Counting Hypergraph Matchings up to Uniqueness Threshold[*]

Renjie Song[1], Yitong Yin[2], and Jinman Zhao[3]

1 **Department of Computer Science and Technology, Nanjing University, China**
   `song.renjie@foxmail.com`
2 **State Key Lab for Novel Software Technology, Nanjing University, China**
   `yinyt@nju.edu.cn`
3 **Department of Computer Science, University of Wisconsin-Madison, USA**
   `jinman.zhao@gmail.com`

―――― **Abstract** ――――

We study the problem of approximately counting matchings in hypergraphs of bounded maximum degree and maximum size of hyperedges. With an activity parameter λ, each matching M is assigned a weight $\lambda^{|M|}$. The counting problem is formulated as computing a partition function that gives the sum of the weights of all matchings in a hypergraph. This problem unifies two extensively studied statistical physics models in approximate counting: the hardcore model (graph independent sets) and the monomer-dimer model (graph matchings).

For this model, the critical activity $\lambda_c = \frac{d^d}{k(d-1)^{d+1}}$ is the threshold for the uniqueness of Gibbs measures on the infinite $(d+1)$-uniform $(k+1)$-regular hypertree. Consider hypergraphs of maximum degree at most $k+1$ and maximum size of hyperedges at most $d+1$. We show that when $\lambda < \lambda_c$, there is an FPTAS for computing the partition function; and when $\lambda = \lambda_c$, there is a PTAS for computing the log-partition function. These algorithms are based on the decay of correlation (strong spatial mixing) property of Gibbs distributions. When $\lambda > 2\lambda_c$, there is no PRAS for the partition function or the log-partition function unless NP=RP.

Towards obtaining a sharp transition of computational complexity of approximate counting, we study the local convergence from a sequence of finite hypergraphs to the infinite lattice with specified symmetry. We show a surprising connection between the local convergence and the reversibility of a natural random walk. This leads us to a barrier for the hardness result: The non-uniqueness of infinite Gibbs measure is not realizable by any finite gadgets.

1 Introduction

Counting problems have long been studied in the context of statistical physics models. Perhaps the two most well studied statistical physics models for approximate counting are the *hardcore model* and the *monomer-dimer model*.

In the hardcore model, given a graph $G = (V, E)$ and a vertex-activity λ, the model assigns each independent set I of G a weight $w_\lambda^{\mathsf{IS}}(I) = \lambda^{|I|}$. A natural probability distribution, the Gibbs distribution, is defined over all independent sets of G as $\mu_\lambda^{\mathsf{IS}}(I) = w_\lambda^{\mathsf{IS}}(I)/Z_\lambda^{\mathsf{IS}}(G)$ where

―――――――――――――――

the normalizing factor $Z_\lambda^{\mathsf{IS}}(G) = \sum_I w_\lambda^{\mathsf{IS}}(I)$ is the *partition function*. In the monomer-dimer model, given a graph $G = (V, E)$ and an edge-activity λ, the model assigns each matching M of G a weight $w_\lambda^{\mathsf{M}}(M) = \lambda^{|M|}$. The Gibbs distribution over all matchings of G is defined accordingly. And the partition function now becomes $Z_\lambda^{\mathsf{M}}(G) = \sum_M w_\lambda^{\mathsf{M}}(M)$. The counting problems are then formulated as computing the partition functions $Z_\lambda^{\mathsf{IS}}(G)$ and $Z_\lambda^{\mathsf{M}}(G)$, or the *log-partition functions* $\log Z_\lambda^{\mathsf{IS}}(G)$ and $\log Z_\lambda^{\mathsf{M}}(G)$.

It was well known that the hardcore model exhibits the following phase transition. For the infinite $(d+1)$-regular tree $\mathbb{T}_d$, there is a critical activity $\lambda_c(\mathbb{T}_d) = d^d/(d-1)^{d+1}$, called the *uniqueness threshold*, such that when $\lambda < \lambda_c$ the correlation between the marginal distribution at the root and any boundary condition on leaves at level t decays exponentially in the depth t, but when $\lambda > \lambda_c$ the boundary-to-root correlation remains substantial even as $t \to \infty$. This property of correlation decay is also called spatial mixing, and was known to be equivalent to the uniqueness of the infinite-volume Gibbs measure on the infinite $(d+1)$-regular tree $\mathbb{T}_d$ [33]. In a seminal work [34], Weitz showed that for all $\lambda < \lambda_c(\mathbb{T}_d)$ the decay of correlation holds for the hardcore model on all graphs of maximum degree bounded by $d+1$ and there is a deterministic FPTAS for approximately computing the partition function on all such graphs. Here the specific notion of decay of correlation established is the strong spatial mixing. The connection of approximability of partition function to the phase transition of the model is further strengthened in a series of works [30, 31, 7, 9] which show that unless NP=RP there is no PRAS for the partition function or the log-partition function of the hardcore model when $\lambda > \lambda_c(\mathbb{T}_d)$ on graphs with maximum degree bounded by $d+1$.

For the monomer-dimer model, it was well known that the model has no such phase transition [13, 14]. And analogously there is an FPRAS due to Jerrum and Sinclair [16] for the partition function of the monomer-dimer model on all graphs. In [1] strong spatial mixing with an exponential rate was established for the model on all graphs with maximum degree bounded by an arbitrary constant and a deterministic FPTAS was also given for the partition function on all such graphs.

In this paper, we study *hypergraph matchings*, a model that unifies both the hardcore model and the monomer-dimer model. A hypergraph $\mathcal{H} = (V, E)$ consists of a vertex set V and a collection E of vertex subsets, called the (hyper)edges. A *matching* of $\mathcal{H}$ is a set $M \subseteq E$ of disjoint hyperedges in $\mathcal{H}$. Given a hypergraph $\mathcal{H}$ and an *activity parameter* $\lambda > 0$, a *configuration* is a matching M of $\mathcal{H}$, and is assigned a weight $w_\lambda(M) = \lambda^{|M|}$. The *Gibbs measure* over all matchings of $\mathcal{H}$ is defined as $\mu(M) = w_\lambda(M)/Z_\lambda(\mathcal{H})$, where the normalizing factor $Z_\lambda(\mathcal{H})$ is the *partition function* for the model, defined as:

$$Z_\lambda(\mathcal{H}) = \sum_{M:\text{ matching of } \mathcal{H}} \lambda^{|M|}.$$

This model represents an interesting subclass of Boolean CSP defined by the matching (packing) constraints. It also unifies the hardcore model and the monomer-dimer model. Consider the family of hypergraphs of maximum edge size $d+1$ and maximum degree $k+1$:

- When $d = 1$, the model becomes the monomer-dimer model on graphs of maximum degree $k+1$.
- When $k = 1$, the partition function takes sum over independent sets in the dual graph, and the model becomes the hardcore model on graphs of maximum degree $d+1$.

For hypergraphs, the study of approximate counting hypergraph matchings was initiated in [17]. In [5], an FPTAS was obtained for counting matchings in 3-uniform hypergraphs of maximum degree at most 3 by considering the correlation decay for the independent sets in claw-free graphs. In [22], an FPTAS was given for 3-uniform hypergraphs of maximum

degree at most 4 by the correlation decay of the original CSP. All these results assumed $\lambda = 1$, i.e. the problem of counting the number of matchings in a hypergraph.

Our results

We show that for hypergraph matchings $\lambda_c = \lambda_c(\mathbb{T}_{d,k}) = \frac{d^d}{k(d-1)^{d+1}}$ is the uniqueness threshold on the infinite $(d+1)$-uniform $(k+1)$-regular hypertree $\mathbb{T}_{d,k}$.

▶ **Proposition 1.** *There is a unique Gibbs measure on matchings of $\mathbb{T}_{d,k}$ if and only if $\lambda \leq \lambda_c$.*

This fact was implicit in the literature. Here we give a formal proof. It subsumes the well-known uniqueness threshold $\lambda_c(\mathbb{T}_{d,1}) = \frac{d^d}{(d-1)^{d+1}}$ for the hardcore model on the infinite $(d+1)$-regular tree and also the lack of phase-transition for the monomer-dimer model.

We then establish the decay of correlation for hypergraph matchings on all hypergraphs with bounded maximum size of hyperedges and bounded maximum degree when the activity λ is in the uniqueness regime for the uniform regular hypertree. The specific notion of decay of correlations that we establish here is the strong spatial mixing [34] (see Section 2 for a formal definition). Consequently, we give an FPTAS for the partition function when λ is in the interior of the uniqueness regime, and a PTAS for the log-partition function when λ is at the critical threshold.

▶ **Theorem 2.** *For every finite integers $d, k \geq 1$, the following holds for matchings with activity λ on all hypergraphs of maximum edge-size at most $d + 1$ and maximum degree at most $k + 1$:*
- *if $\lambda < \lambda_c$, the model exhibits strong spatial mixing at an exponential rate and there exists an FPTAS for computing the partition function;*
- *if $\lambda = \lambda_c$, the model exhibits strong spatial mixing at a polynomial rate and there is a PTAS for computing the log-partition function.*

▶ Remark 3. The theorem unifies the strong spatial mixing and FPTAS for the hardcore model [34] and the monomer-dimer model [1], and also covers as special cases the results for approximate counting non-weighted hypergraph matchings in [17, 5, 22].

For hypergraph matchings, the case of critical threshold is of significance. There is a natural combinatorial problem that corresponds to the threshold case: counting matchings in 3-uniform hypergraphs of maximum degree at most 5. Here $d = 2$, $k = 4$, and the critical $\lambda_c = \frac{d^d}{k(d-1)^{d+1}} = 1$, which corresponds to counting the number of hypergraph matchings without weight.

Unlike most recent correlation-decay-based algorithms, where the strong spatial mixings were established by a potential analysis, we do not use the potential method to analyze the decay of correlation. Instead, we prove the following stronger extremal statement.

▶ **Proposition 4.** *For hypergraph matchings, the worst case of (weak or strong) spatial mixing, in terms of decay rate, among all hypergraphs of maximum edge-size at most $d + 1$ and maximum degree at most $k + 1$, is represented by the weak spatial mixing on $\mathbb{T}_{d,k}$.*

We construct a hypergraph version of Weitz's self-avoiding walk tree. Then we show that weak spatial mixing on the uniform regular hypertree implies strong spatial mixing on all smaller hypertrees by a step-by-step comparison of correlation decay. This was the original approach used by Weitz for the hardcore model [34]. Compared to the more recent potential method [20, 19, 27, 28, 29, 22], this method of analyzing the decay of correlation has the advantage in dealing with the critical case.
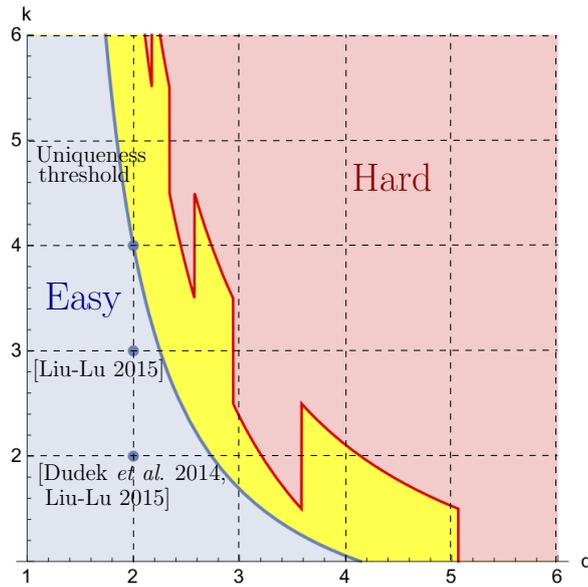
Figure 1 The classification of computational complexity of approximately counting matchings in hypergraphs of max-degree $(k+1)$ and max-edge-size $(d+1)$ when $\lambda = 1$. The blue curve is the uniqueness threshold. The non-continuity of the red curve is due to rounding.

On the other hand, due to a simple reduction from the inapproximability of the hardcore model in the non-uniqueness regime [31], we have the following hardness result.

▶ **Theorem 5.** *If $\lambda > \frac{2k+1+(-1)^k}{k+1}\lambda_c \approx 2\lambda_c$, then there is no PRAS for the partition function or the log-partition function for the family of hypergraphs stated in Theorem 2, unless NP=RP.*

Figure 1 illustrates the classification of approximability of counting hypergraph matchings when $\lambda = 1$. Each integral point (d, k) corresponds to the problem of approximately counting matchings in hypergraphs of max-degree $(k+1)$ and max-edge-size $(d+1)$. The landscape will continuously change when λ changes.

It is worth noticing that in our reduction the hard instances contain many small cycles, while from the algorithmic side the worst cases for the decay of correlation are trees. This obvious inconsistency between upper and lower bounds and the *ad hoc* nature of the simple reduction seem to suggest that the current hardness threshold is not optimal.

We then explore the possibility of bringing the current hardness threshold from $\approx 2\lambda_c$ down to the phase-transition threshold λ_c. We discover a reason why getting the exact transition of approximability could be so challenging for this model on hypergraphs.

To state our discovery, let us first review the current approach for establishing computational phase transition for approximate counting [6, 26, 30, 31, 7, 9, 10], which consists of two main steps:

- (from all infinite measures to finitely many infinite measures) The uniqueness threshold $\lambda_c(\mathbb{T}_d)$ for the Gibbs measure on the infinite regular tree $\mathbb{T}_d$ is achieved by a sub-family of Gibbs measures with simple structure: the Gibbs measures that are invariant under a group $\mathbb{G}$ of automorphisms on $\mathbb{T}_d$. For the hardcore model, these are the so-called *semi-translation invariant* Gibbs measures, which are invariant under parity-preserving automorphisms on $\mathbb{T}_d$, and the threshold $\lambda_c(\mathbb{T}_d)$ for the uniqueness of all Gibbs measures on $\mathbb{T}_d$ is the same as the threshold $\lambda_c(\mathbb{T}_d^{\mathbb{G}})$ for the uniqueness of only those Gibbs measures that are invariant under the group $\mathbb{G}$ of parity-preserving automorphisms.

