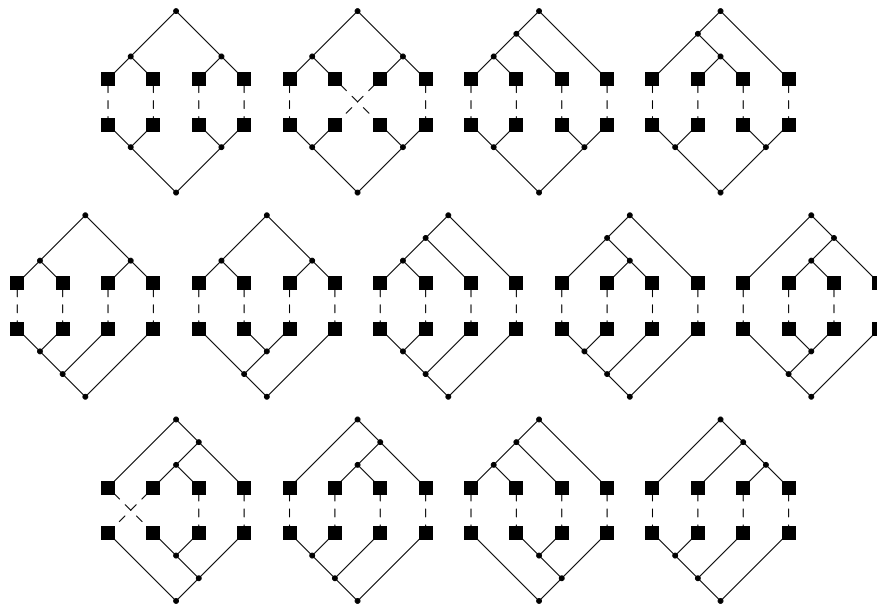
Figure 3 All tanglegrams of size 4. Only the second and tenth tanglegram cannot be drawn without crossing.

▶ **Theorem 1.** *Let $T(x)$ be the (ordinary) generating function for the number of planar tanglegrams, counted up to isomorphism. The function $T(x)$ is uniquely determined by the following system of functional equations involving two auxiliary functions $A(x)$ and $H(x)$:*

$$A(x) = \sum_{r=1}^{\infty} \frac{1}{(r+1)^2} \binom{2r}{r}^2 x^r (1 - A(x))^{r+1}, \tag{1}$$

$$H(x) = \frac{xA(x)}{2}, \tag{2}$$

$$T(x) = H(T(x)) + x + \frac{T(x^2)}{2}. \tag{3}$$

It turns out that $A(x)$ in the theorem above is the generating function for the number of ordered pairs of triangulations of polygons without common diagonals. Moreover, the auxiliary function $H(x)$ that occurs in Theorem 1 has a natural combinatorial interpretation as well: it is the generating function for *irreducible* planar tanglegrams (to be defined in the following). We will see that unordered pairs of triangulations of polygons without common diagonals and irreducible tanglegrams are in one-to-one correspondence. In order to define irreducible planar tanglegrams, we first need the concept of proper subtanglegrams.

A binary subtree T' of a binary tree T is an induced binary tree consisting of a vertex and all its successors. We call the binary subtree T' a proper binary subtree if it is not a leaf and the root of T' is different from the root of T. A subtanglegram of a planar tanglegram consists of a binary subtree of the top tree and a binary subtree of the bottom tree with the same number of leaves, where each leaf of the top subtree is matched to a leaf of the bottom subtree. Moreover, a subtanglegram is called a proper subtanglegram if the two corresponding binary subtrees are proper. Figure 4 shows a proper subtanglegram of a planar tanglegram. An irreducible planar tanglegram is a planar tanglegram which does not contain any proper subtanglegrams and which has more than one leaf in each tree. For example, in Figure 3, the first, third, fifth, seventh, and ninth tanglegram contain proper subtanglegrams
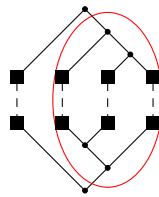
Figure 4 A proper subtanglegram of a planar tanglegram.

Table 1 The first 10 values of T_n and H_n.

n	1	2	3	4	5	6	7	8	9	10
T_n	1	1	2	11	76	649	6 173	63 429	688 898	7 808 246
H_n	0	1	1	5	34	273	2 436	23 391	237 090	2 505 228

of size 2; the seventh and eighth contain proper subtanglegrams of size 3. Only five of the eleven planar tanglegrams shown in the figure are irreducible.

Let T_n be the number of planar tanglegrams of size n, and let H_n be the number of irreducible planar tanglegrams of size n. It is easy (with the help of a computer algebra system) to determine the first few values of T_n and H_n from the functional equations in Theorem 1 – see Table 1. Figure 3 illustrates the values $T_4 = 11$ and $H_4 = 5$. The sequence H_n also occurs in a different context as the number of proper diagonals of the n-dimensional associahedron, see [8] and [11, A257887].

Several ingredients are needed in order to prove Theorem 1. In the next section, we show a bijection between pairs of triangulations of polygons without common diagonals and irreducible planar tanglegrams. Thereafter, we use this bijection to obtain functional equations for the generating function of irreducible planar tanglegrams and related generating functions. Finally, we derive a functional equation relating the generating function of planar tanglegrams with the generating function of irreducible planar tanglegrams. An important feature of irreducible planar tanglegrams is the fact that their embeddings in the plane are almost unique, see Proposition 5. Moreover, every planar tanglegram can be reduced to an irreducible planar tanglegram by contracting maximal proper subtanglegrams.

In order to determine the asymptotic behaviour of T_n, we study the analytic properties of its generating function and eventually apply singularity analysis. This is also done in several steps, starting from the function $A(x)$ that is closely related to an elliptic integral, from which the behaviour of $H(x)$ is derived. As a side result, we also obtain the asymptotic behaviour of the coefficients H_n (see Theorem 11). Our main analytic result regarding the generating function $T(x)$ reads as follows:

▶ **Theorem 2.** *The generating function T enumerating planar tanglegrams satisfies the following properties:*

(i) *Let ρ be the radius of convergence of T. There exist positive real numbers θ and ϵ such that T is analytic in*

$$\Delta = \{x \,:\, |x| < \rho + \epsilon \text{ and } |\operatorname{Arg}(x - \rho)| > \theta\},$$

and for $x \in \Delta$, we have:

$$T(x) = \alpha + C_1(\rho - x) + C_2(\rho - x)^2 + B(\rho - x)^2 \log(\rho - x) + O(|(\rho - x)^3 \log(\rho - x)|). \quad (4)$$

Here, C_1, C_2 and B are constants that can be computed numerically, and $\alpha = \frac{4 - \pi}{4\pi}$.

(ii) *As $n \to \infty$, the nth coefficient T_n of T satisfies the asymptotic formula*

$$T_n \sim C \cdot n^{-3} \cdot \rho^{-n},$$

where $\rho \approx 0.0633892927$, $\rho^{-1} \approx 15.7755349051$ and $C \approx 0.0078873668$.

▶ Remark. From the analytic behaviour of T given in the previous theorem, it follows that T cannot be algebraic (see [6, Theorem VII.7] and [6, Theorem VII.8]).

The first property allows us to use singularity analysis on the generating function $T(x)$ and obtain the asymptotic formula in the second statement. This analysis is outlined in Section 3. We remark that, once the enumeration problem has been solved, it will also be possible to study statistics of planar tanglegrams. This is left as a future project.

2   Deriving the functional equations

2.1   Irreducible tanglegrams and triangulations

First, we work with rooted plane binary trees, which are rooted binary trees with a plane embedding, so that left and right child of every vertex are distinguishable. It is well known that rooted plane binary trees are counted by the Catalan numbers. We denote by $\mathcal{C}_b$ the set of ordered pairs of rooted plane binary trees with the same number of leaves. If we label the leaves canonically (from left to right) and match leaves with the same label, we obtain a planar tanglegram. Every planar tanglegram can be obtained in this way, but of course several pairs of rooted plane binary trees may represent the same planar tanglegram. The next proposition relates pairs of triangulations of a polygon and elements of $\mathcal{C}_b$, based on the well-known bijection between rooted plane binary trees and triangulations.

▶ **Proposition 3.** *To every element (T_1, T_2) of $\mathcal{C}_b$ with n leaves corresponds a unique pair of triangulations of an $(n + 1)$-gon. The tanglegram associated with (T_1, T_2) contains a proper subtanglegram if and only if the corresponding pair of triangulations has a common diagonal.*

Proof. The bijection between binary trees and triangulations is a classical application of the plane dual (see [6, Section I.5.3]): given a rooted plane binary tree with n leaves, draw non-intersecting lines from the root and all leaves to infinity. These lines and the edges of the tree divide the plane into regions. We place a vertex in each of these regions and connect two such vertices by an edge if the corresponding regions share part of their boundaries. The result is a triangulation of an $(n + 1)$-gon, and the correspondence is bijective. A canonical way to label the vertices of the triangulation is to number them clockwise, starting from the root of the tree – see Figure 5.

This also yields a bijection between the elements of $\mathcal{C}_b$ and pairs of triangulations. It is not difficult to see that diagonals of triangulations correspond to proper subtrees, so that a common diagonal in a pair of triangulations corresponds to a pair of proper subtrees whose leaves are matched to each other, i.e. a proper subtanglegram. This is illustrated in Figure 6 for the tanglegram in Figure 4.   ◀

We call an element of $\mathcal{C}_b$ that corresponds to an irreducible planar tanglegram a *representation* of that irreducible tanglegram. The next theorem relates irreducible planar tanglegrams and their representations to pairs of triangulations.
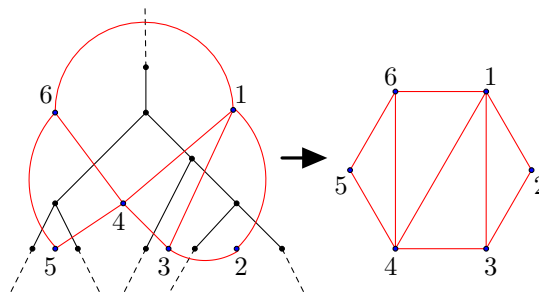
■ **Figure 5** The correspondence between binary trees and triangulations
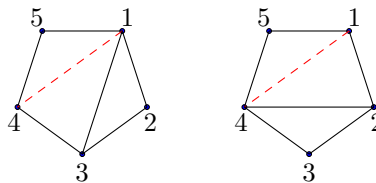
■ **Figure 6** The triangulations corresponding to the two halves of the tanglegram in Figure 4 (the bottom tree is reflected by a horizontal axis before applying the bijection). The common diagonal is indicated in red and dashed.

▶ **Theorem 4.** *The following statements hold:*

(1) *To every representation of an irreducible planar tanglegram of size n corresponds a unique ordered pair of triangulations of an $(n+1)$-gon without common diagonals.*

(2) *There is a bijection between irreducible planar tanglegrams of size n and unordered pairs of triangulations of an $(n+1)$-gon that do not have a common diagonal.*

The first part of Theorem 4 is a consequence of Proposition 3. In order to prove the second part, we first have to show that an irreducible planar tanglegram has a unique representation up to homeomorphism.

▶ **Proposition 5.** *Every irreducible planar tanglegram with more than two leaves in each tree has precisely two possible representations, which are mirror images of each other.*

Proposition 5 is obtained by means of a famous theorem of Whitney:

▶ **Theorem 6** (Whitney [13]). *Every 3-connected planar graph has a unique plane embedding up to homeomorphism.*

The main idea is to apply Whitney's theorem to the graph obtained from a tanglegram by removing the leaves on each side (but leaving the connecting edges) and connecting the roots by an additional edge. Let us call this process *smoothing* – see Figure 7. We have the following proposition, whose proof is given in the appendix.

▶ **Proposition 7.** *The graph obtained by smoothing a tanglegram is 3-regular and 3-connected if the tanglegram is irreducible and has more than 2 leaves in each tree.*

We can now proceed to the proof of Proposition 5.

Proof of Proposition 5. Let I be an irreducible tanglegram with more than two leaves in each tree, and let T_1 and T_2 be the corresponding binary trees. Every representation yields a plane embedding of the graph that is obtained by smoothing. By Whitney's theorem and
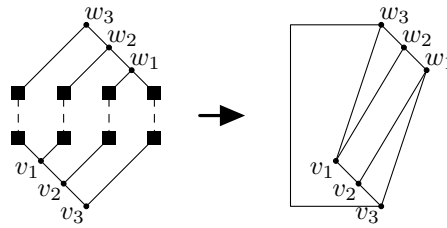
Proposition 7, there are only two possible representations, which are mirror images of each other. Suppose that the mirror images are identical. Then the mirror images of T_1 and T_2 are respectively the same as T_1 and T_2, which implies that the left and right branches of T_1 and T_2 are the same. So the branches of T_1 and T_2 induce proper subtanglegrams since they contain more than one leaf each. This contradicts the assumption that I is irreducible. Thus, we find that an irreducible tanglegram with more than two leaves on each side has precisely two distinct irreducible representations that are mirror images of each other. ◄

We conclude this section with the proof of part (2) of Theorem 4.

Proof of Theorem 4, part (2). For $n = 2$, the statement is clearly true since both sets contain exactly one element. Now suppose that $n > 2$. Let $\mathcal{P}_n$ be the set of pairs of triangulations of an $(n + 1)$-gon without common diagonal, and let $\mathcal{I}_n$ be the set of representations of irreducible tanglegrams of size n. Moreover, denote by $\mathcal{P}'_n$ the set of unordered pairs of triangulations of an $(n + 1)$-gon without common diagonal and $\mathcal{I}'_n$ the set of irreducible tanglegrams.

By Propositon 5, we know that to every element of $\mathcal{I}'_n$, there are two distinct corresponding elements of $\mathcal{I}_n$. Moreover, to every pair of triangulations in $\mathcal{P}'_n$, there are two distinct corresponding ordered pairs in P_n. This is because the two triangulations of an element of $\mathcal{P}'_n$ have to be distinct, as they would otherwise have a common diagonal. By the first part of the theorem, there is a bijection between $\mathcal{P}_n$ and $\mathcal{I}_n$. Since $\mathcal{I}'_n$ and $\mathcal{I}_n$ are in a 2–1 correspondence, as are $\mathcal{P}'_n$ and $\mathcal{P}_n$, it follows that there is a bijection between $\mathcal{P}'_n$ and $\mathcal{I}'_n$. ◄

▶ **Remark.** The only symmetric irreducible tanglegram (equal to its own mirror image) is the tanglegram with two leaves in each tree.

2.2 From bijections to generating functions

Since there is a bijection between irreducible tanglegrams and unordered pairs of triangulations of a polygon without common diagonals, the generating functions of the two combinatorial objects are the same. We derive a functional equation for the generating function of pairs of triangulations without common diagonals using the inclusion-exclusion method described in [6, Section III.7]. We consider ordered pairs of triangulations (of the same polygon) in which some of the common diagonals (not necessarily all and possibly none) are marked. Let $\mathcal{T}$ be the family of ordered pairs of triangulations (of the same polygon) without marked diagonals. For a pair (T_1, T_2) in $\mathcal{T}$, $M(T_1, T_2)$ is the set of all possible configurations of the pair (T_1, T_2) with marked diagonals. For every $m \in M(T_1, T_2)$, we denote by $N(m)$ the

number of marked diagonals in m. Lastly, $n(T_1, T_2)$ is the number of triangles in each of the two triangulations T_1 and T_2. Now define a bivariate generating function $A(x, v)$ by

$$A(x, v) = \sum_{(T_1, T_2) \in \mathcal{T}} \left(\sum_{m \in M(T_1, T_2)} v^{N(m)} \right) x^{n(T_1, T_2)}. \tag{5}$$

We have the following key observation: if T_1 and T_2 are two triangulations of a polygon with $k(T_1, T_2)$ common diagonals, then

$$\sum_{m \in M(T_1, T_2)} v^{N(m)} = (1 + v)^{k(T_1, T_2)}.$$

Indeed, we can choose to mark a common diagonal, which yields a factor v, or not to mark it, which yields a factor 1. Thus,

$$A(x, v) = \sum_{(T_1, T_2) \in \mathcal{T}} \left(\sum_{m \in M(T_1, T_2)} v^{N(m)} \right) x^{n(T_1, T_2)} = \sum_{(T_1, T_2) \in \mathcal{T}} (1 + v)^{k(T_1, T_2)} x^{n(T_1, T_2)}. \tag{6}$$

If we plug in $v = -1$, all pairs with $k(T_1, T_2) \neq 0$ vanish, and we are left precisely with those ordered pairs of triangulations that have no common diagonals. Hence $A(x, -1)$ represents the generating function for ordered pairs of triangulations without common diagonals. Next we prove that $A(x) = A(x, -1)$ satisfies the functional equation (1).

Proof of (1). Let $\mathcal{A}$ be the set of all configurations consisting of two triangulations of a polygon with vertices labelled $1, 2, \ldots, n$ with some of the common diagonals potentially marked. Then $A(x, v)$, as defined in (6), is the bivariate generating function corresponding to $\mathcal{A}$, where the exponents of x and v indicate the number of triangles in each triangulation and the number of common diagonals respectively.

We can decompose an element of $\mathcal{A}$ in the following way: the marked common diagonals divide the polygon into one or more subpolygons. One of them (let us call it P) contains the side from vertex 1 to vertex 2. This polygon P is bounded by edges of the larger polygon and marked diagonals that separate it from smaller elements of $\mathcal{A}$. Thus we have a decomposition into a pair of triangulations without marked diagonals (inside of polygon P) that is surrounded by sides of the larger polygon and elements of $\mathcal{A}$. Let r be the number of triangles in each of the triangulations of P; then P has $r + 2$ sides. The number of possibilities for each of the triangulations is the Catalan number $C_r = \frac{1}{r+1} \binom{2r}{r}$, and each of the $r + 1$ sides of P other than the side between vertices 1 and 2 is either a side of the whole polygon or a marked diagonal that separates off a smaller element of $\mathcal{A}$.

This decomposition can be translated to the functional equation

$$A(x, v) = \sum_{r=1}^{\infty} C_r^2 x^r (1 + v A(x, v))^{r+1}. \tag{7}$$

From (6) we have

$$A(x, v) = \sum_{(T_1, T_2) \in \mathcal{T}} (1 + v)^{k(T_1, T_2)} x^{n(T_1, T_2)},$$

where $n(T_1, T_2)$ is the number of triangles and $k(T_1, T_2)$ is the number of common diagonals in (T_1, T_2). Setting $v = -1$, all pairs of triangulations $(T_1, T_2) \in \mathcal{T}$ such that $k(T_1, T_2) \neq 0$ vanish, as mentioned before. This means that all pairs of triangulations (T_1, T_2) which have a common diagonal will not contribute to the sum for $A(x, -1)$. In other words, only the

pairs of triangulations without common diagonal contribute to $A(x, -1) = A(x)$, i.e. $A(x)$ is the generating function for pairs of triangulations without common diagonals. The equation

$$A(x) = \sum_{r=1}^{\infty} \frac{1}{(r+1)^2} \binom{2r}{r}^2 x^r (1 - A(x))^{r+1} \tag{8}$$

follows immediately from (7). ◀

Recall that there is a 2–1 correspondence between ordered and unordered pairs of triangulations without common diagonals, except for the trivial case of triangulations of a triangle. This and Theorem 4 yield the following proposition.

▶ **Proposition 8.** *The generating function $H(x)$ of irreducible tanglegrams is given by*

$$H(x) = \frac{xA(x)}{2}.$$

Proof. The coefficient of x^r in $A(x)$ corresponds to pairs of triangulations without common diagonal and r triangles in each triangulation. When we transform a triangulation of an $(r+2)$-gon into a planted binary tree, we obtain a planted binary tree with $r+1$ leaves. So, by the first part of Theorem 4, the coefficient of x^r in $A(x)$ is the number of representations of irreducible tanglegrams which have $r+1$ leaves on each side. Multiplying $A(x)$ by x gives us the generating function of representations of irreducible tanglegrams. From Theorem 4, we know that to every irreducible tanglegram with more than two leaves on each side, there are two irreducible representations. The statement of the proposition follows. ◀

▶ Remark. The coefficient of x^2 in $H(x)$ is $\frac{1}{2}$. We maintain it as it will help us later to take symmetries into account when the irreducible tanglegram has two leaves.

Proposition 8 gives us equation (2). It only remains to prove (3) to complete the proof of Theorem 1.

Proof of (3)**.** We would like to prove the identity

$$T(x) = H(T(x)) + x + \frac{T(x^2)}{2}.$$

The term x accounts for the tanglegram with only one leaf in each tree. Now consider an arbitrary planar tanglegram T with more than one leaf in each tree. It has maximal proper subtanglegrams (with respect to inclusion) $T_1, T_2, \ldots, T_k$ for some nonnegative integer k (if all T_j's have size 1, then T is irreducible). For each of these subtanglegrams T_j we have two proper binary subtrees T_j' and T_j'' in the two trees that constitute T. Replace both of them by leaves, and include an inter-tree edge between these two leaves. Contracting each maximal proper subtanglegram to a single pair of leaves in this way, we obtain an irreducible tanglegram (see Figure 8). Conversely, if I is an irreducible planar tanglegram, we can replace each pair of matched leaves in I by some planar tanglegram (possibly of size 1, i.e. the leaves remain as they are) to obtain a new planar tanglegram.

Thus every planar tanglegram T can be decomposed uniquely into an irreducible planar tanglegram I and a collection of planar tanglegrams corresponding to the edges of I. We have two cases to consider:

- The irreducible tanglegram I has size greater than 2. Then I is not symmetric, as we have seen in the proof of Proposition 5. In the monomial x^r in $H(x)$, r represents the number of leaves, and replacing a pair of leaves by a planar tanglegram in the irreducible tanglegram translates to replacing x by $T(x)$ in $H(x) - \frac{x^2}{2}$.
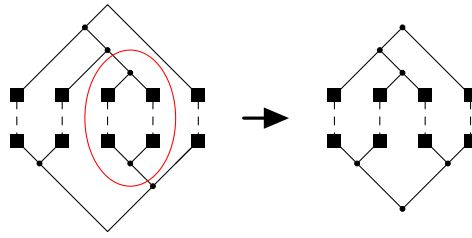
Figure 8 Contracting a proper subtanglegram to obtain an irreducible tanglegram.

▬ The irreducible tanglegram has size 2. We have to replace the two pairs of leaves by two planar tanglegrams; however, in view of the symmetry, the order is irrelevant, so this amounts to taking an unordered pair of tanglegrams. By Pólya's enumeration theorem (see [7, Section 2.4] or [6, Section I.6.1]) the generating function for these unordered pairs is given by $\frac{1}{2}(T(x)^2 + T(x^2))$.

Combining all cases, we get

$$T(x) = x + \left(H(T(x)) - \frac{T(x)^2}{2}\right) + \frac{1}{2}\big(T(x)^2 + T(x^2)\big) = H(T(x)) + x + \frac{T(x^2)}{2}.$$

This completes the proof of (3) and thus of Theorem 1. ◀

3 Asymptotic analysis

In this section, we consider analytic properties of the generating functions in Theorem 1. Since the proofs of the results in this section are all rather technical, they are deferred to the appendix. We will first work with H and deduce the properties of T in Theorem 2 from H. Setting $u(x) = x(1 - A(x))$, Equation (1) can be rewritten in the form

$$x = \sum_{r=0}^{\infty} C_r^2 \cdot u(x)^{r+1}.$$

This motivates the definition of a function ϕ by

$$\phi(u) = \sum_{r=0}^{\infty} C_r^2 \cdot u^{r+1} = \sum_{r=0}^{\infty} \frac{1}{(r+1)^2} \binom{2r}{r}^2 u^{r+1}.$$

The function u becomes the inverse of ϕ. We will obtain the analytic behaviour of the generating functions A and H by studying ϕ. Next, we note that the function ϕ is connected to the complete elliptic integral

$$k(x) = \int_0^{\pi/2} \frac{1}{\sqrt{1 - x \sin^2 t}}\, dt$$

by means of the identity (cf. [4, 19.5.1])

$$\sum_{r=0}^{\infty} \binom{2r}{r}^2 u^r = \frac{2}{\pi} k(16u),$$

which is valid for $|u| < \frac{1}{16}$. We can exploit this connection to obtain the following proposition:

▶ **Proposition 9.** *The function ϕ has an analytic continuation to the slit plane $\mathbb{C} \setminus [\frac{1}{16}, \infty)$. Moreover, when u tends to $\frac{1}{16}$, we have*

$$\phi(u) = \frac{4-\pi}{4\pi} - \frac{1}{4\pi}(1 - 16u) - \frac{1}{64\pi}\Big(5 - 8\log 2 + 2\log(1 - 16u)\Big)(1 - 16u)^2 \qquad (9)$$
$$+ O\Big(\big|(1 - 16u)^3 \log(1 - 16u)\big|\Big).$$

Since we are interested in the inverse of ϕ, we also need the information given in the following lemma:

▶ **Lemma 10.** *The function ϕ is injective in $\mathbb{C} \setminus [\frac{1}{16}, \infty)$, and for all $u \in \mathbb{C} \setminus [\frac{1}{16}, \infty)$ we have $\phi'(u) \neq 0$.*

The branch cut singularity of ϕ at $\frac{1}{16}$ corresponds to a singularity of H at $\phi(\frac{1}{16}) = \frac{4-\pi}{4\pi}$. Inverting the asymptotic expansion of ϕ around the singularity by means of bootstrapping, we derive an asymptotic expansion for the generating function $H(x)$ at its dominant singularity, which also yields an asymptotic formula for the number of irreducible tanglegrams by a typical application of singularity analysis [6, Chapter VI].

▶ **Theorem 11.** *There exist constants $\theta' \in (0, \frac{\pi}{2})$ and $\epsilon' > 0$ such that H is analytic in*

$$\Delta' = \{x \mid |x| < \alpha + \epsilon' \text{ and } |\operatorname{Arg}(x - \alpha)| > \theta'\},$$

and for $x \in \Delta'$, we have

$$H(x) = C_0' + C_1'(\alpha - x) + C_2'(\alpha - x)^2 + B'(\alpha - x)^2 \log(\alpha - x) + O(|(\alpha - x)^3 \log(\alpha - x)|) \quad (10)$$

where $\alpha = \phi(\frac{1}{16}) = \frac{4-\pi}{4\pi}$, $C_0' = \frac{1}{2\pi} - \frac{5}{32}$, $C_1' = \frac{\pi}{8} - \frac{1}{2}$, $C_2' = -\frac{\pi^2}{32}(5 - 4\log 2 + 2\log\pi)$ and $B' = -\frac{\pi^2}{16}$. Thus, the number of irreducible planar tanglegrams is asymptotically given by

$$H_n = [x^n]H(x) \sim \frac{(\pi\alpha)^2}{8} \cdot n^{-3} \cdot \alpha^{-n}.$$

Finally, we move on to the analysis of the generating function $T(x)$ for planar tanglegrams, culminating in the proof of Theorem 2. Details of this proof can be found in the appendix, we focus on the main points. Recall that $T(x)$ satisfies the functional equation

$$T(x) = H(T(x)) + x + \frac{T(x^2)}{2}. \qquad (11)$$

The radius of convergence ρ of $T(x)$ can be bounded above by the radius of convergence of $H(x)$ (since $T_n = [x^n]T(x) \geq [x^n]H(x) = H_n$ for all n in view of the combinatorial interpretation) and is thus less than 1. Pringsheim's Theorem guarantees that ρ is also a singularity. We note that $T(x^2)$ has radius of convergence $\sqrt{\rho} > \rho$, so it is an analytic function in a larger region than $T(x)$ itself.

One also finds that, importantly, $H'(x) \neq 1$ for all x inside the closed disk of convergence of H. This means that the implicit function theorem never fails and the dominant singularity of T is carried over from H: we reach a singularity when $T(x)$ equals the value of H's singularity, which is $\frac{4-\pi}{4\pi}$. This gives us the following characterisation of ρ:

$$T(\rho) = \frac{4-\pi}{4\pi}.$$

For practical computation of ρ, however, it is useful to plug this into (11). Since $H\left(\frac{4-\pi}{4\pi}\right) = \frac{1}{2\pi} - \frac{5}{32}$, we obtain

$$\rho + \frac{T(\rho^2)}{2} = \frac{1}{2\pi} - \frac{3}{32},$$

which can be solved numerically with high accuracy. The singular expansion of $T(x)$ around the singularity ρ can be obtained by means of the same bootstrapping process that is also used to prove Theorem 11. Finally, the asymptotic formula for T_n is another standard application of singularity analysis.

References

1 Sara C. Billey, Matjaž Konvalinka, and Frederick A. Matsen. On the enumeration of tanglegrams and tangled chains. *J. Combin. Theory Ser. A*, 146:239–263, 2017.

2 Kevin Buchin, Maike Buchin, Jaroslaw Byrka, Martin Nöllenburg, Yoshio Okamoto, Rodrigo I. Silveira, and Alexander Wolff. Drawing (complete) binary tanglegrams: hardness, approximation, fixed-parameter tractability. *Algorithmica*, 62(1-2):309–332, 2012.

3 Éva Czabarka, László Székely, and Stephan Wagner. A tanglegram Kuratowski theorem. Preprint, submitted; `arXiv:1708.00309`.

4 NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/, Release 1.0.15 of 2017-06-01. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds. URL: `http://dlmf.nist.gov/`.

5 Henning Fernau, Michael Kaufmann, and Mathias Poths. Comparing trees via crossing minimization. In *Proceedings of the 25th International Conference on Foundations of Software Technology and Theoretical Computer Science*, FSTTCS '05, pages 457–469, Berlin, Heidelberg, 2005. Springer-Verlag.

6 Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University press, 2009.

7 Frank Harary and Edgar M. Palmer. *Graphical Enumeration*. Academic Press, 1973.

8 Danièle Huguet and Dov Tamari. La structure polyédrale des complexes de parenthésages. *J. Combin. Inform. System Sci.*, 3(2):69–81, 1978.

9 Matjaž Konvalinka and Stephan Wagner. The shape of random tanglegrams. *Adv. in Appl. Math.*, 78:76–93, 2016.

10 Frederick A. Matsen IV, Sara C. Billey, Arnold Kas, and Matjaž Konvalinka. Tanglegrams: a reduction tool for mathematical phylogenetics. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 15(1):343–349, 2018.

11 N.J.A. Sloane, editor. The On-line Encyclopedia of Integer sequences. Published electronically at https://oeis.org, 2018.

12 Roderic D. M. Page. *Tangled trees: phylogeny, cospeciation, and coevolution*. University of Chicago Press, 2003.

13 Hassler Whitney. Non-separable and planar graphs. *Trans. Amer. Math. Soc.*, 34(2):339–362, 1932.
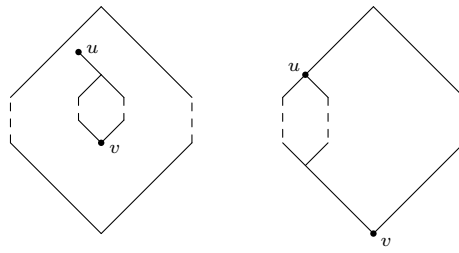
Figure 9 Components of $G \setminus \{u, v\}$ with only 2 edges ending in v

A   Appendix: additional proofs

A.1   Proof of Proposition 7

First, notice that the process of smoothing an irreducible tanglegram does not create any parallel edges since the tanglegram would not be irreducible if that was the case (there would be a proper subtanglegram of size 2). After the process of smoothing, the remaining vertices (except the two roots) are all internal vertices, so they all have degree 3. The two roots are also of degree 3 because of the additional edge joining them. Thus, we have a 3-regular graph, which we will denote by G.

Next, let T_1, T_2 be the two halves of an irreducible tanglegram with more than two leaves on each side. We will prove that removing any pair of vertices u, v of the graph obtained from the smoothing process does not disconnect the graph.

- Suppose u, v are in the same tree, say T_1. Every vertex in T_2 is clearly still connected to T_2's root. Every vertex in T_1 has three connections to the root of T_2 that are disjoint within T_1: via the root of T_1 and via the two children. Removing u, v can only destroy at most two of them, so all vertices of T_1 are also still connected to the root of T_2. This means that $G - \{u, v\}$ is connected.

- Now, suppose that u, v are in different trees. Assume that u is a vertex of T_1, v is a vertex of T_2 and that removing disconnects the graph obtained from the tanglegram by smoothing. $T_1 \setminus u$ has up to three components: two corresponding to the children of u, and one containing the root. Some of these components might be empty. Every non-empty component has at least two edges going to the other half of the tanglegram. Suppose there are only two, and both of them have v as an end. Then we are in one of the following situations:

  Either way, there is a proper subtanglegram. So we can assume that every component of $T_1 \setminus u$ has an edge to $T_2 \setminus v$. The same applies to the components of $T_2 \setminus v$. Now consider the bipartite graph whose vertices are the components of $T_1 \setminus u$ and $T_2 \setminus v$, where we connect two components if there is an edge between them. If this graph is connected, then so is the graph $G \setminus \{u, v\}$. So call this graph G' and suppose it is disconnected. Note that the root components of $T_1 \setminus u$ and $T_2 \setminus v$ (if they exist) are connected in G' by definition (since there is an edge between the roots in G). So there must be a component of G' containing only child components of $T_1 \setminus u$ and $T_2 \setminus v$ respectively. This component must have one of the shapes in Figure 10, each corresponding to a proper subtanglegram, which is impossible. It follows that G' must actually be connected. Therefore, we can conclude that G is 3-connected.
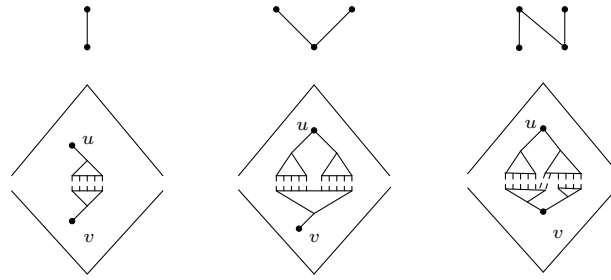
Figure 10 Component of G' containing only child components of $T_1 \setminus u$ and $T_2 \setminus v$.

