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.

2012 ACM Subject Classification Mathematics of computing → Enumeration

Keywords and phrases random mappings with indegree restrictions, Brent-Pollard heuristic, periods of mappings

Digital Object Identifier 10.4230/LIPIcs.AofA.2018.30

Related Version https://arxiv.org/abs/1701.09148

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

[1] This author was partially funded by NSERC of Canada.
[2] This author was funded by FAPESP under grants 2015/26420-1 and 2013/25977-7.

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 $f : [n] \to [n]$ 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 $p \equiv 1 \pmod{k}$. 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 $\mathbf{B}$ or $\mathbf{T}$, 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 $\mathbf{T}$ and $\mathbf{B}$ over different classes of polynomials over finite fields and investigate different classes of mappings as heuristic models for the behavior of $\mathbf{T}$ and $\mathbf{B}$ 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 $\mathbf{T}$ 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 $\mathbf{T}$ 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 $\mathbf{T}$ 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 $\mathbf{T}$ asymptotically in the case of unrestricted mappings. In our case we use Lemma 1 for the distribution of $\mathbf{Z}$ 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*

$$\log \phi_{n,\varepsilon}(x_*) = k_\varepsilon \frac{(\frac{n}{\lambda})^{1/3}}{\log^{2/3}(\frac{n}{\lambda})}(1 + o(1)).$$

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,*

$$\log \mathbb{E}_n^{\{0,k\}}(\mathbf{T}) = k_0 \frac{(\frac{n}{\lambda})^{1/3}}{\log^{2/3}(\frac{n}{\lambda})}(1 + o(1)),$$

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 $\mathbf{B}$ 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 $\mathbf{B}$ 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 $\mathbf{Z}$.

▶ **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

$$\mathbb{P}_n^{\{0,k\}}\left(\log \mathbf{T} \leq \mu_n + x\sigma_n\right) = \zeta_1 + \zeta_2 + \zeta_3, \tag{6}$$

where

$$\zeta_j = \sum_{m \in I_j} \mathbb{P}_n^{\{0,k\}}(\mathbf{Z} = m)\mathbb{P}_n^{\{0,k\}}(\log \mathbf{T} \le \mu_n + x\sigma_n | \mathbf{Z} = m). \tag{7}$$

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 $m > 1$,*

$$\left|\mathbf{Q}_m\left(\log \mathbf{T} \le \alpha_m + x\beta_m\right) - \phi(x)\right| \le \frac{K}{\sqrt{\log m}}.$$

▶ **Lemma 8.** *For $n = kr$ and $m \in I_2$, let*

$$\delta_x(m,n) = \mathbb{P}_n^{\{0,k\}}\left(\log \mathbf{T} \le \mu_n + x\sigma_n | \mathbf{Z} = m\right) - \phi(x),$$

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 $|x| \le c\sqrt{\log n}$, for some positive constant c, then $\Delta_x(n) \le K_4 \log^{-1/4}(\sqrt{n/\lambda})$, 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 $m \in I_2$,

$$|\delta_x(m,n)| \le \left|\mathbf{Q}_m\left(\frac{\log \mathbf{T} - \alpha_m}{\beta_m} \le y\right) - \phi(y)\right| + |\phi(y) - \phi(x)|. \tag{8}$$

We note that Theorem 7 implies that, for some constant $K_1 > 0$,

$$\left|\mathbf{Q}_m\left(\frac{\log \mathbf{T} - \alpha_m}{\beta_m} \le y\right) - \phi(y)\right| \le \frac{K_1}{\sqrt{\log m}}. \tag{9}$$

Using Equations (8) and (9) and $|\phi(y) - \phi(x)| \le |y - x|$ we obtain

$$|\delta_x(m,n)| \le \frac{K_1}{\sqrt{\log m}} + |y - x|. \tag{10}$$

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,*

$$\mathbb{P}_n^{\{0,k\}}\left(\log \mathbf{T} \le \mu_n + x\sigma_n\right) = \phi(x) + o_x(1),$$

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 $|x| \le c\sqrt{\log n}$.

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,*

$$\mathbb{P}_n^{\{0,k\}}\left(\log \mathbf{B} \leq \mu_n + x\sigma_n\right) = \phi(x) + o_x(1),$$

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 $n_j = \#\{y \in [n] \colon |f^{-1}(y)| = j\}$, $j \geq 0$. 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.

```
Pick a permutation σ = σ_1 ··· σ_n ∈ S_n uniformly at random.
Pick a permutation τ = τ_1 ··· τ_n ∈ S_n uniformly at random.
for i = 0, ..., r − 1 do
    for j = 1, ..., k do
        f(τ[ik + j]) = σ[i + 1]          // τ[ℓ] denotes τ_ℓ, same for σ[ℓ].
    end
end
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 $\mathbf{B}$, 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 $\mathbf{B}$.

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.