- (from finitely many infinite measures to finite measures) A sequence of (possibly random) finite graphs G_n is constructed to converge locally to $\mathbb{T}_d^{\mathbb{G}}$, the infinite tree $\mathbb{T}_d$ equipped with the symmetry specified by group $\mathbb{G}$. For the hardcore model, and more generally antiferromagnetic spin systems, G_n are the random regular bipartite graphs [6, 26, 30, 31, 7, 9, 10], which converge locally to the infinite tree $\mathbb{T}_d$ respecting the symmetry between vertices of the same parity. The "random" and "regular" parts in this construction guarantee to preserve the local tree structure in distribution, while the bipartiteness respects the parity of vertices.

For the model of hypergraph matchings, the first step follows. We show that there indeed is a group $\widehat{\mathbb{G}}$ of automorphisms on the infinite $(d + 1)$-uniform $(k + 1)$-regular hypertree $\mathbb{T}_{d,k}$ such that $\lambda_c(\mathbb{T}_{d,k}) = \lambda_c(\mathbb{T}_{d,k}^{\widehat{\mathbb{G}}})$, i.e. the uniqueness of Gibbs measure on $\mathbb{T}_{d,k}$ is represented precisely by the uniqueness of only those Gibbs measures invariant under $\widehat{\mathbb{G}}$. This gives a natural generalization of semi-translation Gibbs measures to the hypergraph model.

However, we show that there does not exist *any* sequence of (deterministic or random) finite hypergraphs that converge locally to $\mathbb{T}_{d,k}^{\widehat{\mathbb{G}}}$ unless $k = 1$ where the model degenerates to the hardcore model on graphs. In fact, we give a complete characterization of the symmetry described by a group $\mathbb{G}$ of automorphisms on $\mathbb{T}_{d,k}$ that there exists a sequence of finite hypergraphs that converge locally to $\mathbb{T}_{d,k}^{\mathbb{G}}$.

▶ **Theorem 6.** *Let $\mathbb{G}$ be a group of automorphisms on $\mathbb{T}_{d,k}$ with finitely many orbits. There exist a sequence of random finite hypergraphs $\mathcal{H}_n$ that converge locally to $\mathbb{T}_{d,k}^{\mathbb{G}}$ if and only if the uniform random walk on $\mathbb{T}_{d,k}$ projected onto the orbits of $\mathbb{G}$ is reversible.*

See Theorem 27 and its proof for more details of Theorem 6.

Discussion

To summarize our discoveries for the model of hypergraph matchings:

- Theorem 2 implicitly but rigorously shows that the worst case for the decay of correlation among a family of hypergraphs with bounded maximum degree and bounded maximum edge-size, is achieved by the infinite uniform regular hypertree.

- However, in the current inapproximability stated by Theorem 5, the hard instances are not locally tree-like, but rather, the gadgets locally converge to an infinite hypergraph which is not a hypertree (see Section 6).

- And finally, Theorem 6 gives an explanation of this inconsistence between upper and lower bounds: the extremal case for the decay of correlation in Theorem 2, which is achieved by an infinite-hypertree measure, can never be realized by *any* finite hypergraphs.[1]

Altogether, these discoveries deliver the following very interesting message: In order to establish a sharp connection between computational complexity of approximate counting and phase transitions for hypergraph matchings or other more general models, a more fine-grained definition of uniqueness on finite graphs is necessary.

---

[1] In fact, aided by numerical simulations, so far we have not encountered any family of measures on the infinite uniform regular hypertree $\mathbb{T}_{d,k}$ realizable by finite hypergraphs, whose uniqueness threshold is below $2\lambda_c$. This seems to provide some empirical evidence for that on finite hypergraphs, the worst case for uniqueness might not be locally tree-like.

Remark on exposition

For convenience of visualizing the results, all our results in the rest of the paper are presented for *independent sets* in the dual hypergraphs. Note that matchings are equivalent to independent sets under hypergraph duality. The only effect of duality on a family of hypergraphs with bounded maximum edge size and bounded maximum degree is to switch the bounds on the edge size and the degree. We emphasize that our notion of hypergraph independent set is different from the more popular definition used in [2, 3]. We call a vertex subset $I \subseteq V$ in a hypergraph $\mathcal{H} = (V, E)$ an independent set if no two vertices in I are contained in the same hyperedge, while in [2, 3], an $I \subseteq V$ is an independent set if it does not contain any hyperedge as subset.

Related works

Approximate counting of hypergraph matchings was studied in [17] for hypergraphs with restrictive structures, and in [22, 5] for hypergraphs with bounded edge size and maximum degree. In [3, 24], approximate counting of a variant of hypergraph independent sets was studied, where the definition of hypergraph independent set is different from ours. In a very recent breakthrough [2], FPTAS for this problem is obtained when there is no strong spatial mixing. In [8], the hardness is established for a class of hypergraph models including ours.

The spatial mixing (decay of correlation) is already a widely studied topic in Computer Science, because it may support FPTAS for #P-hard counting problems. The decay of correlation was established via the self-avoiding walk tree for the hardcore model [34, 29], monomer-dimer model [1, 28], and two-spin systems [20, 19, 28]. Similar tree-structured recursions were employed to prove the decay of correlation for multi-spin systems [11, 25, 12] and more general CSPs [21, 23, 22].

2   Preliminaries

For a *hypergraph* $\mathcal{H} = (V, E)$, the *size* of a hyperedge $e \in E$ is its cardinality $|e|$, and the *degree* of a vertex $v \in V$, denoted by $\deg v = \deg_{\mathcal{H}}(v)$, is the number of hyperedges $e \in E$ *incident to* v, i.e. satisfying $v \in e$. A hypergraph $\mathcal{H}$ is *k-uniform* if all hyperedges are of the same size k, and is *d-regular* if all vertices have the same degree d. The *incidence graph* of a hypergraph $\mathcal{H} = (V, E)$ is a bipartite graph with V and E as vertex sets on the two sides, such that each $(v, e) \in V \times E$ is a bipartite edge if and only if v is incident to e.

A *matching* of hypergraph $\mathcal{H} = (V, E)$ is a set $M \subseteq E$ of disjoint hyperedges in $\mathcal{H}$. Given an *activity parameter* $\lambda > 0$, the *Gibbs measure* is a probability distribution over matchings of $\mathcal{H}$ proportional to the weight $w_\lambda^{\mathsf{M}}(M) = \lambda^{|M|}$, defined as $\mu_\lambda^{\mathsf{M}}(M) = w_\lambda^{\mathsf{M}}(M)/Z_\lambda^{\mathsf{M}}(\mathcal{H})$, where the normalizing factor $Z_\lambda^{\mathsf{M}}(\mathcal{H}) = \sum_M w_\lambda^{\mathsf{M}}(M)$ is the partition function.

Similarly, an *independent set* of hypergraph $\mathcal{H} = (V, E)$ is a set $I \subseteq V$ of vertices satisfying $|I \cap e| \leq 1$ for all hyperedges e in $\mathcal{H}$. The *Gibbs measure* over independent sets of $\mathcal{H}$ with activity $\lambda > 0$ is given by

$$\mu_\lambda^{\mathsf{IS}}(I) = \frac{w_\lambda^{\mathsf{IS}}(I)}{Z_\lambda^{\mathsf{IS}}(\mathcal{H})} = \frac{\lambda^{|I|}}{Z_\lambda^{\mathsf{IS}}(\mathcal{H})}, \tag{1}$$

where the normalizing factor $Z_\lambda^{\mathsf{IS}}(\mathcal{H}) = \sum_I w_\lambda^{\mathsf{IS}}(I)$ is the partition function for independent sets of $\mathcal{H}$ with activity λ.

Independent sets and matchings are equivalent under hypergraph duality. The *dual* of a hypergraph $\mathcal{H} = (V, E)$, denoted by $\mathcal{H}^* = (E^*, V^*)$, is the hypergraph whose vertex set

is denoted by E^* and edge set is denoted by V^*, such that every vertex $v \in V$ (and every hyperedge $e \in E$) in $\mathcal{H}$ is one-to-one corresponding to a hyperedge $v^* \in V^*$ (and a vertex $e^* \in E^*$), such that $e^* \in v^*$ if and only if $v \in e$. Note that under duality, matchings and hypergraphs are the same CSP and hence result in the same Gibbs measure, which remains to be true even with activity λ. Also a family of hypergraphs of bounded maximum edge size and bounded maximum degree is transformed under duality to a family of hypergraphs with the bounds on the edge size and degree exchanged.

▶ **Remark 7.** With the above equivalence under duality, from now on we state all our results in terms of the independent sets in the dual hypergraph and omit the superscript $\cdot^{\mathsf{IS}}$ in notations.

Given the Gibbs measure over independent sets of hypergraph $\mathcal{H}$ and a vertex v, we define the *marginal probability* p_v as

$$p_v = p_{\mathcal{H},v} = \Pr[v \in I]$$

which is the probability that v is in an independent set I sampled from the Gibbs measure (such a vertex is also said to be *occupied*). Given a vertex set $\Lambda \subset V$, a *configuration* is a $\sigma_\Lambda \in \{0,1\}^\Lambda$ which corresponds to an independent set I_Λ partially specified over Λ such that $\sigma_\Lambda(v)$ indicates whether a $v \in \Lambda$ is occupied by the independent set. We further define the marginal probability $p_{\mathcal{H},v}^{\sigma_\Lambda}$ as

$$p_v^{\sigma_\Lambda} = p_{\mathcal{H},v}^{\sigma_\Lambda} = \Pr[v \in I \mid I_\Lambda = \sigma_\Lambda]$$

which is the probability that v is occupied under the Gibbs measure conditioning on the configuration of vertices in $\Lambda \subset V$ being fixed as σ_Λ.

▶ **Definition 8.** The independent sets of a finite hypergraph $\mathcal{H} = (V, E)$ with activity $\lambda > 0$ exhibit *weak spatial mixing (WSM)* with rate $\delta : \mathbb{N} \to \mathbb{R}^+$ if for any $v \in V$, $\Lambda \subseteq V$, and any two configurations $\sigma_\Lambda, \tau_\Lambda \in \{0,1\}^\Lambda$ which correspond to two independent sets partially specified on Λ,

$$|p_v^{\sigma_\Lambda} - p_v^{\tau_\Lambda}| \leq \delta(\mathrm{dist}_{\mathcal{H}}(v, \Lambda)),$$

where $\mathrm{dist}_{\mathcal{H}}(v, \Lambda)$ is the shortest distance between v and any vertex in Λ in hypergraph $\mathcal{H}$.

▶ **Definition 9.** The independent sets of a finite hypergraph $\mathcal{H} = (V, E)$ with activity $\lambda > 0$ exhibit *strong spatial mixing (SSM)* with rate $\delta : \mathbb{N} \to \mathbb{R}^+$ if for any $v \in V$, $\Lambda \subseteq V$, and any two configurations $\sigma_\Lambda, \tau_\Lambda \in \{0,1\}^\Lambda$ which correspond to two independent sets partially specified on Λ,

$$|p_v^{\sigma_\Lambda} - p_v^{\tau_\Lambda}| \leq \delta(\mathrm{dist}_{\mathcal{H}}(v, \Delta)),$$

where $\Delta \subseteq \Lambda$ stands for the subset on which σ_Λ and τ_Λ differ and $\mathrm{dist}_{\mathcal{H}}(v, \Delta)$ is the shortest distance between v and any vertex in Δ in hypergraph $\mathcal{H}$.

The definitions of WSM and SSM extend to infinite hypergraphs with the same conditions to be satisfied for every finite region $\Psi \subset V$ conditioning on the vertices in $\partial\Psi$ being unoccupied.

Gibbs measures on the infinite tree

We follow Remark 7 and state our discoveries in terms of independent sets in the dual hypergraphs. Let $\mathbb{T}_{k,d}$ be the infinite $(k+1)$-uniform $(d+1)$-regular hypertree, whose incidence graph is the infinite tree in which all vertices with parity 0 are of degree $(k+1)$ and all vertices with parity 1 are of degree $(d+1)$. A probability measure μ on hypergraph independent sets of $\mathbb{T}_{k,d}$ is *Gibbs* if for any finite sub-hypertree $\mathcal{T}$, conditioning μ upon the event that all vertices on the outer boundary of $\mathcal{T}$ are unoccupied gives the same distribution on independent sets of $\mathcal{T}$ as defined by (1) with $\mathcal{H} = \mathcal{T}$. We further consider the *simple* Gibbs measures satisfying *conditional independence*: Conditioning μ on a configuration of a subset Λ of vertices results in a measure in which the configurations on the components separated by Λ are independent of each other. The Gibbs distribution on a finite hypergraph is always simple. A Gibbs measure on $\mathbb{T}_{k,d}$ is *translation-invariant* if it is invariant under all automorphisms of $\mathbb{T}_{k,d}$. Fix an automorphism group $\mathbb{G}$ of $\mathbb{T}_{k,d}$. A $\mathbb{G}$-*translation-invariant* Gibbs measure on $\mathbb{T}_{k,d}$ is a measure that is invariant under all automorphisms from $\mathbb{G}$. For example, the *semi-translation-invariant* Gibbs measures on regular tree are invariant under all parity-preserving automorphisms on $\mathbb{T}_{1,d}$. The natural group actions of $\mathbb{G}$ respectively on vertices and hyperedges partition the sets of vertices and hyperedges into orbits. For example, in the semi-translation-invariant symmetry on regular tree, vertices with the same parity form an orbit. We will show that $\lambda_c(\mathbb{T}_{k,d}) = \frac{d^d}{k(d-1)^{d+1}}$ is the uniqueness threshold for the Gibbs measures on hypergraph independent sets of $\mathbb{T}_{k,d}$. Furthermore, this uniqueness threshold is achieved by a family of Gibbs measures with simple structure.

▶ **Theorem 10.** *There is always a unique simple translation-invariant Gibbs measure on independent sets of $\mathbb{T}_{k,d}$. Let $\lambda_c = \lambda_c(\mathbb{T}_{k,d}) = \frac{d^d}{k(d-1)^{d+1}}$. There is a unique Gibbs measure on $\mathbb{T}_{k,d}$ if and only if $\lambda \leq \lambda_c$. Furthermore, there is an automorphism group $\widehat{\mathbb{G}}$ on $\mathbb{T}_{k,d}$ which classifies all vertices of $\mathbb{T}_{k,d}$ into 2 orbits, such that the threshold for the uniqueness of $\widehat{\mathbb{G}}$-translation invariant Gibbs measures on $\mathbb{T}_{k,d}$, denoted as $\lambda_c(\mathbb{T}_{k,d}^{\widehat{\mathbb{G}}})$, is $\lambda_c(\mathbb{T}_{k,d}^{\widehat{\mathbb{G}}}) = \lambda_c(\mathbb{T}_{k,d})$.*

This proves the uniqueness threshold stated in Proposition 1.