A.2   Proofs for Section 3

Proof of Proposition 9. First of all, it is well known that

$$\sum_{r=0}^{\infty} \binom{2r}{r}^2 u^r = \frac{2}{\pi} k(16u) \tag{12}$$

for $|u| < \frac{1}{16}$, with the complete elliptic integral

$$k(x) = \int_0^{\pi/2} \frac{1}{\sqrt{1 - x \sin^2 t}} \, dt. \tag{13}$$

The integral $k(x)$ defines an analytic function on the slit plane $\mathbb{C} \setminus [1, \infty)$. Now, [4, 19.12.1] gives the series

$$k(1 - x) = \sum_{m=0}^{\infty} \binom{m - 1/2}{m}^2 x^m \left(-\frac{1}{2} \log x + d(m) \right)$$

after some rewriting (note that [4] uses a different notation, where $K(u) = k(u^2)$ according to our notation), where $d(0) = 2 \log 2$, $d(m) = d(m - 1) - \frac{1}{m(2m-1)}$, or equivalently $d(m) = \psi(1 + m) - \psi(\frac{1}{2} + m)$ (here ψ is the Digamma function). Now since

$$\sum_{m=0}^{\infty} \binom{m - 1/2}{m}^2 x^m = \frac{2}{\pi} k(x),$$

which is in fact equivalent to (12), we can also write this as

$$k(1 - x) = -\frac{1}{\pi} k(x) \log x + \sum_{m=0}^{\infty} \binom{m - 1/2}{m}^2 d(m) x^m,$$

which provides us with an analytic continuation around the branch cut for $|x| < 1$, $x \notin (-1, 0]$. In particular, we have the following asymptotic expansion around $u = \frac{1}{16}$ (by taking the first term in the series):

$$k(16u) = 2 \log 2 - \frac{1}{2} \log(1 - 16u) + O\left(\left| (1 - 16u) \log(1 - 16u) \right| \right).$$

Now

$$\phi(u) = \frac{2}{\pi} \int_0^u \frac{1}{v} \int_0^v k(16z) \, dz \, dv,$$

which provides an analytic continuation of ϕ to the slit plane $\mathbb{C} \setminus [\frac{1}{16}, \infty)$. The asymptotic expansion can be integrated termwise by writing

$$\int_0^v k(16z)\,dz = \int_0^{1/16} k(16z)\,dz - \int_v^{1/16} k(16z)\,dz,$$

cf. [6, Theorem VI.9]. We only need the values

$$\int_0^{1/16} k(16z)\,dz = \frac{1}{8} \quad \text{and} \quad \phi\Big(\frac{1}{16}\Big) = \frac{2}{\pi}\int_0^{1/16} \frac{1}{v}\int_0^v k(16z)\,dz\,dv = \frac{4-\pi}{4\pi},$$

which can be obtained by plugging in (13) and interchanging the order of integration. This gives us first

$$\int_0^v k(16z)\,dz = \frac{1}{8} - \frac{1}{32}\Big(1 + 4\log 2 - \log(1-16v)\Big)(1-16v) + O\Big(\big|(1-16v)^2\log(1-16v)\big|\Big).$$

Then, by multiplication with $\frac{1}{v} = 16 + 16(1-16v) + O(|1-16v|^2)$, we obtain

$$\frac{1}{v}\int_0^v k(16z)\,dz = 2 + \frac{1}{2}\Big(3 - 4\log 2 + \log(1-16v)\Big)(1-16v) + O\Big(\big|(1-16v)^2\log(1-16v)\big|\Big).$$

One more integration step yields (9). ◄

Proof of Lemma 10. We notice that

$$\phi'(u) = \frac{1}{4\pi u}\int_0^{\frac{\pi}{2}} \frac{1 - \sqrt{1 - 16u\sin^2(t)}}{\sin^2(t)}\,dt = \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \frac{1}{1 + \sqrt{1 - 16u\sin^2(t)}}\,dt,$$

which follows from differentiating (13). Now we can make use of the fact that

$$\mathrm{Re}\left(1 + \sqrt{1 - 16u\sin^2(t)}\right) > 0,$$

since $\mathrm{Re}(\sqrt{z}) > 0$ holds for every $z \in \mathbb{C} \setminus (-\infty, 0]$. Thus,

$$\mathrm{Re}\left(\frac{1}{1 + \sqrt{1 - 16u\sin^2(t)}}\right) > 0,$$

which in turn means that $\mathrm{Re}(\phi'(u)) > 0$ for all $u \in \mathbb{C} \setminus [\infty, \frac{1}{16})$. In particular, $\phi'(u) \neq 0$ for all possible values of u. In the same way, we can show that $\mathrm{Im}(\phi'(u))$ has the same sign as $\mathrm{Im}(u)$ for all u, and the two combined imply that ϕ is injective on its domain of analyticity. Indeed, let $u, v \in \mathbb{C}$ such that $u \neq v$. Since ϕ is analytic in $\mathbb{C} \setminus [\frac{1}{16}, \infty)$, we have $\phi(u) = \phi(v) + \int_u^v \phi'(z)dz$, where we can integrate along any path joining u and v in the slit plane. We have several cases to consider, depending on the location of u and v. In each case, after integration one finds that either $\mathrm{Im}(\phi(u)) \neq \mathrm{Im}(\phi(v))$ or $\mathrm{Re}(\phi(u)) \neq \mathrm{Re}(\phi(v))$, which means $\phi(u) \neq \phi(v)$. ◄

Proof of Theorem 11. In order to simplify computations, we write $y = 1 - 16u$. We let u tend to $\frac{1}{16}$ so y tends to 0. Then, by Proposition 9, we have

$$\phi(u) = x = \alpha - \frac{1}{4\pi}y - \frac{1}{32\pi}y^2\log(y) + O(|y^2|). \tag{14}$$

By Lemma 10 and the implicit function theorem, ϕ is invertible and the inverse ϕ^{-1} is analytic in $\phi(\mathbb{C} \setminus [\frac{1}{16}, \infty))$. The function ϕ comes from the integration of k, which has a
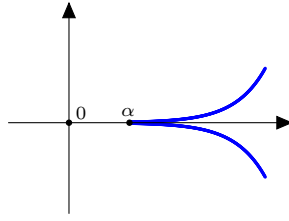
Figure 11 Two branches of ϕ.

branch cut of square root type. The cut $[\frac{1}{16}, \infty)$ is mapped to two branches (see Figure 11), corresponding to the two branches of the square root, and it is easily verified that $\operatorname{Re}\phi(u)$ and $\operatorname{Im}\phi(u)$ are monotone functions of u for both branches. In view of the expansion (14), it is possible to choose θ' in such a way that Δ' lies in the image $\phi(\mathbb{C} \setminus [\frac{1}{16}, \infty))$. Hence, $u = \phi^{-1}$ is well defined and analytic in Δ'.

By means of bootstrapping, we get

$$y = 4\pi(\alpha - x) - 2\pi^2(\alpha - x)^2 \log(\alpha - x) - \pi^2(5 - 4\log 2 + 2\log \pi)(\alpha - x)^2$$
$$+ O\left(\left|(\alpha - x)^3 \log(\alpha - x)\right|\right).$$

Now, since $y = 1 - 16u$, we have

$$H(x) = \frac{x - u}{2} = \frac{x}{2} - \frac{1}{32} + \frac{\pi}{8}(\alpha - x) - \frac{\pi^2}{16}(\alpha - x)^2 \log(\alpha - x)$$
$$- \frac{\pi^2}{32}(K + 2\log(4\pi))(\alpha - x)^2 + O\left(\left|(\alpha - x)^3 \log(\alpha - x)\right|\right)$$
$$= C_0' + C_1'(\alpha - x) + C_2'(\alpha - x)^2 + B'(\alpha - x)^2 \log(\alpha - x)$$
$$+ O\left(\left|(\alpha - x)^3 \log(\alpha - x)\right|\right).$$

When $n > 2$, the coefficient of x^n in $(\alpha - x)^2$ vanishes, so asymptotically only the term $(\alpha - x)^2 \log(\alpha - x)$ contributes to $[x^n]H(x)$. The function H and the region Δ' satisfy the conditions of [6, Theorem VI.4], so we can apply singularity analysis to get

$$[x^n]H(x) \sim -2 \cdot \alpha^2 \cdot B' \cdot \alpha^{-n} \cdot n^{-3} = \frac{(\pi\alpha)^2}{8} \cdot n^{-3} \cdot \alpha^{-n}. \qquad \blacktriangleleft$$

Proof of Theorem 2. For the first property, we investigate each term in the functional equation for $T(x)$, which reads

$$T(x) = H(T(x)) + x + \frac{T(x^2)}{2}.$$

First, the term x on the right side represents an entire function. Next, $\rho \le \alpha < 1$ since the coefficients of T are greater than or equal to the coefficients of H in view of their combinatorial interpretation. Since the radius of convergence of $T(x^2)$ is $\sqrt{\rho}$ it follows that $T(x^2)$ has a radius of convergence greater than $T(x)$. Thus, the dominant singularity of $T(x)$ is inherited from the dominant singularity of $H(x)$. $T(x)$ has non-negative coefficients, so by Pringsheim's theorem, the radius of convergence ρ of $T(x)$ is also a singularity. We know that $H(x)$ has its dominant singularity at $x = \alpha = \frac{4-\pi}{4\pi}$, so $H(T(x))$ has a singularity at any point x for which $T(x) = \alpha = \frac{4-\pi}{4\pi}$. Suppose that there exists a positive real number τ such

that $\tau < \alpha$ and $H(T(x))$ is singular at $T(x_0) = \tau$ for some $x_0 > 0$. We define the bivariate function

$$F(t,x) = H(t) + x + \frac{T(x^2)}{2} - t,$$

and we have

$$\frac{\partial F(t,x)}{\partial t} = H'(t) - 1.$$

Since τ is a singularity of $T(x)$, the implicit function theorem has to fail at $(t,x) = (\tau, x_0)$ for $F(t,x)$. In other words, we must have

$$\frac{\partial F(t,x)}{\partial t}(\tau, x_0) = H'(\tau) - 1 = 0, \text{ i.e. } H'(\tau) = 1.$$

Next, we have

$$H'(x) = \frac{1 - u'(x)}{2}.$$

$H'(x)$ has non-negative coefficients, hence $H'(x)$ is an increasing function in $(0, \alpha]$. Moreover, we have $u'(\alpha) = \frac{\pi}{4}$. Hence, $H'(\alpha) = \frac{1}{2} - \frac{\pi}{8} < 1$. Thus, $H'(\tau) \leq H'(\alpha) < 1$ contradicting the assumption that $H'(\tau) = 1$. We conclude that the dominant singularity of T appears at ρ and $T(\rho) = \alpha$.

Now, we continue with the proof of the second property. Let $x \in B(0, \rho)$. Since T has positive coefficients, we have

$$|T(x)| = \left| \sum_{n=0}^{\infty} T_n x^n \right| \leq \sum_{n=0}^{\infty} T_n |x^n| < \sum_{n=0}^{\infty} T_n \rho^n = \alpha.$$

Hence $T(x) \in B(0, \alpha)$. Moreover, by the implicit function theorem, T can be continued analytically around each point x of the circle $C(0, \rho)$ of center 0 and radius ρ, except perhaps around ρ. However around ρ, T can be continued by Theorem 11. Thus, it is indeed possible to find ϵ and θ such that T is analytic in Δ as required. Let $G(x) = x + \frac{T(x^2)}{2}$. Using the same arguments as in the proof of the first property, $G(x)$ is still analytic around ρ. Hence, the Taylor expansion of $G(x)$ around ρ gives

$$G(x) = G(\rho) + G'(\rho)(x - \rho) + \frac{G''(\rho)}{2}(x - \rho)^2 + O(|x - \rho|^3).$$

For simplicity, we let $D_0 = G(\rho) = \rho + \frac{T(\rho^2)}{2}$, $D_1 = -G'(\rho) = -(1 + \rho T'(\rho^2))$ and $D_2 = \frac{G''(\rho)}{2} = \frac{T'(\rho)}{2} + \rho^2 T''(\rho^2)$. Since $\rho^2 < \rho$, $T'(\rho^2)$ and $T''(\rho^2)$ exist, and the power series for T converges exponentially at ρ^2, which allows for D_0, D_1 and D_2 to be determined with high numerical accuracy. By Theorem 11 and the functional equation for T given in Theorem 1, we have

$$\begin{aligned}
T(x) = {} & C'_0 + C'_1(\alpha - T(x)) + C'_2(\alpha - T(x))^2 \\
& + B'(\alpha - T(x))^2 \log(\alpha - T(x)) + O(|(\alpha - T(x))^3 \log(\alpha - T(x))|) \\
& + D_0 + D_1(\rho - x) + D_2(\rho - x)^2 + O(|(\rho - x)^3|).
\end{aligned} \tag{15}$$

We note that when $x \to \rho$, we have $T(x) \to T(\rho) = \alpha$, hence $C'_0 + D_0 = \alpha$. Again, by means of bootstrapping, we obtain

$$T(x) = \alpha + \frac{D_1}{1 + C'_1}(\rho - x) + \frac{B' \cdot D_1^2}{(1 + C'_1)^3}(\rho - x)^2 \log(\rho - x) + C_2(\rho - x)^2 + O(|(\rho - x)^3 \log(\rho - x)|),$$

which gives us the asymptotic expansion (4) for T by setting $C_1 = \frac{D_1}{1+C_1}$ and $B = \frac{B' \cdot D_1^2}{(1+C_1')^3}$.

We remain with the proof of the asymptotic formula for T_n. As in the proof of Theorem 11, only $(\rho - x)^2 \log(\rho - x)$ contributes to the main term of $[x^n]H(x)$ when n is large. So by singularity analysis, we have

$$T_n = [x^n]T(x) \sim C \cdot \rho^{-n} \cdot n^{-3},$$

where $C = \frac{-2 \cdot B' \cdot D_1^2 \cdot \rho^2}{(1+C_1')^3}$. Here, $C > 0$ because $B' = -\frac{\pi^2}{16} < 0$, and $C_1' = \frac{\pi}{8} - \frac{1}{2} > -1$. ◀

Asymptotic Normality of Almost Local Functionals in Conditioned Galton-Watson Trees

Dimbinaina Ralaivaosaona[1]

Department of Mathematical Sciences, Stellenbosch University, Stellenbosch 7602, South Africa
naina@sun.ac.za

Matas Šileikis[2]

The Czech Academy of Sciences, Institute of Computer Science, Pod Vodárenskou věží 2, 182
07 Prague, Czech Republic. With institutional support RVO:67985807.
matas.sileikis@gmail.com

Stephan Wagner[3]

Department of Mathematical Sciences, Stellenbosch University, Stellenbosch 7602, South Africa
swagner@sun.ac.za

⎯ Abstract ⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯

An additive functional of a rooted tree is a functional that can be calculated recursively as the
sum of the values of the functional over the branches, plus a certain toll function. Janson recently
proved a central limit theorem for additive functionals of conditioned Galton-Watson trees under
the assumption that the toll function is local, i.e. only depends on a fixed neighbourhood of
the root. We extend his result to functionals that are almost local, thus covering a wider range
of functionals. Our main result is illustrated by two explicit examples: the (logarithm of) the
number of matchings, and a functional stemming from a tree reduction process that was studied
by Hackl, Heuberger, Kropf, and Prodinger.

1 Introduction

A functional F that associates a value $F(T)$ with every rooted tree is said to be *additive* if it
satisfies a recursion of the form

$$F(T) = \sum_{i=1}^{k} F(T_i) + f(T), \tag{1}$$

where $T_1, T_2, \ldots, T_k$ are the branches of T and f is a so-called "toll function", another
function that assigns a value to every rooted tree. If T only consists of the root (so that
$k = 0$), we interpret the empty sum as 0 and set $F(T) = f(T)$. Of course, every functional F
is additive in this sense (for a suitable choice of f), so the usefulness of the concept depends
on what is known about the toll function f.

An important special case is the number of occurrences of a prescribed "fringe subtree". A fringe subtree is an induced subtree of a rooted tree that consists of one of the nodes and all its descendants. Now fix a rooted tree S. We say that S occurs on the fringe of T if there is a fringe subtree of T that is isomorphic to S. The number of occurrences of S as a fringe subtree in T (i.e., the number of nodes v of T for which the fringe subtree rooted at v is isomorphic to S) is an additive functional, which we shall denote by $F_S(T)$. Indeed, one has

$$F_S(T) = \sum_{i=1}^{k} F(T_i) + f_S(T),$$

where

$$f_S(T) = \begin{cases} 1 & S \text{ is isomorphic to } T, \\ 0 & \text{otherwise.} \end{cases}$$

This is because an occurrence of S in T is either an occurrence in one of the branches, or comprises the entire tree T. Every additive functional can be expressed as a linear combination of these elementary functionals: it is easy to see (e.g. by induction) that a functional satisfying (1) can be expressed as

$$F(T) = \sum_{S} f(S) F_S(T).$$

Functionals of the form F_S are known to be asymptotically normally distributed in different classes of trees, notably simply generated trees/Galton-Watson trees [6, 14], which will also be the topic of this paper, and classes of increasing trees [4, 10]. In view of this and several other important examples of additive functionals that satisfy a central limit theorem, general schemes have been devised that yield a central limit theorem under different technical assumptions. This includes work on simply generated trees/Galton-Watson trees [6, 14] (labelled trees, plane trees and d-ary trees are well-known special cases) as well as Pólya trees [14] and increasing trees [10, 14] (specifically recursive trees, d-ary increasing trees and generalised plane-oriented recursive trees). It is worth mentioning, however, that there are also many instances of additive functionals that are not normally distributed in the limit, since the toll functions can be quite arbitrary. A well-known example is the case of the path length, i.e. the sum of the distances of all nodes to the root. It satisfies (1) with toll function

$$f(T) = |T| - 1,$$

and, when suitably normalised, its limiting distribution for simply generated trees is the Airy distribution (see [11]).

Previous results [4, 6, 10, 14], while giving rather general conditions on the toll function that imply normality, are unfortunately still insufficient to cover all possible examples one might be interested in. This paper is essentially an extension of Janson's work [6] on local functionals. By weakening the conditions he makes on the toll functions, we arrive at a new general central limit theorem that can be applied to a variety of examples that were not previously covered. Two such examples are presented in detail in this extended abstract: one is concerned with the number of matchings of a tree, the other settles an open problem from a paper of Hackl, Heuberger, Kropf and Prodinger [3] on tree reductions.

A *local* functional (as considered in Janson's paper [6]) is a functional for which the value of the toll function can be determined from the knowledge of a fixed neighbourhood of the root. A typical example is the number of nodes with a given outdegree: the corresponding

toll function (whose value is either 0 or 1) is completely determined by the root degree. We relax this condition somewhat (to what we call "almost local functionals") in our main theorem. Intuitively speaking, functionals that satisfy our conditions have toll functions that can be approximated well from knowledge of a neighbourhood of the root, with the approximation getting better the wider the neighbourhood is chosen.

The model of random trees that we consider here are *conditioned Galton-Watson trees*: these are determined by an offspring distribution ξ, which we will assume to be normalised to satisfy $\mathbb{E}\xi = 1$. We also assume that $\mathrm{Var}\xi$ is finite and nonzero (to avoid a degenerate case). The Galton-Watson process starts from a single node, the root. At time t, all nodes at level t (distance t from the root) generate a number of children according to the offspring distribution ξ. The numbers of children of different nodes on the same level are mutually independent. The outcome of this process, which ends when all nodes at level t generate 0 children, is a random tree $\mathcal{T}$ (almost surely finite). By conditioning the process to "die out" when the total number of nodes is n, we obtain a conditioned Galton-Watson tree, which will be denoted by $\mathcal{T}_n$.

Conditioned Galton-Watson trees are known to be essentially equivalent to so-called *simply generated trees* [2, Section 3.1.4]. Classical examples include rooted labelled trees (corresponding to a Poisson distribution for ξ), plane trees (corresponding to a geometric distribution for ξ) and binary trees (with a distribution whose support is $\{0, 2\}$).

We conclude the introduction with some more notation: for a tree T, we let $T^{(M)}$ be its restriction to the first M levels, i.e. all nodes whose distance to the root is at most M. A local functional as defined above is thus a functional for which the value of $f(T)$ is determined by $T^{(M)}$ for some fixed M (the "cut-off"). The conditioned Galton-Watson tree $\mathcal{T}_n$ is known to converge in the local topology induced by these restrictions to the (infinite) *size-biased* Galton-Watson tree $\hat{\mathcal{T}}$ as defined by Kesten [8]: one has

$$\mathbb{P}(\hat{\mathcal{T}}^{(M)} = T) = w_M(T)\mathbb{P}(\mathcal{T}^{(M)} = T)$$

for all trees T, where $w_M(T)$ is the number of nodes of depth M in T.

For a rooted tree T (possibly infinite), we let $\deg(T)$ denote the degree of the root of T. Finally, it will be convenient for us to use the Vinogradov notation $\ll$ interchangeably with the O-notation, i.e. $f(n) \ll g(n)$ and $f(n) = O(g(n))$ both mean that $|f(n)| \leq Kg(n)$ for a fixed positive constant K and all sufficiently large n.

2 The general theorem

Let us now formulate our main result, which is a central limit theorem for additive functionals under suitable technical conditions on the toll function f.

▶ **Theorem 1.** *Let $\mathcal{T}_n$ be a conditioned Galton-Watson tree of order n with offspring distribution ξ, where ξ satisfies $\mathbb{E}\xi = 1$ and $0 < \sigma^2 := \mathrm{Var}\xi < \infty$. Assume further that $\mathbb{E}\xi^{2\alpha+1} < \infty$ for some integer $\alpha \geq 0$. Consider a functional f of finite rooted ordered trees with the property that there is an absolute constant $C_0 > 0$ such that*

$$|f(T)| \leq C_0 \deg(T)^\alpha. \tag{2}$$

Furthermore, let $(p_M)_{M \geq 1}$ be a sequence of positive real numbers with $p_M \to 0$, and assume that f satisfies the following:

- *for every $M \in \{1, 2, \dots\}$,*

$$\mathbb{E}\left|f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)}) \,\big|\, \hat{\mathcal{T}}^{(M)}\right)\right| \leq p_M \tag{3}$$

  uniformly in N, with $N \geq M$,

■ *there is a sequence of positive integers $(M_n)_{n \geq 1}$ such that for large enough n,*

$$\mathbb{E}\left| f(\mathcal{T}_n) - f\left(\mathcal{T}_n^{(M_n)}\right) \right| \leq p_{M_n}. \tag{4}$$

If $a_n := n^{-1/2}(n^{\max\{\alpha,1\}} p_{M_n} + M_n^2)$ satisfies

$$a_n \to 0, \quad and \quad \sum_{n=1}^{\infty} \frac{a_n}{n} < \infty, \tag{5}$$

then

$$\frac{F(\mathcal{T}_n) - n\mu}{\sqrt{n}} \xrightarrow{\mathrm{d}} \mathcal{N}(0, \gamma^2) \tag{6}$$

where $\mu = \mathbb{E}f(\mathcal{T})$, and $0 \leq \gamma < \infty$.

▶ Remark. The proof of the above theorem is a generalisation of Janson's proof of his theorem for bounded and local functionals in [6]. By slightly weakening the condition on the offspring distribution ξ, we are able to reduce the boundedness condition to (2). However, the main difficulty to overcome is the fact that our toll function is no longer local. To give a simple example, an essential part of the proof is the existence of the expectation $\mathbb{E}f(\hat{\mathcal{T}})$. When f is local with a cut-off M, then $f(\hat{\mathcal{T}}) := f(\hat{\mathcal{T}}^{(M)})$. So, $\mathbb{E}f(\hat{\mathcal{T}})$ is simply defined to be $\mathbb{E}f(\hat{\mathcal{T}}^{(M)})$. In our case, where f is not necessarily local, we can define

$$\mathbb{E}f(\hat{\mathcal{T}}) := \lim_{M \to \infty} \mathbb{E}f(\hat{\mathcal{T}}^{(M)}), \tag{7}$$

which may not exist in general. However, if f satisfies (3), then we can show that $\mathbb{E}f(\hat{\mathcal{T}})$ exists. Indeed,

$$|\mathbb{E}f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}f(\hat{\mathcal{T}}^{(N)})| = \left| \mathbb{E}\left(f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)}) \,|\, \hat{\mathcal{T}}^{(M)} \right) \right) \right|$$

$$\leq \mathbb{E}\left| f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)}) \,|\, \hat{\mathcal{T}}^{(M)} \right) \right| \leq p_M,$$

which tends to zero as $M \to \infty$, uniformly for $N \geq M$. In other words, $(\mathbb{E}f(\hat{\mathcal{T}}^{(M)}))_{M \geq 1}$ is a Cauchy sequence, so the limit (7) exists.

Throughout the rest of the paper, the offspring distribution ξ is assumed to satisfy $\mathbb{E}\xi = 1$, $\mathbb{P}(\xi = 0) > 0$, $0 < \sigma^2 := \mathrm{Var}\,\xi < \infty$, and $\mathbb{E}\xi^{2\alpha+1} < \infty$ for some fixed integer $\alpha \geq 0$. The distribution of the number of nodes at level k, w_k, for the three random trees $\mathcal{T}$, $\hat{\mathcal{T}}$, and $\mathcal{T}_n$ will play an important role in our proof. This parameter has been studied in [5], and in particular, the following results were proved there: for every positive integer $r \leq \max\{2\alpha, 1\}$, we have

$$\mathbb{E}\left(w_k(\mathcal{T})^r \right) = O(k^{r-1}), \ \ \mathbb{E}(w_k(\hat{\mathcal{T}})^r) = O(k^r), \ \ and \ \ \mathbb{E}\left(w_k(\mathcal{T}_n)^r \right) = O(k^r), \tag{8}$$

where the constants in the O-terms depend on the offspring distribution ξ only. Moreover, for a rooted tree T, we know that $|T^{(M)}| = \sum_{k=0}^{M} w_k(T)$. Hence, we can deduce from the estimates in (8), for $r = 1$, that

$$\mathbb{E}|\mathcal{T}^{(M)}| = O(M), \ \mathbb{E}|\hat{\mathcal{T}}^{(M)}| = O(M^2), \ \ and \ \ \mathbb{E}|\mathcal{T}_n^{(M)}| = O(M^2). \tag{9}$$

In fact, it can be shown that $\mathbb{E}|\mathcal{T}^{(M)}| = M + 1$. We are also going to make extensive use of the higher moments of the root degree. By definition, the distribution of $\deg(\mathcal{T})$ is ξ, so we

know the higher moments of $\deg(\mathcal{T})$. On the other hand, note that $\deg(T) = w_1(T)$. So, as particular cases of the estimates in (8), we have

$$\mathbb{E}(\deg(\hat{\mathcal{T}})^r) < \infty \text{ and } \mathbb{E}\left(\deg(\mathcal{T}_n)^r\right) = O(1), \tag{10}$$

for every positive integer $r \leq \max\{2\alpha, 1\}$, where the implied constant in the second estimate is independent of n.

3 Mean and variance

We first look at the expectation $\mathbb{E}f(\mathcal{T}_n)$. As it is also the case in [6], one of the key observations in the proof of Theorem 1 is the fact that $\mathbb{E}f(\mathcal{T}_n)$ is asymptotically equal to $\mathbb{E}f(\hat{\mathcal{T}})$ (which is finite, cf. Remark 2) with an explicit bound on the error term. This is made precise in the following lemma:

▶ **Lemma 2.** *If f satisfies the conditions of Theorem 1, then*

$$\mathbb{E}f(\mathcal{T}_n) = \mathbb{E}f(\hat{\mathcal{T}}) + O(p_{M_n} + n^{-1/2} M_n^2). \tag{11}$$

Proof (sketch). We let M_n be defined as in Theorem 1, but write $M = M_n$ for easy reading. Notice first that

$$
\begin{aligned}
|\mathbb{E}f(\mathcal{T}_n) - \mathbb{E}f(\hat{\mathcal{T}})| \\
\leq |\mathbb{E}f(\mathcal{T}_n) - \mathbb{E}f(\mathcal{T}_n^{(M)})| + |\mathbb{E}f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}f(\hat{\mathcal{T}})| + |\mathbb{E}f(\mathcal{T}_n^{(M)}) - \mathbb{E}f(\hat{\mathcal{T}}^{(M)})|. \quad (12)
\end{aligned}
$$

The first term on the right side is at most p_M by assumption (4). The second term is also bounded above by p_M in view of (3), using the same argument as in Remark 2: we have

$$
\begin{aligned}
|\mathbb{E}f(\hat{\mathcal{T}}^{(N)}) - \mathbb{E}f(\hat{\mathcal{T}}^{(M)})| &= \left|\mathbb{E}\left(f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)})|\hat{\mathcal{T}}^{(M)}\right)\right)\right| \\
&\leq \mathbb{E}\left|f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)})|\hat{\mathcal{T}}^{(M)}\right)\right| \leq p_M,
\end{aligned}
$$

uniformly for $N \geq M$. Therefore,

$$|\mathbb{E}f(\hat{\mathcal{T}}) - \mathbb{E}f(\hat{\mathcal{T}}^{(M)})| = \lim_{N \to \infty} |\mathbb{E}f(\hat{\mathcal{T}}^{(N)}) - \mathbb{E}f(\hat{\mathcal{T}}^{(M)})| \leq p_M.$$

The estimate of the term $|\mathbb{E}f(\mathcal{T}_n^{(M)}) - \mathbb{E}f(\hat{\mathcal{T}}^{(M)})|$ is rather technical and therefore given in the appendix. It can be shown, using the bound (2), that

$$|\mathbb{E}f(\mathcal{T}_n^{(M)}) - \mathbb{E}f(\hat{\mathcal{T}}^{(M)})| = O\left(n^{-1/2}M^2\mathbb{E}(\deg(\hat{\mathcal{T}})^{\alpha+1}) + n^{-1}M^2\mathbb{E}(\deg(\mathcal{T}_n)^{\alpha+1})\right). \tag{13}$$

In view of (10), the moment $\mathbb{E}(\deg(\hat{\mathcal{T}})^{\alpha+1})$ is finite and $\mathbb{E}(\deg(\mathcal{T}_n)^{\alpha+1})$ is $O(1)$. Therefore, we conclude that

$$|\mathbb{E}f(\mathcal{T}_n) - \mathbb{E}f(\hat{\mathcal{T}})| \ll p_M + n^{-1/2}M^2 = p_{M_n} + n^{-1/2}M_n^2,$$

which is equivalent to the statement in the lemma. ◀

Lemma 2 is already enough to show that $\mathbb{E}F(\mathcal{T}_n) = \mu n + o(\sqrt{n})$, where $\mu = \mathbb{E}f(\mathcal{T})$, by simply applying Part (i) of [6, Theorem 1.5] to the shifted toll function $f(T) - \mathbb{E}f(\hat{\mathcal{T}})$. Next, we estimate the variance of $F(\mathcal{T}_n)$.

▶ **Lemma 3.** *Assume that f satisfies the conditions of Theorem 1. Moreover, set $a_k = k^{-1/2}(k^{\max\{\alpha,1\}} p_{M_k} + M_k^2)$ (as in Theorem 1) and $\mu_k = \mathbb{E}f(\mathcal{T}_k)$. Moreover, set $N = \min\{|T| : f(T) \neq 0\}$. Then we have*

$$n^{-1/2}\mathrm{Var}\left(F(\mathcal{T}_n)\right)^{1/2} \ll \left(\sup_{k \geq N} a_k + \sum_{k=N}^{\infty} \frac{a_k}{k}\right)^{1/2} + \sup_{k \geq N} |\mu_k| + \sum_{k=N}^{\infty} \frac{|\mu_k|}{k}. \tag{14}$$

Proof (sketch). We follow the proof of [6, Theorem 6.12]. We start with a decomposition $f(T) = f'(T) + f''(T)$, where $f'(T) = f(T) - \mu_{|T|}$ and $f''(T) = \mu_{|T|}$. In view of Minkowski's inequality $\mathrm{Var}(X+Y)^{1/2} \leq \mathrm{Var}(X)^{1/2} + \mathrm{Var}(Y)^{1/2}$, it suffices to check that (14) holds for the following cases:

(i) if $f(T) = \mu_{|T|}$, that is, f depends on $|T|$ only,

(ii) when $\mathbb{E}f(\mathcal{T}_k) = 0$ for every k.

Case (i) works precisely as in [6, Theorem 6.7] and gives a bound

$$\mathrm{Var}\left(F(\mathcal{T}_n)\right)^{1/2} \ll n^{1/2}\left(\sup_{k \geq N} |\mu_k| + \sum_{k=N}^{\infty} \frac{|\mu_k|}{k}\right). \tag{15}$$

The contribution from $k < N$ is zero, since $\mu_k = 0$ for $k < N$. So we only consider Case (ii), where $\mathbb{E}f(\mathcal{T}_k) = 0$ for every k. By [6, (6.28)], we have

$$\frac{1}{n}\mathrm{Var}\left(F(\mathcal{T}_n)\right) \leq 2\sum_{k=N}^{n} \frac{\mathbb{P}\left(S_{n-k} = n-k\right)}{\mathbb{P}\left(S_n = n-1\right)}\pi_k\mathbb{E}(f(\mathcal{T}_k)F(\mathcal{T}_k)), \tag{16}$$

where $\pi_k = \mathbb{P}\left(|\mathcal{T}| = k\right)$, and S_k is the sum of k independent copies of ξ. From [6, Lemma 5.2], we know that

$$\frac{\mathbb{P}\left(S_{n-k} = n-k\right)}{\mathbb{P}\left(S_n = n-1\right)} \ll \frac{n^{1/2}}{(n-k+1)^{1/2}},$$

uniformly for $1 \leq k \leq n$. Recalling that $\pi_k = O(k^{-3/2})$, which can also be found in [6], we obtain

$$\frac{1}{n}\mathrm{Var}\left(F(\mathcal{T}_n)\right) \ll \sum_{k=N}^{n} \frac{n^{1/2}}{(n-k+1)^{1/2}\,k^{3/2}}\mathbb{E}(f(\mathcal{T}_k)F(\mathcal{T}_k)). \tag{17}$$

So it remains to estimate $\mathbb{E}(f(\mathcal{T}_k)F(\mathcal{T}_k))$. It can be shown (see appendix) that

$$\mathbb{E}(f(\mathcal{T}_k)F(\mathcal{T}_k)) \ll k^{\max\{\alpha,1\}} p_{M_k} + \mathbb{E}(\deg(\mathcal{T}_k)^{2\alpha}) + M_k^2\,\mathbb{E}(\deg(\mathcal{T}_k)^{\alpha+1}). \tag{18}$$

Once again, by means of the second estimate in (10), $\mathbb{E}(\deg(\mathcal{T}_k)^{2\alpha})$ and $\mathbb{E}(\deg(\mathcal{T}_k)^{\alpha+1})$ are both bounded above by constants. Thus, we have

$$\mathbb{E}(f(\mathcal{T}_k)F(\mathcal{T}_k)) \ll k^{\max\{\alpha,1\}} p_{M_k} + M_k^2 = k^{1/2}a_k, \tag{19}$$

where a_k is defined as in Theorem 1. Applying (19) to (17), we get

$$\frac{1}{n}\mathrm{Var}F(\mathcal{T}_n) \ll \sum_{k=N}^{n} \frac{n^{1/2}\,a_k}{(n-k+1)^{1/2}\,k} \ll \sum_{k=N}^{n/2} \frac{a_k}{k} + \sup_{k \geq n/2} a_k \sum_{n/2 \leq k \leq n} \frac{1}{(n-k+1)^{1/2}\,n^{1/2}} \tag{20}$$

Noting that the last sum on the right side is $O(1)$, the result follows by applying Minkowski's inequality to combine the results from the two cases. ◀

4 Central limit theorem

We use a truncation argument as in the proof of [6, Theorem 1.5]. This is formulated in the following lemma:

▶ **Lemma 4.** *Let $(X_n)_{n \geq 1}$ and $(W_{N,n})_{N,n \geq 1}$ be sequences of centred random variables. If we have*

- $W_{N,n} \xrightarrow{d}_n W_N$, and $W_N \xrightarrow{d}_N W$, *for some random variables* $W, W_1, W_2, \ldots$
- $\mathrm{Var}(X_n - W_{N,n}) = O(\sigma_N^2)$ *uniformly in* n, *and* $\sigma_N^2 \to_N 0$,

then $X_n \xrightarrow{d}_n W$.

This lemma is a simple consequence of [7, Theorem 4.28] or [1, Theorem 4.2].

Proof of Theorem 1. We may assume, without loss of generality, that $\mathbb{E}f(\hat{\mathcal{T}}) = 0$, by subtracting $\mathbb{E}f(\hat{\mathcal{T}})$ from f if it is not zero, because shifting f by a constant will only add a deterministic term in $F(\mathcal{T}_n)$. For each k, let μ_k denote the expectation $\mathbb{E}f(\mathcal{T}_k)$. By Lemma 2, we have

$$|\mu_k| = |\mathbb{E}f(\mathcal{T}_k)| \ll p_{M_k} + k^{-1/2}M_k^2 \leq a_k. \tag{21}$$

For a positive integer N, let $f^{(N)}$ be the truncated functional defined by $f^{(N)}(T) = f(T)\mathbb{I}_{\{|T| < N\}}$ and $F^{(N)}$ be the additive functional associated to the toll function $f^{(N)}$. It is important to notice that $f^{(N)}$ is local, for any fixed N. So, if f satisfies the conditions of Theorem 1, then $f^{(N)}$ also satisfies the conditions of Theorem 1. Note further that $\mathbb{E}f^{(N)}(\mathcal{T}_k) = \mu_k$ if $k < N$, and zero otherwise. Hence, we have $|\mathbb{E}f^{(N)}(\mathcal{T}_k)| \leq |\mu_k|$ for every positive integer N. Let

$$W_{N,n} := \frac{F^{(N)}(\mathcal{T}_n) - \mathbb{E}F^{(N)}(\mathcal{T}_n)}{\sqrt{n}}, \text{ and } X_n := \frac{F(\mathcal{T}_n) - \mathbb{E}F(\mathcal{T}_n)}{\sqrt{n}}.$$

Since $f^{(N)}$ has finite support, by [6, Theorem 1.5], we have

$$W_{N,n} \xrightarrow{d}_n \mathcal{N}(0, \gamma_N^2),$$

where

$$\gamma_N^2 = \lim_{n \to \infty} n^{-1}\mathrm{Var}(F^{(N)}(\mathcal{T}_n))$$

$$= 2\mathbb{E}\left(f^{(N)}(\mathcal{T})\left(F^{(N)}(\mathcal{T}) - |\mathcal{T}|\mu^{(N)}\right)\right) - \mathrm{Var}f^{(N)}(\mathcal{T}) - \frac{(\mu^{(N)})^2}{\sigma^2},$$

and $\mu^{(N)} = \mathbb{E}f^{(N)}(\mathcal{T})$.