3.1   Branching matrices

The automorphism group $\mathbb{G}$ on $\mathbb{T}_{k,d}$ can be described conveniently by a notion of branching matrices. For an automorphism group $\mathbb{G}$ on $\mathbb{T}_{k,d}$, the natural group actions of $\mathbb{G}$ respectively on vertices and hyperedges partition the sets of vertices and hyperedges into orbits. Let τ_v and τ_e be the respective numbers of orbits for vertices and hyperedges. For each $i \in [\tau_v]$, we say a vertex is of *type-i* if it is in the i-th orbit for vertices; and the same also applies to hyperedges. Assuming the symmetry on $\mathbb{T}_{k,d}$ given by automorphism group $\mathbb{G}$, the *hypergraph branching matrices*, or just *branching matrices*, are the following two nonnegative integral matrices:

$$\boldsymbol{D} = \boldsymbol{D}^{\tau_v \times \tau_e} = [d_{ij}] \quad \text{and} \quad \boldsymbol{K} = \boldsymbol{K}^{\tau_e \times \tau_v} = [k_{ji}],$$

which satisfy that for any $i \in [\tau_v]$ and $j \in [\tau_e]$:
- every vertex in $\mathbb{T}_{k,d}$ of type-i is incident to precisely d_{ij} hyperedges of type-j;
- every hyperedge in $\mathbb{T}_{k,d}$ of type-j contains precisely k_{ji} vertices of type-i.

The $\boldsymbol{D}$ and $\boldsymbol{K}$ are transition matrices from vertex-types to hyperedge-types and vice versa in $\mathbb{T}_{k,d}$. The definition can be seen as a hypergraph generalization of the branching matrix for

multi-type Galton-Watson tree [27]. Since types (orbits) are invariant under all automorphisms from $\mathbb{G}$, it is clear that the above $\boldsymbol{D}$ and $\boldsymbol{K}$ are well-defined for every automorphism group $\mathbb{G}$ on $\mathbb{T}_{k,d}$ with finitely many orbits.

▶ **Proposition 11.** *Every automorphism group $\mathbb{G}$ on $\mathbb{T}_{k,d}$ with finitely many orbits can be identified by a pair of branching matrices $\boldsymbol{D}$ and $\boldsymbol{K}$ with rules as described above and satisfy: (1) $\sum_j d_{ij} = d + 1$ and $\sum_i k_{ji} = k + 1$; (2) $d_{ij} = 0$ if and only if $k_{ji} = 0$; and (3) $\boldsymbol{DK}$ and $\boldsymbol{KD}$ are irreducible.*

Conversely, any pair of nonnegative integral matrices $\boldsymbol{D}$ and $\boldsymbol{K}$ satisfying these conditions are branching matrices for some automorphism group $\mathbb{G}$ on $\mathbb{T}_{k,d}$.

Proof. Let $\mathbb{G}$ be an automorphism group on $\mathbb{T}_{k,d}$ with finitely many orbits. It is trivial to see that the branching matrices $\boldsymbol{D}$ and $\boldsymbol{K}$ are well-defined and satisfy $\sum_j d_{ij} = d + 1$ and $\sum_i k_{ji} = k + 1$.

A vertex v of type-i is incident to a hyperedge e of type-j if and only if e of type-j contains a vertex v of type i, thus $k_{ji} \neq 0$ if and only if $d_{ij} \neq 0$.

The irreducibility of $\boldsymbol{DK}$ and $\boldsymbol{KD}$ follows that of the matrix $\begin{bmatrix} \boldsymbol{0} & \boldsymbol{D} \\ \boldsymbol{K} & \boldsymbol{0} \end{bmatrix}$, which is a consequence to the that every type of vertex and hyperedge is accessible from all other types of vertices and hyperedges, which follows the simple fact that the incidence graph $\mathbb{T}_{k,d}$ is strongly connected.

Conversely, let $\boldsymbol{D}$ and $\boldsymbol{K}$ be a pair of nonnegative integral matrices satisfying the conditions above. We can start from any vertex (or hyperedges) o of type-i and construct an infinite hypertree rooted at o with each vertex and hyperedge labeled with the respective type according to the rules specified by the branching matrices $\boldsymbol{D}$ and $\boldsymbol{K}$. Since $d_{ij} = 0$ if and only if $k_{ji} = 0$, the construction is always possible. Since $\sum_j d_{ij} = d + 1$ and $\sum_i k_{ji} = k + 1$, the resulting infinite hypertree must be k-uniform and d-regular. Since $\boldsymbol{DK}$ and $\boldsymbol{KD}$ are irreducible, no matter how we choose the type for the root o, the resulting hypertree contains all types of vertices and hyperedges.

We can then construct an automorphism group $\mathbb{G}$ on $\mathbb{T}_{k,d}$ according with orbits being the types just specified. For every pair of vertices (or hyperedges) u, v with the same type, by generating the hypertree according to $\boldsymbol{D}$, $\boldsymbol{K}$ starting from u and v respectively, we obtain an automorphism $\phi_{u \to v}$ on $\mathbb{T}_{k,d}$ which maps u to v and preserves the types of all vertices and hyperedges. Let $\mathbb{G} = \langle \{\phi_{u \to v} \mid \forall u, v \text{ with the same type}\} \rangle$ be the group generated from all such automorphisms. Then $\boldsymbol{D}$ and $\boldsymbol{K}$ are branching matrices for automorphism group $\mathbb{G}$ on $\mathbb{T}_{k,d}$. ◀

3.2   Extremal Gibbs measures

Consider a special automorphism group $\widehat{\mathbb{G}}$ on $\mathbb{T}_{k,d}$ defined by the following branching matrices $(\widehat{\boldsymbol{D}}, \widehat{\boldsymbol{K}})$. Assume that there are two vertex-types and two hyperedge-types, both denoted as $\{+, -\}$, and the branching matrices are defined as $\widehat{\boldsymbol{D}} = \begin{bmatrix} 1 & d \\ d & 1 \end{bmatrix}$ and $\widehat{\boldsymbol{K}} = \begin{bmatrix} k & 1 \\ 1 & k \end{bmatrix}$, i.e.:

1. every '$\pm$'-vertex is incident to a '$\pm$'-hyperedge and d '$\mp$'-hyperedges;
2. every '$\pm$'-hyperedge contains k '$\pm$'-vertices and a '$\mp$'-vertex.

See Figure 2 for an illustration.

Fix a '+'-vertex v in $\mathbb{T}_{k,d}$ as the root. Let μ^+ (resp. μ^-) be the Gibbs measure on $\mathbb{T}_{k,d}$ defined by conditioning on all vertices to be occupied for the t-th '+'-vertices (resp. '-'-vertices) along all path from the root and taking the weak limit as $t \to \infty$. Note that for the 2-coloring given by $\widehat{\boldsymbol{D}}$ and $\widehat{\boldsymbol{K}}$, on any path any '$\pm$'-vertex has a '$\mp$'-vertex within 2 steps,
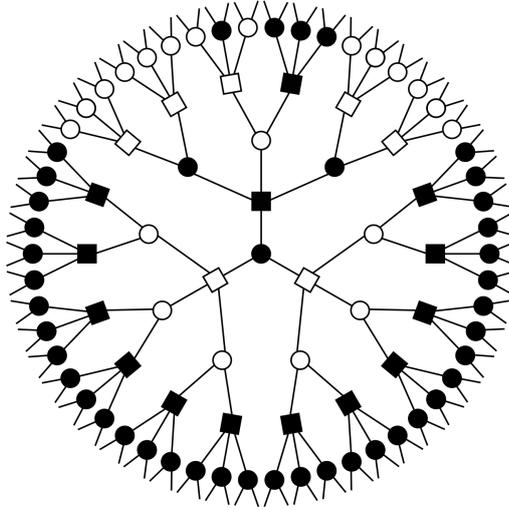
■ **Figure 2** Classifying vertices and hyperedges of $\mathbb{T}_{3,2}$ into two types '+'(black) and '−'(white). The hypergraph is represented as its incidence graph where circles stand for vertices and squares stand for hyperedges.

so the limiting sequence is well-defined. And by symmetry, starting from a root of type-'−' gives the same pair of measures.

The $\mu^{\pm}$ generalize the extremal semi-translation-invariant Gibbs measures on infinite regular trees. For hypertree $\mathbb{T}_{k,d}$ with $k \geq 2$, there are no parity-preserving automorphisms. Nevertheless, the symmetry given by $\widehat{D}$ and $\widehat{K}$ generalizes the parity-preserving automorphisms to hypertrees and has the similar phase-transition as semi-translation-invariant Gibbs measures on trees.

The $\mu^{\pm}$ are simple and are $\widehat{\mathbb{G}}$-translation-invariant for the automorphism group $\widehat{\mathbb{G}}$ with orbits given by $\widehat{D}$ and $\widehat{K}$. In fact, they are extremal $\widehat{\mathbb{G}}$-translation-invariant Gibbs measures on $\mathbb{T}_{k,d}$. We will see that the model has uniqueness if and only if $\mu^+ = \mu^-$.

3.3 Uniqueness of Gibbs measures

▶ **Lemma 12.** *Let μ be a simple Gibbs measure on independent sets of $\mathbb{T}_{k,d}$. Let v be a vertex in $\mathbb{T}_{k,d}$ and v_{ij} the j-th vertex (besides v) in the i-th hyperedge incident to v, for $i = 1, 2, \ldots, d+1$ and $j = 1, 2, \ldots, k$. Let $p_v = \mu[\, v \text{ is occupied}\,]$ and $p_{v_{ij}} = \mu[\, v_{ij} \text{ is occupied}\,]$. It holds that*

$$p_v = \lambda(1 - p_v)^{-d} \prod_{i=1}^{d+1} \left(1 - p_v - \sum_{j=1}^{k} p_{v_{ij}} \right). \tag{2}$$

Proof. Since μ is a Gibbs measure, for any vertex v in $\mathbb{T}_{k,d}$, it holds that

$$p_v = \mu[\, v \text{ is occupied}\,] = \frac{\lambda}{1 + \lambda} \cdot \mu[\, \text{all the neighbors of } v \text{ are unoccupied}\,]$$

On the other hand, since μ is simple, conditioning on the root being unoccupied the sub-

hypertrees are independent of each other, thus

$\mu[\text{all the neighbors of } v \text{ are unoccupied}]$

$=\mu[\,v \text{ is occupied}\,] \cdot \mu[\text{all the neighbors of } v \text{ are unoccupied} \mid v \text{ is occupied}]$

$\qquad + \mu[\,v \text{ is unoccupied}\,] \displaystyle\prod_{i=1}^{d+1} \mu[\forall 1 \le j \le k, v_{ij} \text{ is unoccupied} \mid v \text{ is unoccupied}]$

$=p_v + (1 - p_v) \displaystyle\prod_{i=1}^{d+1} \left(1 - \sum_{j=1}^{k} \mu[\,v_{ij} \text{ is occupied} \mid v \text{ is unoccupied}\,] \right).$

Note that for any two adjacent vertices v, v_{ij}, we have $\mu[\,v_{ij} \text{ is occupied}\,] = \mu[\,v_{ij} \text{ is occupied} \mid v \text{ is unoccupied}\,] \cdot \mu[\,v \text{ is unoccupied}\,]$, thus

$$\mu[\,v_{ij} \text{ is occupied} \mid v \text{ is unoccupied}\,] = \frac{\mu[\,v_{ij} \text{ is occupied}\,]}{1 - \mu[\,v \text{ is occupied}\,]} = \frac{p_{v_{ij}}}{1 - p_v}.$$

The lemma follows by combining everything together. ◀

Equation (2) gives an infinite system involving all vertices in $\mathbb{T}_{k,d}$. If the simple Gibbs measure μ is $\mathbb{G}$-translation-invariant for some automorphism group $\mathbb{G}$ on $\mathbb{T}_{k,d}$, the marginal probability $p_v = \mu[\,v \text{ is occupied}\,]$ depends only on the type (orbit) of v.

▶ **Corollary 13.** *Let μ be a simple $\mathbb{G}$-translation-invariant Gibbs measure on $\mathbb{T}_{k,d}$ with branching matrices $\boldsymbol{D}^{\tau_v \times \tau_e} = [d_{ij}]$ and $\boldsymbol{K}^{\tau_e \times \tau_v} = [k_{ji}]$. For every $i \in [\tau_v]$, let $p_i = \mu[\,v \text{ is occupied}\,]$ for vertex v in $\mathbb{T}_{k,d}$ of type-i. It holds for every $s \in [\tau_v]$ that*

$$p_s = \lambda (1 - p_s)^{-d} \prod_{j \in [\tau_e]} \left(1 - \sum_{i \in [\tau_v]} k_{ji} \cdot p_i \right)^{d_{ij}}.$$

Applying with the branching matrices $\widehat{\boldsymbol{D}}$ and $\widehat{\boldsymbol{K}}$ defined in Section 3.2, the system in Corollary 13 becomes

$$\begin{cases} p_+ = \lambda (1 - p_+)^{-d} (1 - k\,p_+ - p_-)(1 - p_+ - k\,p_-)^d, \\ p_- = \lambda (1 - p_-)^{-d} (1 - k\,p_- - p_+)(1 - p_- - k\,p_+)^d. \end{cases}$$

Let $x = \frac{kp_+}{1 - p_- - k\,p_+}$ and $y = \frac{kp_-}{1 - p_+ - k\,p_-}$. The system becomes $\begin{cases} y = f(x) \\ x = f(y) \end{cases}$, where $f(x) = \frac{k\lambda}{(1+x)^d}$ is the hardcore tree-recursion. Since $f(x)$ is positive and decreasing in x, it follows that there is a unique positive $\hat{x}$ such that $\hat{x} = f(\hat{x})$, which means there is always a unique simple translation-invariant Gibbs measure on $\mathbb{T}_{k,d}$. It is well-known (see [9] and [18, 32]) the system has three distinct solutions $(\hat{x}, \hat{x}), (x^+, x^-)$ and (x^-, x^+) where $0 < x^- < \hat{x} < x^+$, when $k\lambda > d^d/(d-1)^{d+1}$, i.e. $\lambda > \lambda_c(\mathbb{T}_{k,d}) = \frac{d^d}{k(d-1)^{d+1}}$; and the three solutions collide into a unique solution $(\hat{x}, \hat{x})$ when $\lambda \le \lambda_c(\mathbb{T}_{k,d})$, which means there is a unique simple $\widehat{\mathbb{G}}$-translation-invariant Gibbs measure on $\mathbb{T}_{k,d}$ if and only if $\lambda \le \lambda_c(\mathbb{T}_{k,d})$. Recall that $\mu^{\pm}$ are simple and are extremal $\widehat{\mathbb{G}}$-translation-invariant Gibbs measures, and hence it also holds that $\mu^+ = \mu^-$ if and only if $\lambda \le \lambda_c(\mathbb{T}_{k,d})$, therefore, it holds that $\lambda_c(\mathbb{T}_{k,d}) = \lambda_c(\mathbb{T}_{k,d}^{\widehat{\mathbb{G}}})$. In particular if $\lambda > \lambda_c(\mathbb{T}_{k,d})$, then $\mu^+ \ne \mu^-$ and the Gibbs measure on $\mathbb{T}_{k,d}$ is non-unique.