Next we need to show that $\lim_{N \to \infty} \gamma_N$ exists. To that end, we take an arbitrary integer $M \geq N$. We have

$$\gamma_M - \gamma_N = \lim_{n \to \infty} n^{-1/2}\left(\mathrm{Var}(F^{(M)}(\mathcal{T}_n))^{1/2} - \mathrm{Var}(F^{(N)}(\mathcal{T}_n))^{1/2}\right)$$

If we apply Minkowski's inequality to the random variables $F^{(M)}(\mathcal{T}_n) - F^{(N)}(\mathcal{T}_n)$ and $F^{(N)}(\mathcal{T}_n)$, we obtain

$$\mathrm{Var}(F^{(M)}(\mathcal{T}_n))^{1/2} \leq \mathrm{Var}\left(F^{(M)}(\mathcal{T}_n) - F^{(N)}(\mathcal{T}_n)\right)^{1/2} + \mathrm{Var}(F^{(N)}(\mathcal{T}_n))^{1/2}.$$

Consequently,

$$|\gamma_M - \gamma_N| = \lim_{n \to \infty} n^{-1/2} |\text{Var}(F^{(M)}(\mathcal{T}_n))^{1/2} - \text{Var}(F^{(N)}(\mathcal{T}_n))^{1/2}|$$

$$\leq \limsup_{n \to \infty} n^{-1/2} \text{Var}\left(F^{(M)}(\mathcal{T}_n) - F^{(N)}(\mathcal{T}_n)\right)^{1/2}.$$

The toll function associated to the functional $F^{(M)} - F^{(N)}$ is $f^{(M)} - f^{(N)}$, which is zero for all trees of order smaller than N. Hence, the idea of Lemma 3 can be used to estimate the variance $\text{Var}(F^{(M)}(\mathcal{T}_n) - F^{(N)}(\mathcal{T}_n))^{1/2}$, and we obtain

$$|\gamma_M - \gamma_N| \ll \left(\sup_{k \geq N} a_k + \sum_{k=N}^{\infty} \frac{a_k}{k}\right)^{1/2} + \sup_{k \geq N} |\mu_k| + \sum_{k=N}^{\infty} \frac{|\mu_k|}{k}$$

$$\ll \left(\sup_{k \geq N} a_k + \sum_{k=N}^{\infty} \frac{a_k}{k}\right)^{1/2} + \sup_{k \geq N} a_k + \sum_{k=N}^{\infty} \frac{a_k}{k}.$$

The last line follows from (21). By the condition (5) of Theorem 1, we also deduce that $|\gamma_M - \gamma_N| \to_N 0$ uniformly for $M \geq N$. Hence, the sequence $(\gamma_N)_N$ is a Cauchy sequence, which implies that $\gamma := \lim_{N \to \infty} \gamma_N$ exists.

Similarly, we have

$$\text{Var}(X_n - W_{N,n})^{1/2} = n^{-1/2} \text{Var}(F(\mathcal{T}_n) - F^{(N)}(\mathcal{T}_n))^{1/2}$$

$$\ll \left(\sup_{k \geq N} a_k + \sum_{k=N}^{\infty} \frac{a_k}{k}\right)^{1/2} + \sup_{k \geq N} a_k + \sum_{k=N}^{\infty} \frac{a_k}{k},$$

which tends to zero as $N \to \infty$ uniformly in n, so Lemma 4 applies and the proof of Theorem 1 is complete.   ◀

5   Examples

In this section, we give two representative applications of our main theorem (further examples will be provided in the full version). The absolute values of the toll functions in both examples are not bounded by positive constants, but they are both bounded above by the root degree. Hence, we need α to be at least 1, i.e. $\mathbb{E}\xi^3 < \infty$.

5.1   The number of matchings

The number of matchings in random trees has been studied previously, and means and variances have been determined for different classes of trees [9, 12, 13]. However, in order to obtain a limiting distribution, one has to consider the logarithm of this quantity. For a rooted tree T, let $m(T)$ be the total number of matchings of T and $m_0(T)$ be the number of matchings of T that do not cover the root (by this, we mean matchings that do not contain an edge incident to the root). It is easy to see that these parameters can be determined recursively in the following way:

$$m_0(T) = \prod_i m(T_i), \tag{22}$$

$$m(T) = m_0(T) + \sum_i m_0(T_i) \prod_{j \neq i} m(T_j). \tag{23}$$

Defining an additive functional $F(T) := \log m(T)$, we observe from (22) that the associated toll function is

$$f(T) = F(T) - \sum_i F(T_i) = \log m(T) - \sum_i \log m(T_i) = -\log\left(\frac{m_0(T)}{m(T)}\right).$$

It is convenient to define $\rho(T) = \frac{m_0(T)}{m(T)}$, which is the probability that a random matching does not cover the root, when all matchings are equally likely. By (22) and (23), $\rho(T)$ also satisfies a recursion

$$\rho(T) = \frac{1}{1 + \sum_i \rho(T_i)}. \tag{24}$$

It follows immediately that $0 \leq f(T) \leq \log(1 + \deg(T))$. Hence, the condition (2) of Theorem 1 is satisfied by f with $\alpha = 1$. Next, we measure the difference between $f(T)$ and $f(T^{(M)})$. Define the exact bounds on ρ given the first M levels:

$$\rho_{\min}^M(T) := \inf\{\rho(S) : S^{(M)} = T^{(M)}\}, \qquad \rho_{\max}^M(T) := \sup\{\rho(S) : S^{(M)} = T^{(M)}\}.$$

The functions $\rho_{\min}^M(T)$ and $\rho_{\max}^M(T)$, $M = 0, 1, 2, \ldots$ can also be determined recursively from the root branches $T_1, T_2, \ldots$ by observing $\rho_{\min}^0(T) = 0$ and $\rho_{\max}^0(T) = 1$ for any T, and for any $M \geq 1$, we have

$$\rho_{\max}^M(T) = \frac{1}{1 + \sum_i \rho_{\min}^{M-1}(T_i)} \quad \text{and} \quad \rho_{\min}^M(T) = \frac{1}{1 + \sum_i \rho_{\max}^{M-1}(T_i)}. \tag{25}$$

Since $\rho(T), \rho(T^{(M)}) \in [\rho_{\min}^M(T), \rho_{\max}^M(T)]$, we obtain

$$\frac{\rho_{\min}^M(T)}{\rho_{\max}^M(T)} \leq \frac{\rho(T)}{\rho(T^{(M)})} \leq \frac{\rho_{\max}^M(T)}{\rho_{\min}^M(T)}. \tag{26}$$

Writing $\tau^M(T) := \log(\rho_{\max}^M(T)/\rho_{\min}^M(T)) \geq 0$, (26) gives us

$$|f(T) - f(T^{(M)})| \leq \tau^M(T). \tag{27}$$

Using (25), we get

$$\tau^M(T) = -\log\left(\frac{1 + \sum_i \rho_{\min}^{M-1}(T_i)}{1 + \sum_i \rho_{\max}^{M-1}(T_i)}\right) = -\log\left(\frac{1 + \sum_i \rho_{\max}^{M-1}(T_i)\exp(-\tau^{M-1}(T_i))}{1 + \sum_i \rho_{\max}^{M-1}(T_i)}\right).$$

Since the term inside the logarithm on the right side can be regarded as an expectation (of the expression $\exp(-\tau^{M-1}(T_i))$), applying Jensen's inequality to the convex function $-\log x$ yields

$$\tau^M(T) \leq \frac{1}{1 + \sum_i \rho_{\max}^{M-1}(T_i)} \sum_i \rho_{\max}^{M-1}(T_i)\tau^{M-1}(T_i)$$

$$\leq \frac{\max_i \rho_{\max}^{M-1}(T_i)}{1 + \max_i \rho_{\max}^{M-1}(T_i)} \sum_i \tau^{M-1}(T_i) \leq \frac{1}{2}\sum_i \tau^{M-1}(T_i). \tag{28}$$

From (25) it is clear that $\rho_{\max}^1(T) = 1$ and $\rho_{\min}^1(T) = (1 + \deg(T))^{-1}$ for any T. Therefore

$$\tau^1(T) = \log(1 + \deg(T)) \leq \deg(T). \tag{29}$$

Let $v_1, v_2, \ldots, v_{w_{M-1}(T)}$ be the nodes at level $M-1$ of T. By iterating (28) $M-1$ times and applying (29), we arrive at the bound

$$\tau^M(T) \leq 2^{-(M-1)} \sum_{i=1}^{w_{M-1}(T)} \tau^1(T_{v_i}) \leq 2^{-(M-1)} \sum_{i=1}^{w_{M-1}(T)} \deg(T_{v_i}) \leq 2^{-(M-1)} w_M(T). \quad (30)$$

Combining (27) and (30), we obtain

$$|f(T) - f(T^{(M)})| \leq 2^{-M+1} w_M(T). \quad (31)$$

This is essentially enough to show that the remaining conditions of Theorem 1 are satisfied by our toll function. Let us first check (3). Note that for any $N \geq M$, we have

$$\mathbb{E}\left|f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)}) \mid \hat{\mathcal{T}}^{(M)}\right)\right| \leq \mathbb{E}\left(\mathbb{E}\left(|f(\hat{\mathcal{T}}^{(M)}) - f(\hat{\mathcal{T}}^{(N)})| \mid \hat{\mathcal{T}}^{(M)}\right)\right).$$

Using (31), we deduce that for any $N \geq M$,

$$\mathbb{E}\left(|f(\hat{\mathcal{T}}^{(M)}) - f(\hat{\mathcal{T}}^{(N)})| \mid \hat{\mathcal{T}}^{(M)}\right) \leq 2^{-M+1} \mathbb{E}\left(w_M(\hat{\mathcal{T}}^{(N)}) \mid \hat{\mathcal{T}}^{(M)}\right).$$

By taking the expectations, and using $w_M(\hat{\mathcal{T}}^{(N)}) = w_M(\hat{\mathcal{T}})$ as well as the estimate $\mathbb{E} w_M(\hat{\mathcal{T}}) = O(M)$ (see (8)), we get

$$\mathbb{E}\left|f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)}) \mid \hat{\mathcal{T}}^{(M)}\right)\right| \ll M\, 2^{-M}. \quad (32)$$

To check (4) we use (31) and $\mathbb{E} w_M(\mathcal{T}_n) = O(M)$ (see (8)) and get

$$\mathbb{E}|f(\mathcal{T}_n) - f(\mathcal{T}_n^{(M)})| = \mathbb{E}\left(\mathbb{E}\left(|f(\mathcal{T}_n) - f(\mathcal{T}_n^{(M)})| \mid \mathcal{T}_n^{(M)}\right)\right) \leq \mathbb{E}\left(2^{-M+1} w_M(\mathcal{T}_n)\right) \ll M 2^{-M},$$

$$(33)$$

where the implied constant is independent of n. To sum up, (32) and (33) show that the conditions of Theorem 1 are satisfied for the choice $p_M := C_1 M 2^{-M}$ and $M_n := \lfloor C_2 \log n \rfloor$ with sufficiently large positive constants C_1 and C_2.

5.2   Tree reductions

An old leaf is a leaf that is the leftmost child of its parent node, and an old path is a maximal path with the property that its lower endpoint is an old leaf, and its internal nodes are all nodes of outdegree 1 that are leftmost children of their parents. As in [3], consider the process of reducing a tree by cutting off all old paths from the tree at each step. This process is called old path-reduction. For a given positive integer r, and for a tree T, let $X_r(T)$ be the number of nodes in the reduced tree after the first r steps of the old path-reduction process. The authors of [3] proved estimates for the mean and variance of $X_r(\mathcal{T}_n)$ for the special case where $\mathcal{T}_n$ is the random plane (=ordered) tree on n nodes, but they did not derive a limiting distribution. Theorem 1 can be applied to show asymptotic normality for this case. However, we do not need to restrict ourselves to plane trees.

We let

$$F_r(T) = |T| - X_r(T),$$

which corresponds to the number of deleted nodes after r steps in T. The functional F_r is additive with toll function f_r, where

$$f_r(T) = \sum_j \eta_T(T_j)$$

and the sum is over all branches T_j, with

$$\eta_T(T_j) = \begin{cases} 1 & \text{if the root of } T_j \text{ is deleted within the first } r \text{ steps,} \\ 0 & \text{otherwise.} \end{cases}$$

We can immediately see that

$$0 \le f_r(T) \le \deg(T).$$

Next, we show that f_r is almost local. For a tree T, let T^* be the planted tree where the root of T is connected to a new node, which becomes the root of T^*. Let $\kappa = \min\{k \ge 2 : \mathbb{P}(\xi = k) > 0\}$ (this must exist under our current assumptions on ξ), and let T_0 be the complete κ-ary tree of depth r. It is clear that $F_r(T_0^*) \ne 1$, i.e. T_0^* is not reduced to the root in r steps, and

$$\mathbb{P}(\mathcal{T} = T_0) > 0. \tag{34}$$

For each positive integer M, let $\mathcal{B}_M$ be the set of all trees T (not necessarily finite) of height at least $M - 1$ such that $F_r((T^{(M-1)})^*) = 1$ (i.e. the tree $T^{(M-1)}$ vanishes after the first r steps of the reduction). It is important to notice here that a rooted tree T is not reduced to a single node after the first r steps of the reduction if the fixed tree T_0 appears as a *subtree* of T (by *subtree*, we mean a subtree of the form $T_v^{(k)}$ for some integer $k \ge 0$ and some node v of T). This observation is key in the proof of the next lemma, which can be found in the appendix.

▶ **Lemma 5.** *There is a positive constant $c < 1$, that depends only on ξ and r, such that*

$$\mathbb{P}(\mathcal{T} \in \mathcal{B}_M) \ll c^M \quad \text{and} \quad \mathbb{P}\left(\hat{\mathcal{T}} \in \mathcal{B}_M\right) \ll c^M.$$

For a finite tree T, the only possibility for which $f_r(T^{(M)}) \ne f_r(T)$ is when there is a root branch T_j of T such that $T_j^{(M-1)}$ vanishes after the first r steps of the reduction of $T^{(M)}$, but T_j does not vanish after the first r steps of the reduction of T. This means that if $f_r(T^{(M)}) \ne f_r(T)$, then T must have a branch in $\mathcal{B}_M$. Therefore, we have

$$\mathbb{P}\left(f_r(\mathcal{T}^{(M)}) \ne f_r(\mathcal{T})\right) \le \sum_{k=1}^{\infty} k \mathbb{P}(\xi = k) \mathbb{P}(\mathcal{T} \in \mathcal{B}_M) \ll c^M.$$

The estimate on the right follows from Lemma 5. As an immediate consequence of this, we have

$$\mathbb{P}\left(f_r(\mathcal{T}_n^{(M)}) \ne f_r(\mathcal{T}_n)\right) \le \frac{\mathbb{P}\left(f_r(\mathcal{T}^{(M)}) \ne f_r(\mathcal{T})\right)}{\mathbb{P}(|\mathcal{T}| = n)} \ll n^{3/2} c^M.$$

Hence,

$$\mathbb{E}|f_r(\mathcal{T}_n^{(M)}) - f_r(\mathcal{T}_n)| \ll n^{3/2} c^M \max_{|T|=n} |f_r(T^{(M)}) - f_r(T)| \ll n^{5/2} c^M. \tag{35}$$

Let $\mathcal{E}_M$ be the event $\bigcup_{N>M}\{f_r(\hat{\mathcal{T}}^{(M)}) \ne \mathbb{E}(f_r(\hat{\mathcal{T}}^{(N)}) \,|\, \hat{\mathcal{T}}^{(M)})\}$. Then, for any $N \ge M$, we have

$$\left| f_r(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}(f_r(\hat{\mathcal{T}}^{(N)}) \,|\, \hat{\mathcal{T}}^{(M)}) \right| \ll \deg(\hat{\mathcal{T}}^{(M)}) I_{\mathcal{E}_M}.$$

For $\hat{\mathcal{T}}$ to be in $\mathcal{E}_M$, $\hat{\mathcal{T}}$ must have a root branch in $\mathcal{B}_M$. Therefore,

$$\mathbb{E}\left|f_r(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}(f_r(\hat{\mathcal{T}}^{(N)}) \,|\, \hat{\mathcal{T}}^{(M)})\right|$$
$$\ll \sum_{k=1}^{\infty} k\mathbb{P}\left(\deg(\hat{\mathcal{T}}) = k\right)\left((k-1)\mathbb{P}\left(\mathcal{T} \in \mathcal{B}_M\right) + \mathbb{P}\left(\hat{\mathcal{T}} \in \mathcal{B}_M\right)\right). \quad (36)$$

In view of Lemma 5, we have

$$\mathbb{E}\left|f_r(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}(f_r(\hat{\mathcal{T}}^{(N)}) \,|\, \hat{\mathcal{T}}^{(M)})\right| \ll c^M \sum_{k=1}^{\infty} k^2\mathbb{P}\left(\deg(\hat{\mathcal{T}}) = k\right) \ll c^M, \quad (37)$$

since $\mathbb{E}(\deg(\hat{\mathcal{T}})^2) < \infty$ if $\mathbb{E}\xi^3 < \infty$. The estimates (35) and (37) confirm that f_r is indeed almost local where, for example, $p_M = c_2^M$, for some c_2 with $c < c_2 < 1$, and $M_n = \lfloor(\log n)^2\rfloor$.

▶ **Remark.** We only made very little use of the actual definition of the old path-reduction. To be precise, we only used it when we argued that our constructed T_0 does not vanish after the first r steps of the reduction and that all trees that contain T_0 as a subtree will not be reduced after the first r steps. This means that the same proof will work for any tree reduction with a similar property. This includes all tree reductions considered in [3].

───── **References** ─────

1   Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons, Inc., New York-London-Sydney, 1968.

2   Michael Drmota. *Random trees.* SpringerWienNewYork, Vienna, 2009.

3   Benjamin Hackl, Clemens Heuberger, Sara Kropf, and Helmut Prodinger. Fringe Analysis of Plane Trees Related to Cutting and Pruning. *Aequ. Math.*, 92(2):311–353, 2018.

4   Cecilia Holmgren, Svante Janson, and Matas Šileikis. Multivariate normal limit laws for the numbers of fringe subtrees in m-ary search trees and preferential attachment trees. *Electron. J. Combin.*, 24(2):Paper 2.51, 49, 2017.

5   Svante Janson. Random cutting and records in deterministic and random trees. *Random Structures Algorithms*, 29(2):139–179, 2006.

6   Svante Janson. Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees. *Random Structures Algorithms*, 48(1):57–101, 2016.

7   Olav Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

8   Harry Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(4):425–487, 1986.

9   Martin Klazar. Twelve countings with rooted plane trees. *European J. Combin.*, 18(2):195–210, 1997.

10   Dimbinaina Ralaivaosaona and Stephan Wagner. Additive functionals of d-ary increasing trees. *arXiv 1605.03918*, 2016.

11   Lajos Takács. Conditional limit theorems for branching processes. *J. Appl. Math. Stochastic Anal.*, 4(4):263–292, 1991.

12   Stephan Wagner. Correlation of graph-theoretical indices. *SIAM J. Discrete Math.*, 21(1):33–46, 2007.

13   Stephan Wagner. On the number of matchings of a tree. *European J. Combin.*, 28(4):1322–1330, 2007.

14   Stephan Wagner. Central limit theorems for additive tree parameters with small toll functions. *Combin. Probab. Comput.*, 24(1):329–353, 2015.

A   Appendix

Proof of Estimate (13). From the proof of [6, Lemma 5.9] (see (5.42) there), we have, for T with $|T| \leq n/2$, that

$$\mathbb{P}\left(\mathcal{T}_n^{(M)} = T\right) = \mathbb{P}\left(\hat{\mathcal{T}}^{(M)} = T\right)\left(1 + O\left(\frac{|T|}{n^{1/2}}\right)\right). \tag{38}$$

Using (38), we infer

$$|\mathbb{E}f(\mathcal{T}_n^{(M)}) - \mathbb{E}f(\hat{\mathcal{T}}^{(M)})|$$

$$= \left|\sum_T f(T)\mathbb{P}\left(\mathcal{T}_n^{(M)} = T\right) - \sum_T f(T)\mathbb{P}\left(\hat{\mathcal{T}}^{(M)} = T\right)\right|$$

$$\leq \sum_{|T|\leq n/2} |f(T)| \left|\mathbb{P}\left(\mathcal{T}_n^{(M)} = T\right) - \mathbb{P}\left(\hat{\mathcal{T}}^{(M)} = T\right)\right| +$$

$$\sum_{|T|>n/2} |f(T)| \left(\mathbb{P}\left(\mathcal{T}_n^{(M)} = T\right) + \mathbb{P}\left(\hat{\mathcal{T}}^{(M)} = T\right)\right)$$

$$\ll \sum_T \mathbb{P}\left(\hat{\mathcal{T}}^{(M)} = T\right) \frac{\deg(T)^\alpha |T|}{n^{1/2}}$$

$$+ \sum_{|T|>n/2} \mathbb{P}\left(\hat{\mathcal{T}}^{(M)} = T\right) \deg(T)^\alpha + \sum_{|T|>n/2} \mathbb{P}\left(\mathcal{T}_n^{(M)} = T\right) \deg(T)^\alpha.$$

We can now estimate each of the three terms in the last two lines separately. First, we have

$$\sum_T \mathbb{P}\left(\hat{\mathcal{T}}^{(M)} = T\right) \frac{\deg(T)^\alpha |T|}{n^{1/2}} = n^{-1/2}\mathbb{E}(\deg(\hat{\mathcal{T}}^{(M)})^\alpha |\hat{\mathcal{T}}^{(M)}|)$$

$$= n^{-1/2}\mathbb{E}\left(\deg(\hat{\mathcal{T}}^{(M)})^\alpha \,\mathbb{E}\left(|\hat{\mathcal{T}}^{(M)}| \,\Big|\, \deg(\hat{\mathcal{T}}^{(M)})\right)\right).$$

Conditioning on $\deg(\hat{\mathcal{T}}^{(M)})$ (which is the same as $\deg(\hat{\mathcal{T}})$ for $M \geq 1$), $\hat{\mathcal{T}}$ consists of a root, a copy of $\hat{\mathcal{T}}$ and $\deg(\hat{\mathcal{T}}) - 1$ independent copies of $\mathcal{T}$. Thus, by the estimates in (9), we have

$$\mathbb{E}\left(|\hat{\mathcal{T}}^{(M)}| \,\Big|\, \deg(\hat{\mathcal{T}}^{(M)})\right) = O(M^2 \deg(\hat{\mathcal{T}}^{(M)})).$$

Therefore,

$$\mathbb{E}\left(\deg(\hat{\mathcal{T}}^{(M)})^\alpha \,\mathbb{E}\left(|\hat{\mathcal{T}}^{(M)}| \,\Big|\, \deg(\hat{\mathcal{T}}^{(M)})\right)\right) \ll M^2\mathbb{E}(\deg(\hat{\mathcal{T}}^{(M)})^{\alpha+1}),$$

which yields

$$\sum_T \mathbb{P}\left(\hat{\mathcal{T}}^{(M)} = T\right) \frac{\deg(T)^\alpha |T|}{n^{1/2}} \ll n^{-1/2}M^2\mathbb{E}(\deg(\hat{\mathcal{T}}^{(M)})^{\alpha+1}). \tag{39}$$

For the second term, we have

$$\sum_{|T|>n/2} \mathbb{P}\left(\hat{\mathcal{T}}^{(M)} = T\right) \deg(T)^\alpha$$

$$= \sum_{k\geq 1} k^\alpha \,\mathbb{P}\left(|\hat{\mathcal{T}}^{(M)}| > n/2 \text{ and } \deg(\hat{\mathcal{T}}) = k\right)$$

$$= \sum_{k\geq 1} k^\alpha \,\mathbb{P}\left(\deg(\hat{\mathcal{T}}) = k\right) \mathbb{P}\left(|\hat{\mathcal{T}}^{(M)}| > n/2 \,\Big|\, \deg(\hat{\mathcal{T}}) = k\right).$$

By Markov's inequality and a similar argument as before, we obtain

$$\mathbb{P}\left(|\hat{\mathcal{T}}^{(M)}| > n/2 \,\Big|\, \deg(\hat{\mathcal{T}}) = k\right) \leq \frac{2\mathbb{E}(|\hat{\mathcal{T}}^{(M)}| \,|\, \deg(\hat{\mathcal{T}}) = k)}{n} \ll \frac{kM^2}{n}.$$

Thus,

$$\sum_{|T|>n/2} \mathbb{P}\left(\hat{\mathcal{T}}^{(M)} = T\right) \deg(T)^\alpha \ll n^{-1} M^2 \mathbb{E}(\deg(\hat{\mathcal{T}})^{\alpha+1}). \tag{40}$$

Finally, for the last term, we proceed in a similar fashion:

$$\sum_{|T|>n/2} \mathbb{P}\left(\mathcal{T}_n^{(M)} = T\right) \deg(T)^\alpha$$

$$\leq \sum_{k\geq 1} k^\alpha \, \mathbb{P}\left(|\mathcal{T}_n^{(M)}| > n/2 \text{ and } \deg(\mathcal{T}_n) = k\right)$$

$$\leq \sum_{k\geq 1} k^\alpha \, \mathbb{P}\left(\deg(\mathcal{T}_n) = k\right) \mathbb{P}\left(|\mathcal{T}_n^{(M)}| > n/2 \,\Big|\, \deg(\mathcal{T}_n) = k\right).$$

If $\mathcal{T}_{n,1}, \mathcal{T}_{n,2}, \ldots, \mathcal{T}_{n,k}$ are the branches of $\mathcal{T}_n$, given that $\deg(\mathcal{T}_n) = k$, then, conditioning on their sizes $n_1, n_2, \ldots, n_k$, they are k independent conditioned Galton-Watson trees $\mathcal{T}_{n_1}, \mathcal{T}_{n_2}, \ldots, \mathcal{T}_{n_k}$. On the other hand, we have

$$|\mathcal{T}_n^{(M)}| = 1 + \sum_{i=1}^{k} |\mathcal{T}_{n,i}^{(M-1)}|.$$

Thus,

$$\mathbb{E}\left(|\mathcal{T}_n^{(M)}| \,|\, \deg(\mathcal{T}_n) = k\right) = \mathbb{E}\left(\mathbb{E}\left(|\mathcal{T}_n^{(M)}| \,\Big|\, n_1, n_2, \cdots, n_k\right)\right)$$

$$= 1 + \sum_{i=1}^{k} \mathbb{E}\left(\mathbb{E}\left(|\mathcal{T}_{n_i}^{(M-1)}| \,\Big|\, n_1, n_2, \cdots, n_k\right)\right) \ll kM^2,$$

which again follows from the last estimate in (9). Now, Markov's inequality yields

$$\mathbb{P}\left(|\mathcal{T}_n^{(M)}| > n/2 \,\Big|\, \deg(\mathcal{T}_n) = k\right) \ll n^{-1} kM^2.$$

Therefore, making use of (10) once again, we have

$$\sum_{|T|>n/2} \mathbb{P}\left(\mathcal{T}_n^{(M)} = T\right) \deg(T)^\alpha$$

$$\ll n^{-1} M^2 \sum_{k\geq 1} k^{\alpha+1} \, \mathbb{P}\left(\deg(\mathcal{T}_n) = k\right) = n^{-1} M^2 \mathbb{E}(\deg(\mathcal{T}_n)^{\alpha+1}). \tag{41}$$

Combining the estimates (39), (40), and (41), we finally arrive at the estimate

$$|\mathbb{E}f(\mathcal{T}_n^{(M)}) - \mathbb{E}f(\hat{\mathcal{T}}^{(M)})| \ll n^{-1/2} M^2 \mathbb{E}(\deg(\hat{\mathcal{T}})^{\alpha+1}) + n^{-1} M^2 \mathbb{E}(\deg(\mathcal{T}_n)^{\alpha+1}), \tag{42}$$

which is exactly as we claimed in (13). ◄

Proof of Estimate (18). We decompose $F(\mathcal{T}_k)$ according to the depth $d(v)$ of the nodes:

$$F(\mathcal{T}_k) = \sum_{v\in\mathcal{T}_k} f(\mathcal{T}_{k,v}) = \sum_{d(v)<M} f(\mathcal{T}_{k,v}) + \sum_{d(v)\geq M} f(\mathcal{T}_{k,v}) =: S_1 + S_2. \tag{43}$$

We first observe that

$$\mathbb{E}|f(\mathcal{T}_k)S_1| \ll \mathbb{E}\Big(\deg(\mathcal{T}_k)^\alpha \sum_{d(v)<M} \deg(\mathcal{T}_{k,v})^\alpha\Big)$$

$$= \mathbb{E}\Big(\deg(\mathcal{T}_k)^\alpha \, \mathbb{E}\Big(\sum_{d(v)<M} \deg(\mathcal{T}_{k,v})^\alpha \,\Big|\, \deg(\mathcal{T}_k)\Big)\Big).$$

Next, for any positive integer $m \le M$, we have

$$\mathbb{E}\Big(\sum_{d(v)<m} \deg(\mathcal{T}_{k,v})^\alpha \,\Big|\, \mathcal{T}_k^{(m-1)}\Big) = \sum_{d(v)<m-1} \deg(\mathcal{T}_{k,v})^\alpha + O\left(w_{m-1}(\mathcal{T}_k)\right).$$

This is because the $w_{m-1}(\mathcal{T}_k)$ fringe subtrees with roots at level $m-1$, conditioned on their sizes, are conditioned Galton-Watson trees and thus by (10) the root degrees are $O(1)$. Taking the expectation conditioned on $\deg(\mathcal{T}_k)$, again by the same argument, and by the estimate $\mathbb{E}w_{m-1}(\mathcal{T}_k) = O(m)$ as in (8), we have

$$\mathbb{E}\Big(\sum_{d(v)<m} \deg(\mathcal{T}_{k,v})^\alpha \,\Big|\, \deg(\mathcal{T}_k)\Big) = \mathbb{E}\Big(\sum_{d(v)<m-1} \deg(\mathcal{T}_{k,v})^\alpha \,\Big|\, \deg(\mathcal{T}_k)\Big) + O(m\deg(\mathcal{T}_k)).$$

Thus, iterating from M, we obtain

$$\mathbb{E}\Big(\sum_{d(v)<M} \deg(\mathcal{T}_{k,v})^\alpha \,\Big|\, \deg(\mathcal{T}_k)\Big) \ll \deg(\mathcal{T}_k)^\alpha + M^2 \deg(\mathcal{T}_k).$$

Therefore,

$$\mathbb{E}|f(\mathcal{T}_k)S_1| \ll \mathbb{E}(\deg(\mathcal{T}_k)^{2\alpha}) + M^2\mathbb{E}(\deg(\mathcal{T}_k)^{\alpha+1}). \tag{44}$$

For S_2, we condition on $\mathcal{T}_k^{(M)}$ and the sizes of the fringe subtrees $\mathcal{T}_{k,v_i}$, $i = 1, \ldots, w_M(\mathcal{T}_k)$, induced by nodes at level M. Conditionally, each $\mathcal{T}_{k,v_i}$ is distributed as $\mathcal{T}_{n_i}$. From the assumption that for every n we have $\mathbb{E}f(\mathcal{T}_n) = 0$ it follows (see [6, (6.25)]) that $\mathbb{E}F(T_{k,v_i}) = 0$ and therefore

$$\mathbb{E}\left(S_2 \,|\, \mathcal{T}_k^{(M)}\right) = 0. \tag{45}$$

Let us define $\tilde{f}_M(\mathcal{T}_k) := \mathbb{E}(f(\mathcal{T}_k) \,|\, \mathcal{T}_k^{(M)})$. Note that

$$\mathbb{E}(\tilde{f}_M(\mathcal{T}_k)S_2) = \mathbb{E}\Big(\mathbb{E}\left(\tilde{f}_M(\mathcal{T}_k)S_2 \,|\, \mathcal{T}_k^{(M)}\right)\Big) = \mathbb{E}\left(\tilde{f}_M(\mathcal{T}_k)\mathbb{E}\left(S_2 \,|\, \mathcal{T}_k^{(M)}\right)\right) = 0.$$

Hence,

$$|\mathbb{E}(f(\mathcal{T}_k)S_2)| = |\mathbb{E}(S_2(f(\mathcal{T}_k) - \tilde{f}_M(\mathcal{T}_k)))| \le \max|S_2|\,\mathbb{E}|f(\mathcal{T}_k) - \tilde{f}_M(\mathcal{T}_k)|.$$

It is important to notice here that the expectation in the last term remains unchanged if $f(T)$ is shifted by $\mu_{|T|}$ (this is the reason why we can assume that $\mathbb{E}f(\mathcal{T}_k) = 0$ and f still satisfies the conditions of Theorem 1). By the triangle inequality and the definition of $\tilde{f}_M(\mathcal{T}_k)$, we have

$$|f(\mathcal{T}_k) - \tilde{f}_M(\mathcal{T}_k)| \le |f(\mathcal{T}_k) - f(\mathcal{T}_k^{(M)})| + |f(\mathcal{T}_k^{(M)}) - \tilde{f}_M(\mathcal{T}_k)|$$

$$\le |f(\mathcal{T}_k) - f(\mathcal{T}_k^{(M)})| + \mathbb{E}(|f(\mathcal{T}_k^{(M)}) - f(\mathcal{T}_k)| \,|\, \mathcal{T}_k^{(M)}).$$

Taking the expectation again, and using our condition (4), we obtain

$$\mathbb{E}|f(\mathcal{T}_k) - \tilde{f}_M(\mathcal{T}_k)| \leq 2p_M.$$

Here, we are assuming that $M = M_k$. On the other hand, we have

$$|S_2| \leq \sum_{d(v)\geq M} |f(\mathcal{T}_{k,v})| \leq \sum_{v\in\mathcal{T}_k} \deg(\mathcal{T}_{k,v})^{\alpha}.$$

Since α is a nonnegative integer, the last term is bounded above by $(\sum_{v\in\mathcal{T}_k} \deg(\mathcal{T}_{k,v}))^{\alpha}$ (which is equal to $(k-1)^{\alpha}$) except for $\alpha = 0$. Hence, we get

$$\max |S_2| \leq k^{\max\{\alpha,1\}}.$$

Therefore,

$$\mathbb{E}(f(\mathcal{T}_k)F(\mathcal{T}_k)) \ll k^{\max\{\alpha,1\}}p_M + \mathbb{E}(\deg(\mathcal{T}_k)^{2\alpha}) + M^2\mathbb{E}(\deg(\mathcal{T}_k)^{\alpha+1}),$$

as claimed. ◀

Proof of Lemma 5. We start with the first estimate. We notice that for $\mathcal{T}$ to be in $\mathcal{B}_M$, $\mathcal{T}^{(r)}$ must not be equal to T_0. So

$$\mathbb{P}(\mathcal{T} \in \mathcal{B}_M) = \sum_{T \neq T_0} \mathbb{P}\left(\mathcal{T}^{(r)} = T\right) \mathbb{P}\left(\mathcal{T} \in \mathcal{B}_M \mid \mathcal{T}^{(r)} = T\right).$$

Conditioning on the event $\{\mathcal{T}^{(r)} = T\}$, the rest of $\mathcal{T}$ is a forest consisting of $w_r(T)$ independent copies of $\mathcal{T}$. Hence, by the union bound, we obtain

$$\mathbb{P}(\mathcal{T} \in \mathcal{B}_M) \leq \sum_{T \neq T_0} \mathbb{P}\left(\mathcal{T}^{(r)} = T\right) w_r(T)\mathbb{P}(\mathcal{T} \in \mathcal{B}_{M-r})$$

$$= \mathbb{P}(\mathcal{T} \in \mathcal{B}_{M-r}) \sum_{T \neq T_0} \mathbb{P}\left(\mathcal{T}^{(r)} = T\right) w_r(T)$$

$$\leq \mathbb{P}(\mathcal{T} \in \mathcal{B}_{M-r}) \sum_{T \neq T_0} \mathbb{P}(\mathcal{T} = T) w_r(T).$$

If we let $q = \sum_{T \neq T_0} \mathbb{P}(\mathcal{T} = T) w_r(T)$, then

$$\mathbb{P}(\mathcal{T} \in \mathcal{B}_M) \leq q\,\mathbb{P}(\mathcal{T} \in \mathcal{B}_{M-r}). \tag{46}$$

On the other hand, we know that

$$q + w_r(T_0)\mathbb{P}(\mathcal{T} = T_0) = \sum_{T} \mathbb{P}(\mathcal{T} = T) w_r(T) = \mathbb{E}w_r(\mathcal{T}) = 1.$$

In view of (34), we deduce that $q < 1$. Therefore, iterating (46) yields

$$\mathbb{P}(\mathcal{T} \in \mathcal{B}_M) \leq q^{\lfloor M/r \rfloor} \leq c_1^M, \tag{47}$$

where $c_1 := q^{1/r} < 1$, proving the first estimate.

For the second estimate, we also begin in a similar fashion, i.e. we have

$$\mathbb{P}\left(\hat{\mathcal{T}} \in \mathcal{B}_M\right) = \sum_{T \neq T_0} \mathbb{P}\left(\hat{\mathcal{T}}^{(r)} = T\right) \mathbb{P}\left(\hat{\mathcal{T}} \in \mathcal{B}_M \mid \hat{\mathcal{T}}^{(r)} = T\right).$$

Here, when conditioning on the event $\{\hat{\mathcal{T}}^{(r)} = T\}$, the rest of $\hat{\mathcal{T}}$ is a forest consisting of $w_r(T) - 1$ independent copies of $\mathcal{T}$ and a copy of $\hat{\mathcal{T}}$. Thus,

$$\mathbb{P}\left(\hat{\mathcal{T}} \in \mathcal{B}_M\right) \leq \sum_{T \neq T_0} \mathbb{P}\left(\hat{\mathcal{T}}^{(r)} = T\right) \left((w_r(T) - 1)\mathbb{P}\left(\mathcal{T} \in \mathcal{B}_{M-r}\right) + \mathbb{P}\left(\hat{\mathcal{T}} \in \mathcal{B}_{M-r}\right)\right).$$

Using (47), letting $q_2 = \sum_{T \neq T_0} \mathbb{P}\left(\hat{\mathcal{T}}^{(r)} = T\right)(w_r(T) - 1)$ (which is finite since it is bounded above by $\mathbb{E}w_r(\hat{\mathcal{T}}) < \infty$), and noting that $q \geq \sum_{T \neq T_0} \mathbb{P}\left(\hat{\mathcal{T}}^{(r)} = T\right)$ by the definition of $\hat{\mathcal{T}}$, we obtain

$$\mathbb{P}\left(\hat{\mathcal{T}} \in \mathcal{B}_M\right) \ll q_2 \, c_1^{M-r} + q \, \mathbb{P}\left(\hat{\mathcal{T}} \in \mathcal{B}_{M-r}\right).$$

Iterating this, we have

$$\mathbb{P}\left(\hat{\mathcal{T}} \in \mathcal{B}_M\right) \ll q^{\lfloor M/r \rfloor} + q_2 \sum_{j=1}^{\lfloor M/r \rfloor} q^{j-1} \, c_1^{M-jr}.$$

Since we set $q = c_1^r$, the latter estimate becomes

$$\mathbb{P}\left(\hat{\mathcal{T}} \in \mathcal{B}_M\right) \ll M \, c_1^M.$$

The proof is completed by choosing any constant $c > 0$ such that $c_1 < c < 1$. ◀

Local Limits of Large Galton–Watson Trees Rerooted at a Random Vertex

Benedikt Stufler[1]

University of Zurich, Institute of Mathematics, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
benedikt.stufler@math.uzh.ch

───── **Abstract** ─────

We prove limit theorems describing the asymptotic behaviour of a typical vertex in random simply generated trees as their sizes tends to infinity. In the standard case of a critical Galton–Watson tree conditioned to be large, the limit is the invariant random sin-tree constructed by Aldous (1991). Our main contribution lies in the condensation regime where vertices of macroscopic degree appear. Here we describe in complete generality the asymptotic local behaviour from a random vertex up to its first ancestor with "large" degree. Beyond this distinguished ancestor, different behaviours may occur, depending on the branching weights. In a subregime of complete condensation, we obtain convergence toward a novel limit tree, that describes the asymptotic shape of the vicinity of the full path from a random vertex to the root vertex. This includes the important case where the offspring distribution follows a power law up to a factor that varies slowly at infinity.

1 Introduction

A Galton–Watson branching process is a classical stochastic model for population evolution. The process starts with a single individual. All individuals reproduce asexually and independently from each other according to the same offspring distribution. The genealogical tree corresponding to such a process is called a Galton–Watson tree. We call such a tree critical if the average number of children of a node equals 1, and subcritical if it is strictly less than one. In this context we call the number of offspring of node its outdegree. The tree obtained by conditioning the total population size to be equal to an integer n is a special case of a simply generated tree $\mathcal{T}_n$. That is, there is a fixed sequence $(\omega_k)_{k \geq 0}$ of so-called branching weights and the tree $\mathcal{T}_n$ assumes a plane tree T with n vertices with probability proportional to the product $\prod_{v \in T} \omega_{d_T^+(v)}$ of weights corresponding to the vertex outdegrees $d_T^+(v)$. The present work aims to describe the vicinity of a typical vertex in $\mathcal{T}_n$ as n tends to infinity. We refer the reader to Section 3 for a brief introduction of this model of random trees and to Drmota's book [7, Sec. 1.2.7] for a more detailed discussion.

The asymptotic shape of the vicinity of the fixed root vertex in random trees has received considerable attention in recent literature. Jonsson and Stefánsson [11] described a phase transition between an infinite spine case and a condensation setting for large Galton–Watson trees with a power-law offspring distribution. A third regime for random

───────────────

simply generated trees with superexponential branching weights was studied by Janson, Jonsson and Stefánsson [10]. The asymptotic shape of simply generated trees as their size tends to infinity was later described in complete generality by Janson [9]. Abraham and Delmas [3, 4] classified the limits of conditioned Galton–Watson trees as the total number of vertices with outdegree in a given fixed set tends to infinity. Limits of Galton–Watson trees having a large number of protected nodes were established by Abraham, Bouaziz, and Delmas [1]. The asymptotic shape of conditioned multi-type Galton–Watson trees was studied by Stephenson [14], Abraham, Delmas, and Guo [2], and Pénisson [13].

Clearly considerable effort and progress is being made in understanding local limits of random trees that describe the asymptotic behaviour near the fixed root vertex, and for random simply generated trees even a complete classification is available. As for the question of the asymptotic shape of the vicinity of a random vertex, Aldous [5] studied in his pioneering work asymptotic fringe distributions for general families of random trees. For the case of critical Galton–Watson trees, he established, at least when the offspring distribution has finite variance, convergence of the tree obtained by rerooting at a random vertex. Janson [9, Thm. 7.12] described the asymptotic behaviour of the fringe subtree rooted at a random vertex of a simply generated tree. Here a fringe subtree at a vertex v refers to the subtree formed by v and all its descendants. A fringe subtree of some ancestor of v is also called an extended fringe subtree. A recent work by Holmgren and Janson [8] studied fringe subtrees and extended fringe subtrees of models of random trees that may be described by the family tree of a Crump–Mode–Jagers branching process stopped at a suitable time, including random recursive trees, preferential attachment trees, fragmentation trees and m-ary search trees.

Janson [9] distinguishes three types of simply generated trees, numbered I, II and III, and for each the local limit exhibits a distinguishing characteristic. See Subsection 3.1 for a brief review of this notation. We use this terminology in our study of the vicinity of a random vertex. In the type I setting, the simply generated tree $\mathcal{T}_n$ is distributed like a critical Galton–Watson tree conditioned on having n vertices. Thus the height of a random vertex in $\mathcal{T}_n$ is typically large and extended fringe trees are typically small. In this regime, the limit is given by the random sin-tree constructed by Aldous [5]. Here the word *sin* refers to the fact that, like the Kesten tree, this tree has almost surely up to finite initial segments only a single infinite path. When the offspring distribution has finite variance, we may even verify total variational convergence of the extended fringe subtree up to $o(\sqrt{n})$-distant ancestors.

While trees in the type I regime usually have small maximum degree, the types II and III are characterized by the appearance of vertices with large degree, which may be viewed as a form of condensation. Specifically, type II simply generated trees correspond to subcritical Galton–Watson trees with a heavy-tailed offspring distribution, and type III simply generated trees have superexponential branching weights such that no equivalent conditioned Galton–Watson tree exists. Our main contribution is in this condensation setting, where contrary to the type I regime a random vertex may be near to the root, and extended fringe trees may have size comparable to the total number of vertices of $\mathcal{T}_n$, as we are likely to encounter an ancestor with large degree. This is also a major difference to the settings addressed in the mentioned works by Aldous [5] and Holmgren and Janson [8].

We set up a compact space that encodes rooted plane trees that are centered around a second distinguished vertex, and establish several limit theorems. For arbitrary weight-sequences having type II or III, we establish a limit that describes the vicinity of a random vertex up to and including its first ancestor with large degree. Here *large* means having outdegree bigger than a deterministic sequence that tends to infinity sufficiently slowly. The

asymptotic shape of what lies beyond this ancestor appears to depend on the branching weights. In a way, the vertex with large degree obstructs the view to older generations.

We describe a novel limit object $\mathcal{T}^*$ given by a random pointed plane tree, in which the pointed vertex has random distance from its first ancestor with infinite degree, and this ancestor again has a random number of ancestors with finite degree before the construction stops. For arbitrary weight-sequences, the asymptotic probability for the vicinity of a random vertex of $\mathcal{T}_n$ to have a specific shape that admits at most one single ancestor of large degree, but allows ancestors with small degrees afterwards, coincides with the corresponding probability for the tree $\mathcal{T}^*$. Our approach is based on a heavily modified depth-first-search to explore the tree $\mathcal{T}_n$. This yields information on how parts of a limit tree for the complete vicinity, that is not truncated at the first large ancestor, must look, if the simply generated tree $\mathcal{T}_n$ pointed at a random vertex converges weakly (along a subsequence). Note also that the compactness of the space, in which we formulate our limits, guarantees the existence of such subsequences. Thus the obstruction by the ancestor with large degree, that prevents us from seeing older generations, is not a complete blockage. However, this is not yet sufficient to deduce convergence in the space of pointed plane trees. In general, the tip of the backwards growing spine, where the construction of $\mathcal{T}^*$ breaks off, may correspond to the root vertex of $\mathcal{T}_n$, but just as well to a second ancestor with large degree.

If the branching weights belong to a general regime of complete condensation, we manage to surpass the blockage and deduce weak convergence toward $\mathcal{T}^*$. There are two main steps involved. First, we show that convergence toward $\mathcal{T}^*$ is in fact equivalent to weak convergence of the height of a random vertex in $\mathcal{T}_n$ to the height of the pointed vertex in the tree $\mathcal{T}^*$, which in the type II regime is distributed like 1 plus the sum of two independent identically distributed geometric random variables, and in the type III regime equals 1. In this case, the root of $\mathcal{T}^*$ really corresponds to the root of $\mathcal{T}_n$. The second step verifies this property in the case of complete condensation, where the maximum degree of $\mathcal{T}_n$ has the correct order.

In particular, Kortchemski's central limit theorem for the maximum degree [12, Theorem 1] allows us to deduce convergence toward $\mathcal{T}^*$ in the general case of a subcritical Galton–Watson tree conditioned on having n vertices, if the offspring distribution ξ satisfies $\mathbb{P}(\xi = k) = f(k)k^{-\alpha}$ for a constant $\alpha > 2$ and a function f that varies slowly at infinity. In the type III regime where branching weights grow superexponentially fast, we consider the specific case where $\omega_k = k!^\alpha$ for $\alpha > 0$. It is known that for these weights the maximum degree of $\mathcal{T}_n$ has order $n + o_p(n)$, which may also be seen as complete condensation, see Janson, Jonsson, and Stefánsson [10] and Janson [9, Example 19.36]. Thus here the tree $\mathcal{T}^*$ is also the weak limit of the simply generated tree $\mathcal{T}_n$ pointed at a random vertex. There are, however, also examples of superexponential branching weights that exhibit a more irregular behaviour [9, Example 19.38], in which weak convergence towards $\mathcal{T}^*$ does not hold.

The present work is an extended abstract of [15], where detailed justifications of all results are provided.

2 The space of pointed plane trees

2.1 Centering at a specified vertex

A *plane tree* is a rooted tree in which the offspring set of each vertex is endowed with a linear order. (Such trees are also sometimes referred to as planted plane trees or corner rooted plane trees, in order to distinguish them from related planar structures [7].) In the present work we will also encounter plane trees that have no root, but whose vertex sets are

endowed with a partial order that specifies the ancestry relations, and whose offspring sets are endowed with a linear order that is not required to have a smallest element.

Given a plane tree T and a vertex $v \in T$ we let $d_T^+(v)$ denote its *outdegree*, that is, the number of offspring. Its *height* $\mathrm{h}_T(v)$ is its distance from the root-vertex. Traditionally, plane trees are encoded as subtrees of the *Ulam–Harris tree*. The Ulam–Harris tree $\mathcal{U}_\infty$ is an infinite plane tree with vertex set $\mathcal{V}_\infty = \mathbb{N}^{(\mathbb{N})}$ given by the space of finite sequences of non-negative integers. Its root vertex is the unique sequence with length zero, and the ordered offspring of a vertex v are the concatenations (v, i) for $i \geq 1$. Thus a plane tree is a subtree of the Ulam–Harris tree that contains its root, such that the offspring set of each vertex is an initial segment of the offspring of the corresponding vertex in $\mathcal{U}_\infty$. Here we explicitly allow trees with infinitely many vertices, and vertices with countably infinite outdegree. If all outdegrees of a plane tree are finite, we say that it is *locally finite*. The tree is *finite*, if its total number of vertices is. We will usually let o denote the root-vertex of a plane tree.

Subtrees of the Ulam–Harris tree are however not an adequate form to represent the vicinity of a specified vertex in a plane tree. If this vertex does not coincide with the root of the tree, then it has an ordered sequence of ancestors and possibly also siblings that lie to the left and right of it. If we look at a random vertex of the simply generated $\mathcal{T}_n$, then it may happen that the number of siblings to the left and/or right of it is asymptotically large, or that its distance from the root vertex is large. A sensible space in which we may describe the limit of the vicinity of the random vertex in $\mathcal{T}_n$ must hence contain trees with a center that may have infinitely many ancestors, such that each may have infinitely many siblings to the left and/or right of it, including the center vertex itself.

For this reason, we describe the construction of an infinite tree $\mathcal{U}_\infty^\bullet$ that is embedded in the plane and has a spine $(u_i)_{i \geq 0}$ that grows "backwards". That is, we construct the tree $\mathcal{U}_\infty^\bullet$ by starting with an infinite path $u_0, u_1, \ldots u_i$ of abstract vertices and define u_i to be a parent of u_{i-1} for all $i \geq 1$. Additionally, any vertex u_i with $i \geq 1$ receives an infinite number of vertices to the left and to the right of its distinguished offspring u_{i-1}, and each of these "non-centered" offspring vertices is the root of a copy of the Ulam–Harris tree $\mathcal{U}_\infty$. To conclude the construction, the start-vertex u_0 of the spine also gets identified with the root of a copy of $\mathcal{U}_\infty$. We let $\mathcal{V}_\infty^\bullet$ denote the vertex-set of the tree $\mathcal{U}_\infty^\bullet$. The tree $\mathcal{U}_\infty^\bullet$ is illustrated in Figure 1.

Note that the vertex set $\mathcal{V}_\infty^\bullet$ carries a natural partial order (given by the transitive hull of the parent-child relations specified in the construction of $\mathcal{U}_\infty^\bullet$), and the offspring set of any given vertex carries a natural linear order. This allows us to continue using the terms *ancestor* and *offspring* in this context.

A plane tree T together with a distinguished vertex v_0 is called a *pointed* plane tree, and may be interpreted in a canonical way as a subtree of $\mathcal{U}_\infty^\bullet$. To do so, let $v_0, v_1, \ldots, v_k$ denote the path from v_0 to the root of T. This way, any vertex v_i for $i \geq 1$ may have offspring to the left and to the right of v_{i-1}. Thus there is a unique order-preserving and outdegree preserving embedding of T into $\mathcal{U}_\infty^\bullet$ such that v_i corresponds to u_i for all $0 \leq i \leq k$. Compare with Figure 1.

2.2 Topological properties

Any plane tree T may be identified with its family of outdegrees $(d_T^+(v))_{v \in \mathcal{V}_\infty} \in \overline{\mathbb{N}}_0^{\mathcal{V}_\infty}$, where we set $\overline{\mathbb{N}}_0 = \mathbb{N}_0 \cup \{\infty\}$. Here we use the convention $d_T^+(v) = 0$ if $v \in \mathcal{V}_\infty$ is not a vertex of the tree T. We endow $\overline{\mathbb{N}}_0$ with the one-point compactification topology of the discrete space
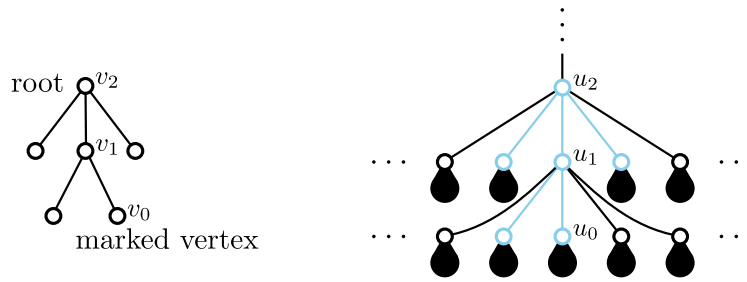
Figure 1 Embedding of a pointed plane tree into the tree $\mathcal{U}_\infty^\bullet$. Each black blob represents a copy of the Ulam–Harris tree.

$\mathbb{N}_0$. Thus plane trees are elements of the compact product space $\overline{\mathbb{N}}_0^{\mathcal{V}_\infty}$. It is not hard to see that the subspace $\mathfrak{T} \subset \overline{\mathbb{N}}_0^{\mathcal{V}_\infty}$ of all plane trees is closed.

Similarly, we may identify a pointed plane tree $T^\bullet = (T, v_0)$ with the corresponding family of outdegrees $(d_{T^\bullet}^+(v))_{v \in \mathcal{V}_\infty^\bullet}$, such that $d_{T^\bullet}^+(v) \in \overline{\mathbb{N}}_0$ for $v \notin \{u_1, u_2, \ldots\}$, and $d_{T^\bullet}^+(u_i) \in \{*\} \sqcup (\overline{\mathbb{N}}_0 \times \overline{\mathbb{N}}_0)$ for $i \geq 1$. Here the two numbers represent the number of offspring vertices to the left and right of the distinguished son u_{i-1}, and the $*$-placeholder represents the fact that the vertex does not belong to the tree.

Since $\overline{\mathbb{N}}_0$ is a compact Polish space, so are the product $\overline{\mathbb{N}}_0 \times \overline{\mathbb{N}}_0$ and the disjoint union topology on $\{*\} \sqcup (\overline{\mathbb{N}}_0 \times \overline{\mathbb{N}}_0)$. Hence the space of all families $(d^+(v))_{v \in \mathcal{V}_\infty^\bullet}$ satisfying

$$d^+(v) \in \begin{cases} \overline{\mathbb{N}}_0 & \text{for } v \notin \{u_1, u_2, \ldots\} \\ \{*\} \sqcup (\overline{\mathbb{N}}_0 \times \overline{\mathbb{N}}_0) & \text{for } v \in \{u_1, u_2, \ldots\} \end{cases}$$

is the product of countably many compact Polish spaces, and hence also compact and Polish. It is easy to verify that the subset $\mathfrak{T}^\bullet$ of all elements that correspond to trees (that is, connected acyclic graphs) is closed, and hence also a compact Polish space with respect to the subspace topology.

3 Simply generated trees

We let $\mathbf{w} = (\omega_i)_{i \geq 0}$ denote a sequence of non-negative weights satisfying $\omega_0 > 0$ and $\omega_k > 0$ for at least one $k \geq 2$. The weight of a plane tree T is defined by $\omega(T) = \prod_{v \in T} \omega_{d_T^+(v)}$. The simply generated tree $\mathcal{T}_n$ with n vertices gets drawn from the set of all n-vertex plane trees with probability proportional to its weight. Galton–Watson trees conditioned on having a fixed number of vertices are encompassed by this model of random plane trees. Of course, the tree $\mathcal{T}_n$ is only well-defined if there is at least one plane tree with n vertices that has positive weight. We set $\mathsf{span}(\mathbf{w}) = \gcd\{i \geq 0 \mid \omega_i > 0\}$. As argued in [9, Corollary 15.6], n-sized trees with positive weight may only exist for $n \equiv 1 \mod \mathsf{span}(\mathbf{w})$, and conversely, they always exist if n is large enough and belongs to this congruence class. We tacitly only consider such n throughout this paper.

3.1 Three types of weight-sequences

Janson [9, Chapter 8] distinguishes three types of weight-sequences. The classification is as follows. Let ρ_ϕ denote the radius of convergence of the generating series $\phi(z) = \sum_{k \geq 0} \omega_k z^k$. As argued in [9, Lemma 3.1], if $\rho_\phi > 0$ then the function $\psi(t) = \phi'(t)t/\phi(t)$ admits a limit $\nu = \lim_{t \nearrow \rho_\phi} \psi(t) \in]0, \infty]$ with the following properties. If $\nu \geq 1$, then there is a unique

number τ with $\psi(\tau) = 1$ and we say the weight sequence $\mathbf{w}$ has type I. If $0 < \nu < 1$, then we set $\tau := \rho_\phi < \infty$ and say $\mathbf{w}$ has type II. If $\rho_\phi = 0$, we say $\mathbf{w}$ has type III and set $\nu = 0$ and $\tau = 0$. The constant ν has a natural interpretation as the supremum of the means of all probability weight sequences equivalent to $\mathbf{w}$. The inclined reader may see [9, Remark 4.3] for details.

3.2 An associated Galton–Watson tree

We define the probability distribution $(\pi_k)_k$ on $\mathbb{N}_0$ by $\pi_k = \tau^k \omega_k / \phi(\tau)$. The mean and variance of the distribution $(\pi_k)_k$ are given by $\mu = \min(\nu, 1)$ and $\sigma^2 = \tau \psi'(\tau) \leq \infty$. We let ξ denote a random non-negative integer with density $(\pi_k)_k$, and $\mathcal{T}$ a Galton–Watson tree with offspring distribution ξ. Note that if $\mathbf{w}$ has type III, then $\xi = 0$ almost surely and the tree $\mathcal{T}$ consists of a single deterministic vertex. As detailed in [9, Section 4], if $\mathbf{w}$ has type I or II then the simply generated tree $\mathcal{T}_n$ is distributed like the Galton–Watson tree $\mathcal{T}$ conditioned on having n vertices.

4 The limit theorems

As discussed in Section 3 there is a probability distribution $(\pi_k)_k$ associated with the weight sequence $\mathbf{w}$. Let ξ be distributed according to $(\pi_k)_k$ and let $\mathcal{T}$ be a ξ-Galton–Watson tree. Thus $\mu := \mathbb{E}[\xi] \leq 1$. We may consider the *size-biased* random variable $\hat{\xi}$ with values in $\overline{\mathbb{N}}_0$ and distribution given by

$$\mathbb{P}(\hat{\xi} = k) = k\pi_k \quad \text{and} \quad \mathbb{P}(\hat{\xi} = \infty) = 1 - \mu.$$

For any tree T and any vertex $v \in T$ we let $f(T, v)$ denote the *fringe-subtree* of T at v. That is, the maximal subtree of T that is rooted at the vertex v.

   Throughout the following, we let v_0 denote a uniformly at random selected vertex of the simply generated plane tree $\mathcal{T}_n$, that in the type I and II regime is distributed like the Galton–Watson tree $\mathcal{T}$ conditioned on having n vertices.

4.1 The type I regime

If the weight-sequence $\mathbf{w}$ has type I, then $\hat{\xi} < \infty$ almost surely, and we define the random pointed tree $\mathcal{T}^*$ as follows. Let u_0 be the root of an independent copy of the Galton–Watson tree $\mathcal{T}$. For each $i \geq 1$, we let u_i receive offspring according to an independent copy of $\hat{\xi}$. The vertex u_{i-1} gets identified with an uniformly at random chosen offspring of u_i. All other offspring vertices of u_i become the root of an independent copy of the Galton–Watson tree $\mathcal{T}$. Compare with Figure 2.

▶ **Theorem 1.** *If the weight-sequence* $\mathbf{w}$ *has type I, then*

$$(\mathcal{T}_n, v_0) \xrightarrow{d} \mathcal{T}^*$$

in the space $\mathfrak{T}^\bullet$.

Let T be a plane tree, $v \in T$ a vertex, and $k \geq 0$ an integer. If the vertex v has a kth ancestor v_k, then we may define the pointed plane tree $H_k(T, v)$ as the fringe tree $f(T, v_k)$ that is rooted at the vertex v_k and pointed at the vertex v. Here we use the term vertex in the graph-theoretic sense, since the coordinates of the vertex v as node of the Ulam–Harris tree depend on whether we talk about $v \in T$ or $v \in f(T, v_k)$. If the vertex v has height $h_T(v) < k$, we set $H_k(T, v) = \diamond$ for some placeholder symbol $\diamond$.
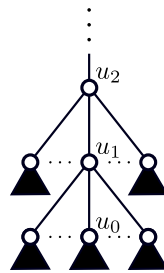
Figure 2 The limit tree $\mathcal{T}^*$ in the type I regime. Each triangle represents an independent copy of the Galton–Watson tree $\mathcal{T}$. For each $i \geq 1$ the vertex u_i receives offspring according to an independent copy of $\hat{\xi}$, and the location of u_{i-1} within that offspring set is chosen uniformly at random.
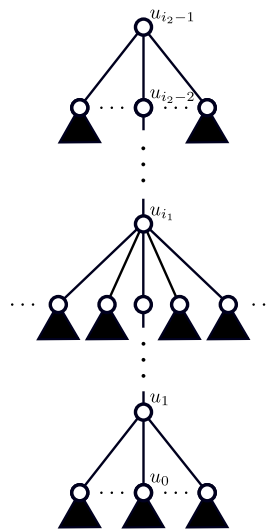
Figure 3 The limit tree $\mathcal{T}^*$ in the complete condensation regime. The vertex u_{i_1} is the only one having infinite degree, and each triangle represents an independent copy of the Galton–Watson tree $\mathcal{T}$.

▶ **Theorem 2.** *Suppose that weight-sequence has type I and the offspring distribution ξ has finite variance. Let k_n be an arbitrary sequence of non-negative integers that satisfies $k_n/\sqrt{n} \to 0$. Then*

$$d_{\mathrm{TV}}(H_{k_n}(\mathcal{T}_n, v_0), H_{k_n}(\mathcal{T}^*, u_0)) \to 0$$

as n becomes large.

Here we use the redundant notation $(\mathcal{T}^*, u_0)$ to emphasize that the tree $\mathcal{T}^*$ is marked at the vertex u_0.

4.2 Complete condensation in the type II regime

If the weight-sequence $\mathbf{w}$ has type II or III, then we construct $\mathcal{T}^*$ similarly as in the type I case, letting u_0 become the root of an independent copy of the Galton–Watson tree $\mathcal{T}$, and letting for $i = 1, 2, \ldots$ the vertex u_i receives offspring according to an independent copy $\hat{\xi}_i$ of $\hat{\xi}$, where a uniformly at random chosen son gets identified with u_{i-1} (specifying the number

of siblings to the left and right of u_{i-1}) and the rest become roots of independent copies of $\mathcal{T}$. We proceed in this way for $i = 1, 2, \ldots$ *until* it occurs for the first time i_1 that $\hat{\xi}_{i_1} = \infty$. When $\hat{\xi}_1, \ldots, \hat{\xi}_{i_1-1} < \infty$ and $\hat{\xi}_{i_1} = \infty$, then u_{i_1} receives infinitely many offspring to the left and right of its son u_{i_1-1}. Each of these vertices (except u_{i_1-1} of course) gets identified with an independent copy of the Galton–Watson tree $\mathcal{T}$. We then proceed as before for $i = i_1, i_1 + 1, \ldots$, such that u_i receives offspring according to an independent copy $\hat{\xi}_i$ of $\hat{\xi}$, with a random son being identified with u_{i-1} and the rest becoming roots of independent copies of $\mathcal{T}$, *until* it happens for the second time i_2 that $\hat{\xi}_{i_2} = \infty$. When $\hat{\xi}_{i_1} = \infty = \hat{\xi}_{i_2}$ for $i_1 < i_2$ and $\hat{\xi}_i < \infty$ for all $i < i_2$ with $i \neq i_1$, then we stop the construction. The spine of the resulting tree is then given by the ordered path $u_0, \ldots, u_{i_2-1}$. Compare with Figure 3.

Note that this construction also works in the type I regime and yields the tree $\mathcal{T}^*$ as defined in Section 4.1, since $\mathbf{w}$ having type I implies that almost surely $\hat{\xi}_i < \infty$ for all $i \geq 1$.

▶ **Theorem 3.** *Suppose that the weight-sequence $\mathbf{w}$ has type II. If the maximum degree $\Delta(\mathcal{T}_n)$ satisfies $\Delta(\mathcal{T}_n) = (1 - \mu)n + o_p(n)$, then it holds that*

$$(\mathcal{T}_n, v_0) \xrightarrow{d} \mathcal{T}^*$$

in the space $\mathfrak{T}^\bullet$. In particular, this is the case when there is a constant $\alpha > 2$ and a slowly varying function f such that for all k $\mathbb{P}(\xi = k) = f(k)k^{-\alpha}$.

Here we make use of a result by Kortchemski [12, Theorem 1] who established a central limit theorem for the maximum degree, that ensures that $\Delta(\mathcal{T}_n)$ has the correct order if the offspring distribution ξ has a power law up to a slowly varying factor. There are also examples of offspring distributions with a more irregular behaviour. Janson [9, Example 19.37] constructed a weight sequence such that along a subsequence $n = n_k$ it holds that $\Delta(\mathcal{T}_n) = o_p(n)$, and along another subsequence several vertices with degree comparable to n exist. This may be seen as incomplete condensation.

The proof idea of Theorem 3 is to deduce the asymptotic distribution of the height $\mathrm{h}_{\mathcal{T}_n}(v_0)$ by localizing the vertex of $\mathcal{T}_n$ having maximum degree at a position, that was also given in [12, Theorem 2] using results by Armendáriz and Loulakis [6] concerning conditioned random walks having a subexponential jump distribution. To do so, we employ results of Janson [9, Chapter 20] that (partially) use

$$\Delta(\mathcal{T}_n) = (1 - \mu)n + o_p(n),$$

but do not assume the offspring distribution to be subexponential. The following main lemma, which characterizes convergence toward the tree $\mathcal{T}^*$ in terms of weak convergence of the height $\mathrm{h}_{\mathcal{T}_n}(v_0)$, then finalizes the proof of Theorem 3.

▶ **Lemma 4.** *If the weight-sequence $\mathbf{w}$ has type II or III, then the following three conditions are equivalent.*
1. $(\mathcal{T}_n, v_0) \xrightarrow{d} \mathcal{T}^*$ *in $\mathfrak{T}^\bullet$.*
2. $\mathrm{h}_{\mathcal{T}_n}(v_0) \xrightarrow{d} \mathrm{h}_{\mathcal{T}^*}(u_0)$.
3. $\limsup_{n \to \infty} \mathbb{P}(\mathrm{h}_{\mathcal{T}_n}(v_0) \geq k) \leq \mu^k + k(1 - \mu)\mu^{k-1}$ *for all $k \geq 1$.*
Note that $\mathrm{h}_{\mathcal{T}^*}(u_0)$ is distributed like 1 plus the sum of two independent identically distributed geometric random variables that assume an integer i with probability $\mu^i(1 - \mu)$.

4.3 Complete condensation in the type III regime

If the weight-sequence $\mathbf{w}$ has type III, then it holds that $\mu = 0$ and almost surely $\xi = 0$ and $\hat{\xi} = \infty$. Here the Galton–Watson tree $\mathcal{T}$ is always equal to a single point. Hence the tree

$\mathcal{T}^*$ is obtained by letting u_1 have infinitely many offspring to the left and right of u_0, all of which (including u_0) are leaves.

▶ **Proposition 5.** *If the weight-sequence* **w** *has type III, then the following claims are equivalent.*
1. $(\mathcal{T}_n, v_0) \xrightarrow{d} \mathcal{T}^*$ *in* $\mathfrak{T}^\bullet$.
2. $\mathrm{h}_{\mathcal{T}_n}(v_0) \xrightarrow{p} 1$.
3. *The maximum degree* $\Delta(\mathcal{T}_n)$ *satisfies* $\Delta(\mathcal{T}_n) = n + o_p(n)$.
A general class of weight-sequences that demonstrate this behaviour is given by $\omega_k = k!^\alpha$ *with* $\alpha > 0$ *a constant.*

Here we use that if $\omega_k = k!^\alpha$ with $\alpha > 0$ a constant, then it is known [9, page 226, Example 19.36], that the largest degree in $\mathcal{T}_n$ has size $n + o_p(n)$. But there are also other examples that exhibit a more irregular behaviour. In [9, page 227, Example 19.38] a weight-sequence is constructed such that along a subsequence $n = n_k$, for each $j \geq 1$ the jth largest degree $Y_{(j)}$ in $\mathcal{T}_{n_k}$ satisfies $Y_{(j)} = 2^{-j}n$ with high probability. This may be seen as incomplete condensation. It is clear that in this case the limit of $(\mathcal{T}_n, v_0)$, if it exists at all, must have a different shape than $\mathcal{T}^*$.

4.4 Large nodes and truncated limits

Suppose that the weight sequences **w** has type II or type III. The limit theorems in Subsections 4.2 and 4.3 work in settings of complete condensation, where the maximum degree of the tree $\mathcal{T}_n$ satisfies

$$\Delta(\mathcal{T}_n) = (1 - \mu)n + o_p(n).$$

If we content ourselves with the vicinity of the vertex v_0 up to and including the first vertex having large degree, we may obtain a limit theorem in complete generality. We are also going to construct a coupling to demonstrate how the vertex with infinite degree in the limit corresponds to a vertex with large degree in the simply generated tree $\mathcal{T}_n$.

Janson [9, Lemma 19.32] showed that there is a deterministic sequence Ω_n that tends to infinity sufficiently slowly, such that for any sequence $K_n \to \infty$ with $K_n \leq \Omega_n$ it holds that the numbers N_k of vertices with outdegree k in the tree $\mathcal{T}_n$ satisfy

$$\sum_{k \leq K_n} kN_k = \mu n + o_p(n) \qquad \text{and} \qquad \sum_{k > K_n} kN_k = (1 - \mu)n + o_p(n) \tag{1}$$

The sequence Ω_n may be replaced by any sequence that tends to infinity more slowly. Hence we may assume without loss of generality that Ω_n additionally satisfies $\Omega_n = o(n)$. Let $\tilde{D}_n$ denote a random positive integer, that is independent from all previously considered random variables, with distribution given by $\tilde{D}_n \overset{d}{=} (d^+_{\mathcal{T}_n}(o) \mid d^+_{\mathcal{T}_n}(o) > \Omega_n)$. Here we let o denote the root-vertex of $\mathcal{T}_n$. That is, $\tilde{D}_n$ is distributed like the root-degree conditioned to be "large". We form the random tree $\bar{\mathcal{T}}_n^*$ in a similar manner as the random tree $\mathcal{T}^*$. The vertex u_0 becomes the root of an independent copy of the Galton–Watson tree $\mathcal{T}$. For $i = 1, 2, \ldots$ the vertex u_i receives offspring according to independent copy $\hat{\xi}_i$ of $\hat{\xi}$, where a randomly chosen son gets identified with u_{i-1} and the rest become roots of independent copies of $\mathcal{T}$. We proceed in this way for $i = 1, 2, \ldots$ until it occurs that $\hat{\xi}_i = \infty$. When $\hat{\xi}_1, \ldots, \hat{\xi}_{i-1} < \infty$ and $\hat{\xi}_i = \infty$, then u_i receives $\tilde{D}_n$ offspring vertices, such that a uniformly at random chosen one gets identified with u_{i-1}, and the rest get identified with the roots of independent copies

of $\mathcal{T}$. Rather than continuing with the spine as in the construction of the tree $\mathcal{T}^*$, we stop at this point, so that u_i becomes the root of this tree.

Given a pointed tree $T^\bullet = (T, v)$ and an ancestor a of v we may consider the fringe subtree of T at a as pointed at the vertex v. Let v_0 denote a uniformly at random selected vertex of the simply generated tree $\mathcal{T}_n$. Let $H(\mathcal{T}_n, v_0, \Omega_n)$ denote the pointed fringe subtree of $(\mathcal{T}_n, v_0)$ at the youngest ancestor of v_0 that has outdegree bigger than Ω_n. If no such vertex exists (which is unlikely to happen, as we are going to verify), set $H(\mathcal{T}_n, v_0, \Omega_n) = \diamond$ for some fixed placeholder value $\diamond$.

▶ **Theorem 6.** *Suppose that the weight sequence* **w** *has type II or III. Let* $\bar{\mathcal{T}}^*$ *denote the pointed fringe subtree of the tree* $\mathcal{T}^*$ *at its unique vertex with infinite degree. Then it holds that*

$$H(\mathcal{T}_n, v_0, \Omega_n) \xrightarrow{d} \bar{\mathcal{T}}^*.$$

in the space $\mathfrak{T}^\bullet$.

The strength of this theorem is its generality, as we make no additional assumptions on the weight-sequence at all. It is suitable for applications where it is not necessary to look behind the large vertex.