To complete the proof of Theorem 10, we only need to show the Gibbs measure on $\mathbb{T}_{k,d}$ is unique if $\lambda \le \lambda_c(\mathbb{T}_{k,d})$. This is implied by the weak spatial mixing on $\mathbb{T}_{k,d}$ when $\lambda \le \lambda_c$, proved later in Theorem 21. With the weak spatial mixing on $\mathbb{T}_{k,d}$, the uniqueness of the Gibbs measure is implied by a generic equivalence between weak spatial mixing and uniqueness of Gibbs measure (see e.g. [33]).

4 The hypergraph self-avoiding walk tree

We call a hypergraph a hypertree if its incidence graph has no cycles. Let $\mathcal{T} = (V, E)$ be a rooted hypertree with vertex v as its root. We assume that root v is incident to d distinct hyperedges $e_1, e_2, \ldots, e_d$, such that for $i = 1, 2, \ldots, d$,

- $|e_i| = k_i + 1$; and
- $e_i = \{v, v_{i1}, v_{i2}, \ldots, v_{ik_i}\}$.

For $1 \leq i \leq d$ and $1 \leq j \leq k_i$, let $\mathcal{T}_{ij}$ be the sub-hypertree rooted at v_{ij}. Recall that all hypertrees considered by us satisfy the property that any two hyperedges share at most one common vertex, thus all v_{ij} are distinct and the sub-hypertrees $\mathcal{T}_{ij}$ are disjoint.

Let $\Lambda \subset V$. Let $\sigma_\Lambda \in \{0, 1\}^\Lambda$ be a configuration indicating an independent set partially specified on vertex set Λ, and for each $1 \leq i \leq d$ and $1 \leq j \leq k_i$, let $\sigma_{\Lambda_{ij}}$ be the restriction of σ_Λ on the sub-hypertree $\mathcal{T}_{ij}$. Consider the ratios of marginal probabilities:

$$R_{\mathcal{T}}^{\sigma_\Lambda} = p_{\mathcal{T},v}^{\sigma_\Lambda} / \left(1 - p_{\mathcal{T},v}^{\sigma_\Lambda}\right) \quad \text{and} \quad R_{\mathcal{T}_{ij}}^{\sigma_{\Lambda_{ij}}} = p_{\mathcal{T}_{ij},v_{ij}}^{\sigma_{\Lambda_{ij}}} / \left(1 - p_{\mathcal{T}_{ij},v_{ij}}^{\sigma_{\Lambda_{ij}}}\right).$$

The following recursion can be easily verified due to the disjointness between sub-hypertrees:

$$R_{\mathcal{T}}^{\sigma_\Lambda} = \lambda \prod_{i=1}^{d} \frac{1}{1 + \sum_{j=1}^{k_i} R_{\mathcal{T}_{ij}}^{\sigma_{\Lambda_{ij}}}}. \tag{3}$$

This is the "tree recursion" for hypergraph independent sets. The tree recursions for the hardcore model [34] and the monomer-dimer model [1] can both be interpreted as special cases.

For general hypergraphs which are not trees, we construct a hypergraph version of *self-avoiding-walk tree*, which allows computing marginal probabilities in arbitrary hypergraphs with the tree recursion. Moreover, we show that the uniform regular hypertree is the worst case for SSM among all hypergraphs of bounded maximum edge-size and bounded maximum degree.

▶ **Theorem 14.** *For any positive integers k, d and any positive λ, if the independent sets of $\mathbb{T}_{k,d}$ with activity λ exhibit strong spatial mixing with rate $\delta(\cdot)$, then the independent sets of any hypergraph of maximum edge size at most $(k + 1)$ and maximum degree at most $(d + 1)$, with activity λ, exhibit strong spatial mixing with the same rate $\delta(\cdot)$.*

Under duality, the same holds for the hypergraph matchings.

We then define the hypergraph self-avoiding walk tree. A walk in a hypergraph $\mathcal{H} = (V, E)$ is a sequence $(v_0, e_1, v_1, \ldots, e_\ell, v_\ell)$ of alternating vertices and hyperedges such that every two consecutive vertices v_{i-1}, v_i are incident to the hyperedge e_i between them. A walk $w = (v_0, e_1, v_1, \ldots, e_\ell, v_\ell)$ is called *self-avoiding* if:

- $w = (v_0, e_1, v_1, \ldots, e_\ell, v_\ell)$ forms a simple path in the incidence graph of $\mathcal{H}$; and
- for every $i = 1, 2, \ldots, \ell$, vertex v_i is incident to none of $\{e_1, e_2, \ldots, e_{i-1}\}$.

Note that the second requirement is new to the hypergraphs.

A self-avoiding walk $w = (v_0, e_1, v_1, \ldots, e_\ell, v_\ell)$ can be extended to a *cycle-closing* walk $w' = (v_0, e_1, v_1, \ldots, e_\ell, v_\ell, e', v')$ so that the suffix $(v_i, e_{i+1}, v_{i+1}, \ldots, e_\ell, v_\ell, e', v')$, for some $0 \leq i \leq \ell - 1$, of the walk forms a simple cycle in the incidence graph of $\mathcal{H}$. We call v' the cycle-closing vertex.

Given a hypergraph $\mathcal{H} = (V, E)$, an ordering of incident hyperedges at every vertex can be arbitrarily fixed, so that for any two hyperedges e_1, e_2 incident to a vertex u we use $e_1 <_u e_2$ to denote that e_1 is ranked higher than e_2 according to the ordering of hyperedges incident
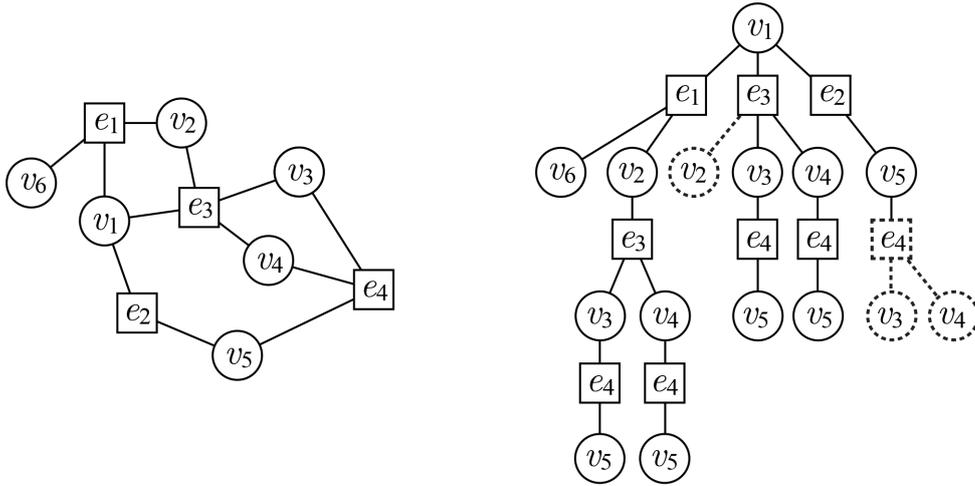
Figure 3 The construction of $\mathcal{T}_{\mathrm{SAW}}$. On the left is a hypergraph $\mathcal{H}$ and on the right is $\mathcal{T}_{\mathrm{SAW}}(\mathcal{H}, v_1)$, both drawn as incident graphs. The ordering of the hyperedges incident to each vertex in $\mathcal{H}$ is given by the subscripts. Each vertex or hyperedge in $\mathcal{T}_{\mathrm{SAW}}$ is labeled by the name of the vertex or hyperedge to which it is identified in $\mathcal{H}$. Dashed vertices are the ones deleted according to the ordering of incident hyperedges at the cycle-closing vertices. Dashed hyperedge is deleted because its size becomes 1.

to u. With this local ordering of hyperedges, given any vertex $v \in V$, a rooted hypertree $\mathcal{T} = \mathcal{T}_{\mathrm{SAW}}(\mathcal{H}, v)$, called the *self-avoiding walk (SAW) tree*, is constructed as follows:
1. Every vertex of $\mathcal{T}$ corresponds to a distinct self-avoiding walk in $\mathcal{H}$ originating from v, where the root corresponds to the trivial walk (v).
2. For any vertex u in $\mathcal{T}$, which corresponds to a self-avoiding walk $w = (v, e_1, v_1 \ldots, e_\ell, v_\ell)$, we partition all self-avoiding walks $w' = (v, e_1, v_1 \ldots, e_\ell, v_\ell, e', v')$ in $\mathcal{H}$ which extends w, into sets according to which hyperedge they use to extend the original walk w, so that self-avoiding walks within the same sets extends w with the same hyperedge e'. For every set, we create a distinct hyperedge in $\mathcal{T}$ incident to u which contains the children of u corresponding to the self-avoiding walks within that set.
3. We further modify the hypertree $\mathcal{T}$ obtained from the above two steps according to how cycles are closed. For any vertex u in $\mathcal{T}$ corresponding to a self-avoiding walk $w = (v, e_1, v_1 \ldots, e_\ell, v_\ell)$ which can be extended to a cycle-closing walk $w' = (v, e_1, v_1 \ldots, e_\ell, v_\ell, e', v')$ such that $v' \in \{v, v_1, \ldots, v_{\ell-1}\}$, denoted by e'' the hyperedge in w starting that cycle, if it holds that $e' <_{v'} e''$, i.e. the hyperedge ending the cycle is ranked higher than the hyperedge starting the cycle by the cycle-closing vertex, then vertex u along with all its descendants in $\mathcal{T}$ are deleted from $\mathcal{T}$. Any hyperedges whose size becomes 1 because of this step are also deleted from $\mathcal{T}$.

The construction is illustrated in Figure 3.

We consider the Gibbs measure of a rooted hypertree $\mathcal{T}$ with activity λ, and use $\mathbb{P}_{\mathcal{T}}^{\sigma_\Lambda}$ to denote the marginal probability of the root of $\mathcal{T}$ being occupied conditioning on σ_Λ.

Note that each vertex u in $\mathcal{T}_{\mathrm{SAW}}(\mathcal{H}, v)$ can be naturally identified (many-to-one) to the vertex in $\mathcal{H} = (V, E)$ at which the self-avoiding walk corresponding to u ends, thus a configuration σ_Λ partially specified on a subset $\Lambda \subset V$ of vertices in $\mathcal{H}$ can be directly translated to a partially specified configuration in $\mathcal{T}_{\mathrm{SAW}}(\mathcal{H}, v)$ through the one-to-many association. We abuse the notation and still denote the resulting configuration in $\mathcal{T} = \mathcal{T}_{\mathrm{SAW}}(\mathcal{H}, v)$ as σ_Λ, thus $\mathbb{P}_{\mathcal{T}}^{\sigma_\Lambda}$ is well-defined.

▶ **Theorem 15.** *Let $\mathcal{H} = (V, E)$ be a hypergraph and $\lambda > 0$. For any $v \in V$, $\Lambda \subseteq V$ and $\sigma_\Lambda \in \{0,1\}^\Lambda$, it holds that $p_{\mathcal{H},v}^{\sigma_\Lambda} = \mathbb{P}_{\mathcal{T}}^{\sigma_\Lambda}$ where $\mathcal{T} = \mathcal{T}_{\mathrm{SAW}}(\mathcal{H}, v)$.*

Proof. The proof follows the same routine as that of Weitz [34], with some extra cares to be taken to avoid the complications caused by hypergraphs.

Denote $R_{\mathcal{H},v}^{\sigma_\Lambda}(\lambda) = p_{\mathcal{H},v}^{\sigma_\Lambda}/(1 - p_{\mathcal{H},v}^{\sigma_\Lambda})$ for the ratio between the probability that v in $\mathcal{H}$ is occupied and unoccupied conditioning on configuration σ_Λ of $\Lambda \subset V$. We write $R_{\mathcal{T}}^{\sigma_\Lambda} = R_{\mathcal{T},v}^{\sigma_\Lambda}$ when v is unambiguously the root of $\mathcal{T}$.

Let d be the degree of the root of $\mathcal{T}$. Suppose that there are k_i children contained in i-th child-edge, where the order is determined during the construction of $\mathcal{T}_{\mathrm{SAW}}(\mathcal{H}, v)$. $\mathcal{T}_{ij}$ is the subtree rooted at the j-th child in the i-th child-edge. Let $\Lambda_{ij} = \Lambda \cap \mathcal{T}_{ij}$ and $\sigma_{\Lambda_{ij}}$ be the restriction of σ_Λ on Λ_{ij}. Applying the tree recursion (3) for the self-avoiding walk tree $\mathcal{T}$, we have

$$R_{\mathcal{T}}^{\sigma_\Lambda} = \lambda \prod_{i=1}^{d} \frac{1}{1 + \sum_{i=1}^{k_i} R_{\mathcal{T}_{ij}}^{\sigma_{\Lambda_{ij}}}}, \tag{4}$$

This defines a recursive procedure for calculating $R_{\mathcal{T}}^{\sigma_\Lambda}$. The base cases are naturally defined when v lies in Λ, in which case $R_{\mathcal{T}}^{\sigma_\Lambda} = 0$ if v is fixed unoccupied or $R_{\mathcal{T}}^{\sigma_\Lambda} = \infty$ if it is fixed occupied, or when v has no child, in which case $R_{\mathcal{T}}^{\sigma_\Lambda} = \lambda$.