We may still improve upon this. For each n, let $\bar{\mathcal{T}}_n^*$ be constructed from $\bar{\mathcal{T}}^*$ by pruning at its root vertex such that its outdegree becomes $\tilde{D}_n$. Of course we have to select one of the $\tilde{D}_n$ ways of how much we prune from the left and right so that the total outdegree becomes $\tilde{D}_n$, and we choose an option uniformly at random.

For each integer $m \geq 0$ we let $\bar{V}^{[m]} \subset \mathcal{V}_\infty^\bullet$ denote the vertex set of the tree obtained from $\mathcal{U}_\infty^\bullet$ by deleting all vertices with distance larger than m from the center vertex u_0 and pruning so that the vertices u_i, $1 \leq i \leq m$ have outdegree (m, m) and the remaining vertices all have outdegree equal to m. The topology on the subspace $\mathfrak{T}_{1f}^\bullet \subset \mathfrak{T}^\bullet$ of locally finite trees is induced by the metric

$$d_{\mathfrak{T}_{1f}^\bullet}(T_1^\bullet, T_2^\bullet) = 1/\sup\{m \geq 0 \mid d_{T_1^\bullet}^+(v) = d_{T_2^\bullet}^+(v) \text{ for all } v \in \bar{V}^{[m]}\}.$$

This can be verified using the fact that a sequence $(T_n)_n$ in $\mathfrak{T}^\bullet$ converges towards an element $T \in \mathfrak{T}^\bullet$ if and only if $d_{T_n}^+(v)$ converges towards $d_T^+(v)$ for each $v \in \mathcal{V}_\infty^\bullet$.

▶ **Theorem 7.** *Suppose that the weight sequence* **w** *has type II or III. For any finite set of vertices* $x_1, \ldots, x_r \in \mathcal{V}_\infty^\bullet$ *it holds that*

$$d_{\mathrm{TV}}((d_{H(\mathcal{T}_n, v_0, \Omega_n)}^+(x_i))_{1 \leq i \leq r}, (d_{\bar{\mathcal{T}}_n^*}^+(x_i))_{1 \leq i \leq r}) \to 0.$$

Equivalently, there is a coupling of $(\mathcal{T}_n, v_0)$ *and* $\bar{\mathcal{T}}_n^*$ *such that* $d_{\mathfrak{T}_{1f}^\bullet}(H(\mathcal{T}_n, v_0, \Omega_n), \bar{\mathcal{T}}_n^*) \xrightarrow{p} 0$.

In Equation (20.4) and the subsequent paragraph of [9], Janson also argues that if

$$\Delta(\mathcal{T}_n) = (1 - \mu)n + o_p(n), \tag{2}$$

then $d_{\mathrm{TV}}(\Delta(\mathcal{T}_n), \tilde{D}_n) \to 0$. Hence in the complete condensation regime where (2) is assumed to hold, we may choose $\tilde{D}_n$ in the coupling of Theorem 7 such that $\tilde{D}_n = \Delta(\mathcal{T}_n)$ with probability tending to 1 as n becomes large. This yields the asymptotic location of the vertex with maximum degree with respect to the random vertex v_0.

───── **References** ─────

1   R. Abraham, A. Bouaziz, and J.-F. Delmas. Local limits of Galton-Watson trees conditioned on the number of protected nodes. *ArXiv e-prints*, 2015. `arXiv:1509.02350`.

2   R. Abraham, J.-F. Delmas, and H. Guo. Critical Multi-Type Galton-Watson Trees Conditioned to be Large. *ArXiv e-prints*, 2015. `arXiv:1511.01721`.

3   Romain Abraham and Jean-François Delmas. Local limits of conditioned Galton-Watson trees: the condensation case. *Electron. J. Probab.*, 19:no. 56, 29, 2014.

4   Romain Abraham and Jean-François Delmas. Local limits of conditioned Galton-Watson trees: the infinite spine case. *Electron. J. Probab.*, 19:no. 2, 19, 2014.

5   David Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.*, 1(2):228–266, 1991. URL: `http://links.jstor.org/sici?sici=1050-5164(199105)1:2<228:AFDFGF>2.0.CO;2-8&origin=MSN`.

6   Inés Armendáriz and Michail Loulakis. Conditional distribution of heavy tailed random variables on large deviations of their sum. *Stochastic Process. Appl.*, 121(5):1138–1147, 2011. `doi:10.1016/j.spa.2011.01.011`.

7   Michael Drmota. *Random trees*. SpringerWienNewYork, Vienna, 2009. An interplay between combinatorics and probability. `doi:10.1007/978-3-211-75357-6`.

8   C. Holmgren and S. Janson. Fringe trees, Crump-Mode-Jagers branching processes and m-ary search trees. *ArXiv e-prints*, 2016. `arXiv:1601.03691`.

9   Svante Janson. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.*, 9:103–252, 2012. `doi:10.1214/11-PS188`.

10 Svante Janson, Thordur Jonsson, and Sigurdur Örn Stefánsson. Random trees with superexponential branching weights. *J. Phys. A*, 44(48):485002, 16, 2011. `doi:10.1088/1751-8113/44/48/485002`.

11 Thordur Jonsson and Sigurdur Örn Stefánsson. Condensation in nongeneric trees. *J. Stat. Phys.*, 142(2):277–313, 2011. `doi:10.1007/s10955-010-0104-8`.

12 Igor Kortchemski. Limit theorems for conditioned non-generic Galton-Watson trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(2):489–511, 2015. `doi:10.1214/13-AIHP580`.

13 Sophie Pénisson. Beyond the Q-process: various ways of conditioning the multitype Galton-Watson process. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(1):223–237, 2016.

14 Robin Stephenson. Local convergence of large critical multi-type galton–watson trees and applications to random maps. *Journal of Theoretical Probability*, pages 1–47, 2016. `doi:10.1007/s10959-016-0707-3`.

15 B. Stufler. Local limits of large Galton-Watson trees rerooted at a random vertex. To appear in Annales de l'Institut Henri Poincaré Probabilités et Statistiques.

The Depoissonisation Quintet:
Rice–Poisson–Mellin–Newton–Laplace

Brigitte Vallée

GREYC Laboratory, CNRS and University of Caen, France
brigitte.vallee@unicaen.fr

──── **Abstract** ────

This paper is devoted to the Depoissonisation process which is central in various analyses of the AofA domain. We first recall in Section 1 the two possible paths that may be used in this process, namely the Depoissonisation path and the Rice path. The two paths are rarely described for themselves in the literature, and general methodological results are often difficult to isolate amongst particular results that are more directed towards various applications. The main results for the Depoissonisation path are scattered in at least five papers, with a chronological order which does not correspond to the logical order of the method. The Rice path is also almost always presented with a strong focus towards possible applications. It is often very easy to apply, but it needs a tameness condition, which appears *a priori* to be quite restrictive, and is not deeply studied in the literature. This explains why the Rice path is very often undervalued.

Second, the two paths are not precisely compared, and the situation creates various "feelings": some people see the tools that are used in the two paths as quite different, and strongly prefer one of the two paths; some others think the two paths are almost the same, with just a change of vocabulary. It is thus useful to compare the two paths and the tools they use. This will be done in Sections 2 and 3. We also "follow" this comparison on a precise problem, related to the analysis of tries, introduced in Section 1.7.

The paper also exhibits in Section 4 a new framework, of practical use, where the tameness condition of Rice path is proven to hold. This approach, perhaps of independent interest, deals with the shifting of sequences and then the inverse Laplace transform, which does not seem of classical use in this context. It performs very simple computations. This adds a new method to the Depoissonisation context and explains the title of the paper. We then conclude that the Rice path is both of easy and practical use: even though (much?) less general than the Depoissonisation path, it is easier to apply.

1   General framework

This section first recalls the two probabilistic models, the Bernoulli model and the Poisson model, together with the two main objects attached to a sequence f: the classical Poisson transform P_f, and another sequence, denoted as $\Pi[f]$ and called here the Poisson sequence. We insist on the involutive characteristic of the mapping Π and introduce two new notions, shift and canonical sequences. After the description of the two paths of interest, we present analyses on tries which strongly motivate the work, and will be performed within each path.

1.1   Probabilistic settings

Many algorithms deal with inputs that are finite sequences of data. We give some examples: (a) for text algorithms, data are words, and inputs are finite sequences of words; (b) for geometric algorithms, data are points, and inputs are finite sequences of points; (c) for a source, data are symbols, and inputs are finite sequences of symbols, namely finite words. The cardinality of the input sequence is often chosen as the input size, and, as usual, one is interested in the asymptotic behaviour of the algorithm for large input size.

The probabilistic framework is as follows: Each data (word or point) is produced along a distribution, and the set of data is thus a probabilistic space $(\mathcal{X}, \mathbb{P})$. Very often, the data are independently chosen with the same distribution and the set $(\mathcal{X}^n, \mathbb{P}_{[n]})$ is the product of order n of the space $(\mathcal{X}, \mathbb{P})$. The space of all the inputs is thus the set $\mathcal{X}^\star := \sum_{n \geq 0} \mathcal{X}^n$ of finite sequences $\boldsymbol{x}$ of elements of $\mathcal{X}$, and there are two main probabilistic models:

  (i) The Bernoulli model $\mathcal{B}_n$ (more natural in algorithmics), where the cardinality $N(\boldsymbol{x})$ of $\boldsymbol{x}$ is fixed and equal to n (then tends to ∞);

  (ii) The Poisson model $\mathcal{P}_z$ of parameter z, where the cardinality $N(\boldsymbol{x})$ is a random variable that follows a Poisson law of parameter z, where the fixed parameter z tends also to ∞,

$$\mathbb{P}[N(\boldsymbol{x}) = n] = e^{-z} \frac{z^n}{n!} \,.$$

This model has very nice probabilistic properties, notably properties of independence.

1.2   The Poisson transform and the Poisson sequence

There is a variable (or a cost) $R : \mathcal{X}^\star \to \mathbb{N}$ which describes the behaviour of the algorithm on the input; for instance, for $\boldsymbol{x} \in \mathcal{X}^\star$, $R(\boldsymbol{x})$ is the path length of a tree (trie or dst) built on the sequence $\boldsymbol{x} := (x_1, \ldots, x_n)$ of words x_i. Our final aim is the analysis of R in the Bernoulli model $\mathcal{B}_n$, i.e., the asymptotic study of the sequence $f : n \mapsto f(n)$, where $f(n) := \mathbb{E}_{[n]}[R]$ is the expectation of R in the model $\mathcal{B}_n$. We begin with the easier Poisson model $\mathcal{P}_z$, and study the expectation $\mathbb{E}_z[R]$ in the model $\mathcal{P}_z$ that satisfies

$$\mathbb{E}_z[R] = \sum_{n \geq 0} \mathbb{E}_z[R \,|\, N = n] \, \mathbb{P}_z[N = n] = \sum_{n \geq 0} \mathbb{E}_{[n]}[R] \, \mathbb{P}_z[N = n] = e^{-z} \sum_{n \geq 0} f(n) \frac{z^n}{n!} \,.$$

This leads us to the Poisson transform P_f of the sequence $f : n \mapsto f(n)$, that is written as an exponential generating function (with "signs")[1] and thus defines another sequence p,

$$P_f(z) := e^{-z} \sum_{n \geq 0} f(n) \frac{z^n}{n!} = \sum_{k \geq 0} (-1)^k \frac{z^k}{k!} p(k), \quad \text{with} \quad p(k) := (-1)^k k! [z^k] P_f(z) \,. \qquad (1)$$

---

[1]   The Poisson transform is often called the Poisson generating function. The signs are added in order to get an involutive formula in (2).

▶ **Definition 1.** Consider a sequence $f : n \mapsto f(n)$. Then,

(a) the series P_f defined in (1) is the *Poisson transform* of the sequence f;

(b) the sequence $p : k \mapsto p(k)$ defined in (1) is the *Poisson sequence* of the sequence f.

The following holds:

▶ **Lemma 2.** *Consider a cost R defined on $\mathcal{X}^\star$, its expectation $f(n)$ in the model $\mathcal{B}_n$. Then,*

(a) *its expectation in the model $\mathcal{P}_z$ is the Poisson transform $P_f(z)$;*

(b) *there are binomial relations between the sequences f and p, namely*

$$p(n) = \sum_{k=0}^{n}(-1)^k \binom{n}{k} f(k), \quad and \quad f(n) = \sum_{k=0}^{n}(-1)^k \binom{n}{k} p(k)\,; \tag{2}$$

(c) *the map Π which associates with the sequence f the sequence p is involutive.*

1.3 Description of the two paths.

We only deal here with a sequence f of polynomial growth, for which the Poisson transform $z \mapsto P_f(z)$ is entire. The sequence f is often given in an implicit way, and we assume that we have some knowledge on $P_f(z)$, which may be of two different types

(a) about the Poisson transform $P_f(z)$ itself,

(b) about its coefficients, namely the sequence $\Pi[f]$.

The main question is now: *Is it possible to return to the initial sequence f and obtain some knowledge about its asymptotics?* There are two main return paths, one for each framework: the Depoissonisation path for (a), and the Rice path for (b). We first describe in Section 1.4 the classical toolbox, then, in Section 1.5, a new useful tool. As we aim to provide a precise comparison between the two paths, we perform a kind of "test" on a particular instance which arises when analysing the *trie* structure and is introduced in Sections 1.6 and 1.7.

1.4 Toolbox and main definitions

This section gathers various definitions about domains of the plane, behaviours of functions. Itthen presents the Mellin transform.

Cones and vertical strips. There are two important types of domains of the complex plane we deal with.

(i) The cones built on the real line $\mathbb{R}^+$, with two possible definitions,

$$\mathcal{C}(a,\theta) := \{z \mid |\arg(z-a)| < \theta\} \text{ for } \theta < \pi$$
$$\widehat{\mathcal{C}}(a,\gamma) = \{z \mid \Re(z-a) > \gamma|z-a|\} \text{ for } |\gamma| \leq 1\,,$$

related by the relation $\widehat{\mathcal{C}}(a, \cos\theta) = \mathcal{C}(a,\theta)$. When $a = 0$, it is omitted.

(ii) The vertical strips, or halfplanes: $\mathcal{S}(a,b) := \{z \mid a < \Re z < b\}, \quad \mathcal{S}(a) := \{z \mid \Re z > a\}\,.$

Polynomial growth. This notion plays a fundamental role: A function $s \mapsto \varpi(s)$ defined in an unbounded domain $\Omega \subset \mathbb{C}$ is said to be of polynomial growth if there exists r for which the estimate $|\varpi(s)| = O(|s|^r)$ holds as $s \to \infty$ on Ω. When $\Omega \subset \mathcal{S}(a,b)$, this means: $|\varpi(s)| = O(|\Im s|^r)$; when $\Omega \subset \mathcal{C}(\theta)$ with $\theta < \pi/2$, this means: $|\varpi(s)| = O(|\Re s|^r)$;

Tameness.   A function $s \mapsto \varpi(s)$ is tame on $\Re s > c$ when it is analytic and of polynomial growth there. This notion is extended when $\varpi(s)$ stops being analytic on $\Re s = c$. We will say that ϖ is tame at $s = c$ if it is meromorphic and of polynomial growth in a larger region $\mathcal{R}$ on the left of the line $\Re s = c$ delimited by a frontier curve $\mathcal{F}$. (see [4] and the Annex).

Mellin transform.   The Mellin transform of a function Q defined in $[0, +\infty]$ is defined as

$$Q^{\star}(s) := \int_0^{+\infty} Q(u)\, u^{s-1} du\,.$$

The Mellin transform plays a central role in each of the two paths (see its main properties in the survey paper [7]). In particular, the transform has a nice behaviour on harmonic sums:

$$Q(z) = \sum_k g(\mu_k z) \quad \Longrightarrow \quad Q^{\star}(s) = \left(\sum_k \mu_k^{-s} \right) g^{\star}(s)\,. \tag{3}$$

Moreover, the following lemma[2] proves that the function $\Gamma(s)$ and its derivatives $\Gamma^{(m)}(s)$ are exponentially small along vertical lines (when $|\Im(s)| \to \infty$).

▶ **Lemma 3** (Exponential Smallness Lemma, [7])**.** *If, inside the closure of the cone* $\mathcal{C}(\theta)$ *with* $\theta > 0$*, one has* $Q(z) = O(|z|^{-\alpha})$ *as* $z \to 0$ *and* $Q(z) = O(|z|^{-\beta}$ *as* $|z| \to \infty$*, then the estimate* $Q^*(s) = O(\exp[-\theta|\Im(s)|])$ *uniformly holds in the vertical strip* $\mathcal{S}(\alpha, \beta)$*.*

1.5   A first new tool: Shift and canonical sequences

The notions that are presented here are not introduced in this way in the literature, and, in particular, the notion of canonical sequence appears to be new (and useful), notably in Section 4.

▶ **Definition 4.** Consider a non zero real sequence $n \mapsto f(n)$.
(a) Its degree $\deg(f)$ and its valuation $\mathrm{val}(f)$ are defined as

$$\mathtt{deg}(f) := \inf\{c \mid f(k) = O(k^c)\} \qquad \mathtt{val}(f) := \min\{k \mid f(k) \neq 0\}\,.$$

A sequence f with finite degree is said to be of polynomial growth.
(b) A sequence $n \mapsto f(n)$ satisfies the Valuation-Degree Condition (VD), iff $\mathtt{val}(f) > \mathtt{deg}(f) + 1$.
(c) It is reduced if it satisfies $\mathtt{val}(f) = 0$ and $\mathtt{deg}(f) < -1$.

The VD Condition is essential in the Rice path. As we are (only) interested in the asymptotics of the sequence f, the VD condition is easy to ensure, as we now show: With a sequence F of polynomial growth, we associate the integer

$$\sigma(F) := 1 \quad (\text{if } \mathtt{deg}(F) < 0), \quad \sigma(F) := 2 + \lfloor \mathtt{deg}(F) \rfloor \quad (\text{if } \mathtt{deg}(F) \geq 0)\,, \tag{4}$$

that satisfies the inequality $\sigma(F) > \mathtt{deg}(F) + 1$; we only modify the first terms of the sequence F : we put zeroes for indices $k < \sigma(F)$ together with one 1 for $k = \sigma(F)$, and obtain a new sequence $\tau(F)$ of valuation $\sigma(F)$ that keeps the same asymptotics as the initial sequence F and now satisfies the VD condition.

We will need here to deal with the stronger notion of reduced sequences, and we now explain how to associate with a sequence F, and in a canonical way, a reduced sequence.

---

[2] It is called the Exponential Smallness Lemma in the paper [13], and we keep the same terminology.

▶ **Lemma 5.** *Consider the shifting map T which associates with a sequence f the sequence $T(f)$ defined, for any $n \geq 0$ as*

$$T[f](n) = \frac{f(n+1)}{n+1} \qquad \text{and thus, for } m \geq 1, \text{ as} \qquad T^m[f](n) = \frac{f(n+m)}{(n+1)\ldots(n+m)}.$$

For $m \geq 1$, the inverse mapping T^{-m} associates with a sequence g the sequence f defined as $f(n) = n(n-1)\ldots(n-m+1)\,g(n-m)$, for $n \geq m$.
(a) *The shifting T^m anti-commutes with the involution Π, namely $T^m \circ \Pi = (-1)^m \Pi \circ T^m$.*
(b) *The sequence $\rho(F) := T^{\sigma(F)}(\tau(F))$ associated with F is reduced, with a degree equal to $\deg(F) - \sigma(F)$. It is called the canonical sequence associated to F.*

Proof. Start with the sequence f with valuation ℓ. Then the Poisson transform $P_f(z)$ has itself valuation ℓ and is written as

$$P_f(z) = z^\ell Q(z) \quad \text{with} \quad Q(z) = e^{-z} \sum_{k \geq 0} g(k) \frac{z^k}{k!} = \sum_{k \geq 0} (-1)^k \frac{z^k}{k!} q(k). \tag{5}$$

Then, the two sequences g and $q := \Pi[g]$ associated with f via Eqn (5) are expressed with the iterate of T of order ℓ, namely $g = T^\ell[f]$, $\qquad q := \Pi[g] = (-1)^\ell T^\ell[\Pi[f]]$. ◀

In the sequel, it will be then *sufficient* to deal with the *canonical* sequence $\rho(F)$, and its Poisson sequence $\Pi(\rho(F)) = (-1)^{\sigma(F)} \rho(\Pi(F))$. Then, the results on the asymptotics on $\rho(F)$ will be easily transfered on the initial Poisson pair of F with Properties (a) and (b).

Example. In Section 1.6, we will deal with the following sequences F_0, F_1, F_2, all of valuation 2, which satisfy moreover $F_0(k) = 1$, $\quad F_1(k) = k$, $\quad F_2(k) = k \log k$, $\qquad$ for $k \geq 2$. Their canonical sequences are defined for $k \geq 0$, as

$$f_0(k) = f_1(k) = \frac{1}{(k+1)(k+2)}, \qquad f_2(k) = \frac{\log(k+3)}{(k+1)(k+2)}.$$

1.6 An instance of the context. Probabilistic analysis of tries

A source $\mathcal{S}$ is a probabilistic process which produces infinite words on the (finite) alphabet $\Sigma := [0..r-1]$. A trie is a tree structure, used as a dictionary, which compares words via their prefixes. Given a finite sequence $\boldsymbol{x}$ of (infinite) words emitted by the source $\mathcal{S}$, the trie $\mathcal{T}(\boldsymbol{x})$ built on the sequence[3] $\boldsymbol{x}$ is defined recursively by the following three rules which involve the cardinality $N(\boldsymbol{x})$ of the sequence $\boldsymbol{x}$:
(a) If $N(\boldsymbol{x}) = 0$, then $\mathcal{T}(\boldsymbol{x}) = \emptyset$
(b) If $N(\boldsymbol{x}) = 1$, with $\boldsymbol{x} = (x)$, then $\mathcal{T}(\boldsymbol{x})$ is a leaf labeled by x.
(c) If $N(\boldsymbol{x}) \geq 2$, then $\mathcal{T}(\boldsymbol{x})$ is formed with an internal node and r subtries equal to

$$\mathcal{T}(\boldsymbol{x}_{\langle 0 \rangle}), \ldots, \mathcal{T}(\boldsymbol{x}_{\langle r-1 \rangle}),$$

where $\boldsymbol{x}_{\langle \sigma \rangle}$ denotes the sequence consisting of words of $\boldsymbol{x}$ which begin with symbol σ, stripped of their initial symbol σ. If the set $\boldsymbol{x}_{\langle \sigma \rangle}$ is non empty, the edge which links the subtrie $\mathcal{T}(\boldsymbol{x}_{\langle \sigma \rangle})$ to the internal node is labelled with the symbol σ.

---

[3] The trie depends only on the underlying set $\{x_1, x_2, \ldots, x_n\}$.

Iterating the process, we consider, for a finite prefix $\boldsymbol{w}$, the sequence $\boldsymbol{x}_{\langle\boldsymbol{w}\rangle}$ consisting of words of $\boldsymbol{x}$ which begin with prefix $\boldsymbol{w}$, stripped of their initial prefix $\boldsymbol{w}$, and denote by $N_{\boldsymbol{w}}(\boldsymbol{x}) := N(\boldsymbol{x}_{\langle\boldsymbol{w}\rangle})$ the cardinality of such a sequence. Then, the internal nodes are used for directing the search: they are labelled by prefixes $\boldsymbol{w}$ with $N_{\boldsymbol{w}}(\boldsymbol{x}) \geq 2$. The leaves contain suffixes of $\boldsymbol{x}$, and there are as many leaves as words in $\boldsymbol{x}$.

Trie analysis aims at describing the average shape of a trie. We focus here on *additive* parameters, whose (recursive) definition exactly copies the (recursive) definition of the trie. With a sequence $f : \mathbb{N} \to \mathbb{R}$ – called a *toll* – which satisfies $f(0) = f(1) = 0$ and $f(k) \geq 0$ for $k \geq 2$, we associate a random variable R defined on the set $\mathcal{X}^{\star}$ as follows:
(ab) If $N(\boldsymbol{x}) \leq 1$, then $\quad R(\boldsymbol{x}) = 0$;
(c) if $N(\boldsymbol{x}) \geq 2$, then $\quad R(\boldsymbol{x}) = f(N(\boldsymbol{x})) + \sum_{\sigma \in \Sigma} R(\boldsymbol{x}_{\langle\sigma\rangle})$.
Iterating the recursion leads to the expression $\quad R(\boldsymbol{x}) := \sum_{\boldsymbol{w} \in \Sigma^{\star}} f(N_{\boldsymbol{w}}(\boldsymbol{x}))$. $\qquad$ (6)

The probabilistic properties of R will depend both on the toll f and the source $\mathcal{S}$:

– The probabilistic properties of the source $\mathcal{S}$ are encapsulated in the Dirichlet series $\Lambda(s)$ of the source, introduced in [19], and defined with the fundamental probabilities $\pi_{\boldsymbol{w}}$,

$$\Lambda(s) := \sum_{\boldsymbol{w} \in \Sigma^{\star}} \pi_{\boldsymbol{w}}^{s}, \quad \text{with} \quad \pi_{\boldsymbol{w}} := \mathbb{P}[\text{a word emitted by } \mathcal{S} \text{ begins with the prefix } \boldsymbol{w}]. \text{ (7)}$$

The series $\Lambda(s)$ mainly intervenes via its behaviour near $s = 1$. We consider here a *tame* source, for which $s \mapsto \Lambda(s)$ is tame at $s = 1$, with a simple pole at $s = 1$ whose residue equals $1/h(\mathcal{S})$ where $h(\mathcal{S})$ is the entropy of the source. (See [4] about tameness of sources.)

– Here are some instances of natural tolls : the size is associated to the toll $f(k) = 1$ (for $k \geq 2$) and the path length to the toll $f(k) = k$ (for $k \geq 2$). A version of the `QuickSort` algorithm on words [4] leads to the *sorting toll* $f(k) = k \log k$ (for $k \geq 2$).

We focus here on this last toll, and are interested in the analysis of the associated cost R. The analysis was already performed in [4] with Depoissonisation path (a). We would have wished there to use the Rice path (b) (as we got used in our previous analyses) but we did not succeed. This failure was a strong motivation for the present study, and we now present here two proofs for the following result, each of them using one path.

▶ **Theorem 6.** *Consider a trie built on n words emitted by a* tame *source $\mathcal{S}$. Then the mean value of parameter R associated with the sorting toll f satisfies in the Bernoulli model $\mathcal{B}_n$*

$$r(n) \sim \frac{1}{2\,h(\mathcal{S})}\, n \log^2 n \qquad (n \to \infty)\,.$$

1.7   Main principles of trie analysis

We begin to deal with the Poisson model, that presents the following advantage: In the model $\mathcal{P}_z$, the cardinality $N_{\boldsymbol{w}}$ which appears in Eqn (6) follows a Poisson law of rate $z\,\pi_{\boldsymbol{w}}$ that involves the fundamental probability $\pi_{\boldsymbol{w}}$ defined in (7). We then adapt the general framework defined in Subsection 1.3, both for the initial sequence f and for the sequence r, and consider the two paths:

– Path (a) deals with the Poisson transforms $P_r(z)$ and $P_f(z)$; averaging Relation (6) in the model $\mathcal{P}_z$ entails a relation between the two Poisson transforms

$$P_r(z) = \sum_{\boldsymbol{w} \in \Sigma^{\star}} \mathbb{E}_z[f(N_{\boldsymbol{w}})] = \sum_{\boldsymbol{w} \in \Sigma^{\star}} P_f(z\,\pi_{\boldsymbol{w}}). \tag{8}$$

Then, the function $P_r(z)$ writes as a harmonic sum, and, with (3), its Mellin transform $P_r^*(s)$ factorises and involves the Λ function defined in (7), namely $P_r^*(s) = \Lambda(-s) \cdot P_f^*(s)$.

– Path (b) deals with the Poisson sequences $q = \Pi[r]$ and $p = \Pi[f]$. Then, Relation (8) entails the equality which also involves the Λ function, namely: $q(n) = \Lambda(n) \, p(n)$ for $n \geq 2$.

2 The Depoissonization path

We first provide a general description of the path; then, we apply it to the analysis of tries and obtain a first proof of Theorem 6.

2.1 General description

We first describe the main steps of the path in an informal way.

Main steps. The Depoissonization path deals with the Poisson transform $P_f(z)$:
(a) It compares $f(n)$ and $P_f(n)$ via the Poisson–Charlier expansion.
(b) It uses the tameness of the Mellin transform $P_f^*(s)$ for the asymptotics of $P_f(n)$.
(c) Under Conditions $(\mathcal{JS})$ on the Poisson transform $P_f(z)$, the Poisson-Charlier expansion may be truncated and provides the asymptotic of $f(n)$ with a good remainder.
(d) Moreover, there exists a Condition $(\mathcal{DP})$ on the input sequence f under which the Conditions $(\mathcal{JS})$ hold.

We then describe more precisely the main objects that are involved.

The Poisson-Charlier expansion. Using the Taylor expansion of $P_f(z)$ at $z = n$, the term $f(n)$ admits an (infinite) expansion,

$$f(n) := n![z^n] \, (e^z P(z)) = \sum_{j \geq 0} \frac{P^{(j)}(n)}{j!} \tau_j(n),$$

where the coefficient $\tau_j(n) := n![z^n] \left((z-n)^j e^z \right)$ is a polynomial in n of degree $\lfloor j/2 \rfloor$, closely related to the (classical) Charlier polynomial.

Conditions $(\mathcal{JS})$. *An entire function $P(z)$ satisfies the Conditions $\mathcal{JS}(\alpha, \beta)$ if there exist $\theta \in]0, \pi/2[$, and $\delta < 1$ for which one has, for $z \to \infty$:*
(I) Inside cone $\mathcal{C}(\theta)$, one has $|P(z)| = O\left(|z|^\alpha \log^\beta(1 + |z|) \right)$.
(O) Outside cone $\mathcal{C}(\theta)$, one has $|P(z)e^z| = O\left(e^{\delta|z|} \right)$.

Condition $(\mathcal{DP})$. *There exists an analytic lifting φ for the sequence f which is of polynomial growth inside horizontal cones.*

We now state the two main results of the Depoissonisation path.

▶ **Theorem 7** ([15, 13]). *If the Poisson transform $P_f(z)$ satisfies the $\mathcal{JS}(\alpha, \beta)$ conditions, then the first terms of the Poisson-Charlier expansion provide the beginning of the asymptotic expansion of $f(n)$. More precisely, for any $k > 0$, one has:*

$$f(n) = \sum_{0 \leq j < 2k} P^{(j)}(n) \frac{\tau_j(n)}{j!} + O(n^{\alpha-k} \log^\beta n).$$

▶ **Theorem 8** ([16, 14]). *The two conditions are equivalent :*
  (i) *the sequence f satisfies the Condition $(\mathcal{DP})$;*
  (ii) *the Poisson transform P_f satisfies the Conditions $(\mathcal{JS})$.*

Bibliographic references. The Depoissonisation path is based on five main contributions, that are scattered in the literature. The path, together with its name, was systemized in 1998 by Jacquet and Szpankowski in [15]. They compare the asymptotics of the two sequences, the sequence $f(n)$ and the sequence $P_f(n)$. There were previous results of the same vein, notably a paper due to Hayman [11] in 1956, but they were not known by the AofA community. Jacquet and Szpankowski did not use the Poisson-Charlier expansion which was later introduced in 2010 into the AofA domain by Hwang, Fuchs and Zacharovas in [13]. Jacquet and Szpankowski also introduced conditions on the Poisson transform that we call (following the proposal of [13]) the Conditions $(\mathcal{JS})$. In [15], the authors prove that, under Conditions $(\mathcal{JS})$, it is possible to compare the two sequences $P_f(n)$ and $f(n)$. Later on, in 2010, using the Poisson Charlier expansion, the authors of [13] obtain a direct and natural proof of this comparison, with a more explicit remainder term. Finally, in two other papers, Jacquet and Szpankowski show that the two conditions – Condition $(\mathcal{DP})$ on the sequence f and Conditions $(\mathcal{JS})$ on P_f – are equivalent. The paper [16] deals with the necessary condition whereas the very recent paper [14] deals with the sufficient condition.

2.2 Application to the sorting toll in tries. First proof of Theorem 6.

This section ends with an example of the Depoissonisation path in the study of trie parameters. The Mellin transform of $P_f(z)$ satisfies,

$$P_f^*(s) = \sum_{k \geq 2} \frac{f(k)}{k!} \int_0^\infty e^{-z} z^k z^{s-1} dz = \sum_{k \geq 2} \frac{f(k)}{k!} \Gamma(k+s) = \sum_{k \geq 2} \frac{f(k)}{k} \frac{\Gamma(k+s)}{\Gamma(k)}.$$

The ratio of Gamma Functions can be estimated with the Stirling Formula,

$$\frac{\Gamma(k+s)}{\Gamma(k)} = \frac{(k+s)^{k+s}}{k^k} \frac{e^{-k-s}}{e^{-k}} \sqrt{\frac{k+s}{k}} \left[1 + O\left(\frac{1}{k}\right)\right] = k^s \left[1 + O\left(\frac{|s|}{k}\right)\right], \tag{9}$$

with a O-term uniform in k. Then, the Mellin transform of P_f satisfies, for $f(k) = k \log k$,

$$P_f^*(s) = \sum_{k \geq 2} k^s \log k \left[1 + O\left(\frac{|s|}{k}\right)\right] = -\zeta'(-s) + H_1(s), \quad H_1(s) \text{ analytic on } \Re s < 0. \tag{10}$$

Then $P_f^*(s)$ has a pole at $s = -1$ of order 2, and, together with the tameness of $\Lambda(s)$ at $s = 1$, this entails the following singular expressions for $P_f^*(s)$ and $P_r^*(s)$ at $s = -1$,

$$P_f^*(s) \asymp \frac{1}{(s+1)^2}, \qquad P_r^*(s) \asymp \frac{1}{h(\mathcal{S})} \frac{1}{(s+1)^3}.$$

The tamenesses of $P_f^*(s)$ and $\Lambda(-s)$ at $s = -1$ are enough to deduce, using standard Mellin inverse transform [7], the estimates, for $z \to \infty$,

$$P_f(z) = z \log z \, (1 + o(1)), \qquad P_r(z) = \frac{1}{2h(\mathcal{S})} z \log^2 z \, (1 + o(1)). \tag{11}$$

We now return to the Bernoulli model; we prove that $P_r(z)$ satisfies the Conditions $(\mathcal{JS})$. This will entail the estimate $r(n) \sim P_r(n)$ and end the proof. Assertion (I) is deduced from (11) in some cone $\mathcal{C}(\theta_1)$. For Assertion (O), we write $P_f(z)$ as

$$P_f(z) = z^2 e^{-z} G(z) \quad \text{with} \quad G(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!} g(k) \quad \text{and} \quad g(k) := \frac{1}{k+1} \log(k+2). \quad (12)$$

As the sequence $g := T^2[f]$ satisfies Condition $(\mathcal{DP})$, Theorem 8 entails good behaviour for $G(z)$ outside horizontal cones. Namely, for some θ_2, and for all linear cones $\mathcal{C}(\theta)$ with $\theta < \theta_2$, there exist $\delta < 1$ and $A > 0$ such that the exponential generating function $G(z)$ of g satisfies

$$z \notin \mathcal{C}(\theta) \implies (\forall w \in \Sigma^\star), \quad |G(p_w z)| \leq A \exp(\delta |p_w z|) \quad (13)$$

We now consider, for $\gamma < 1$, a cone $\widehat{\mathcal{C}}(\gamma)$ defined in Section 1.4, with γ large enough to ensure the inclusions $\widehat{\mathcal{C}}(\gamma) \subset \mathcal{C}(\theta_1)$ (with θ_1 relative to Assertion (I) for $P_r(z)$) and $\widehat{\mathcal{C}}(\gamma) \subset \mathcal{C}(\theta_2)$ (with θ_2 relative to Eqn (13) for $G(z)$). With (8) and (12), and $\alpha := \max(\delta, \gamma)$, one has

$$\text{for } z \notin \widehat{\mathcal{C}}(\gamma), \quad |G(p_w z) \exp(z - p_w z)| \leq A \exp\left[\delta p_w |z| + \Re(z)(1 - p_w)\right]$$
$$\leq A \exp[|z|(\delta p_w + \gamma(1 - p_w))] \leq A \exp(\alpha |z|).$$

We then transfer the bounds on

$$P_r(z) e^z = e^z \sum_{w \in \Sigma^\star} P_f(z p_w) = z^2 \sum_{w \in \Sigma^\star} p_w^2 \, G(p_w z) \, \exp(z - p_w z).$$

and obtain, with $B := A\Lambda(2)$, and for $|z|$ large enough

$$z \notin \widehat{\mathcal{C}}(\gamma) \implies |P_r(z) e^z| \leq B|z|^2 \exp(\alpha |z|) \leq C \exp(\alpha' |z|)$$

with $\alpha' \in \,]\alpha, 1[$ and a given constant C. Finally, Assertion (O) of Condition $(\mathcal{JS})$ holds for $P_r(z)$ and this ends the proof.

3 The Rice path

In the Rice path, we deal with the Poisson sequence $\Pi[f]$. We assume the following condition, denoted as Condition $\mathcal{RM}$ [Rice-Mellin], to hold on the sequence $\Pi[f]$

Condition $(\mathcal{RM})$: *There is an analytic lifting $\psi(s)$ for the sequence $\Pi[f]$ which is tame.*

Then the binomial recurrence (2) is transfered into a relation which expresses the term $f(n)$ as an integral along a vertical line which involves the analytic lifting $\psi(s)$. With tameness of ψ, we obtain the asymptotics of the sequence f.

3.1 The three steps of the Rice path

The Rice path performs three steps. It deals with a sequence f which satisfies the (VD) conditions, but we describe it in the stronger case when f is reduced. The complete proofs are in the Annex.

Step 1. It proves the existence of an analytical lifting ψ of the sequence $\Pi[f]$, on a halfplane $\Re s > c$ (for some c). It uses the (direct) Mellin transform and the Newton interpolation, without any other condition on the sequence f.

▶ **Proposition 9** (Nördlund-Rice). *The sequence $\Pi[f]$ associated with a reduced sequence f of degree $c < -1$ admits as an analytic lifting on $\Re s > c$ a function ψ, which is also an analytic extension of $P_f^*(-s)/\Gamma(-s)$ there.*

Step 2.   If moreover ψ is of *polynomial growth* "on the right", the binomial relation (2) is transfered into a Rice integral expression

▶ **Proposition 10.** *Assume that the analytic lifting ψ of $\Pi[f]$ is of polynomial growth on the halfplane $\Re s > c$, with $c < -1$. Then, for any $a \in]c, 0[$ and $n \geq n_0$, the sequence $f(n)$ admits an integral representation of the form*

$$f(n) = \sum_{k=0}^{n}(-1)^k \binom{n}{k}p(k) = \frac{1}{2i\pi}\int_{a-i\infty}^{a+i\infty} L_n(s) \cdot \psi(s)\, ds\,,$$

where the Rice kernel $L_n(s) := \dfrac{(-1)^{n+1}\, n!}{s(s-1)(s-2)\ldots(s-n)} = \dfrac{\Gamma(n+1)\Gamma(-s)}{\Gamma(n+1-s)}$

involves the Beta Function B with the equality $L_n(s) = B(n+1, -s)$.

This integral representation is valid for any abscissa a which belongs to the interval $]c, 0[$. We now shift the vertical line $\Re s = a$ to the left, and thus use *tameness conditions* on ψ at $s = c$, as defined in Section 1.4.

Step 3.   If moreover ψ is *tame* "on the left", the integral is *shifted* to the left; this provides the asymptotics of of the sequence f.

▶ **Proposition 11.** *Consider a reduced sequence $f : n \mapsto f(n)$ with $\deg(f) = c < -1$. If the lifting ψ of $\Pi[f]$ is tame at $s = c$ with a region $\mathcal{R}$ of tameness and a left frontier $\mathcal{F}$, then*

$$f(n) = -\left[\sum_{k|s_k\in\mathcal{R}} \mathrm{Res}\left[L_n(s)\cdot\psi(s); s = s_k\right] + \frac{1}{2i\pi}\int_{\mathcal{F}} L_n(s)\cdot\psi(s)\, ds\right],$$

where the sum is over the poles s_k of ψ inside $\mathcal{R}$.

3.2   The main question about the Rice method: Tameness of ψ

The main results are due to Nörlund [18, 17], then to Rice who popularized them. Later on, with the paper [10], Flajolet and Sedgewick brought this methodology into the AofA domain. The Rice-Mellin method is also well described in [6]. There exist many analyses of various data structures or algorithms that are based on the application of the method: tries ([9, 8, 3, 1]), digital trees ([9, 12]), or fine complexity analyses of sorting or searching algorithms on sources ([4, 2]).

The situation for applying the Rice method is not the same as in Section 2: previously, with Condition $(\mathcal{DP})$, we know exactly when the Depoissonisation method may be applied. This is not the case for the Rice method. Even though the literature well explains how to use this method in various cases of interest, the following question is never asked: *What are sufficient conditions on the sequence f that would entail tameness of ψ?*

As $\psi(s)$ is closely related to the Mellin transform $P_f^*(-s)$, meromorphy is often easy to prove, and the poles often easy to find. In many natural contexts, the polynomial growth and the tameness of the Mellin transform $P_f^*(s)$ generally hold, and are often used in the Depoissonisation approach [see Section 2.2]. But the main difference between the Rice method and the Depoissonisation method is the division by $\Gamma(s)$.

Sometimes, and this is often the case in classical tries problems, the factor $\Gamma(s)$ already appears in $P_f^*(s)$, and $\psi(s)$ has an explicit form, from which its polynomial growth may be easily proven. For instance, for the toll $f = f_1$ associated to the path length, then

$P_f^*(s) = \Gamma(s + 1)$ and $\psi(-s)$ is explicit, and equal to s. This is also the case for polynomial tolls f of the form $f = T^{-m}[f_1]$ with $m \geq 1$.

But what about other sequences, for instance the sorting toll $f(k) = k \log k$, and more generally, the basic sequence $f(k) = k^d \log^b k$ (with $d \in \mathbb{R}$ and an integer $b \geq 1$)? In this case, the following expansion holds for $P_f^*(-s)$, that involves the b-th derivative of the Riemann ζ function and generalizes (10),

$$P_f^*(-s) = (-1)^b \zeta^{(b)}(s - (d - 1)) + H_1(s), \tag{14}$$

where $H_1(s)$ is analytic on $\Re s > d - 1$. Then principles of Depoissonisation apply in this case, due to good properties of the Riemann function. Now, in the Rice method, the function ψ satisfies $\psi(s) = P_f^*(-s)/\Gamma(-s)$, and the function $1/\Gamma(-s)$, even though it is analytic on the half-plane $\Re(s) \geq 0$, is of exponential growth along vertical lines. The Stirling formula indeed entails the estimate

$$\frac{1}{\Gamma(x + iy)} \sim \frac{1}{\sqrt{2\pi}} e^{\pi|y|/2} |y|^{1/2-x}, \qquad \text{as } |y| \to \infty.$$

It is thus not clear whether $\psi(s)$ attached to the sorting toll is tame at $s = 1$. Then, the Rice method seems to have a more restrictive use than the Depoissonisation method. As we wish to compare the power of the two paths [Depoissonisation path and Rice path], we ask the two (complementary) questions: *Is the Rice path only useful for very specific tolls, where the Mellin transform $P_f^*(s)$ of the Poisson transform $P_f(s)$ factorizes with the factor $\Gamma(s)$, or is it useful for more general tolls?*

This leads us to study sufficient conditions under which the analytic lifting ψ may be proven to be tame. We now propose to use the (inverse) Laplace transform. With this tool, we prove the tameness of ψ for basic sequences (see Theorem 13).

4 The Rice–Laplace approach.

As in the previous Section, we deal with the Poisson sequence $\Pi[F]$. Our main result proves the tameness of the analytic continuation $\Psi(s)$ of $\Pi[F]$, when F is a basic sequence.

▶ **Definition 12.** Consider a pair (d, b) with a real d and an *integer* $b \geq 0$.
 (i) A sequence F is *basic* with pair (d, b) if it writes as $F_{b,d}(k) = k^d \log^b k$ for any $k \geq 2$.
 (ii) A sequence F is *extended basic* with pair (d, b) if it has an analytic extension Φ on some halplane $\Re s > a$, of the form $\Phi(z) = F_{d,b}(z) W(1/z)$, with W analytic at 0, and $W(0) = 1$.

▶ **Remark.** In the proof, an integral exponent $b \geq 0$ is needed to relate $F_{d,b}$ to a b-th derivative.

Our main result is as follows:

▶ **Theorem 13.** *Consider a basic extended sequence F with pair (d, b). Then, for some $\sigma_0 > 0$, the analytic continuation $\Psi(s)$ of the $\Pi[F]$ sequence is of polynomial growth on any halfplane $\Re s \geq a > d$. Moreover, it writes in terms of the integer $\ell := \sigma(F)$ defined in (4),*

$$\Psi(s) = \left[s(s - 1) \ldots (s - \ell + 1) \sum_{m=0}^{b} a_m \Gamma_\ell^{(m)}(s - d) \right] + B(s) \tag{15}$$

on the halfplane $\Re s > d - \sigma_0$, for $\sigma_0 \in]0,1[$. Here, $B(s)$ is of polynomial growth, and $\Gamma_\ell^{(m)}$ is the m-th derivative of the twisted Γ function that is defined for $\Re s > 0$, integers $m \geq 0$ as

$$\Gamma_\ell^{(m)}(s) := \int_0^\infty e^{-\ell u} u^{s-1} \log^m u \, du \,. \tag{16}$$

The coefficients a_m involve the derivatives of order $k \leq b$ of $s \mapsto 1/\Gamma(s)$ at $s = \ell - d$.

Remarks.
(a) With Lemma 3, the twisted function Γ_ℓ and its derivatives are of exponential decrease along the vertical lines. This entails the tameness of Ψ at $s = d$.
(b) We already know the singular part of Ψ at $s = d$ which is given by the expansion (14), and the singular expansion given in (15) is just an alternative and complicate expression. What is new is the tameness, not the singular expansion.

4.1 Plan of the proof

We first recall the principles of Section 1.5 : with a initial sequence F, we associate its canonical sequence $f := \rho(F)$, and now deal with this new sequence f; it is easy to return (later) to the initial sequence F with Lemma 5. If the initial F admits an analytic lifting of polynomial growth on $\Re s > 0$, then the sequence $f = \rho(F)$ is reduced and admits an analytic lifting φ on $\Re s > -1$ that satisfies $\varphi(s) = O(|s+1|^c)$ there, with $c < -1$.

The first step performed in Section 4.2 deals with any reduced sequence f which admits an analytic lifting φ on $\Re s > -1$ that satisfies $\varphi(s) = O(|s+1|^c)$ there, with $c < -1$. With a strong use of the involutive character of Π, we first exhibit a new expression of the analytical extension ψ of $\Pi[f]$ which deals with the inverse Laplace transform $\widehat{\varphi}$ of the extension φ of the sequence f. The proof is then applied to the canonical sequence $\rho[F]$ of the initial sequence F.

Then, the sequel of the present section focuses on (extended) basic sequences $F_{d,b}$. Here, in this Section, we only deal with exact basic sequences. The extension to *extended basic* sequences will be done in the Annex. We first obtain in Section 4.3 a precise expression of the inverse Laplace transform $\widehat{\varphi}$ of extension φ of the canonical sequence $f_{d,b} := \rho(F_{d,b})$, that is transfered into a precise estimate of $\Pi[f_{d,b}]$. This leads to the proof of Theorem 13.

4.2 A new general expression for ψ with the inverse Laplace transform

This section is of independent interest and provides a new expression of the extension of the sequence $\Pi[f]$ in the case when f is reduced.

▶ **Proposition 14.** *Consider a sequence f which admits an analytic lifting φ on $\Re s > -1$, with the estimate $\varphi(s) = O(|s+1|^c)$ with $c < -1$. Then:*
(i) *The function φ admits an inverse Laplace transform $\widehat{\varphi}$ whose restriction to the real line $[0, +\infty[$ is written as the Bromwich integral for $a \in]-1, 0[$,*

$$\widehat{\varphi}(u) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) \, e^{su} ds \,, \qquad \text{and satisfies} \quad |\widehat{\varphi}(u)| \leq K e^{au} \,.$$

(ii) *There is an analytical lifting ψ of the sequence $\Pi[f]$ that admits an integral form*

$$\psi(s) = \mathcal{I}_s[\widehat{\varphi}] \qquad \text{with} \quad \mathcal{I}_s[h] := \int_0^\infty h(u)(1 - e^{-u})^s du \qquad \text{for } \Re s > -1 \,. \tag{17}$$

Sketch of the proof. The complete proof is in the Annex.

(i) In a general context, where the analytic lifting $\varphi(s)$ is only defined on $\Re s > 0$, the Bromwich integral is written as an integral on a vertical line $\Re s = a$ with $a > 0$. Here, the hypotheses on φ are stronger and the Bromwich integral may be shifted to the left with $a \in]-1, 0[$. Moreover, the Bromwich integral is absolutely convergent, and the exponential bound on $\widehat{\varphi}(u)$ holds.

(ii) As φ is polynomial growth, we use the involutive character of Π and apply Proposition 10 to the pair $(p := \Pi[f], f = \Pi^2[f])$. It transfers the binomial expression of $\Pi[f]$ in terms of $\Pi^2[f] = f$ into a Rice integral, with $a \in]-1, 0[$,

$$p(n) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) L_n(s) ds, \qquad L_n(s) = \frac{\Gamma(n+1)\Gamma(-s)}{\Gamma(n+1-s)} = B(n+1, -s).$$

We now use the integral expression of the Beta function, and "exchange" the two integrals. ◄

The integral representation (17) leads us to introduce the two functions, defined on $[0, +\infty]$,

$$N_s(u) := \left(\frac{1 - e^{-u}}{u}\right)^s, \quad M_s(u) := \left[\left(\frac{1 - e^{-u}}{u}\right)^s - 1\right], \tag{18}$$

that satisfy the two estimates, with $\sigma := \Re s$,

$$N_s(u) = \Theta(1), \quad (u \to 0), \qquad N_s(u) = O(u^{-\sigma}) \quad (u \to \infty),$$

$$M_s(u) = \Theta(u) \quad (u \to 0), \qquad M_s(u) = O(u^{-\sigma}) \quad (u \to \infty, \sigma > 0).$$

Then, for "good" functions h, the integral $\mathcal{I}_s[h]$ may be compared to the Mellin transform $h^\star(s+1)$. We now apply this idea to the particular cases where the behaviour of $h = \widehat{\varphi}$ is well-known. Then, there are two steps which deal with a reduced sequence f, and aim at studying the tameness of the analytical extension ψ of $\Pi[f]$:

(a) transfer properties of φ into properties of its inverse Laplace transform $\widehat{\varphi}$;

(b) with properties of $\widehat{\varphi}$, study the tameness of ψ, via the representation (17).

We now perform these two steps. The (inverse) Laplace transform is not well studied, and we do not know a general transfer result of type (a). This is why we only perform the two steps for canonical sequences related to basic sequences. The proofs of the following section are in the Annex.

4.3 Dealing with basic sequences

Step (a). We obtain first an expression for φ, then an expression for $\widehat{\varphi}$.

▶ **Proposition 15.**

(i) *The sequence $f_{b,d}$ is extended in a function φ defined on $\Re s > -1$*

$$\varphi(s) = (s + \ell)^{d-\ell} \log^b(s + \ell) \, U\left(\frac{1}{s + \ell}\right) ; \quad \ell := \sigma(d)$$

Here U satisfies $U(u) = 1$ for $d < 0$. For $d \geq 0$, it is defined as

$$U(u) = (1 - u)^{-1}(1 - 2u)^{-1} \ldots (1 - (\ell - 1)u)^{-1} \quad (with \ \ell = 2 + \lfloor d \rfloor)$$

For $d \geq 0$, the coefficient $a_j := [u^j]U(u)$ satisfies $a_j = \Theta_d(\ell - 1)^j$.

(ii) *The inverse Laplace transform $\widehat{\varphi}(u)$ is a linear combination of functions, for $m \in [0..b]$,*

$$e^{-\ell u} \, u^{-c-1} \log^m u \, \left[1 + V^{\langle m \rangle}(u)\right], \qquad with \quad |V^{\langle m \rangle}(u)| \leq A_{(d,b)} \, u \, e^{(\ell-1)u}.$$

Step (b). The previous expression of $\widehat{\varphi}$ together with the representation (17) entail a decomposition for ψ. Using the estimates of functions defined in (18), each term is compared to the twisted version of the Γ function and its m-th derivative, defined in (16). This provides the estimate for the initial function $\Psi := \Pi[f]$.

▶ **Proposition 16.**

(i) *The extension ψ of the sequence $\Pi[f_{b,d}]$ is a linear combination of functions, for $m \in [0..b]$, each term being the sum of a main term $A^{\langle m \rangle}(s)$ and a remainder term $O(B^{\langle m \rangle}(s))$, with $c := d - \sigma(d) < -1$ and*

$$A^{\langle m \rangle}(s) := \mathcal{I}_s \left[e^{-\ell u} \, u^{-c-1} \, \log^m u \right], \quad B^{\langle m \rangle}(s) := \mathcal{I}_s \left[e^{-u} \, u^{-c} \, \log^m u \right] .$$

(ii) *For $\Re s \geq 0$, the two functions $A^{\langle m \rangle}(s)$ and $B^{\langle m \rangle}(s)$ are bounded on the halfplane $\Re s \geq 0$. For any integer $m \geq 0$ and any integer $\ell \geq 1$, the two functions*

$$A^{\langle m \rangle}(s) - \Gamma_\ell^{(m)}(s - c), \quad B^{\langle m \rangle}(s)$$

are analytic and of bounded growth on the vertical strip $\Re s > c - \sigma_0$, with $\sigma_0 \in]0, 1[$.

4.4 A second proof for Theorem 6.

Within the framework of Section 1.7, we deal with the sorting toll $F_{1,1}$. The singular part of the extension $\Psi := \Pi[F_{1,1}]$ at $s = 1$ is obtained in (14). The tameness of Ψ at $s = 1$ is proven in Theorem 13. Together with the tameness of the Dirichet series Λ at $s = 1$, this entails the tameness of $\Pi[r]$ and gives a three-lines proof of Theorem 6. We prefer this proof!

5 Final comparison between the two paths.

The Annex describes a formal comparison between the two paths. From analytical properties, the Rice-Laplace path remains of more restrictive use than the Depoissonisation path:

(a) We need the analytic extension φ of f to hold on a halfplane, whereas the Depoissonisation path needs it only on a horizontal cone.

(b) The analytic extension φ of f involves a precise expansion in terms of an analytic series W, whereas the Depoissonisation path only needs a rough asymptotic estimate of φ.

(c) The exponent of the log term must be an *integer b*, whereas the Depoissonisation path deals with any real exponent. The need of an integer exponent b is related to the interpretation in terms of b-derivatives, and this is a restriction which is also inherent in the method used by Flajolet in [5] in a similar context.

These are strong restrictions... However, most of the Depoissonisation analyses (at least for mean values) deal with extended basic sequences, where the Rice-Laplace path may be used. We let the final conclusion to the reader !

─── **References** ───

1   Eda Cesaratto and Brigitte Vallée. Gaussian distribution of trie depth for strongly tame sources. *Combinatorics, Probability and Computing*, 24, issue 01:54–103, 2014.

2   Julien Clément, Jim Fill, Thu Hien Nguyen Thi, and Brigitte Vallée. Towards a realistic analysis of the Quickselect algorithm. *Theory of Computing Systems*, 58(4):528–578, 2016.

3   Julien Clément, Philippe Flajolet, and Brigitte Vallée. Dynamical sources in information theory: a general analysis of trie structures. *Algorithmica*, 29, 1-2:307–369, 2001.

4   Julien Clément, Thu Hien Nguyen Thi, and Brigitte Vallée. Towards a realistic analysis of some popular sorting algorithms. *Combinatorics, Probability and Computing*, 24 issue 01:104–144, 2014.

5   Philippe Flajolet. Singularity analysis and asymptotics of Bernoulli sums. *Theoret. Comput. Sci.*, 215 1-2:371–381, 1999.

6   Philippe Flajolet. A journey to the outskirts of Quickbits. *Personal Communication*, 2010.

7   Philippe Flajolet, Xavier Gourdon, and Philippe Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoretical Computer Science*, 144 1-2:3–58, 1995.

8   Philippe Flajolet, Mireille Régnier, and Wojtek Szpankowski. Some uses of the Mellin integral transform in the analysis of algorithms. *In Combinatorial Algorithms on Words*, 12:241–254, 1995.

9   Philippe Flajolet and Robert Sedgewick. Digital search trees revisited. *SIAM J. Comput.*, 15 (3):48–67, 1986.

10   Philippe Flajolet and Robert Sedgewick. Mellin transforms and asymptotics: finite differences and Rice's integrals. *Theoretical Computer Science*, 144 (1-2):101–124, 1995.

11   Walter Kurt Hayman. A generalisation of Stirling's formula. *J. Reine Angew. Math.*, 196:67–95, 1956.

12   Kanal Hun and Brigitte Vallée. Typical depth of a digital search tree built on a general source. In *Proceedings of ANALCO*, 2014.

13   Hsien-Kuei Hwang, Michael Fuchs, and Vytas Zacharovas. Asymptotic variance of random symmetric digital search trees. *Discrete Mathematics and Theoretical Computer Science*, 12(2):103–166, 2010.

14   Philippe Jacquet. Trie structure for graphs sequences. In *Proceedings of AofA 2014*, volume 950 of *DMTCS Proceedings*, pages 181–192, 2014.

15   Philippe Jacquet and Wojtek Szpankowski. An analytic approach to the asymptotic variance of trie statistics and related structures. *Theoretical Computer Science*, 201 1-2:1–62, 1998.

16   Philippe Jacquet and Wojtek Szpankowski. Entropy computations via Analytic depoissonization. *IEEE Transactions on Information Theory*, 45 (4):1072–1081, 1999.

17   Niels Erik Nörlund. *Vorlesungen über Differenzenrechnung*. Chelsea Publishing Company, New-York, 1954.

18   Nils Erik Nörlund. Leçons sur les équations linéaires aux différences finies. In *Collection de monographies sur la théorie des fonctions*. Gauthier-Villars, Paris, 1929.

19   Brigitte Vallée. Dynamical sources in information theory: Fundamental intervals and word prefixes. *Algorithmica*, 29 (1/2):262–306, 2001.

A   Annex

A.1   More on tameness

▶ **Definition 17** (Tameness). A function ϖ analytic and of polynomial growth on $\Re s > c$ is tame at $s = c$ if one of the three following properties holds:

(a) [S-shape] (shorthand for Strip shape) there exists a vertical strip $\Re(s) > c - \delta$ for some $\delta > 0$ where $\varpi(s)$ is meromorphic, has a sole pole (of order $b + 1 \geq 1$) at $s = c$ and is of polynomial growth as $|\Im s| \to +\infty$.

(b) [H-shape] (shorthand for Hyperbolic shape) there exists an hyperbolic region $\mathcal{R}$, defined as, for some $A, B, \rho > 0$

$$\mathcal{R} := \{s = \sigma + it; \ \ |t| \geq B, \ \ \sigma > c - \frac{A}{|t|^\rho}\} \bigcup \{s = \sigma + it; \ \ \sigma > c - \frac{A}{B^\rho}, |t| \leq B\},$$

where $\varpi(s)$ is meromorphic, with a sole pole (of order $b+1$) at $s = c$ and is of polynomial growth in $\mathcal{R}$ as $|\Im s| \to +\infty$.

(c) [*P*-shape] (shorthand for Periodic shape) there exists a vertical strip $\Re(s) > c - \delta$ for some $\delta > 0$ where $\varpi(s)$ is meromorphic, has only a pole (of order $b + 1 \geq 1$) at $s = c$ and a family (s_k) (for $k \in \mathbb{Z} \setminus \{0\}$) of simple poles at points $s_k = c + 2ki\pi t$ with $t \neq 0$, and is of polynomial growth as $|\Im s| \to +\infty^4$.

A.2   Proofs of the Rice path

Proof of Proposition 9. In the strip $\mathcal{S}(0, -c)$, the Mellin transform $P_f^*(s)$ of $P_f(z)$ exists and satisfies

$$\frac{P_f^*(s)}{\Gamma(s)} = \frac{1}{\Gamma(s)} \sum_{k \geq 0} \frac{f(k)}{k!} \int_0^\infty e^{-z} z^k z^{s-1} dz = \sum_{k \geq 0} \frac{f(k)}{k!} \frac{\Gamma(k+s)}{\Gamma(s)}$$

where the exchange of integration and summation is justified by the estimates given in (9). On the strip $\mathcal{S}(c, 0)$, the series is a Newton interpolation series,

$$\psi(s) := \frac{P_f^*(-s)}{\Gamma(-s)} = \sum_{k \geq 0} (-1)^k \frac{f(k)}{k!} s(s-1) \ldots (s-k+1) \,. \tag{19}$$

Such series converge in right halfplanes and thus the previous series converges on $\Re s > c$. Moreover, Relation (19), together the binomial relation (2), entails the equality

$$\psi(n) = \sum_{k=0}^n (-1)^k \frac{f(k)}{k!} n(n-1) \ldots (n-k+1) = \sum_{k=0}^n (-1)^k \binom{n}{k} f(k) = \Pi[f](n) \,.$$

This proves that ψ provides an analytic lifting of the sequence $\Pi[f]$ on $\Re s > c$ which is also an analytic extension of $P_f^*(-s)/\Gamma(-s)$. ◀

Proof of Proposition 10. (Sketch) Use the Residue Theorem and the polynomial growth of $\psi(s)$ "on the right". First, we consider the rectangle $\mathcal{A}_M$ delimited by the contour τ_M defined by the two vertical lines $\Re s = a$ and $\Re s = n + M$ and the two horizontal lines $\Im s = \pm M$. If the contour τ_M is taken counterclockwise, then the Residue Theorem applies,

$$\frac{1}{2i\pi} \int_{\tau_M} L_n(s) \cdot \psi(s) \, ds = \sum_{k=0}^n \mathrm{Res}[L_n(s) \cdot \psi(s); s = k] = -\sum_{k=0}^n (-1)^k \binom{n}{k} \Pi[f](k) = -f(n).$$

Next, the integral on the curve τ_M is the sum of four integrals. Let now M tend to ∞. The integrals on the right, top and bottom lines tend to 0, due to the polynomial growth of the function $\psi(s)$. The integral on the left becomes

$$-\int_{a-i\infty}^{a+i\infty} L_n(s) \cdot \psi(s) \, ds,$$

and we have proven the result. For details on the proof, we may refer to papers [18, 17, 10]. ◀

---

4   More precisely, this means that $\varpi(s)$ is of polynomial growth on a family of horizontal lines $t = t_k$ with $t_k \to \infty$, and on vertical lines $\Re(s) = \sigma_0 - \delta'$ with some $\delta' < \delta$.

Proof of Proposition 11. (Sketch) The proof is similar to the previous proof. With the tameness of $\psi(s)$ at $s = c$ with a tameness domain $\mathcal{R}$, we now deal with the residues of ψ; we consider the domains

$$\widehat{\mathcal{R}} := \mathcal{R} \cap \{\Re s < a\} \text{ and } \mathcal{R}_M := \widehat{\mathcal{R}} \cap \{|\Im s| \leq M\},$$

and denote $\mathcal{L}_M$ the curve (taken counterclockwise) which borders the region $\mathcal{R}_M$. As $\psi(s)$ is meromorphic in $\mathcal{R}_M$ and $L_n(s)$ analytic there, we apply the Residue Theorem to the function $L_n(s) \cdot \psi(s)$ inside $\mathcal{R}_M$, and obtain

$$\frac{1}{2i\pi} \int_{\mathcal{L}_M} L_n(s) \cdot \psi(s) \, ds = \sum_{s_k \in \mathcal{R}_M} \text{Res}\left[L_n(s) \cdot \psi(s); s = s_k\right]$$

where the sum is taken over the poles s_k of $\psi(s)$ inside $\mathcal{R}$. Now, when $M \to \infty$, the integrals on the two horizontal segments tend to 0, since $\psi(s)$ is of polynomial growth, and

$$\lim_{M \to \infty} \int_{\mathcal{R}_M} L_n(s) \cdot \psi(s) \, ds = \int_{a-i\infty}^{a+i\infty} L_n(s) \cdot \psi(s) \, ds - \int_{\mathcal{F}} L_n(s) \cdot \psi(s) ds$$
$$= 2i\pi \sum_{s_k \in \mathcal{R}} \text{Res}\left[L_n(s) \cdot \psi(s); s = s_k\right],$$

where the sum is taken over the poles s_k of $\psi(s)$ inside the domain $\mathcal{R}$. ◄

A.3 Proofs of the Rice-Laplace path

Proof of Proposition 14.

 (i) In a general context, where the analytic lifting $\varphi(s)$ is only defined on $\Re s > 0$, the Bromwich integral is written as

$$\widehat{\varphi}(u) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) \, e^{su} ds, \qquad (\text{with } a > 0).$$

 Here, the hypotheses on φ are stronger: we can shift the integral on the left and choose $a \in]-1, 0[$. Moreover, the Bromwich integral is ansolutely convergent, and the exponential bound on $\widehat{\varphi}(u)$ holds.

 (ii) We use the involutive character of Π and apply Proposition 10 to the pair $(p := \Pi[f], f = \Pi^2[f])$. In the classical Rice path, it is applied to the pair $(f, \Pi[f])$, when $\Pi[f]$ is of polynomial growth, and it transfers the binomial expression of f in terms of $\Pi[f]$ into an integral expression. Here, due to the polynomial growth of $f = \Pi^2[f]$ on $\Re s > -1$, it transfers the binomial expression of $\Pi[f]$ in terms of $\Pi^2[f] = f$ into a Rice integral, with $a \in]-1, 0[$,

$$p(n) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) L_n(s) ds, \qquad L_n(s) = \frac{\Gamma(n+1)\Gamma(-s)}{\Gamma(n+1-s)}.$$

 We now deal with the Beta function

$$B(t+1, -s) = \frac{\Gamma(t+1)\Gamma(-s)}{\Gamma(t+1-s)},$$

 that is well defined for $\Re t > -1$ and $\Re s < 0$, and admits an integral expression

$$B(t+1, -s) = \int_0^\infty e^{su} (1 - e^{-u})^t du, \qquad (\text{for } \Re t > -1, \Re s < 0).$$

Together with the equality $L_n(s) = B(n+1, -s)$, this entails an analytic extension ψ of the sequence $\Pi[f]$ on the halfplane $\Re t > -1$,

$$\psi(t) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) B(t+1, -s) ds, \quad (b < 0)$$

with an integral expression,

$$\psi(t) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) \left[\int_0^\infty e^{su} (1 - e^{-u})^t du \right] ds.$$

With properties of φ, it is possible to exchange the integrals: then, the equality holds

$$\psi(t) = \int_0^\infty (1 - e^{-u})^t \left[\frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) e^{su} ds \right] du,$$

and the second integral is the inverse Laplace transform $\widehat{\varphi}$ of φ. This ends the proof. ◀

Proof of Proposition 15. For (i), letting $\ell := \sigma(d)$ with σ defined in (4), the canonical sequence f associated with F can be extended to a function φ defined on $]-1, +\infty[$ as

$$\varphi(x) = \log^b(x+\ell) \frac{(x+\ell)^d}{(x+1)(x+2)\dots(x+\ell)} = \log^b(x+\ell)(x+\ell)^{d-\ell} U\left(\frac{1}{x+\ell}\right),$$

and involves a function U defined as $U(u) = 1$ for $d < 0$ and, for $d \geq 0$ as

$$U(u) = (1-u)^{-1}(1-2u)^{-1}\dots(1-(\ell-1)u)^{-1} \quad (\text{with } \ell = 2 + \lfloor d \rfloor). \tag{20}$$

Then, for $d \geq 0$, the coefficient $a_j := [u^i]U(u)$ satisfies $a_j = \Theta(\ell-1)^j$.

For (ii), there are three main steps, according to the type of the basic sequence.

Step 1. We begin with the particular case when $\varphi(s)$ is of the form $\varphi(s) = (s+\ell)^c$ (with $c < -1$). Its inverse Laplace transform $\widehat{\varphi}$ is then

$$\widehat{\varphi}(u) = \frac{1}{\Gamma(-c)} e^{-\ell u} u^{-c-1}.$$

Step 2. We now consider a function (without logarithmic factor) of the form

$$\varphi(s) = \varphi_c(s) = (s+\ell)^c U\left(\frac{1}{s+\ell}\right) = \sum_{j \geq 0} a_j (s+\ell)^{c-j}. \tag{21}$$

Then φ is a linear combination of functions of Step 1 and the inverse Laplace transform $\widehat{\varphi}$ of φ is written as

$$\widehat{\varphi}_c(u) = e^{-\ell u} \frac{u^{-c-1}}{\Gamma(-c)} [1 + V_c(u)], \quad \text{with} \quad V_c(u) := \sum_{j \geq 1} a_j u^j G_j(c), \tag{22}$$

where the function G_j is the rational fraction which associates with c the ratio

$$G_j(c) := \frac{\Gamma(-c)}{\Gamma(j-c)} = \frac{1}{-c(1-c)\dots(j-1-c)}. \tag{23}$$

As $c < -1$, the inequality $G_j(c) \leq (1/j!)$ holds and this entails the inequality $|V_c(u)| \leq A u e^{(\ell-1)u}$, where the constant A only depends on d.

Step 3. We add finally a logarithmic factor and consider a function of the form

$$\varphi(s) = (s + \ell)^c \log^b(s + \ell) \, U\left(\frac{1}{s + \ell}\right) \tag{24}$$

which is written as a b-th derivative. Indeed, the equality holds

$$U\left(\frac{1}{s + \ell}\right)(s + \ell)^c \log^b(s + \ell) = \frac{\partial^b}{\partial t^b}\left[(s + \ell)^{c+t} U\left(\frac{1}{s + \ell}\right)\right]\bigg|_{t=0},$$

and we can take the derivative "under the Laplace integral": we then deduce that the inverse Laplace transform $\widehat{\varphi}$ of the function φ defined in (24) is equal to

$$\frac{\partial^b}{\partial t^b}\widehat{\varphi}_{c+t}(u)\bigg|_{t=0} = e^{-\ell u}\frac{\partial^b}{\partial c^b}\left[\frac{u^{-c-1}}{\Gamma(-c)}(1 + V_c(u))\right].$$

The coefficient of u^j in the k-th derivative of $c \mapsto V_c(u)$ involves the k-th derivative of the function $c \mapsto G_j(c)$, defined in (23) which satisfies the inequality

$$|G_j^{(k)}(c)| \le A_k \log^k(j + c)\, G_j(c) \quad \text{for some constant } A_k.$$

Then, the inequality holds,

$$\left|\frac{\partial^k}{\partial c^k}V_c(u)\right| \le A_{(d,b)}\, u\, e^{(\ell-1)u},$$

and involves a constant $A_{(d,b)}$ which depends on the pair (d, b). On the other hand, the following m-th derivative is a linear combination of the form

$$\frac{\partial^m}{\partial c^m}\left[\frac{u^{-c-1}}{\Gamma(-c)}\right] = u^{-c-1}\left[(-1)^m\sum_{a=0}^{m}\binom{m}{a}(\log^a u)\, H^{(m-a)}(c)\right],$$

where H is the function defined as $H(c) = 1/\Gamma(-c)$. This ends the proof. ◄

Proof of Proposition 16.

(a) is clear : For $\Re s \ge 0$, the result follows from the inequalities $(1 - e^{-u})^\sigma \le 1$, $c < -1$, together with the integrability of the function $u \mapsto e^{-\ell u}u^{-c-1}\log^m u$ on the interval $[0, +\infty]$.

(b) The difference $A^{\langle m\rangle}(s) - \Gamma_\ell^{(m)}(s - c)$ is expressed with M_s, whereas $B^{\langle m\rangle}(s)$ is expressed with N_s, both defined in (18). Together with their estimates, this leads to the following bounds, for any $\rho > 0$,

$$A^{\langle m\rangle}(s) - \Gamma_\ell^{(m)}(s - c) = O_\rho\Big(\Gamma_\ell(\sigma - c + 1 - \rho)\Big), \quad B^{\langle m\rangle}(s) = O_\rho\Big(\Gamma(\sigma - c + 1 - \rho)\Big)$$

and also to the analyticity of the functions of interest on the vertical strip $\Re s > c - \sigma_0$, with $\sigma_0 \in\,]0, 1[$. ◄

Extension to extended basic sequences. It is easy to extend the proof of Theorem 13 to this more general case: We denote by r the convergence radius of W, and we thus choose a shift T^ℓ with an integer which now satisfies

$$\ell \ge \max\Big[2 + \lfloor d\rfloor, a + 1, (1/r) + 1\Big],$$

and deal with the sequence $f := T^\ell[F]$. We replace the previous series U defined in Proposition 15 by the series $U \cdot W$ which has now a convergence radius $\tilde{r} := \min(r, 1/(\ell - 1))$ for which the bound $1/\tilde{r} < \ell$ holds. We choose $\widehat{r} \in\,]1/\tilde{r}, \ell[$, and the new series V_c defined in (22) satisfies $|V_c(u)| \le A u\, e^{\widehat{r}u}$ and indeed gives rise to a remainder term. ◄

A.4   Description of a formal comparison between the two paths

As it is observed in the paper [13], there are formal manipulations which allow us to compare the two paths.