In the following we describe our procedure for calculating $R_{\mathcal{H},v}^{\sigma_\Lambda}$ at v in the original hypergraph $\mathcal{H}$. The problem comes that the ratio at different neighbors of v may still depend on each other when we fix the value at v since there may exist cycles in $\mathcal{H}$. We resolve this problem by editing the original hypergraph around v and imposing appropriate conditions for each neighbor of v.

Let $\mathcal{H}^v$ be the same hypergraph as $\mathcal{H}$ except that vertex $v \in V$ is substituted by d vertices $v_1, v_2, ..., v_d$, where d is the degree of v. Each vertex v_i is contained into a single hyperedge e_i, where e_i is the i-th hyperedge connecting v, and the order here is the same as the one determined in the definition of $\mathcal{T}_{\mathrm{SAW}}(\mathcal{H}, v)$. At the same time, we associated each v_i with an activity of $\lambda^{1/d}$ rather than λ. It is now clear to see that an independent set in $\mathcal{H}$ with v occupied has the same weight as the corresponding independent set in $\mathcal{H}^v$ with all the v_i occupied, and so is the case when v is unoccupied. Therefore, $R_{\mathcal{H},v}^{\sigma_\Lambda}$ equals to the ratio between the probabilities in $\mathcal{H}^v$ with all v_i ($1 \le i \le d$) being occupied and unoccupied, conditioning on σ_Λ. Let τ_i be the configuration for vertex v_i in which the values of v_j are fixed to occupied if $j < i$ and unoccupied if $j > i$. We can then write this in a form of telescopic product:

$$R_{\mathcal{H},v}^{\sigma_\Lambda} = \prod_{i=1}^{d} R_{\mathcal{H}^v,v_i}^{\sigma_\Lambda \tau_i},$$

where $\sigma_\Lambda \tau_i$ means the combination of the two configurations σ_Λ and τ_i.

We can obtain the value of $R_{\mathcal{H}^v,v_i}^{\sigma_\Lambda \tau_i}$ by further fix vertices in e_i, the hyperedge containing v_i. Since now v_i is contained only in e_i, we can see that

$$R_{\mathcal{H}^v,v_i}^{\sigma_\Lambda \tau_i} = \frac{\lambda^{1/d}}{1 + \sum_{j=1}^{k_i} R_{\mathcal{H}^v/v_i,u_{ij}}^{\sigma_\Lambda \tau_i \rho_{ij}}},$$

where k_i is the number of the vertices other than v_i which is incident to e_i and ρ_{ij} is the configuration at vertices of e_i in which all the vertices $u_{ij'}$ other than u_{ij} are fixed to unoccupied.

Combining above two equations, we get a recursive procedure for calculating $R^{\sigma_\Lambda}_{\mathcal{H},v}$ in the same manner that equation (4) has:

$$R^{\sigma_\Lambda}_{\mathcal{H},v} = \lambda \prod_{i=1}^{d} \frac{1}{1 + \sum_{j=1}^{k_i} R^{\sigma_\Lambda \tau_i \rho_{ij}}_{\mathcal{H}^v/v_i, u_{ij}}}. \tag{5}$$

Notice that the recursion does terminate, since the number of unfixed vertices reduces at least by one in each step because in calculating $R^{\sigma_\Lambda \tau_i \rho_{ij}}_{\mathcal{H}^v/v_i, u_{ij}}$ all copies $v_{i'}$ of v is either fixed (when $i' \neq i$) or erased (when $i' = i$) from the hypergraph $\mathcal{H}^v/v_i$.

We now show that the procedure described above for calculating $R^{\sigma_\Lambda}_{\mathcal{H},v}$ results in the same value as using the hypertree procedure for $\mathcal{T}_{\text{SAW}}(\mathcal{H}, v)$ with corresponding condition of σ_Λ imposed on it. First notice that the calculation carried out by the two procedure is the same, since they share the same function (Equation (4) and (5)) when we view them as recursive calls. Furthermore, we have the same stopping values for the both recursive procedures. During constructing $\mathcal{T}_{\text{SAW}}(\mathcal{H}, v)$, if node u corresponding to walk is not included in the hypertree, which is equivalent to fix u to unoccupied in the sense of causing the same effect on the ratio of occupation to its parent node. And when node u in the hypertree corresponding to a self-avoiding walk $w = (v, e_1, v_1 \ldots, e_\ell, v_\ell)$, with that w can be extended as $w' = (w, e_{\ell+1}, v_{\ell+1})$ to a cycle-closing vertex $v_{\ell+1} = v_i$ for some $0 \leq i < \ell$ via a new hyperedge $e_{\ell+1} \notin \{e_0, e_1, \ldots, e_\ell\}$, and $e_{\ell+1} <_{v_i} e_i$, then the node u along with all its descendants are deleted. This gives the equivalent effect to parent node of u as if u is fixed to unoccupied, or one of the children of u (i.e. the node corresponding to w') to occupied, which is what we did to fix the vertices v_j for $j < i$ in τ_i. Eliminating a hyperedge with no child also does not affect the final value of $R^{\sigma_\Lambda}_{\mathcal{T}}$.

Thus, what is left to complete the proof is to show that the hypertree $\mathcal{T}_{\text{SAW}}(\mathcal{H}^v/v_i, u_{ij})$ with $(\sigma_\Lambda \tau_i \rho_{ij})$'s corresponding condition imposed on it is exactly the same as the subtree of $\mathcal{T}_{\text{SAW}}(\mathcal{H}, v)$ rooted at the j-th child vertex of the i-th child-edge of the root with σ_Λ's corresponding condition imposed on it. This is enough because then the resulting values are the same for both procedures by induction. The observation is that both trees are the hypertree of all self-avoiding walks in $\mathcal{H}$ starting at u_{ij}, except that $\mathcal{T}_{\text{SAW}}(\mathcal{H}^v/v_i, u_{ij})$ has some extra vertices which are fixed to be occupied or unoccupied depending on whether the corresponding walk reaches v via a higher or lower ranked hyperedge, or reaches i-th hyperedge of v, which results in the same probability of occupation at the root. ◀

A hypergraph $\mathcal{H}$ is a sub-hypergraph of another hypergraph $\mathcal{G}$ if the incidence graph of $\mathcal{H}$ is a subgraph of that of $\mathcal{G}$, and for hypertrees this is samely defined. Note that for hypergraphs, a subgraph is not necessarily formed by a sub-collection of hyperedges, but maybe also by sub-hyperedges. The $\mathcal{T}_{\text{SAW}}$ of a hypergraph $\mathcal{H}$ with maximum edge-size at most $k + 1$ and maximum degree at most $d + 1$ is sub-hypertree of $\mathbb{T}_{k,d}$.

▶ **Proposition 16.** *Let $\mathcal{T}_0 = (V_0, E_0)$ be a rooted hypertree and $\mathcal{T} = (V, E)$ its sub-hypertree with the same root. For any $\Lambda \subseteq V$ and any $\sigma_\Lambda \in \{0,1\}^\Lambda$, there exists a configuration $\sigma_{\Lambda_0} \in \{0,1\}^{\Lambda_0}$ for $\Lambda \subseteq \Lambda_0 \subseteq V_0$, extending the configuration σ_Λ, such that $\mathbb{P}^{\sigma_\Lambda}_{\mathcal{T}} = \mathbb{P}^{\sigma_{\Lambda_0}}_{\mathcal{T}_0}$.*

The configuration σ_{Λ_0} just extends σ_Λ by fixing all the vertices missing in $\mathcal{T}$ (actually only those who are closest to the root along each path) to be unoccupied.

Theorem 14 follows immediately from Theorem 15 and Proposition 16.

Proof of Theorem 14. Given any hypergraph $\mathcal{H}$ of maximum edge-size at most $(k + 1)$ and maximum degree at most $(d + 1)$, by Theorem 15 we have $|p^{\sigma_\Lambda}_{\mathcal{H},v} - p^{\tau_\Lambda}_{\mathcal{H},v}| = |\mathbb{P}^{\sigma_\Lambda}_{\mathcal{T}} - \mathbb{P}^{\tau_\Lambda}_{\mathcal{T}}|$ where

$\mathcal{T} = \mathcal{T}_{\text{SAW}}(\mathcal{H}, v)$. The distance from the root v to any vertex u in $\mathcal{T}$ is no shorter than the distance $\mathcal{H}$ between v and the vertex in $\mathcal{H}$ to which u is identified. So the SSM with rate $\delta(\cdot)$ on $\mathcal{T}$ implies that on the hypergraph $\mathcal{H}$.

Since $\mathcal{H}$ has maximum edge-size at most $k + 1$ and maximum degree at most $d + 1$, its SAW-tree $\mathcal{T} = \mathcal{T}_{\text{SAW}}(\mathcal{H}, v)$ is a sub-hypertree of $\mathbb{T}_{k,d}$. Thus by Proposition 16, we have $|\mathbb{P}_{\mathcal{T}}^{\sigma_\Lambda} - \mathbb{P}_{\mathcal{T}}^{\tau_\Lambda}| = |\mathbb{P}_{\mathbb{T}_{k,d}}^{\sigma_{\Lambda_0}} - \mathbb{P}_{\mathbb{T}_{k,d}}^{\tau_{\Lambda_0}}|$ for some $\sigma_{\Lambda_0}, \tau_{\Lambda_0}$ extending $\sigma_\Lambda, \tau_\Lambda$. The SSM on $\mathbb{T}_{k,d}$ with rate $\delta(\cdot)$ implies that on $\mathcal{T}$, which implies the same on the original hypergraph $\mathcal{H}$.   ◀

5   Strong spatial mixing

In this section, we show that for independent sets of the infinite $(k+1)$-uniform $(d+1)$-regular hypertree $\mathbb{T}_{k,d}$, weak spatial mixing implies strong spatial mixing at almost the same rate.

▶ **Theorem 17.** *For every positive integers d, k and any λ, if the independent sets of the infinite $(k+1)$-uniform $(d+1)$-regular hypertree $\mathbb{T}_{k,d}$ with activity λ exhibits weak spatial mixing with rate $\delta(\cdot)$ then it also exhibits strong spatial mixing with rate $\frac{(1+\lambda)\left(\lambda+(1+k\lambda)^{d+1}\right)}{\lambda}\delta(\cdot)$.*

By Theorem 14, this implies the strong spatial mixing with the same rate on all hypergraphs of maximum degree at most $d + 1$ and maximum size of hyperedges at most $k + 1$.

Unlike most known strong spatial mixing results, where the spatial mixing is usually established by an analytic approach with help of potential functions, our proof of Theorem 17 adopts the combinatorial argument used in Weitz's original proof of SSM for the hardcore model [34]. Weitz's approach gives us a stronger result: It explicitly gives the extremal case for WSM as well as SSM among a family of hypergraphs with bounded maximum degree and bounded maximum edge-size. It can also easily give us the SSM behavior when at the critical threshold.

Assume the hypertree $T = \mathbb{T}_{k,d}$ is rooted at some vertex v. For $\ell > 0$, let R_ℓ^+ and R_ℓ^- denote the respective maximum and minimum values of R_T^σ achieved by a boundary condition σ that fixes the states of all vertices at level ℓ. By the monotonicity of the tree recursion, it is easy to see that R_ℓ^+ (or R_ℓ^-) is computed by the tree recursion with initial values at all vertices at level ℓ to be ∞ (or 0) if ℓ is even, and 0 (or ∞) if ℓ is odd, with the root v being at level 0.[2]

It is easy to see that fixing a vertex u in T to be occupied has the same effect as fixing u's parent to be unoccupied, therefore to prove SSM, it is sufficient to prove the decay of correlation conditioning on a subset of vertices in T fixed to be unoccupied. Another key observation from the tree recursion is that fixing a vertex u in T to be unoccupied has the same effect as having a local activity $\lambda_u = 0$ at vertex u. Now consider a vector $\vec{\lambda}$ that assigns every vertex u in $T = \mathbb{T}_{k,d}$ a local activity λ_u. Let $R_\ell^+(\vec{\lambda})$ and $R_\ell^-(\vec{\lambda})$ be accordingly defined as the respective extremal values of $R_T^\sigma(\vec{\lambda})$ achieved by boundary conditions σ fixing all vertices at level ℓ in the tree $T = \mathbb{T}_{k,d}$ equipped with the nonuniform activities $\vec{\lambda}$. Clearly, by the same monotonicity, $R_\ell^\pm(\vec{\lambda})$ can be computed from the tree recursion with a nonuniform activities $\vec{\lambda}$ with the same settings of initial values as the uniform case $R_\ell^\pm = R_\ell^\pm(\lambda)$.

The following theorem shows that basically the decay of correlation is dominated by the uniform activity case.

---

2   Note that although the all-∞ initial values corresponds to a boundary condition σ that fixes all vertices at level ℓ to be occupied, which may no longer be a valid independent set in the hypertree, the $R_\ell^\pm$ achieved by this choice of initial values is actually the same as the R_T^σ with a boundary condition σ that fixes exactly one vertex per hyperedge to be occupied at level ℓ.

▶ **Theorem 18.** *Fix an arbitrary $\lambda \geq 0$. Let $\vec{\lambda}$ be an assignment of activities to vertices of $\mathbb{T}_{k,d}$ such that $0 \leq \lambda_v \leq \lambda$ for every $v \in \mathbb{T}_{k,d}$. For every $\ell \geq 1$ we have*

$$\frac{R_\ell^+(\vec{\lambda})}{R_\ell^-(\vec{\lambda})} \leq \frac{R_\ell^+}{R_\ell^-} .$$

Translated to the language of subtrees, the theorem means that the extremal case of WSM among a family hypertrees with bounded maximum degree and maximum edge-size, is given by the uniform regular tree with the highest degree and edge-size in the family. Technically, Theorem 18 measures the decay of correlation in terms of $\log R = \log \frac{p}{1-p}$. Note that for $\ell \geq 2$, it always holds that $p_\ell^+(\lambda) \leq \frac{\lambda}{1+\lambda}$ and $p_\ell^-(\lambda) \geq \frac{\lambda}{\lambda+(1+k\lambda)^{d+1}}$, where $R_\ell^\pm = \frac{p_\ell^\pm}{1-p_\ell^\mp}$. Theorem 18 implies Theorem 17.