In the Depoissonisation path, the asymptotics of $f(n)$ is manipulated in two steps: first use the Cauchy integral formula

$$f(n) = \frac{n!}{2i\pi} \int_{|z|=r} P_f(z) \, e^z \frac{1}{z^{n+1}} dz \,. \tag{25}$$

then derive asymptotics of $P_f(z)$ for large $|z|$ by the inverse Mellin integral

$$P_f(z) = \frac{1}{2i\pi} \int_{\uparrow} P_f^*(s) z^{-s} ds = \frac{1}{2i\pi} \int_{\uparrow} P_f^*(-s) z^s ds \,, \tag{26}$$

where the integration path is some vertical line. This two-stage Mellin-Cauchy formula is the beginning point of the Depoissonization path.

We now compare the formula obtained by this two stage approach with the Nórdlund-Rice formula. First remark that, as the function $P_f(z) e^z$ is entire, we can replace the contour $\{|z| = r\}$ in (25) by a Hankel contour $\mathcal{H}$ starting at $-\infty$ in the upper halfplane, winding clockwise around the origin and proceeding towards $-\infty$ in the lower halfplane. Then (25) becomes

$$f(n) = \frac{n!}{2i\pi} \int_{\mathcal{H}} P_f(z) \, e^z \frac{1}{z^{n+1}} dz \tag{27}$$

Now, if we *formally substitute* (26) into (27), *interchange* the order of integration and use the equality

$$\frac{1}{\Gamma(n+1-s)} = \frac{1}{2i\pi} \int_{\mathcal{H}} e^z \frac{z^s}{z^{n+1}} dz \,,$$

we obtain the representation

$$f(n) = \frac{n!}{2i\pi} \int_{\uparrow} P_f^*(-s) \frac{1}{\Gamma(n+1-s)} ds \,, \tag{28}$$

and we recognize in (28) the Rice integral

$$f(n) = \frac{n!}{2i\pi} \int_{\uparrow} \frac{P_f^*(-s)}{\Gamma(-s)} \frac{\Gamma(-s)}{\Gamma(n+1-s)} ds = \frac{1}{2i\pi} \int_{\uparrow} \psi(s) \frac{(-1)^{n+1} \, n!}{s(s-1)\dots(s-n)} ds \,.$$

This exhibits a formal comparison between the two paths. However, this comparison is *only formal* because the previous manipulations may be meaningless due to the divergence of the integrals.

Average Cost of QuickXsort with Pivot Sampling

Sebastian Wild

University of Waterloo, Canada
wild@uwaterloo.ca
 https://orcid.org/0000-0002-6061-9177

───── **Abstract** ─────

QuickXsort is a strategy to combine Quicksort with another sorting method X so that the result
has essentially the same comparison cost as X in isolation, but sorts in place even when X
requires a linear-size buffer. We solve the recurrence for QuickXsort precisely up to the linear
term including the optimization to choose pivots from a sample of k elements. This allows to
immediately obtain overall average costs using only the average costs of sorting method X (as if
run in isolation). We thereby extend and greatly simplify the analysis of QuickHeapsort and
QuickMergesort with practically efficient pivot selection, and give the first *tight* upper bounds
including the linear term for such methods.

1 Introduction

In QuickXsort [5], we use the recursive scheme of ordinary Quicksort, but instead of doing
two recursive calls after partitioning, we first sort one of the segments by some other sorting
method X. Only the second segment is recursively sorted by QuickXsort. The key insight is
that X can use the second segment as a temporary buffer for elements. By that, QuickXsort
is sorting in-place (using $O(1)$ words of extra space) even when X itself is not.

Not every method makes a suitable 'X'; it must use the buffer in a swap-like fashion: After
X has sorted its segment, the elements originally stored in our buffer must still be intact,
i.e., they must still be stored in the buffer, albeit in a different order. Two possible examples
are Mergesort (see Section 6 for details) and a comparison-efficient Heapsort variant [1] with
an output buffer. With QuickXsort we can make those methods sort in-place while retaining
their comparison efficiency. (We lose stability, though.)

While other comparison-efficient in-place sorting methods are known (e.g. [18, 12, 9]), the
ones based on QuickXsort and elementary methods X are particularly easy to implement[1]

---

[1] See for example the code for QuickMergesort that was presented for discussion on code review stack
exchange, `https://codereview.stackexchange.com/q/149443`, and the succinct C++ code in [6].

since one can adapt existing implementations for X. In such an implementation, the tried and tested optimization to choose the pivot as the median of a small sample suggests itself to improve QuickXsort. In previous works [1, 5, 3, 6], the influence of QuickXsort on the performance of X was either studied by ad-hoc techniques that do not easily apply with general pivot sampling or it was studied for the case of very good pivots: exact medians or medians of a sample of $\sqrt{n}$ elements. Both are typically detrimental to the average performance since they add significant overhead, whereas most of the benefit of sampling is realized already for samples of very small constant sizes like 3, 5 or 9. Indeed, in a very recent manuscript [6], Edelkamp and Weiß describe an optimized median-of-3 QuickMergesort implementation in C++ that outperformed the library Quicksort in `std::sort`.

The contribution of this paper is **a general transfer theorem (Theorem 1) that expresses the costs of QuickXsort with median-of-k sampling (for any odd constant k) directly in terms of the costs of X,** (i.e., the costs that X needs to sort n elements in isolation). We thereby obtain the first analyses of QuickMergesort and QuickHeapsort with best possible constant-coefficient bounds on the linear term under realistic sampling schemes.

Since Mergesort only needs a buffer for one of the two runs, QuickMergesort should not simply give Mergesort the smaller of the two segments to sort, but rather the *largest one for which the other segments still offers sufficient buffer space.* (This will be the larger segment of the two if the smaller one contains at least a third of the elements; see Section 6 for details.) Our transfer theorem covers this refined version of QuickMergesort, as well, which had not been analyzed before.[2]

The rest of the paper is structured as follows: In Section 2, we summarize previous work on QuickXsort with a focus on contributions to its analysis. Section 3 collects mathematical facts and notations used later. In Section 4 we define QuickXsort and formulate a recurrence for its cost. Its solution is stated in Section 5. Section 6 presents the QuickMergesort as our stereotypical instantiation of QuickXsort. The proof of the transfer spreads over Sections 7 and 8. In Section 9, we apply our result to QuickHeapsort and QuickMergesort and discuss some algorithmic implications.

2   Previous Work

The idea to combine Quicksort and a secondary sorting method was suggested by Contone and Cincotti [2, 1]. They study Heapsort with an output buffer (external Heapsort),[3] and combine it with Quicksort to QuickHeapsort. They analyze the average costs for external Heapsort in isolation and use a differencing trick for dealing with the QuickXsort recurrence; however, this technique is hard to generalize to median-of-k pivots.

Diekert and Weiß [3] suggest optimizations for QuickHeapsort (some of which need extra space again), and they give better upper bounds for QuickHeapsort with random pivots and median-of-3. Their results are still not tight since they upper bound the total cost of all Heapsort calls together (using ad hoc arguments on the form of the costs for one Heapsort

---

[2] Edelkamp and Weiß do consider this version of QuickMergesort [5], but only analyze it for median-of-$\sqrt{n}$ pivots. In this case, the behavior coincides with the simpler strategy to always sort the smaller segment by Mergesort since the segments are of almost equal size with high probability.

[3] Not having to store the heap in a consecutive prefix of the array allows to save comparisons over classic in-place Heapsort: After a delete-max operation, we can fill the gap at the root of the heap by promoting the largest child and recursively moving the gap down the heap. (We then fill the gap with a $-\infty$ sentinel value). That way, each delete-max needs exactly $\lfloor \lg n \rfloor$ comparisons.

round), without taking the actual subproblem sizes into account that Heapsort is used on. In particular, their bound on the overall contribution of the Heapsort calls does not depend on the sampling strategy.

Edelkamp and Weiß [5] explicitly describe QuickXsort as a general design pattern and, among others, consider using Mergesort as 'X'. They use the median of $\sqrt{n}$ elements in each round throughout to guarantee good splits with high probability. They show by induction that when X uses at most $n \lg n + cn + o(n)$ comparisons on average for some constant c, the number of comparisons in QuickXsort is also bounded by $n \lg n + cn + o(n)$. By combining QuickMergesort with Ford and Johnson's MergeInsertion [8] for subproblems of logarithmic size, Edelkamp and Weiß obtained an in-place sorting method that uses on the average a close to minimal number of comparisons of $n \lg n - 1.3999n + o(n)$. In a recent follow-up manuscript [6], Edelkamp and Weiß investigated the practical performance of QuickXsort and found that a tuned median-of-3 QuickMergesort variant indeed outperformed the C++ library Quicksort. They also derive an upper bound for the average costs of their algorithm using an inductive proof; their bound is not tight.

3 Preliminaries

A comprehensive list of used notation is given in Appendix A; we mention the most important here. We use Iverson's bracket $[stmt]$ to mean 1 if $stmt$ is true and 0 otherwise. $\mathbb{P}[E]$ denotes the probability of event E, $\mathbb{E}[X]$ the expectation of random variable X. We write $X \stackrel{\mathcal{D}}{=} Y$ to denote equality in distribution.

We heavily use the beta distribution: For $\alpha, \beta \in \mathbb{R}_{>0}$, $X \stackrel{\mathcal{D}}{=} \mathrm{Beta}(\alpha, \beta)$ if X admits the density $f_X(z) = z^{\alpha-1}(1-z)^{\beta-1}/\mathrm{B}(\alpha, \beta)$ where $\mathrm{B}(\alpha, \beta) = \int_0^1 z^{\alpha-1}(1-z)^{\beta-1}\,dz$ is the beta function. Moreover, we use the beta-binomial distribution, which is a conditional binomial distribution with the success probability being a beta-distributed random variable. If $X \stackrel{\mathcal{D}}{=} \mathrm{BetaBin}(n, \alpha, \beta)$ then $\mathbb{P}[X = i] = \binom{n}{i}\mathrm{B}(\alpha + i, \beta + (n - i))/\mathrm{B}(\alpha, \beta)$. For a collection of its properties see [23], Section 2.4.7; one property that we use here is a local limit law showing that the normalized beta-binomial distribution converges to the beta distribution. It is reproduced as Lemma 3 in the appendix.

For solving recurrences, we build upon Roura's master theorems [20]. The relevant continuous master theorem is restated in the appendix (Theorem 2).

4 QuickXsort

Let X be a sorting method that requires buffer space for storing at most $\lfloor \alpha n \rfloor$ elements (for $\alpha \in [0, 1]$) to sort n elements. The buffer may only be accessed by swaps so that once X has finished its work, the buffer contains the same elements as before, but in arbitrary order. Indeed, we will assume that X does not compare any buffer contents; then QuickXsort preserves randomness: if the original input is a random permutation, so will be the segments after partitioning and so will be the buffer after X has terminated.[4]

We can then combine[5] X with Quicksort as follows: We first randomly choose a pivot and partition the input around that pivot. This results in two contiguous segments containing the J_1 elements that are smaller than the pivot and the J_2 elements that are larger than

---

[4] We assume in this paper throughout that the input contains pairwise distinct elements.
[5] Depending on details of X, further precautions might have to be taken, e.g., in QuickHeapsort [1]. We assume here that those have already been taken care of and solely focus on the analysis of QuickXsort.

the pivot, respectively. We exclude the space for the pivot, so $J_1 + J_2 = n - 1$; note that since the rank of the pivot is random, so are the segment sizes J_1 and J_2. We then sort one segment by X using the other segment as a buffer, and afterwards sort the buffer segment recursively by QuickXsort.

To guarantee a sufficiently large buffer for X when it sorts J_r ($r = 1$ or 2), we must make sure that $J_{3-r} \geq \alpha J_r$. In case both segments could be sorted by X, we use the larger one. The motivation behind this is that we expect an advantage from reducing the subproblem size for the recursive call as much as possible.

We consider the practically relevant version of QuickXsort, where we use as pivot the median of a sample of $k = 2t + 1$ elements, where $t \in \mathbb{N}_0$ is constant w.r.t. n. Setting $t = 0$ corresponds to choosing pivots uniformly at random.

4.1   Recurrence for Expected Costs

Let $c(n)$ be the expected number of comparisons in QuickXsort on arrays of size n and $x(n)$ be (an upper bound for) the expected number of comparisons in X. We will assume that $x(n)$ fulfills $x(n) = an \lg n + bn \pm O(n^{1-\varepsilon})$ as $n \to \infty$ for constants a, b and $\varepsilon \in (0, 1]$.

For $\alpha < 1$, we obtain two cases: When the split induced by the pivot is "uneven" – namely when $\min\{J_1, J_2\} < \alpha \max\{J_1, J_2\}$, i.e., $\max\{J_1, J_2\} > \frac{n-1}{1+\alpha}$ – the smaller segment is not large enough to be used as buffer. Then we can only assign the large segment as a buffer and run X on the *smaller* segment. If however the split is about "even", i.e., both segments are $\leq \frac{n-1}{1+\alpha}$ we can sort the *larger* of the two segments by X. These cases also show up in the recurrence of costs:

$$
\begin{aligned}
c(0) &= c(1) = 0 \\
c(n) &= (n-1) + \mathbb{E}\Big[[J_1, J_2 \leq \tfrac{1}{1+\alpha}(n-1)][J_1 > J_2](x(J_1) + c(J_2))\Big] \\
&\quad + \mathbb{E}\Big[[J_1, J_2 \leq \tfrac{1}{1+\alpha}(n-1)][J_1 \leq J_2](x(J_2) + c(J_1))\Big] \\
&\quad + \mathbb{E}\Big[[J_2 > \tfrac{1}{1+\alpha}(n-1)](x(J_1) + c(J_2))\Big] \\
&\quad + \mathbb{E}\Big[[J_1 > \tfrac{1}{1+\alpha}(n-1)](x(J_2) + c(J_1))\Big] \qquad\qquad (n \geq 2) \\
&= \sum_{r=1}^{2} \mathbb{E}[A_r(J_r)c(J_r)] + t(n) \qquad \text{where} \hspace{3cm} (1) \\
A_1(J) &= [J, J' \leq \tfrac{1}{1+\alpha}(n-1)] \cdot [J \leq J'] + [J > \tfrac{1}{1+\alpha}(n-1)] \quad \text{with } J' = (n-1) - J \\
A_2(J) &= [J, J' \leq \tfrac{1}{1+\alpha}(n-1)] \cdot [J < J'] + [J > \tfrac{1}{1+\alpha}(n-1)] \\
t(n) &= (n-1) + \mathbb{E}[A_2(J_2)x(J_1)] + \mathbb{E}[A_1(J_1)x(J_2)]
\end{aligned}
$$

The expectation here is taken over the choice for the random pivot, i.e., over the segment sizes J_1 resp. J_2. Note that we use both J_1 and J_2 to express the conditions in a convenient form, but actually either one is fully determined by the other via $J_1 + J_2 = n - 1$. Note how A_1 and A_2 change roles in recursive calls and toll functions since we always sort one segment recursively and the other segment by X.

Distribution of Subproblem Sizes

If pivots are chosen as the median of a random sample of size $k = 2t + 1$, the subproblem sizes have the same distribution, $J_1 \overset{\mathcal{D}}{=} J_2$. Without pivot sampling, we have $J_1 \overset{\mathcal{D}}{=} \mathcal{U}[0..n-1]$, a discrete uniform distribution. If we choose pivots as medians of a sample of $k = 2t + 1$

elements, the value for J_1 consists of two summands: $J_1 = t + I_1$. The first summand, t, accounts for the part of the sample that is smaller than the pivot. Those t elements do not take part in the partitioning round (but they have to be included in the subproblem). I_1 is the number of elements that turned out to be smaller than the pivot during partitioning.

This latter number I_1 is random, and its distribution is $I_1 \stackrel{\mathcal{D}}{=} \mathrm{BetaBin}(n - k, t + 1, t + 1)$, a so-called *beta-binomial distribution*. The connection to the beta distribution is best seen by assuming n independent and uniformly in $(0, 1)$ distributed reals as input. They are almost surely pairwise distinct and their relative ranking is equivalent to a random permutation of $[n]$, so this assumption is w.l.o.g. for our analysis. Then, the *value P* of the pivot in the first partitioning step has a $\mathrm{Beta}(t + 1, t + 1)$ distribution *by definition*. *Conditional* on that *value $P = p$*, $I_1 \stackrel{\mathcal{D}}{=} \mathrm{Bin}(n - k, p)$ has a binomial distribution; the resulting mixture is the so-called beta-binomial distribution. For $t = 0$, i.e., no sampling, we have $t + \mathrm{BetaBin}(n - k, t + 1, t + 1) = \mathrm{BetaBin}(n - 1, 1, 1)$, so we recover the uniform case $\mathcal{U}[0..n - 1]$.

5 The Transfer Theorem

We now state the main result of the paper: an asymptotic approximation for $c(n)$.

▶ **Theorem 1** (Total Cost of QuickXsort)**.** *The expected number of comparisons needed to sort a random permutation with QuickXsort using median-of-k pivots, $k = 2t + 1$, and a sorting method X that needs a buffer of $\lfloor \alpha n \rfloor$ elements for some constant $\alpha \in [0, 1]$ to sort n elements and requires on average $x(n) = an \lg n + bn \pm O(n^{1-\varepsilon})$ comparisons to do so as $n \to \infty$ for some $\varepsilon \in (0, 1]$ is*

$$c(n) = an \lg n + \left(\frac{1}{H} - a \cdot \frac{H_{k+1} - H_{t+1}}{H \ln 2} + b \right) \cdot n \pm O(n^{1-\varepsilon} + \log n),$$

$$\textit{where} \quad H = I_{0, \frac{\alpha}{1+\alpha}}(t + 2, t + 1) + I_{\frac{1}{2}, \frac{1}{1+\alpha}}(t + 2, t + 1)$$

is the expected relative subproblem size that is sorted by X.

Here $I_{x,y}(\alpha, \beta)$ is the regularized incomplete beta function

$$I_{x,y}(\alpha, \beta) = \int_x^y \frac{z^{\alpha-1}(1 - z)^{\beta-1}}{\mathrm{B}(\alpha, \beta)} \, dz, \qquad (\alpha, \beta \in \mathbb{R}_+, 0 \le x \le y \le 1).$$

We prove Theorem 1 in Sections 7 and 8. To simplify the presentation, we will restrict ourselves to a stereotypical algorithm for X and its value $\alpha = \frac{1}{2}$; the given arguments, however, immediately extend to the general statement above.

6 QuickMergesort

A natural candidate for X is Mergesort: It is comparison-optimal up to the linear term (and quite close to optimal in the linear term), and needs a $\Theta(n)$-element-size buffer for practical implementations of merging.[6] To be usable in QuickXsort, we use a swap-based merge procedure as given in Algorithm 1. Note that it suffices to move the *smaller* of the two runs to a buffer; we use a symmetric version of Algorithm 1 when the second run is shorter. Using classical top-down or bottom-up Mergesort as described in any algorithms textbook (e.g. [22]), we thus get along with $\alpha = \frac{1}{2}$.

---

[6] Merging can be done in place using more advanced tricks (see, e.g., [15]), but those tend not to be competitive in terms of running time with other sorting methods. By changing the global structure, a pure in-place Mergesort variant [13] can be achieved using part of the input as a buffer (as in QuickMergesort) at the expense of occasionally having to merge runs of very different lengths.

$\text{MERGE}(A[\ell..r], m, B[b..e])$

// Merges runs $A[\ell, m-1]$ and $A[m..r]$ in-place into $A[l..r]$ using scratch space $B[b..e]$

1   $n_1 := r - \ell + 1;\quad n_2 := r - \ell + 1$

// Assumes $A[\ell, m-1]$ and $A[m..r]$ are sorted, $n_1 \le n_2$ and $n_1 \le e - b + 1$.

2   **for** $i = 0, \ldots, n_1 - 1$ **do** $\text{SWAP}(A[\ell + i], B[b + i])$ **end for**

3   $i_1 := b;\quad i_2 := m;\ o := \ell$

4   **while** $i_1 < b + n_1$ and $i_2 \le r$

5   **if** $B[i_1] \le A[i_2]$ **then** $\text{SWAP}(A[o], B[i_2]);\quad o := o + 1;\quad i_1 := i_1 + 1$

6   **else** $\text{SWAP}(A[o], A[i_1]);\quad o := o + 1;\quad i_2 := i_2 + 1$ **end if**

7   **while** $i_1 < b + n_1$ **do** $\text{SWAP}(A[o], B[i_2]);\quad o := o + 1;\quad i_1 := i_1 + 1$ **end while**

■ **Algorithm 1** A simple merging procedure that uses the buffer only by swaps. We move the first run $A[\ell..r]$ into the buffer $B[b..b + n_1 - 1]$ and then merge it with the second run $A[m..r]$ (still stored in the original array) into the empty slot left by the first run. By the time this first half is filled, we either have consumed enough of the second run to have space to grow the merged result, or the merging was trivial, i.e., all elements in the first run were smaller.

6.1   Average Case of Mergesort

The average number of comparisons for Mergesort has the same – optimal – leading term $n \lg n$ as in the worst and best case; and this is true for both the top-down and bottom-up variants. The coefficient of the *linear* term of the asymptotic expansion, though, is not a constant, but a bounded periodic function with period $\lg n$, and the functions differ for best, worst, and average case and the variants of Mergesort [21, 7, 17, 10, 11].

In this paper, we will confine ourselves to an *upper bound* for the average case $x(n) = an \lg n + bn \pm O(n^{1-\varepsilon})$ with *constant* b valid for all n, so we will set b to the supremum of the periodic function. We leave the interesting challenge open to trace the precise behavior of the fluctuations through the recurrence, where Mergesort is used on a logarithmic number of subproblems with random sizes.

We use the following upper bounds for t̲op-d̲own [11] and b̲ottom-u̲p [17] Mergesort[7]

$$x_{\text{td}}(n) \;=\; n \lg n - 1.24n + 2 \qquad \text{and} \qquad x_{\text{bu}}(n) \;=\; n \lg n - 0.26n \,\pm\, O(1). \tag{2}$$

7   Solving the Recurrence: Leading Term

We start with Equation (1). Since $\alpha = \frac{1}{2}$ for our Mergesort, we have $\frac{\alpha}{1+\alpha} = \frac{1}{3}$ and $\frac{1}{1+\alpha} = \frac{2}{3}$. (The following arguments are valid for general α, including the extreme case $\alpha = 1$, but in an attempt to de-clutter the presentation, we stick to $\alpha = \frac{1}{2}$ here.) We rewrite $A_1(J_1)$ and $A_2(J_2)$ explicitly in terms of the *relative subproblem size*:

$$A_1(J_1) \;=\; \left[\frac{J_1}{n-1} \in \left[\frac{1}{3}, \frac{1}{2}\right] \cup \left(\frac{2}{3}, 1\right]\right], \quad A_2(J_2) \;=\; \left[\frac{J_2}{n-1} \in \left[\frac{1}{3}, \frac{1}{2}\right) \cup \left(\frac{2}{3}, 1\right]\right].$$

Graphically, if we view $J_1/(n-1)$ as a point in the unit interval, the following picture shows which subproblem is sorted recursively; (the other subproblem is sorted by Mergesort).

---

[7] Edelkamp and Weiß [5] use $x(n) = n \lg n - 1.26n \pm o(n)$; Knuth [14, 5.2.4–13] derived this formula for n a power of 2 (a general analysis is sketched, but no closed result for general n is given). Flajolet and Golin [7] and Hwang [11] continued the analysis in more detail; they find that the average number of comparisons is $n \lg n - (1.25 \pm 0.01)n \pm O(1)$, where the linear term oscillates in the given range.
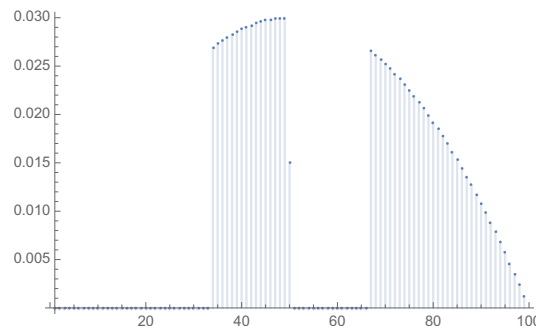
Figure 1 The weights $w_{n,j}$ for $n = 101$, $t = 1$; note the singular point at $j = 50$.
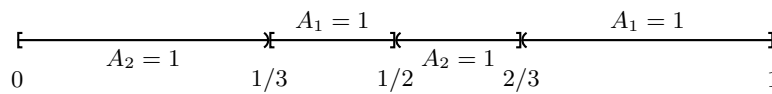
Obviously, we have $A_1 + A_2 = 1$ for any choice of J_1, which corresponds to having exactly one recursive call in QuickMergesort.

7.1 The Shape Function

The expectations $\mathbb{E}[A_r(J_r)c(J_r)]$ in Equation (1) are actually finite sums over the values $0, \ldots, n-1$ that $J := J_1$ can attain. Recall that $J_2 = n - 1 - J_1$ and $A_1(J_1) + A_2(J_2) = 1$ for any value of J. With $J = J_1 \overset{\mathcal{D}}{=} J_2$, we find

$$\sum_{r=1}^{2} \mathbb{E}[A_r(J_r)c(J_r)] \;=\; \sum_{j=0}^{n-1} w_{n,j} \cdot c(j), \qquad \text{where}$$

$$
\begin{aligned}
w_{n,j} \;=\; & \quad \mathbb{P}[J = j] \cdot \left[\tfrac{j}{n-1} \in [\tfrac{1}{3}, \tfrac{1}{2}] \cup (\tfrac{2}{3}, 1] \right] \\
& + \mathbb{P}[J = j] \cdot \left[\tfrac{j}{n-1} \in [\tfrac{1}{3}, \tfrac{1}{2}) \cup (\tfrac{2}{3}, 1] \right] \\
=\; & \begin{cases} 2 \cdot \mathbb{P}[J = j] & \text{if } \tfrac{j}{n-1} \in [\tfrac{1}{3}, \tfrac{1}{2}) \cup (\tfrac{2}{3}, 1] \\ 1 \cdot \mathbb{P}[J = j] & \text{if } \tfrac{j}{n-1} = \tfrac{1}{2} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

We thus have a recurrence of the form required by the Roura's *continuous master theorem (CMT)* (see Theorem 2 in Appendix B) with the weights $w_{n,j}$ from above (Figure 1 shows an example how these weights look like).

It remains to determine $\mathbb{P}[J = j]$. Recall that we choose the pivot as the median of $k = 2t + 1$ elements for a fixed constant $t \in \mathbb{N}_0$, and the subproblem size J fulfills $J = t + I$ with $I \overset{\mathcal{D}}{=} \text{BetaBin}(n - k, t + 1, t + 1)$. So we have for $i \in [0, n - 1 - t]$ by definition

$$
\begin{aligned}
\mathbb{P}[I = i] \;&=\; \binom{n-k}{i} \frac{\mathrm{B}\big(i + t + 1, (n - k - i)\big) + t + 1)}{\mathrm{B}(t+1, t+1)} \\
&=\; \binom{n-k}{i} \frac{(t+1)^{\overline{i}}(t+1)^{\overline{n-k-i}}}{(k+1)^{\overline{n-k}}}
\end{aligned}
$$

(For details, see [23, Section 2.4.7].) Now the local limit law for beta binomials (Lemma 3 in Appendix C) says that the normalized beta binomial I/n converges to a beta variable "in

density", and the convergence is uniform. With the beta density $f_P(z) = z^t(1-z)^t/B(t+1, t+1)$, we thus find by Lemma 3 that

$$\mathbb{P}[J = j] \;=\; \mathbb{P}[I = j - t] \;=\; \frac{1}{n} f_P(j/n) \;\pm\; O(n^{-2}), \qquad (n \to \infty).$$

The shift by the small constant t from $(j-t)/n$ to j/n only changes the function value by $O(n^{-1})$ since f_P is Lipschitz continuous on $[0, 1]$. (Details of that calculation are also given in [23], page 208.)

The first step towards applying the CMT is to identify a shape function $w(z)$ that approximates the relative subproblem size probabilities $w(z) \approx n w_{n,\lfloor zn \rfloor}$ for large n. With the above observation, a natural choice is

$$w(z) \;=\; 2 \left[\tfrac{1}{3} < z < \tfrac{1}{2} \;\vee\; z > \tfrac{2}{3} \right] \frac{z^t(1-z)^t}{B(t+1, t+1)}. \tag{3}$$

We show in Appendix D that this is indeed a suitable shape function, i.e., it fulfills Equation (10) from the CMT.

7.2 Computing the Toll Function

The next step in applying the CMT is a leading-term approximation of the toll function. We consider a general function $x(n) = an \lg n + bn \pm O(n^{1-\varepsilon})$ where the error term holds for any constant $\varepsilon > 0$ as $n \to \infty$. We start with the simple observation that

$$J \lg J \;=\; J\big(\lg(\tfrac{J}{n}) + \lg n\big) \;=\; n \cdot \left(\tfrac{J}{n} \lg \tfrac{J}{n} + \tfrac{J}{n} \lg n \right) \;=\; \tfrac{J}{n} n \lg n \;+\; \tfrac{J}{n} \lg\!\left(\tfrac{J}{n}\right) n. \tag{4}$$

$$=\; \tfrac{J}{n} n \lg n \;\pm\; O(n). \tag{5}$$

For the leading term of $\mathbb{E}[x(J)]$, we thus only have to compute the expectation of J/n, which is essentially a relative subproblem size. In $t(n)$, we also have to deal with the conditionals $A_1(J)$ resp. $A_2(J)$, though. By approximating $\tfrac{J}{n}$ with a beta distributed variable, the conditionals translate to bounds of an integral. Details are given in Lemma 4 (see Appendix E). This yields

$$\begin{aligned}
t(n) \;&=\; n - 1 + \mathbb{E}[A_2(J_2)x(J_1)] + \mathbb{E}[A_1(J_1)x(J_2)] \\
&=\; a\,\mathbb{E}[A_2(J_2)J_1 \lg J_1] + a\,\mathbb{E}[A_1(J_1)J_2 \lg J_2)] \;\pm\; O(n) \\
&\overset{\text{Lemma a}}{=}\; 2a \cdot \frac{t+1}{2t+2} \cdot \left(I_{0,\frac{1}{3}}(t+2, t+1) + I_{\frac{1}{2},\frac{2}{3}}(t+2, t+1) \right) \cdot n \lg n \;\pm\; O(n) \\
&=\; \underbrace{a\left(I_{0,\frac{1}{3}}(t+2, t+1) + I_{\frac{1}{2},\frac{2}{3}}(t+2, t+1) \right)}_{\tilde{a}} \cdot n \lg n \;\pm\; O(n), \qquad (n \to \infty). \tag{6}
\end{aligned}$$

Here we use the *incomplete regularized beta function*

$$I_{x,y}(\alpha, \beta) \;=\; \int_x^y \frac{z^{\alpha-1}(1-z)^{\beta-1}}{B(\alpha, \beta)} \, dz, \qquad (\alpha, \beta \in \mathbb{R}_+, 0 \le x \le y \le 1)$$

for concise notation. ($I_{x,y}(\alpha, \beta)$ is the probability that a $\text{Beta}(\alpha, \beta)$ distributed random variable falls into $(x, y) \subset [0, 1]$, and $I_{0,x}(\alpha, \beta)$ is its cumulative distribution function.)

7.3 Which Case of the CMT?

We are now ready to apply the CMT (Theorem 2). As shown in Section 7.2, our toll function is $\Theta(n \log n)$, so we have $\alpha = 1$ and $\beta = 1$. We hence compute

$$
\begin{aligned}
H &= 1 - \int_0^1 z\, w(z)\, dz \\
&= 1 - \int_0^1 2\left[\tfrac{1}{3} < z < \tfrac{1}{2} \ \vee\ z > \tfrac{2}{3}\right] \frac{z^{t+1}(1-z)^t}{\mathrm{B}(t+1, t+1)}\, dz \\
&= 1 - 2\frac{t+1}{k+1} \int_0^1 \left[\tfrac{1}{3} < z < \tfrac{1}{2} \ \vee\ z > \tfrac{2}{3}\right] \frac{z^{t+1}(1-z)^t}{\mathrm{B}(t+2, t+1)}\, dz \\
&= 1 - \left(I_{\frac{1}{3}, \frac{1}{2}}(t+2, t+1) + I_{\frac{2}{3}, 1}(t+2, t+1)\right) \\
&= I_{0, \frac{1}{3}}(t+2, t+1) + I_{\frac{1}{2}, \frac{2}{3}}(t+2, t+1)
\end{aligned}
\tag{7}
$$

For any sampling parameters, we have $H > 0$, so the overall costs satisfy by Case 1 of Theorem 2

$$
c(n) \ \sim\ \frac{t(n)}{H} \ \sim\ \frac{\bar{a}n \lg n}{H}, \qquad (n \to \infty). \tag{8}
$$

7.4 Cancellations

Combining Equations (6) and (8), we find $c(n) \sim an \lg n$, as $(n \to \infty)$, since $I_{0, \frac{1}{3}} + I_{\frac{1}{3}, \frac{1}{2}} + I_{\frac{1}{2}, \frac{2}{3}} + I_{\frac{2}{3}, 1} = 1$. The leading term of the number of comparisons in QuickXsort is the *same* as in X itself, regardless of how the pivot elements are chosen! This is not as surprising as it might first seem. We are typically sorting a constant fraction of the input by X and thus only do a logarithmic number of recursive calls on a geometrically decreasing number of elements, so the linear contribution of Quicksort (partitioning and recursion cost) is dominated by even the first call of X, which has linearithmic cost. This remains true even if we allow asymmetric sampling, e.g., by choosing the pivot as the *smallest* (or any other order statistic) of a random sample.

Edelkamp and Weiß [5] give the above result for the case of using the median of $\sqrt{n}$ elements, where we effectively have exact medians from the perspective of analysis. In this case, the informal reasoning given above is precise, and in fact, in this case the same form of cancellations also happen for the linear term [5, Thm. 1]. (See also the "exact ranks" result in Section 9.) We will show in the following that for practical schemes of pivot sampling, i.e., with fixed sample sizes, these cancellations happen *only for the leadings-term approximation*. The pivot sampling scheme does affect the linear term significantly; and to measure the benefit of sampling, the analysis thus has to continue to the next term of the asymptotic expansion of $c(n)$.

Relative Subproblem Sizes

The integral $\int_0^1 zw(z)\, dz$ is precisely the expected relative subproblem size for the recursive call, whereas for $t(n)$ we are interested in the subproblem that is sorted using X whose relative size is given by $\int_0^1 (1-z)w(z)\, dz = 1 - \int_0^1 zw(z)\, dz$. We can thus write $\bar{a} = aH$.

The quantity $\int_0^1 zw(z)\, dz$, the average relative size of the recursive call is of independent interest. While it is intuitively clear that for $t \to \infty$, i.e., the case of exact medians as pivots, we must have a relative subproblem size of exactly $\frac{1}{2}$, this convergence is not apparent from the behavior for finite t: the mass of the integral $\int_0^1 zw(z)\, dz$ concentrates at $z = \frac{1}{2}$, a point of discontinuity in $w(z)$. It is also worthy of note that the expected subproblem size is initially larger than $\frac{1}{2}$ ($0.69\overline{4}$ for $t = 0$), then decreases to ≈ 0.449124 around $t = 20$ and then starts to slowly increase again.

8   Solving the Recurrence: The Linear Term

Since $c(n) \sim an \lg n$ for any choice of t, the leading term alone does not allow to make distinctions to judge the effect of sampling schemes. To compute the next term in the asymptotic expansion of $c(n)$, we consider the values $c'(n) = c(n) - an \lg n$. $c'(n)$ has essentially the same recursive structure as $c(n)$, only with a different toll function:

$$
\begin{aligned}
c'(n) &= c(n) - an \lg n \\[1mm]
&= \sum_{r=1}^{2} \mathbb{E}\big[A_r(J_r)c(J_r)\big] - an \lg n + t(n) \\[1mm]
&= \sum_{r=1}^{2} \left(\mathbb{E}\Big[A_r(J_r)\big(c(J_r) - aJ_r \lg J_r\big)\Big] + a\,\mathbb{E}\big[A_r(J_r)J_r \lg J_r\big] \right) - an \lg n \\
&\quad + (n-1) + \mathbb{E}\big[A_2(J_2) \cdot x(J_1)\big] + \mathbb{E}\big[A_1(J_1) \cdot x(J_2)\big] \\[1mm]
&= \sum_{r=1}^{2} \mathbb{E}\big[A_r(J_r)c'(J_r)\big] + (n-1) - an \lg n \\
&\quad + a\,\mathbb{E}\Big[\big(A_1(J_1) + A_2(J_2)\big)J_1 \lg J_1\Big] + b\,\mathbb{E}[A_2(J_2)J_1] \\
&\quad + a\,\mathbb{E}\Big[\big(A_2(J_2) + A_1(J_1)\big)J_2 \lg J_2\Big] + b\,\mathbb{E}[A_1(J_1)J_2] \;\pm\; O(n^{1-\varepsilon})
\end{aligned}
$$

Since $J_1 \overset{\mathcal{D}}{=} J_2$ we can simplify

$$
\begin{aligned}
&\mathbb{E}\Big[\big(A_1(J_1) + A_2(J_2)\big)J_1 \lg J_1\Big] + \mathbb{E}\Big[\big(A_2(J_2) + A_1(J_1)\big)J_2 \lg J_2\Big] \\
&= \mathbb{E}\Big[\big(A_1(J_1) + A_2(J_2)\big)J_1 \lg J_1\Big] + \mathbb{E}\Big[\big(A_2(J_1) + A_1(J_2)\big)J_1 \lg J_1\Big] \\
&= \mathbb{E}\Big[J_1 \lg J_1 \cdot \big(\big(A_1(J_1) + A_1(J_2)\big) + \big(A_2(J_1) + A_2(J_2)\big)\big)\Big] \\
&= 2\mathbb{E}[J \lg J] \\
&\underset{(4)}{=} 2\,\mathbb{E}[\tfrac{J}{n}] \cdot n \lg n + 2 \cdot \tfrac{1}{\ln 2}\mathbb{E}[\tfrac{J}{n} \ln \tfrac{J}{n}] \cdot n \\
&\underset{\text{Lemma b}}{=} n \lg n - \tfrac{1}{\ln 2}\big(H_{k+1} - H_{t+1}\big)n \;\pm\; O(n^{1-\varepsilon}).
\end{aligned}
$$

Plugging this back into our equation for $c'(n)$, we find

$$
\begin{aligned}
c'(n) &= \sum_{r=1}^{2} \mathbb{E}\big[A_r(J_r)c'(J_r)\big] + (n-1) - an \lg n \\
&\quad + a\left(n \lg n - \tfrac{1}{\ln 2}\big(H_{k+1} - H_{t+1}\big)n\right) \\
&\quad + b\left(I_{0,\frac{1}{3}}(t+2, t+1) + I_{\frac{1}{2},\frac{2}{3}}(t+2, t+1)\right) \cdot n \;\pm\; O(n^{1-\varepsilon}) \\
&= \sum_{r=1}^{2} \mathbb{E}\big[A_r(J_r)c'(J_r)\big] + t'(n)
\end{aligned}
$$

where

$$
t'(n) = b'n \pm O(n^{1-\varepsilon}) \quad \text{with} \quad b' = 1 - \tfrac{a}{\ln 2}\big(H_{k+1} - H_{t+1}\big) + b \cdot H
$$

Apart from the smaller toll function $t'(n)$, this recurrence has the very same shape as the original recurrence for $c(n)$; in particular, we obtain the same shape function $w(z)$ and the same $H > 0$ and obtain

$$
c'(n) \;\sim\; \frac{t'(n)}{H} \;\sim\; \frac{b'n}{H}.
$$

■ **Table 1** QuickXsort penalty. QuickXsort with $x(n) = n \lg n + bn$ yields $c(n) = n \lg n + (q + b)n$ where q, the QuickXsort penalty, is given in the table.

	$t = 0$	$t = 1$	$t = 2$	$t = 3$	$t = 10$	$t \to \infty$
$\alpha = 1$	1.1146	0.5070	0.3210	0.2328	0.07705	0
$\alpha = \frac{1}{2}$	0.9120	0.4050	0.2526	0.1815	0.05956	0

8.1 Error Bound

Since our toll function is not given precisely, but only up to an error term $O(n^{1-\varepsilon})$ for a given fixed $\varepsilon \in (0, 1]$, we also have to estimate the overall influence of this term. For that we consider the recurrence for $c(n)$ again, but replace $t(n)$ (entirely) by $C \cdot n^{1-\varepsilon}$. If $\varepsilon > 0$, $\int_0^1 z^{1-\varepsilon} w(z)\, dz < \int_0^1 w(z)\, dz = 1$, so we still find $H > 0$ and apply case 1 of the CMT. The overall contribution of the error term is then $O(n^{1-\varepsilon})$. For $\varepsilon = 0$, $H = 0$ and case 2 applies, giving an overall error term of $O(\log n)$.

This completes the proof of Theorem 1.

9 Discussion

Since all our choices for X are leading-term optimal, so will QuickXsort be. We can thus fix $a = 1$ in Theorem 1; only b (and the allowable α) still depend on X. We then basically find that going from X to QuickXsort adds a "penalty" q in the linear term that depends only on the sampling size (and α), but not on X. Table 1 shows that this penalty is $\approx n$ without sampling, but can be reduced drastically when choosing pivots from a sample of 3 or 5 elements. (Note that the overall costs for pivot sampling are $O(\log n)$ for constant t.)

As we increase the sample size, we converge to the situation studied by Edelkamp and Weiß using median-of-$\sqrt{n}$, where no linear-term penalty is left [5]. Given that q is less than 0.08 already for a sample of 21 elements, these large sample versions are mostly of theoretical interest. It is noteworthy that the improvement from no sampling to median-of-3 yields a reduction of q by more than 50%, which is much more than its effect on Quicksort itself (where it reduces the leading term of costs by 15 % from $2n \ln n$ to $\frac{12}{7} n \ln n$).

We now apply our transfer theorem to the two most well-studied choices for X, Heapsort and Mergesort, and compare the results to analyses and measured comparison counts from previous work. The results confirm that solving the QuickXsort recurrence exactly yields much more accurate predictions for the overall number of comparisons than previous bounds that circumvented this.

9.1 QuickHeapsort

The basic external Heapsort of Cantone and Cincotti [1] always traverses one path in the heap from root to bottom and does one comparison for each edge followed, i.e., $\lfloor \lg n \rfloor$ or $\lfloor \lg n \rfloor - 1$ many per deleteMax. By counting how many leaves we have on each level, Diekert and Weiß found [3, Eq. 1]

$$n\big(\lfloor \lg n \rfloor - 1\big) + 2\big(n - 2^{\lfloor \lg n \rfloor}\big) \;\pm\; O(\log n) \;\;\leq\;\; n \lg n - 0.913929n \;\pm\; O(\log n)$$

comparisons for the sort-down phase. (The constant of the linear term is $1 - \frac{1}{\ln 2} - \lg(2 \ln 2)$, the supremum of the periodic function at the linear term). Using the classical heap construction method adds on average $1.8813726n$ comparisons [4], so here

$$x(n) \;=\; n \lg n + 0.967444n \;\pm\; O(n^\varepsilon) \qquad \text{for any } \varepsilon > 0.$$

Table 2 Comparison of estimates from this paper (W), Theorem 6 of [1] (CC) and Theorem 1 of [3] (DW); shown is the difference between the estimate and the observed average.

Instance	observed	W	CC	DW
Fig. 4 [1], $n = 10^2$, $k = 1$	806	$+67$	$+158$	$+156$
Fig. 4 [1], $n = 10^2$, $k = 3$	714	$+98$	—	$+168$
Fig. 4 [1], $n = 10^5$, $k = 1$	1 869 769	-600	$+90\,795$	$+88\,795$
Fig. 4 [1], $n = 10^5$, $k = 3$	1 799 240	$+9\,165$	—	$+79\,324$
Fig. 4 [1], $n = 10^6$, $k = 1$	21 891 874	$+121\,748$	$+1\,035\,695$	$+1\,015\,695$
Fig. 4 [1], $n = 10^6$, $k = 3$	21 355 988	$+49\,994$	—	$+751\,581$
Tab. 2 [3], $n = 10^4$, $k = 1$	152 573	$+1\,125$	$+10\,264$	$+10\,064$
Tab. 2 [3], $n = 10^4$, $k = 3$	146 485	$+1\,136$	—	$+8\,152$
Tab. 2 [3], $n = 10^6$, $k = 1$	21 975 912	$+37\,710$	$+951\,657$	$+931\,657$
Tab. 2 [3], $n = 10^6$, $k = 3$	21 327 478	$+78\,504$	—	$+780\,091$

Both [1] and [3] report averaged comparison counts from running time experiments. We compare them in Table 2 against the estimates from our result and previous analyses. While the approximation is not very accurate for $n = 100$ (for all analyses), for larger n, our estimate is correct up to the first three digits, whereas previous upper bounds have almost one order of magnitude bigger errors. Note that it is expected for our bound to still be on the conservative side since we used the supremum of the periodic linear term for Heapsort.

9.2   QuickMergesort

For QuickMergesort, Edelkamp and Weiß [5, Fig. 4] report measured average comparison counts for a median-of-3 version using top-down Mergesort: the linear term is shown to be between $-0.8n$ and $-0.9n$. In a recent manuscript [6], they also analytically consider the simplified median-of-3 QuickMergesort which *always* sorts the *smaller* segment by Mergesort (i.e., $\alpha = 1$). It uses $n \lg n - 0.7330n + o(n)$ comparisons on average (using $b = -1.24$). They use this as a (conservative) upper bound for the original QuickMergesort.

Our transfer theorem shows that this bound is off by roughly $0.1n$: median-of-3 Quick-Mergesort uses at most $c(n) = n \lg n - 0.8350n \pm O(\log n)$ comparisons on average. Going to median-of-5 reduces the linear term to $-0.9874n$, which is better than the worst-case for top-down Mergesort for most n.

Skewed Pivots for Mergesort?

For Mergesort with $\alpha = \frac{1}{2}$ the largest fraction of elements we can sort by Mergesort in one step is $\frac{2}{3}$; this suggests that using a slightly skewed pivot might be beneficial since it will increase the subproblem size for Mergesort and decrease the size for recursive calls. Indeed, Edelkamp and Weiß allude to this variation: *"With about 15% the time gap, however, is not overly big, and may be bridged with additional efforts like skewed pivots and refined partitioning."* (the statement appears in the arXiv version of [5], `arxiv.org/abs/1307.3033`). And the above mentioned StackExchange post actually chooses pivots as the second tertile.

Our analysis above can be extended to skewed sampling schemes (omitted due to space constraints), but to illustrate this point it suffices to pay a short visit to "wishful-thinking land" and assume that we can get exact quantiles for free. We can show (e.g., with Roura's

discrete master theorem [20]) that if we always pick the exact ρ-quantile of the input, for $\rho \in (0, 1)$, the overall costs are

$$
c_\rho(n) \;=\; \begin{cases} n \lg n + \left(\dfrac{1 + h(\rho)}{1 - \rho} + b \right) n \;\pm\; O(n^{1-\varepsilon}) & \text{if } \rho \in (\tfrac{1}{3}, \tfrac{1}{2}) \cup (\tfrac{2}{3}, 1) \\[3mm] n \lg n + \left(\dfrac{1 + h(\rho)}{\rho} + b \right) n \;\pm\; O(n^{1-\varepsilon}) & \text{otherwise} \end{cases}
$$

for $h(x) = x \lg x + (1 - x) \lg(1 - x)$. The coefficient of the linear term has a *strict minimum* at $\rho = \frac{1}{2}$: Even for $\alpha = \frac{1}{2}$, the best choice is to use the median of a sample. (The result is the same for fixed-size samples.) For QuickMergesort, skewed pivots turn out to be a pessimization, despite the fact that we sort a larger part by Mergesort. A possible explanation is that skewed pivots significantly decrease the amount of information we obtain from the comparisons during partitioning, but do not make partitioning any cheaper.

9.3 Future Work

More promising than skewed pivot sampling is the use of *several* pivots. The resulting MultiwayQuickXsort would be able to sort all but one segment using X and recurse on only one subproblem. Here, determining the expected subproblem sizes becomes a challenge, in particular for $\alpha < 1$; we leave this for future work.

We also confined ourselves to the expected number of comparisons here, but more details about the distribution of costs are possible to obtain. The variance follows a similar recurrence as the one studied in this paper and a distributional recurrence for the costs can be given. The discontinuities in the subproblem sizes add a new facet to these analyses.

Finally, it is a typical phenomenon that constant-factor optimal sorting methods exhibit periodic linear terms. QuickXsort inherits these fluctuations but also smooths them through the random subproblem sizes. Explicitly accounting for these effects is another interesting challenge for future work.

References

1. D. Cantone and G. Cincotti. Quickheapsort, an efficient mix of classical sorting algorithms. *Theoretical Computer Science*, 285(1):25–42, 2002. `doi:10.1016/S0304-3975(01)00288-2`.

2. Domenico Cantone and Gianluca Cincotti. QuickHeapsort, an efficient mix of classical sorting algorithms. In *Italian Conference on Algorithms and Complexity (CIAC)*, pages 150–162, 2000. `doi:10.1007/3-540-46521-9_13`.

3. Volker Diekert and Armin Weiß. QuickHeapsort: Modifications and improved analysis. *Theory of Computing Systems*, 59(2):209–230, aug 2016. `doi:10.1007/s00224-015-9656-y`.

4. Ernst E. Doberkat. An average case analysis of Floyd's algorithm to construct heaps. *Information and Control*, 61(2):114–131, may 1984. `doi:10.1016/S0019-9958(84)80053-4`.

5. Stefan Edelkamp and Armin Weiß. QuickXsort: Efficient sorting with $n \log n - 1.399n + o(n)$ comparisons on average. In *International Computer Science Symposium in Russia*, pages 139–152. Springer, 2014. `doi:10.1007/978-3-319-06686-8_11`.

6. Stefan Edelkamp and Armin Weiß. QuickMergesort: Practically efficient constant-factor optimal sorting, 2018. `arXiv:1804.10062`.

7. Philippe Flajolet and Mordecai Golin. Mellin transforms and asymptotics. *Acta Informatica*, 31(7):673–696, 1994. `doi:10.1007/BF01177551`.

8. Lester R. Ford and Selmer M. Johnson. A tournament problem. *The American Mathematical Monthly*, 66(5):387, may 1959. `doi:10.2307/2308750`.

9   Viliam Geffert and Jozef Gajdoš. In-place sorting. In *SOFSEM 2011: Theory and Practice of Computer Science*, pages 248–259. Springer, 2011. `doi:10.1007/978-3-642-18381-2_21`.

10   Hsien-Kuei Hwang. Limit theorems for mergesort. *Random Structures and Algorithms*, 8(4):319–336, jul 1996. `doi:10.1002/(sici)1098-2418(199607)8:4<319::aid-rsa3>3.0.co;2-0`.

11   Hsien-Kuei Hwang. Asymptotic expansions of the mergesort recurrences. *Acta Informatica*, 35(11):911–919, 1998. `doi:10.1007/s002360050147`.

12   Jyrki Katajainen. The ultimate heapsort. In *Proceedings of the Computing: The 4th Australasian Theory Symposium*, Australian Computer Science Communications, pages 87–96. Springer-Verlag Singapore Pte. Ltd., 1998. URL: `http://www.diku.dk/~jyrki/Myris/Kat1998C.html`.

13   Jyrki Katajainen, Tomi Pasanen, and Jukka Teuhola. Practical in-place mergesort. *Nordic Journal of Computing*, 3(1):27–40, 1996. URL: `http://www.diku.dk/~jyrki/Myris/KPT1996J.html`.

14   Donald E. Knuth. *The Art Of Computer Programming: Searching and Sorting*. Addison Wesley, 2nd edition, 1998.

15   Heikki Mannila and Esko Ukkonen. A simple linear-time algorithm for in situ merging. *Information Processing Letters*, 18(4):203–208, 1984. `doi:10.1016/0020-0190(84)90112-1`.

16   Conrado Martínez and Salvador Roura. Optimal sampling strategies in Quicksort and Quickselect. *SIAM Journal on Computing*, 31(3):683–705, 2001. `doi:10.1137/S0097539700382108`.

17   Wolfgang Panny and Helmut Prodinger. Bottom-up mergesort—a detailed analysis. *Algorithmica*, 14(4):340–354, 1995. `doi:10.1007/BF01294131`.

18   Klaus Reinhardt. Sorting in-place with a worst case complexity of $n \log n - 1.3n + O(\log n)$ comparisons and $\varepsilon n \log n + O(1)$ transports. In *International Symposium on Algorithms and Computation (ISAAC)*, pages 489–498, 1992. `doi:10.1007/3-540-56279-6_101`.

19   Salvador Roura. *Divide-and-Conquer Algorithms and Data Structures*. Tesi doctoral (Ph. D. thesis, Universitat Politècnica de Catalunya, 1997.

20   Salvador Roura. Improved master theorems for divide-and-conquer recurrences. *Journal of the ACM*, 48(2):170–205, 2001. `doi:10.1145/375827.375837`.

21   Robert Sedgewick and Philippe Flajolet. *An Introduction to the Analysis of Algorithms*. Addison-Wesley-Longman, 2nd edition, 2013.

22   Robert Sedgewick and Kevin Wayne. *Algorithms*. Addison-Wesley, 4th edition, 2011.

23   Sebastian Wild. *Dual-Pivot Quicksort and Beyond: Analysis of Multiway Partitioning and Its Practical Potential*. Doktorarbeit (Ph. D. thesis), Technische Universität Kaiserslautern, 2016. ISBN 978-3-00-054669-3. URL: `http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:hbz:386-kluedo-44682`.

A   Notation

A.1   Generic Mathematics

$\mathbb{N}, \mathbb{N}_0, \mathbb{Z}, \mathbb{R}$ . . . . natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, integers $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, real numbers $\mathbb{R}$.

$\mathbb{R}_{>1}, \mathbb{N}_{\geq 3}$ etc. . . . . restricted sets $X_{\mathrm{pred}} = \{x \in X : x \text{ fulfills pred}\}$.

$0.\overline{3}$ . . . . repeating decimal; $0.\overline{3} = 0.333\ldots = \frac{1}{3}$; numerals under the line form the repeated part of the decimal number.

$\ln(n), \lg(n)$ . . . . natural and binary logarithm; $\ln(n) = \log_e(n)$, $\lg(n) = \log_2(n)$.

X . . . . to emphasize that X is a random variable it is Capitalized.

$[a, b)$ . . . . real intervals, the end points with round parentheses are excluded, those with square brackets are included.

$[m..n]$, $[n]$ . . . . integer intervals, $[m..n] = \{m, m+1, \ldots, n\}$; $[n] = [1..n]$.

$[\text{stmt}]$, $[x = y]$ . . . . . . Iverson bracket, $[\text{stmt}] = 1$ if stmt is true, $[\text{stmt}] = 0$ otherwise.

H_n . . . . nth harmonic number; $H_n = \sum_{i=1}^{n} 1/i$.

$x \pm y$ . . . . x with absolute error $|y|$; formally the interval $x \pm y = [x - |y|, x + |y|]$; as with O-terms, we use one-way equalities $z = x \pm y$ instead of $z \in x \pm y$.

$\mathrm{B}(\alpha, \beta)$ . . . . the beta function, $\mathrm{B}(\alpha, \beta) = \int_0^1 z^{\alpha-1}(1-z)^{\beta-1} \, dz$

$I_{x,y}(\alpha, \beta)$ . . . . the regularized incomplete beta function; $I_{x,y}(\alpha, \beta) = \int_x^y \frac{z^{\alpha-1}(1-z)^{\beta-1}}{\mathrm{B}(\alpha,\beta)} \, dz$ for $\alpha, \beta \in \mathbb{R}_+$, $0 \le x \le y \le 1$.

$a^{\underline{b}}$, $a^{\overline{b}}$ . . . . factorial powers; "a to the b falling resp. rising."

A.2 Stochastics-related Notation

$\mathbb{P}[E]$, $\mathbb{P}[X = x]$ . . . . . . probability of an event E resp. probability for random variable X to attain value x.

$\mathbb{E}[X]$ . . . . expected value of X; we write $\mathbb{E}[X \mid Y]$ for the conditional expectation of X given Y, and $\mathbb{E}_X[f(X)]$ to emphasize that expectation is taken w.r.t. random variable X.

$X \stackrel{\mathcal{D}}{=} Y$ . . . . equality in distribution; X and Y have the same distribution.

$\mathcal{U}(a, b)$ . . . . uniformly in $(a, b) \subset \mathbb{R}$ distributed random variable.

$\mathrm{Beta}(\alpha, \beta)$ . . . . Beta distributed random variable with shape parameters $\alpha \in \mathbb{R}_{>0}$ and $\beta \in \mathbb{R}_{>0}$.

$\mathrm{Bin}(n, p)$ . . . . binomial distributed random variable with $n \in \mathbb{N}_0$ trials and success probability $p \in [0, 1]$.

$\mathrm{BetaBin}(n, \alpha, \beta)$ . . . . . beta-binomial distributed random variable; $n \in \mathbb{N}_0$, $\alpha, \beta \in \mathbb{R}_{>0}$;

A.3 Notation for the Algorithm

n . . . . length of the input array, i.e., the input size.

k, t . . . . sample size $k \in \mathbb{N}_{\ge 1}$, odd; $k = 2t + 1$, $t \in \mathbb{N}_0$.

$x(n)$, a, b . . . . Average costs of X, $x(n) = an \lg n + bn \pm O(n^{1-\varepsilon})$.

$t(n)$, $\bar{a}$, $\bar{b}$ . . . . toll function $t(n) = \bar{a}n \lg n + \bar{b}n \pm O(n^{1-\varepsilon})$

J_1, J_2 . . . . (random) subproblem sizes; $J_1 + J_2 = n - 1$; $J_1 = t + I_1$;

I_1, I_2 . . . . (random) segment sizes in partitioning; $I_1 \stackrel{\mathcal{D}}{=} \mathrm{BetaBin}(n - k, t + 1, t + 1)$; $I_2 = n - k - I_1$; $J_1 = t + I_1$

B The Continuous Master Theorem

We restate Roura's CMT here for convenience.

▶ **Theorem 2** (Roura's Continuous Master Theorem (CMT))**.** *Let* F_n *be recursively defined by*

$$
F_n = \begin{cases} b_n, & \text{for } 0 \le n < N; \\ t_n + \displaystyle\sum_{j=0}^{n-1} w_{n,j} F_j, & \text{for } n \ge N, \end{cases} \tag{9}
$$

where t_n, the toll function, satisfies $t_n \sim Kn^\alpha \log^\beta(n)$ as $n \to \infty$ for constants $K \neq 0$, $\alpha \geq 0$ and $\beta > -1$. Assume there exists a function $w : [0,1] \to \mathbb{R}_{\geq 0}$, the shape function, with $\int_0^1 w(z)dz \geq 1$ and

$$\sum_{j=0}^{n-1} \left| w_{n,j} - \int_{j/n}^{(j+1)/n} w(z)\,dz \right| = O(n^{-d}), \qquad (n \to \infty), \tag{10}$$

for a constant $d > 0$. With $H := 1 - \int_0^1 z^\alpha w(z)\,dz$, we have the following cases:

1. *If $H > 0$, then $F_n \sim \dfrac{t_n}{H}$.*

2. *If $H = 0$, then $F_n \sim \dfrac{t_n \ln n}{\tilde{H}}$ with $\tilde{H} = -(\beta + 1) \displaystyle\int_0^1 z^\alpha \ln(z)\,w(z)\,dz$.*

3. *If $H < 0$, then $F_n = O(n^c)$ for the unique $c \in \mathbb{R}$ with $\displaystyle\int_0^1 z^c w(z)\,dz = 1$.*

Theorem 2 is the "reduced form" of the CMT, which appears as Theorem 1.3.2 in Roura's doctoral thesis [19], and as Theorem 18 of [16]. The full version (Theorem 3.3 in [20]) allows us to handle sublogarithmic factors in the toll function, as well, which we do not need here.

C Local Limit Law for the Beta-Binomial Distribution

Since the binomial distribution is sharply concentrated, one can use Chernoff bounds on beta-binomial variables after conditioning on the beta distributed success probability. That already implies that $\mathrm{BetaBin}(n, \alpha, \beta)/n$ converges to $\mathrm{Beta}(\alpha, \beta)$ (in a specific sense). We can obtain stronger error bounds, though, by directly comparing the PDFs. Doing that gives the following result; a detailed proof is given in [23], Lemma 2.38.

▶ **Lemma 3** (Local Limit Law for Beta-Binomial, [23], Lemma 2.38).
Let $(I^{(n)})_{n \in \mathbb{N}_{\geq 1}}$ be a family of random variables with beta-binomial distribution, $I^{(n)} \overset{\mathcal{D}}{=} \mathrm{BetaBin}(n, \alpha, \beta)$ where $\alpha, \beta \in \{1\} \cup \mathbb{R}_{\geq 2}$, and let $f_B(z)$ be the density of the $\mathrm{Beta}(\alpha, \beta)$ distribution. Then we have uniformly in $z \in (0,1)$ that

$$n \cdot \mathbb{P}\big[I = \lfloor z(n+1) \rfloor\big] = f_B(z) \pm O(n^{-1}), \qquad (n \to \infty).$$

That is, $I^{(n)}/n$ converges to $\mathrm{Beta}(\alpha, \beta)$ in distribution, and the probability weights converge uniformly to the limiting density at rate $O(n^{-1})$.

D Smoothness of the Shape Function

In this appendix we show that $w(z)$ as given in Equation (3) on page 8 fulfills Equation (10) on page 16, the approximation-rate criterion of the CMT. We consider the following ranges for $\frac{\lfloor zn \rfloor}{n-1} = \frac{j}{n-1}$ separately:

- $\frac{\lfloor zn \rfloor}{n-1} < \frac{1}{3}$ and $\frac{1}{2} < \frac{\lfloor zn \rfloor}{n-1} < \frac{2}{3}$.
  Here $w_{n,\lfloor zn \rfloor} = 0$ and so is $w(z)$. So actual value and approximation are exactly the same.
- $\frac{1}{3} < \frac{\lfloor zn \rfloor}{n-1} < \frac{1}{2}$ and $\frac{\lfloor zn \rfloor}{n-1} > \frac{2}{3}$.
  Here $w_{n,j} = 2\mathbb{P}[J = j]$ and $w(z) = 2f_P(z)$ where $f_P(z) = z^t(1-z)^t/\mathrm{B}(t+1, t+1)$ is twice the density of the beta distribution $\mathrm{Beta}(t+1, t+1)$. Since f_P is Lipschitz-continuous on the bounded interval $[0,1]$ (it is a polynomial) the uniform pointwise convergence from above is enough to bound the sum of $\left| w_{n,j} - \int_{j/n}^{(j+1)/n} w(z)\,dz \right|$ over all j in the range by $O(n^{-1})$.

- $\frac{\lfloor zn \rfloor}{n-1} \in \{\frac{1}{3}, \frac{1}{2}, \frac{2}{3}\}$.

  At these boundary points, the difference between $w_{n,\lfloor zn \rfloor}$ and $w(z)$ does not vanish (in particularly $\frac{1}{2}$ is a singular point for $w_{n,\lfloor zn \rfloor}$), but the absolute difference is bounded. Since this case only concerns 3 out of n summands, the overall contribution to the error is $O(n^{-1})$.

Together, we find that Equation (10) is fulfilled as claimed:

$$\sum_{j=0}^{n-1} \left| w_{n,j} - \int_{j/n}^{(j+1)/n} w(z)\, dz \right| = O(n^{-1}) \qquad (n \to \infty). \tag{11}$$

E Approximation by (Incomplete) Beta Integrals

▶ **Lemma 4.** *Let $J \overset{\mathcal{D}}{=} \mathrm{BetaBin}(n - c_1, \alpha, \beta) + c_2$ be a random variable that differs by fixed constants c_1 and c_2 from a beta-binomial variable with parameters $n \in \mathbb{N}$ and $\alpha, \beta \in \mathbb{N}_{\geq 1}$. Then the following holds*

(a) *For fixed constants $0 \leq x \leq y \leq 1$ holds*

$$\mathbb{E}\big[[xn \leq J \leq yn] \cdot J \lg J\big] = \frac{\alpha}{\alpha + \beta}\, I_{x,y}(\alpha + 1, \beta) \cdot n \lg n \,\pm\, O(n), \qquad (n \to \infty).$$

The result holds also when any or both of the inequalities in $[xn \leq J \leq yn]$ are strict.

(b) $\mathbb{E}[\frac{J}{n} \ln \frac{J}{n}] = \frac{\alpha}{\alpha+\beta}(H_\alpha - H_{\alpha+\beta}) \pm O(n^{-h})$ *for any $h \in (0, 1)$.*

Proof. We start with part (a). By the local limit law for beta binomials (Lemma 3) it is plausible to expect a reasonably small error when we replace $\mathbb{E}\big[[xn \leq J \leq yn] \cdot J \lg J\big]$ by $\mathbb{E}\big[[x \leq P \leq y] \cdot (Pn) \lg(Pn)\big]$ where $P \overset{\mathcal{D}}{=} \mathrm{Beta}(\alpha, \beta)$ is beta distributed. We bound the error in the following.

We have $\mathbb{E}\big[[xn \leq J \leq yn] \cdot J \ln J\big] = \mathbb{E}\big[[xn \leq J \leq yn] \cdot \frac{J}{n}\big] \cdot n \ln n \pm O(n)$ by Equation (4); it thus suffices to compute $\mathbb{E}\big[[xn \leq J \leq yn] \cdot \frac{J}{n}\big]$. We first replace J by $I \overset{\mathcal{D}}{=} \mathrm{BetaBin}(n, \alpha, \beta)$ and argue later that this results in a sufficiently small error. We expand

$$\mathbb{E}\big[[x \leq \tfrac{I}{n} \leq y] \cdot \tfrac{I}{n}\big] = \sum_{i=\lceil xn \rceil}^{\lfloor yn \rfloor} \tfrac{i}{n} \cdot \mathbb{P}[I = i]$$

$$= \frac{1}{n} \sum_{i=\lceil xn \rceil}^{\lfloor yn \rfloor} \tfrac{i}{n} \cdot n\mathbb{P}[I = i]$$

$$\overset{\text{Lemma 3}}{=} \frac{1}{n} \sum_{i=\lceil xn \rceil}^{\lfloor yn \rfloor} \tfrac{i}{n} \cdot \left(\frac{(i/n)^{\alpha-1}(1 - (i/n))^{\beta-1}}{\mathrm{B}(\alpha, \beta)} \pm O(n^{-1}) \right)$$

$$= \frac{1}{\mathrm{B}(\alpha, \beta)} \cdot \frac{1}{n} \sum_{i=\lceil xn \rceil}^{\lfloor yn \rfloor} f(i/n) \,\pm\, O(n^{-1}),$$

where $f(z) = z^\alpha (1 - z)^{\beta-1}$.

Note that $f(z)$ is Lipschitz-continuous on the bounded interval $[x, y]$ since it is continuously differentiable (it is a polynomial). Integrals of Lipschitz functions are well-approximated by finite Riemann sums; see Lemma 2.12 (b) of [23] for a formal statement. We use that on the sum above

$$\frac{1}{n} \sum_{i=\lceil xn \rceil}^{\lfloor yn \rfloor} f(i/n) = \int_x^y f(z)\, dz \,\pm\, O(n^{-1}), \qquad (n \to \infty).$$

Inserting above and using $B(\alpha + 1, \beta)/B(\alpha, \beta) = \alpha/(\alpha + \beta)$ yields

$$
\begin{aligned}
\mathbb{E}\big[[x \le \tfrac{I}{n} \le y] \cdot \tfrac{I}{n}\big] &= \frac{\int_x^y z^\alpha (1 - z)^{\beta-1}\, dz}{B(\alpha, \beta)} \quad \pm \quad O(n^{-1}) \\
&= \frac{\alpha}{\alpha + \beta} I_{x,y}(\alpha + 1, \beta) \quad \pm \quad O(n^{-1});
\end{aligned}
\tag{12}
$$

recall that

$$
I_{x,y}(\alpha, \beta) = \int_x^y \frac{z^{\alpha-1}(1 - z)^{\beta-1}}{B(\alpha, \beta)}\, dz = \mathbb{P}\big[x < P < y\big]
$$

denotes the regularized incomplete beta function.

Changing from I back to J has no influence on the given approximation. To compensate for the difference in the number of trials ($n - c_1$ instead of n), we use the above formulas for with $n - c_1$ instead of n; since we let n go to infinity anyway, this does not change the result. Moreover, replacing I by $I + c_2$ changes the value of the argument $z = I/n$ of f by $O(n^{-1})$; since f is smooth, namely Lipschitz-continuous, this also changes $f(z)$ by at most $O(n^{-1})$. The result is thus not affected by more than the given error term:

$$
\mathbb{E}\big[[x \le \tfrac{J}{n} \le y] \cdot \tfrac{J}{n}\big] = \mathbb{E}\big[[x \le \tfrac{I}{n} \le y] \cdot \tfrac{I}{n}\big] \pm O(n^{-1})
$$

We obtain the claim by multiplying with $n \lg n$.

Versions with strict inequalities in $[xn \le J \le yn]$ only affect the bounds of the sums above by one, which again gives a negligible error of $O(n^{-1})$.

This concludes the proof of part (a).

For part (b), we follow a similar route. The function we integrate is no longer Lipschitz continuous, but a weaker form of smoothness is sufficient to bound the difference between the integral and its Riemann sums. Indeed, the above cited Lemma 2.12 (b) of [23] is formulated for the weaker notion of Hölder-continuity: A function $f : I \to R$ defined on a bounded interval I is called Hölder-continuous with exponent $h \in (0, 1]$ when

$$
\exists C \,\forall x, y \in I \; : \; \big|f(x) - f(y)\big| \le C|x - y|^h.
$$

This generalizes Lipschitz-continuity (which corresponds to $h = 1$).

As above, we replace J by $I \stackrel{\mathcal{D}}{=} \mathrm{BetaBin}(n, \alpha, \beta)$, which affects the overall result by $O(n^{-1})$. We compute

$$
\begin{aligned}
\mathbb{E}\big[\tfrac{I}{n} \ln \tfrac{I}{n}\big] &= \sum_{i=0}^n \tfrac{i}{n} \ln \tfrac{i}{n} \cdot \mathbb{P}[I = i] \\
&\underset{\text{Lemma 3}}{=} \frac{1}{n} \sum_{i=0}^n \tfrac{i}{n} \ln \tfrac{i}{n} \cdot \left(\frac{(i/n)^{\alpha-1}(1 - (i/n))^{\beta-1}}{B(\alpha, \beta)} \pm O(n^{-1})\right) \\
&= -\frac{1}{B(\alpha, \beta)} \cdot \frac{1}{n} \sum_{i=0}^n f(i/n) \quad \pm \quad O(n^{-1}),
\end{aligned}
$$

where now $f(z) = \ln(1/z) \cdot z^\alpha (1 - z)^{\beta-1}$. Since the derivative is ∞ for $z = 0$, f cannot be Lipschitz-continuous, but it is Hölder-continuous on $[0, 1]$ for any exponent $h \in (0, 1)$: $z \mapsto \ln(1/z)z$ is Hölder-continuous (see, e.g., [23], Prop. 2.13.), products of Hölder-continuous function remain such on bounded intervals and the remaining factor of f is a polynomial in z, which is Lipschitz- and hence Hölder-continuous.

By Lemma 2.12 (b) of [23] we then have

$$\frac{1}{n}\sum_{i=0}^{n} f(i/n) \;=\; \int_0^1 f(z)\,dz \;\pm\; O(n^{-h})$$

Recall that we can choose h as close to 1 as we wish; this will only affect the constant hidden by the $O(n^{-h})$. It remains to actually compute the integral; fortunately, this "logarithmic beta integral" has a well-known closed form (see, e.g., [23], Eq. (2.30)).

$$\int_0^1 \ln(z)\cdot z^{\alpha}(1-z)^{\beta-1} \;=\; \mathrm{B}(\alpha+1,\beta)\big(H_\alpha - H_{\alpha+\beta}\big)$$

Inserting above, we finally find

$$\begin{aligned}
\mathbb{E}[\tfrac{J}{n}\ln\tfrac{J}{n}] \;&=\; \mathbb{E}[\tfrac{I}{n}\ln\tfrac{I}{n}] \;\pm\; O(n^{-1})\\
&=\; \frac{\alpha}{\alpha+\beta}\big(H_\alpha - H_{\alpha+\beta}\big) \;\pm\; O(n^{-h})
\end{aligned}$$

for any $h \in (0,1)$. ◄