We now consider a slightly different hypertree which is exactly the same as $\mathbb{T}_{k,d}$ except that the degree of root is d. Denote this hypertree as $\widehat{\mathbb{T}}_{k,d}$.

▶ **Lemma 19.** *For every integer $\ell \geq 1$ and any assignment of activities $\vec{\lambda}$ to vertices of $\widehat{\mathbb{T}}_{k,d}$ such that $0 \leq \lambda_v \leq \lambda$ for every vertex v, the following two inequalities hold:*

$$\frac{R_\ell^+(\vec{\lambda})}{R_\ell^-(\vec{\lambda})} \leq \frac{R_\ell^+}{R_\ell^-}, \tag{6}$$

$$\frac{1 + kR_\ell^+(\vec{\lambda})}{1 + kR_\ell^-(\vec{\lambda})} \leq \frac{1 + kR_\ell^+}{1 + kR_\ell^-}, \tag{7}$$

with the convention $0/0 = 1$ and $\infty = \infty$.

Proof of Lemma 19. The proof is by an induction on ℓ. The proof is similar to that of Weitz [34] except for the parts dealing with hyperedges.

First consider the exceptional cases when the denominators in (6) may be zero. Assume $R_\ell^+(\vec{\lambda}) = R_\ell^-(\vec{\lambda}) = 0$, which only happens when the activity of the root is zero. We adopt the convention that $\frac{R_\ell^+(\vec{\lambda})}{R_\ell^-(\vec{\lambda})} = 1$. Assume $R_\ell^- = 0$, which only occurs when $\ell = 1$. Then $R_\ell^-(\vec{\lambda}) = 0$ also holds, and by convention we have $\frac{R_\ell^+(\vec{\lambda})}{R_\ell^-(\vec{\lambda})} = \frac{R_\ell^+}{R_\ell^-} = \infty$. Note that these conventions are consistent with our induction, such that assuming the induction hypothesis $\frac{R_\ell^+(\vec{\lambda})}{R_\ell^-(\vec{\lambda})} \leq \frac{R_\ell^+(\lambda)}{R_\ell^-(\lambda)}$, for any k assignments of activities $0 \leq \vec{\lambda}_1, \vec{\lambda}_2, ..., \vec{\lambda}_k \leq \lambda$, there exists $\alpha \geq 0$ such that $\sum_{i=1}^k R_\ell^-(\vec{\lambda}_i) = \alpha k R_\ell^-$ and $\sum_{i=1}^k R_\ell^+(\vec{\lambda}_i) \leq \alpha k R_\ell^+$.

For the basis, $\ell = 1$. We have $R_\ell^-(\vec{\lambda}) \geq R_\ell^- = 0$, $R_\ell^+ = \lambda$, and $R_\ell^+(\vec{\lambda}) = \lambda_r$ where λ_r is the activity of the root. The hypotheses (6) and (7) are true since $\lambda_r \leq \lambda$.

Assume (6) and (7) are true for an $\ell \geq 1$. We will show that they are true for $\ell + 1$. The following recursion holds

$$\frac{R_{\ell+1}^+(\vec{\lambda})}{R_{\ell+1}^-(\vec{\lambda})} = \prod_{i=1}^d \frac{1 + \sum_{j=1}^k R_\ell^+(\vec{\lambda}_{ij})}{1 + \sum_{j=1}^k R_\ell^-(\vec{\lambda}_{ij})} = \prod_{i=1}^d \frac{\sum_{j=1}^k (1 + kR_\ell^+(\vec{\lambda}_{ij}))}{\sum_{j=1}^k (1 + kR_\ell^-(\vec{\lambda}_{ij}))},$$

where $\vec{\lambda}_{ij}$ stands for the restriction of the assignment $\vec{\lambda}$ to the subtree of $\mathbb{T}_{k,d}$ rooted at the j-th child in the i-th edge incident to the root. By induction hypothesis (7), we have $\frac{1+kR_\ell^+(\vec{\lambda}_{ij})}{1+kR_\ell^-(\vec{\lambda}_{ij})} \leq \frac{1+kR_\ell^+}{1+kR_\ell^-}$, so immediately,

$$\frac{R_{\ell+1}^+(\vec{\lambda})}{R_{\ell+1}^-(\vec{\lambda})} = \prod_{i=1}^d \frac{\sum_{j=1}^k (1 + kR_\ell^+(\vec{\lambda}_{ij}))}{\sum_{j=1}^k (1 + kR_\ell^-(\vec{\lambda}_{ij}))} \leq \left(\frac{1 + kR_\ell^+}{1 + kR_\ell^-}\right)^d = \frac{R_{\ell+1}^+}{R_{\ell+1}^-},$$

where the inequality is due to the simple fact that if $a_i \geq b_i > 0$ and $\frac{a_i}{b_i} \leq t$ for all i, then $\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \leq t$.

This proves (6) for $\ell + 1$. Next we will prove (7). Recall the tree recursions:

$$R_{\ell+1}^{\pm}(\vec{\lambda}) = \lambda_r \prod_{i=1}^d \frac{1}{1 + \sum\limits_{j=1}^k R_\ell^{\mp}(\vec{\lambda}_{ij})} \quad \text{and} \quad R_{\ell+1}^{\pm} = \lambda \prod_{i=1}^d \frac{1}{1 + kR_\ell^{\mp}},$$

where λ_r is the local activity assigned by $\vec{\lambda}$ to the root r. Observe that if $\sum_{j=1}^k R_\ell^-(\vec{\lambda}_{ij}) \geq kR_\ell^-$ for all $i \in [d]$, then $R_{\ell+1}^+(\vec{\lambda}) \leq R_{\ell+1}^+$, which combined with (6) for $\ell + 1$ that we just proved above, would give us that $\frac{1+kR_{\ell+1}^+(\vec{\lambda})}{1+kR_{\ell+1}^-(\vec{\lambda})} \leq \frac{1+kR_{\ell+1}^+}{1+kR_{\ell+1}^-}$. In this good case, the hypothesis (7) easily holds for $\ell + 1$. We then show that the opposite case where $\sum_{j=1}^k R_\ell^-(\vec{\lambda}_{ij}) \leq kR_\ell^-$ for all i represents the worst possible case, and it is enough to prove the hypothesis (7) under this condition. To see this, assume to the contrary that for some i_0, $\sum_{j=1}^k R_\ell^-(\vec{\lambda}_{i_0 j}) > kR_\ell^-$. We then construct a $\vec{\lambda}'$ that satisfies $\sum_{j=1}^k R_\ell^-(\vec{\lambda}'_{i_0 j}) \leq kR_\ell^-$ and has an even worse ratio between $1 + kR_{\ell+1}^+$ and $1 + kR_{\ell+1}^-$. Let $\vec{\lambda}'$ be the same as $\vec{\lambda}$ except that for every $j \in [k]$, $\vec{\lambda}'_{i_0 j}$ is uniform and is equal to λ everywhere. Clearly, it holds that $\sum_{j=1}^k R_\ell^-(\vec{\lambda}'_{i_0 j}) = kR_\ell^-$. On the other hand, by the induction hypothesis, for every j we have $\frac{1+kR_\ell^+(\vec{\lambda}_{i_0 j})}{1+kR_\ell^-(\vec{\lambda}_{i_0 j})} \leq \frac{1+kR_\ell^+}{1+kR_\ell^-}$, and hence

$$\frac{1 + \sum_{j=1}^k R_\ell^+(\vec{\lambda}_{i_0 j})}{1 + \sum_{j=1}^k R_\ell^-(\vec{\lambda}_{i_0 j})} = \frac{\sum_{j=1}^k (1 + kR_\ell^+(\vec{\lambda}_{i_0 j}))}{\sum_{j=1}^k (1 + kR_\ell^-(\vec{\lambda}_{i_0 j}))} \leq \frac{1 + kR_\ell^+}{1 + kR_\ell^-},$$

where again the inequality uses the fact that if $a_i \geq b_i > 0$ and $\frac{a_i}{b_i} \leq t$ for all i, then $\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \leq t$.

Note that $\vec{\lambda}'$ only changes the activities of all the subtrees rooted by the the children in the i_0-th edge of the root. So we have

$$\frac{R_{\ell+1}^+(\vec{\lambda})}{R_{\ell+1}^-(\vec{\lambda})} = \prod_{i=1}^d \frac{1 + \sum_{j=1}^k R_\ell^+(\vec{\lambda}_{ij})}{1 + \sum_{j=1}^k R_\ell^-(\vec{\lambda}_{ij})} \leq \frac{R_{\ell+1}^+(\vec{\lambda}')}{R_{\ell+1}^-(\vec{\lambda}')},$$

$$\text{and} \quad R_{\ell+1}^+(\vec{\lambda}) = \lambda_r \prod_{i=1}^d \frac{1}{1 + \sum\limits_{j=1}^k R_l^-(\vec{\lambda}_{ij})} \leq R_{l+1}^+(\vec{\lambda}').$$

Combine the two inequalities, we have $\frac{1+kR_{\ell+1}^+(\vec{\lambda})}{1+kR_{\ell+1}^-(\vec{\lambda})} \leq \frac{1+kR_{\ell+1}^+(\vec{\lambda}')}{1+kR_{\ell+1}^-(\vec{\lambda}')}$, an even worse case. So for the rest we only need to consider the case in which for every i, $\sum_{j=1}^k R_\ell^-(\vec{\lambda}_{ij}) \leq kR_\ell^-$.

For every $1 \leq i \leq d$, we can choose $0 \leq \alpha_i \leq 1$ so that $\sum_{j=1}^k R_\ell^-(\vec{\lambda}_{ij}) = \alpha_i kR_\ell^-$. Fix i and by the induction hypothesis, for every $1 \leq j \leq k$ we have $\frac{R_\ell^+(\vec{\lambda}_{ij})}{R_\ell^-(\vec{\lambda}_{ij})} \leq \frac{R_\ell^+}{R_\ell^-}$. If all $R_\ell^-(\vec{\lambda}_{ij})$ equal zero, then $\sum_{j=1}^k R_\ell^+(\vec{\lambda}_{ij}) \leq \alpha_i kR_\ell^+$ trivially holds as we argued in the beginning. Otherwise, note that since not all $R_\ell^-(\vec{\lambda}_{ij})$ are zero, we must have $\ell > 1$, so if $R_\ell^-(\vec{\lambda}_{ij}) = 0$ then $R_\ell^+(\vec{\lambda}_{ij}) = 0$. Thus we also have $\frac{\sum_{j=1}^k R_\ell^+(\vec{\lambda}_{ij})}{\sum_{j=1}^k R_\ell^-(\vec{\lambda}_{ij})} \leq \frac{R_\ell^+}{R_\ell^-}$. In conclusion, in both cases we have $\sum_{j=1}^k R_\ell^+(\vec{\lambda}_{ij}) \leq \alpha_i kR_\ell^+$.

Observe that $\lambda_r \leq \lambda$ and $\prod_{i=1}^{d} \frac{1 + \sum_{j=1}^{k} R_\ell^+(\vec{\lambda}_{ij})}{1 + \sum_{j=1}^{k} R_\ell^-(\vec{\lambda}_{ij})} \geq 1$, it holds that

$$\frac{1 + kR_{\ell+1}^+(\vec{\lambda})}{1 + kR_{\ell+1}^-(\vec{\lambda})} = \frac{1 + k\lambda_r \prod_{i=1}^{d} \frac{1}{1+\sum_{j=1}^{k} R_\ell^-(\vec{\lambda}_{ij})}}{1 + k\lambda_r \prod_{i=1}^{d} \frac{1}{1+\sum_{j=1}^{k} R_\ell^+(\vec{\lambda}_{ij})}} \leq \frac{1 + k\lambda \prod_{i=1}^{d} \frac{1}{1+\alpha_i k R_l^-}}{1 + k\lambda \prod_{i=1}^{d} \frac{1}{1+\alpha_i k R_\ell^+}}.$$

Now it is enough to show that for every $\vec{\alpha}$ such that $0 \leq \alpha_i \leq 1$ for all $1 \leq i \leq d$, it holds that

$$\frac{1 + k\lambda \prod_{i=1}^{d} \frac{1}{1+\alpha_i k R_l^-}}{1 + k\lambda \prod_{i=1}^{d} \frac{1}{1+\alpha_i k R_\ell^+}} \leq \frac{1 + kR_{\ell+1}^+}{1 + kR_{\ell+1}^-},$$

which is equivalent to the following:

$$1 + k\lambda \prod_{i=1}^{d} \frac{1}{1+\alpha_i k R_\ell^-} - \frac{1 + kR_{\ell+1}^+}{1 + kR_{\ell+1}^-} - k\lambda \frac{1 + kR_{\ell+1}^+}{1 + kR_{\ell+1}^-} \prod_{i=1}^{d} \frac{1}{1+\alpha_i k R_\ell^+} \leq 0. \tag{8}$$

If $\alpha_i = 1$ for every i then the inequality (8) trivially holds. By symmetry, we only need to show the LHS of (8) is increasing in α_1. In fact, the partial derivative with respect to α_1 of LHS in (8) is:

$$-\frac{k^2 \lambda R_\ell^-}{1 + \alpha_1 k R_\ell^-} \prod_{i=1}^{d} \frac{1}{1+\alpha_i k R_\ell^-} + \frac{k^2 \lambda \left(1 + kR_{\ell+1}^+\right) R_\ell^+}{\left(1 + kR_{\ell+1}^-\right)\left(1 + \alpha_1 k R_\ell^+\right)} \prod_{i=1}^{d} \frac{1}{1+\alpha_i k R_\ell^+}.$$

To prove it is nonnegative, it is equivalent to show that

$$\frac{\left(1 + kR_{\ell+1}^+\right) R_\ell^+}{\left(1 + kR_{\ell+1}^-\right) R_\ell^-} \geq \frac{1 + \alpha_1 k R_\ell^+}{1 + \alpha_1 k R_\ell^-} \prod_{i=1}^{d} \frac{1 + \alpha_i k R_\ell^+}{1 + \alpha_i k R_\ell^-}. \tag{9}$$

To prove (9), we first observe that R_ℓ^- is increasing in ℓ and R_ℓ^+ is decreasing in ℓ, which is exactly the same to prove as the same property of the hardcore model. This gives us the so-called sandwich condition:

$$R_\ell^- \leq R_{\ell+1}^- \leq R_{\ell+1}^+ \leq R_\ell^+,$$

therefore $\frac{R_\ell^+}{R_\ell^-} \geq \frac{R_{\ell+1}^+}{R_{\ell+1}^-}$. We are now ready to prove (9):

$$\begin{aligned}
\frac{\left(1 + kR_{\ell+1}^+\right) R_\ell^+}{\left(1 + kR_{\ell+1}^-\right) R_\ell^-} &\geq \frac{\left(1 + kR_\ell^+\right) R_{\ell+1}^+}{\left(1 + kR_\ell^-\right) R_{\ell+1}^-} \\
&= \frac{1 + kR_\ell^+}{1 + kR_\ell^-} \prod_{i=1}^{d} \frac{1 + kR_\ell^+}{1 + kR_\ell^-} \\
&\geq \frac{1 + \alpha_1 k R_\ell^+}{1 + \alpha_1 k R_\ell^-} \prod_{i=1}^{d} \frac{1 + \alpha_i k R_\ell^+}{1 + \alpha_i k R_\ell^-}.
\end{aligned}$$

The last inequality uses the fact that $R_\ell^+ \geq R_\ell^-$ and $0 \leq \alpha_i \leq 1$. So (9) is proved, which finishes our proof of Lemma 19. ◀

Observe that the subtree rooted at the child of the root of $\mathbb{T}_{k,d}$ is isomorphic to $\widehat{\mathbb{T}}_{k,d}$. While at the root of $\mathbb{T}_{k,d}$, we have

$$\frac{R_\ell^+(\vec{\lambda})}{R_\ell^-(\vec{\lambda})} = \prod_{i=1}^{d+1} \frac{1 + \sum_{j=1}^{k} R_{\ell-1}^+(\vec{\lambda}_{ij})}{1 + \sum_{j=1}^{k} R_{\ell-1}^-(\vec{\lambda}_{ij})} \leq \left(\frac{1 + kR_{\ell-1}^+}{1 + kR_{\ell-1}^-} \right)^{d+1} = \frac{R_\ell^+(\mathbb{T}_{k,d})}{R_\ell^-(\mathbb{T}_{k,d})}.$$

Together with Lemma 19, this completes our proof of Theorem 18.

Calculation of the decay rate

The WSM rate of our model on the infinite $(k+1)$-uniform $(d+1)$-regular hypertree $\mathbb{T}_{k,d}$ is the same as the hardcore model on the infinite $(d+1)$-regular tree with activity $k\lambda$. The WSM rate on regular tree has been addressed implicitly in the literature [18, 32]. Here we provide an analysis for the decay rate for the completeness of the paper.

Let $f_{d,k}(x) \triangleq \frac{k\lambda}{(1+x)^d}$ denote the symmetric version of the tree recursion on $\widehat{\mathbb{T}}_{k,d}$ and substituting $x = kR$. Since $f_{d,k}(x)$ is decreasing in x, it follows that there is a unique positive fixed point $\hat{x}$ such that $\hat{x} = f_{d,k}(\hat{x})$. Let $f_{d,k}'(\hat{x}) = -\frac{d\hat{x}}{1+\hat{x}}$ be the derivative of $f_{d,k}(x)$ evaluated at the fixed point $x = \hat{x}$. The following proposition is well known for hardcore model (see e.g. [18, 32]).

▶ **Proposition 20.** $|f_{d,k}'(\hat{x})| = \frac{d\hat{x}}{1+\hat{x}} \leq 1$ if and only if $\lambda \leq \lambda_c$. And $|f_{d,k}'(\hat{x})| < 1$ if $\lambda < \lambda_c$.

We write $f(x) = f_{d,k}(x)$ if k and d are clear in the context. The main result of this part is the following theorem.

▶ **Theorem 21.** For any positive integers d, k, assuming $\lambda \leq \lambda_c$, the model on $\mathbb{T}_{k,d}$ exhibits weak spatial mixing with rate $\delta(\ell)$ such that for all sufficiently large ℓ:
- if $\lambda < \lambda_c$, then $\delta(\ell) \leq C_1 |f'(\hat{x})|^{\ell-4}$;
- if $\lambda = \lambda_c$, then $\delta(\ell) \leq \frac{C_2}{\sqrt{\ell-\ell_0}}$;
where $C_1, C_2, \ell_0 > 0$ are finite constants depending only on k, d and λ.

Theorem 14, Theorem 17 and 21 together prove the SSM part of Theorem 2.

Denote $g(x) = f(f(x)) = k\lambda \left(1 + \frac{k\lambda}{(1+x)^d} \right)^{-d}$. It is easy to see that $\hat{x} = g(\hat{x})$.

▶ **Lemma 22.** If $\lambda \leq \lambda_c$ then for any $x > \hat{x}$ we have $g(x) - g(\hat{x}) \leq f'(\hat{x})^2(x - \hat{x})$.

Proof. By the mean value theorem, for any $x > \hat{x}$, there exists a $z \in [\hat{x}, x]$ such that

$$g(x) - g(\hat{x}) = \alpha(z)(x - \hat{x}), \tag{10}$$

where $\alpha(z) = g'(z) = \frac{d^2 k\lambda g(z)}{(1+z)^{d+1} + (1+z)k\lambda}$. We will bound the maximum value of $\alpha(z)$ when $\lambda \leq \lambda_c$. Consider the derivative of $\alpha(z)$,

$$\alpha'(z) = A(z) \left[(d-1)k\lambda - (1+z)^d \right],$$

where $A(z) = \frac{d^2(d+1)k\lambda g(z)}{[(1+z)^{d+1} + (1+z)k\lambda]^2} > 0$. Let $z^* = ((d-1)k\lambda)^{1/d} - 1$ be the solution of $(d-1)k\lambda = (1+z)^d$. Note that $\left[(d-1)k\lambda - (1+z)^d \right]$ is decreasing in z, therefore $\alpha(z) \leq \alpha(z^*)$ for all $z > 0$. Due to proposition 20, if $\lambda \leq \lambda_c$ then $|f'(\hat{x})| = \frac{d\hat{x}}{1+\hat{x}} \leq 1$ and hence $\hat{x} \leq \frac{1}{d-1}$, thus $\alpha'(\hat{x}) = A(\hat{x})[(d-1)k\lambda - \frac{k\lambda}{\hat{x}}] \leq 0$, which means $\hat{x} \geq z^*$ and $\alpha(z)$ is decreasing in z for any $z \geq \hat{x}$. On the other hand, we have $\alpha(\hat{x}) = f'(\hat{x})^2$. Thus for any $z \geq \hat{x}$, we have $\alpha(z) \leq \alpha(\hat{x}) = f'(\hat{x})^2$. Plug it into (10). The lemma is proved. ◀

Proof of Theorem 21. It holds that $R_2^+ = R_1^+ = \lambda > \hat{x}/k$. Note that $kR_\ell^+ = g(kR_{\ell-2}^+)$. Due to the monotonicity of $g(x)$, we have $\hat{x} < kR_\ell^+ \leq k\lambda$ for every $\ell \geq 1$.

Consider the case $\lambda < \lambda_c$. First consider the $(k+1)$-uniform d-ary hypertree $\widehat{\mathbb{T}}_{k,d}$. By the mean value theorem and Lemma 22 we have

$$kR_\ell^+ - \hat{x} = g(kR_{\ell-2}^+) - g(\hat{x}) \leq f'(\hat{x})^2 \left(kR_{\ell-2}^+ - \hat{x}\right).$$

We apply this inequality recursively. Since $kR_\ell^+ - \hat{x} < k\lambda$, for any $\ell \geq 2$ we have

$$kR_\ell^+ - \hat{x} \leq k\lambda|f'(\hat{x})|^{\ell-2}. \tag{11}$$

To bound R_ℓ^- we apply the mean value theorem again. There exists a $z \in [\hat{x}, kR_\ell^+]$ such that

$$\hat{x} - kR_\ell^- = f(\hat{x}) - f(kR_{\ell-1}^+) = |f'(z)|(kR_{\ell-1}^+ - \hat{x}).$$

Since $|f'(z)| \leq kd\lambda$ for all $z > 0$, combined with (11) we have

$$\hat{x} - kR_\ell^- \leq kd\lambda(kR_{\ell-1}^+ - \hat{x}) \leq k^2 d\lambda^2|f'(\hat{x})|^{\ell-3}.$$

At last, $R_\ell^+ - R_\ell^- = \frac{1}{k}(kR_\ell^+ - \hat{x} + \hat{x} - kR_\ell^-) \leq C_1'|f'(\hat{x})|^{\ell-3}$ for some $C_1' > 0$ depending only on d, k and λ. This only gives us the desired decay rate at the $(k+1)$-uniform d-ary hypertree $\widehat{\mathbb{T}}_{k,d}$. Move to the $(k+1)$-uniform $(d+1)$-regular hypertree $\mathbb{T}_{k,d}$. The only difference is that the root has $d+1$ children instead of d. By the mean value theorem, this will multiply at most a finite constant factor C_1'' to the gap $R_\ell^+ - R_\ell^-$ at the root of $\mathbb{T}_{k,d}$, where $C_1'' > 0$ depends only on d, k and λ. Overall, this gives us that

$$p_\ell^+ - p_\ell^- \leq R_\ell^+ - R_\ell^- \leq C_1|f'(\hat{x})|^{\ell-4}$$

for some $C_1 > 0$ depending only on d, k and λ. This finishes the case that $\lambda < \lambda_c$.

Now we consider the critical case that $\lambda = \lambda_c = \frac{d^d}{k(d-1)^{d+1}}$. We still start by considering the $(k+1)$-uniform d-ary hypertree $\widehat{\mathbb{T}}_{k,d}$. It is easy to verify that in this case $\hat{x} = \frac{1}{d-1}$, $\alpha(\hat{x}) = f'(\hat{x})^2 = 1$, $z^* = \hat{x}$ and $\alpha'(\hat{x}) = 0$, where $\alpha(z)$ and z^* are defined in the proof of Lemma 22. And we have $\alpha''(\hat{x}) = -\frac{(d+1)(d-1)^3}{d^2}$. By Taylor's expansion for $g(x)$ at the fixed point $x = \hat{x}$, we have that for any constant $c > 0$ there exists a constant $x_0 > \hat{x}$ such that for any $\hat{x} < x < x_0$, it holds that

$$g(x) = g(\hat{x}) + \alpha(\hat{x})(x - \hat{x}) + \frac{\alpha'(\hat{x})}{2}(x - \hat{x})^2 + \frac{\alpha''(\hat{x})}{6}(x - \hat{x})^3 + o\left((x - \hat{x})^3\right)$$

$$\leq \frac{1}{d-1} + x - \hat{x} - \frac{(d+1)(d-1)^3}{6d^2}(x - \hat{x})^3 + c(x - \hat{x})^3.$$

We define a sequence $x_1 = kR_1^+, x_3 = kR_3^+ = g(x_1), \ldots$ and generally $x_{2t+1} = g(x_{2t-1})$. The sequence is strictly decreasing because R_ℓ^+ is decreasing in ℓ. Furthermore, $\lim_{t\to\infty} x_{2t+1} = \hat{x}$. This is due to $\alpha(x) < 1$ for any $x > \hat{x}$.

Denote $\epsilon_{2t+1} \triangleq x_{2t+1} - \hat{x}$. Let c be some positive constant such that $\frac{(d+1)(d-1)^3}{6d^2} - c > 0$. Denote $\beta = \frac{(d+1)(d-1)^3}{6d^2} - c$ and $\gamma = \sqrt{\frac{1}{2\beta}}$. There must be some sufficiently large t_0 such that $\epsilon_{2t_0+1} \leq \frac{\gamma}{\sqrt{2}}$ and for any $t > t_0$, it holds that

$$\epsilon_{2t+3} = g(x_{2t+1}) - \frac{1}{d-1} \leq \epsilon_{2t+1} - \beta\epsilon_{2t+1}^3.$$

We apply an induction to complete the proof. For the basis, when $t = t_0$ we have $\epsilon_{2t_0+1} \leq \frac{\gamma}{\sqrt{2}}$. Assume the hypothesis

$$\epsilon_{2t+1} \leq \frac{\gamma}{\sqrt{t - t_0 + 2}} \tag{12}$$

for some $t \geq t_0$ and we will prove it holds for $t + 1$. First, notice that $h(x) \triangleq x - \beta x^3$ is strictly increasing when $0 \leq x \leq \frac{\gamma}{\sqrt{2}}$. Thus, we have

$$\epsilon_{2t+3} \leq \epsilon_{2t+1} - \beta\epsilon_{2t+1}^3 \leq \frac{\gamma}{\sqrt{t - t_0 + 2}} - \beta\frac{\gamma^3}{(t - t_0 + 2)^{\frac{3}{2}}}.$$

We only need to prove that $\frac{\gamma}{\sqrt{t-t_0+2}} - \beta\frac{\gamma^3}{(t-t_0+2)^{\frac{3}{2}}} \leq \frac{\gamma}{\sqrt{t-t_0+3}}$. Let $t' \triangleq t - t_0 + 2$. It is equivalently to show that

$$t'^{\frac{3}{2}}\left(\frac{1}{\sqrt{t'}} - \frac{1}{\sqrt{t'+1}}\right) \leq \beta\gamma^2. \tag{13}$$

Note that

$$\begin{aligned}
t'^{\frac{3}{2}}\left(\frac{1}{\sqrt{t'}} - \frac{1}{\sqrt{t'+1}}\right) &= t'^{\frac{3}{2}}\left(\frac{\sqrt{t'+1} - \sqrt{t'}}{\sqrt{t'(t'+1)}}\right) \\
&\leq \sqrt{t'}(\sqrt{t'+1} - \sqrt{t'}) \\
&\leq \sqrt{t'}\frac{1}{\sqrt{t'+1} + \sqrt{t'}} \\
&\leq \frac{1}{2}.
\end{aligned}$$

Since $\beta\gamma^2 = \frac{1}{2}$, we just prove the inequality (13), and finishes the induction (12) for all $t \geq t_0$. In conclusion, for any $t \geq t_0$, it holds that

$$kR_{2t+1}^+ - \hat{x} \leq \frac{\gamma}{\sqrt{t - t_0 + 2}}.$$

The rest of the proof is exactly the same as our proof of the case $\lambda < \lambda_c$. ◀

6 Approximation algorithms and inapproximability

For $0 < \varepsilon < 1$, a value $\hat{Z}$ is an ε-approximation of Z if $(1 - \epsilon)Z \leq \hat{Z} \leq (1 + \epsilon)Z$. Recall that $\hat{x}$ is the unique fixed point solution to $\hat{x} = f_{d,k}(\hat{x}) = k\lambda(1 + \hat{x})^{-d}$.

▶ **Theorem 23.** *If $\lambda < \lambda_c = \frac{d^d}{k(d-1)^{d+1}}$, then there exists an algorithm such that given any $\varepsilon > 0$, and any hypergraph $\mathcal{H}$ of n vertices, of maximum edge-size at most $(k + 1)$ and maximum degree at most $(d + 1)$, the algorithm returns an ε-approximation of the partition function for the independent sets of $\mathcal{H}$ with activity λ, within running time $\left(\frac{n}{\varepsilon}\right)^{O\left(\frac{1}{\kappa} \ln kd\right)}$, where $\kappa = \ln\left(\frac{1+\hat{x}}{d\hat{x}}\right)$.*
For the critical case where $\lambda = \lambda_c$, there exists an algorithm that for the above $\mathcal{H}$ returns an ε-approximation of the log-partition function within running time $n(kd)^{O\left(\left(\frac{1}{\varepsilon} \ln \frac{1}{\varepsilon}\right)^2\right)}$.

By duality, the same algorithm with the same approximation ratio and running time works for the matchings of hypergraphs of maximum edge size at most $(d + 1)$ and maximum degree at most $(k + 1)$. By Proposition 20, $|f'_{d,k}(\hat{x})| = \frac{d\hat{x}}{1+\hat{x}} < 1$ if $\lambda < \lambda_c$, therefore, when

$\lambda < \lambda_c$, the running time of the algorithm is $\text{Poly}(n, \frac{1}{\epsilon})$ for any bounded k and d, so the algorithm is an FPTAS for the partition function. And when $\lambda = \lambda_c$, the algorithm is a PTAS for the log-partition function. The algorithmic part of the main theorem Theorem 2 is proved.

In particular, when $d = 1$, the model becomes matchings of graphs of maximum degree $(k + 1)$, and the uniqueness condition $\lambda < \lambda_c(\mathbb{T}_{k,d})$ is always satisfied even for unbounded k since $\lambda_c(\mathbb{T}_{k,1}) = \infty$. In this case, the fixed point $\hat{x}$ for $f_{1,k}(x) = \frac{k\lambda}{1+x}$ can be explicitly solved as $\hat{x} = \frac{-1+\sqrt{1+4k\lambda}}{2}$. We have the following corollary for matchings of graphs with unbounded maximum degree, which achieves the same bound as the algorithm in [1].

▶ **Corollary 24.** *There exists an algorithm which given any graph G of maximum degree at most Δ, and any $\epsilon > 0$, returns an ε-approximation of the partition function for the matchings of G with activity λ, within running time $\left(\frac{n}{\epsilon}\right)^{O(\sqrt{\lambda\Delta}\log\Delta)}$.*

With the construction of hypergraph self-avoiding walk tree and the SSM, the algorithm follows the framework by Weitz [34]. We will describe an algorithm of approximating the partition function for independent sets in hypergraphs with activity λ. Under duality this is the same as approximately counting matchings with activity λ.

By the standard self-reduction, approximately computing the partition function is reduced to approximately computing the marginal probabilities. Let $\mathcal{H} = (V, E)$ be a hypergraph and $V = \{v_1, \ldots, v_n\}$. To calculate $Z = Z_{\mathcal{H}}(\lambda)$, it suffices to calculate the probability of the emptyset $\mu(\varnothing)$ as it is exactly $1/Z$. Let $\varnothing_i$ be the configuration on vertices v_1 up to v_i where all of them are unoccupied, and $p_{v_i}^{\varnothing_{i-1}}$ the probability of v_i being occupied conditioning on all vertices v_1 up to v_{i-1} being unoccupied. Then we have $1/Z = \prod_{i=1}^{n}(1 - p_{v_i}^{\varnothing_{i-1}})$ and $\log Z = -\sum_{i=1}^{n} \log(1 - p_{v_i}^{\varnothing_{i-1}})$. Note that $(1 - p_{v_i}^{\varnothing_{i-1}}) \geq \frac{1}{1+\lambda}$ for the probability of vertex unoccupied by an independent set and $\lambda_c \leq 4$ for any $d \geq 2$ and $k \geq 1$. To get an ε-approximation of Z, it suffices to approximate each of $p_{v_i}^{\varnothing_{i-1}}$ within an additive error $\frac{\varepsilon}{2(1+\lambda)n}$. And to get an ε-approximation of $\log Z$, which can be obtained by getting an ε-approximation of every $-\log(1 - p_{v_i}^{\varnothing_{i-1}})$, it is sufficient to approximate each of $p_{v_i}^{\varnothing_{i-1}}$ within an additive error $\Theta\left(\frac{\varepsilon}{\ln\frac{1}{\varepsilon}}\right)$.

By Theorem 14, we have $p_v^{\sigma} = \mathbb{P}_{\mathcal{T}}^{\sigma}$ where $\mathcal{T} = \mathcal{T}_{\text{SAW}}(\mathcal{H}, v)$, i.e. the marginal probability of v being occupied is preserved in the SAW tree of $\mathcal{H}$ expanded at v. And the value of $\mathbb{P}_{\mathcal{T}}^{\sigma}$ can be computed by the tree recursion (3). To make the algorithm efficient we can run this recursion up to depth t and assume initial value 0 for the variables at depth t as the vertices they represent being unoccupied. The overall running time of the algorithm is clearly $O(n(kd)^t)$ where t is the depth of the recursion. By the strong spatial mixing guaranteed by Theorem 17 and Theorem 21, if $\lambda < \lambda_c$, then the additive error of such estimation of p_v^{σ} is bounded by $C_1 \cdot \left(\frac{d\hat{x}}{1+\hat{x}}\right)^{t-4}$ for some constant $C_1 > 0$ depending only on k, d and λ. We shall choose an integer t so that $C_1 \cdot \left(\frac{d\hat{x}}{1+\hat{x}}\right)^{t-4} \leq \frac{\varepsilon}{2(1+\lambda)n}$, which gives us the suitable time complexity required by the FPTAS for the partition function. And when $\lambda = \lambda_c$, the additive error of p_v^{σ} is bounded by $C_2/\sqrt{t-t_0}$ for some constants $C_2, t_0 > 0$ depending only on k, d. We shall choose an integer $t = O((\frac{1}{\varepsilon}\ln\frac{1}{\varepsilon})^2)$ to get the desirable additive error for every marginal probability, which gives us the PTAS for the log-partition function. This completes the proof of Theorem 23.
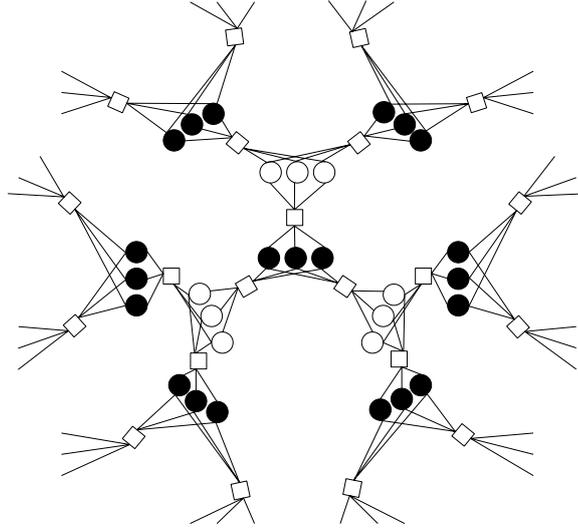
Figure 4 The infinite hypergraph that achieves the uniqueness threshold $\frac{2k+1+(-1)^k}{k+1}\lambda_c$.

Inapproximability

For the inapproximability, by applying an AP-reduction [3] from the inapproximability of the hardcore model [31, 9], we have the following theorem.

▶ **Theorem 25.** *If* $\lambda > \frac{2k+1+(-1)^k}{k+1}\lambda_c$, *there is no PRAS for the partition function or log-partition function of independent sets of hypergraphs with maximum degree at most* $d+1$, *maximum edge-size at most* $k+1$ *and activity* λ, *unless NP=RP.*

Proof. The reduction is as described in [3], which is reduced from the hardcore model. Given a graph $G(V, E)$ with maximum degree at most $(d+1)$, we construct a hypergraph $\mathcal{H}(V_{\mathcal{H}}, E_{\mathcal{H}})$ as follows. For each $v \in V$, we create $t = \left\lfloor \frac{k+1}{2} \right\rfloor$ distinct vertices $w_{v,1}, w_{v,2}, \ldots, w_{v,t}$ and let $V_{\mathcal{H}} = \{w_{v,i} \mid v \in V, 1 \le i \le t\}$. And for every edge $e = (u, v) \in E$, we create a hyperedge $S_e = \{w_{u,1}, \ldots, w_{u,t}, w_{v,1}, \ldots, w_{v,t}\}$ and let $E_{\mathcal{H}} = \{S_e \mid e \in E\}$. Clearly, the maximum degree of $\mathcal{H}$ is at most $d+1$ and the maximum edge-size of $\mathcal{H}$ is at most $2t \le k+1$. We define

$$Z_{\mathcal{H}}(\lambda) = \sum_{I:\ \text{IS of } \mathcal{H}} \lambda^{|I|} \quad \text{and} \quad Z_G(\lambda) = \sum_{I:\ \text{IS of } G} \lambda^{|I|}.$$

Note that by the above reduction every independent set I of $\mathbb{G}$ is naturally identified to $t^{|I|}$ distinct independent sets of hypergraph $\mathcal{H}$ such that a $v \in V$ is occupied by I if and only if one of $w_{v,i}$ is occupied by the corresponding independent set of $\mathcal{H}$. Thus $Z_{\mathcal{H}}(\lambda) = Z_G(\lambda')$ where $\lambda' = t\lambda$.

Recall that G is an arbitrary graph of maximum degree at most $d+1$. According to Sly and Sun [31], when $\lambda' > \frac{d^d}{(d-1)^{d+1}}$, there exists a constant c such that unless NP=RP, the partition function $Z_G(\lambda')$ can not be approximated within a factor of c^n in polynomial time, which means there is no PRAS for the log-partition function $\log Z_G(\lambda')$ when $\lambda' > \frac{d^d}{(d-1)^{d+1}}$, i.e. when $\lambda > \frac{d^d}{\lfloor(k+1)/2\rfloor(d-1)^{d+1}} = \frac{2k+1+(-1)^k}{k+1}\lambda_c$. ◀

The reduction in Theorem 25 transforms a hardcore model on a graph with maximum degree $d+1$ and activity $\frac{2k+1+(-1)^k}{k+1}\lambda$ to an instance of hypergraph independent sets with

maximum degree at most $d + 1$, maximum edge-size at most $k + 1$, and activity λ. In particular, it transforms the infinite $(d + 1)$-regular tree $\mathbb{T}_{d,1}$ to the infinite $2\lfloor(k + 1)/2\rfloor$-uniform hypergraph as shown in Figure 4. This infinite hypergraph has the uniqueness threshold $\frac{d^d}{\lfloor(k+1)/2\rfloor(d-1)^{d+1}} = \frac{2k+1+(-1)^k}{k+1}\lambda_c$.

7 Local convergence of hypergraphs

For the infinite $(k + 1)$-uniform $(d + 1)$-regular hypertree $\mathbb{T}_{k,d}$, a group $\mathbb{G}$ of automorphisms on $\mathbb{T}_{k,d}$ classifies the vertices and hyperedges in $\mathbb{T}_{k,d}$ into orbits (equivalent classes). We consider only $\mathbb{G}$ with finitely many orbits. By Proposition 11, such group $\mathbb{G}$ can be uniquely identified by a pair of branching matrices $(\boldsymbol{D}, \boldsymbol{K})$ defined in Section 3 that classifies vertices and hyperedges in $\mathbb{T}_{k,d}$ into finitely many types (labels), where the incidence relation between vertices and hyperedges with each type is specified by $(\boldsymbol{D}, \boldsymbol{K})$. We use $\mathbb{T}_{k,d}^{\mathbb{G}}$ to denote this resulting labeled hypertree.

For a finite hypergraph $\mathcal{H} = (V, E)$, we also consider the classification of vertices $V = \biguplus_{i \in [\tau_v]} V_i$ and hyperedges $E = \biguplus_{j \in [\tau_e]} E_j$ into disjoint types.

Given a hypergraph $\mathcal{H}$ and a vertex v in $\mathcal{H}$, write $B_t(v) = B_{\mathcal{H},t}(v)$ for the *t-neighborhood* around v in $\mathcal{H}$, that is, the sub-hypergraph induced by the vertices in $\mathcal{H}$ at distance at most t from v. For the labeled hypertree $\mathbb{T}_{k,d}^{\mathbb{G}}$, since once the type of the root is fixed the neighborhoods are identical (in terms of types), for each $i \in [\tau_v]$, we can denote $\mathbb{T}_{k,d}^{\mathbb{G}}(t, i) = B_{T,t}(v)$ where $T = \mathbb{T}_{k,d}^{\mathbb{G}}$ and v is any vertex in T of type-i.

The following definition is inspired by those of [31] and [4] for spin systems. Intuitively, a sequence of finite structures locally resemble the infinite tree structure along with the suitable symmetry which exhibits the uniqueness/nonuniqueness phase transition at the critical threshold, so the measures on the sequence of finite structures may have *local weak convergence* to that on the infinite tree. The existence of such local convergence profoundly leads to several most important phase-transition-based inapproximability results [6, 26, 30, 31, 7, 9, 10] and is a key to the success of random regular bipartite graph as a gadget for anti-ferromagnetic spin systems.

▶ **Definition 26** (local convergence). Let $\mathcal{H}_n = (V_n, E_n)$ be a sequence of random finite hypergraphs, whose vertices $V_n = \biguplus_{i \in [\tau_v]} V_{n,i}$ and hyperedges $E_n = \biguplus_{j \in [\tau_e]} E_{n,j}$ are classified into disjoint types, and for each $i \in [\tau_v]$, let $I_{n,i} \in V_{n,s}$ denote a uniformly random vertex in V_n of type-i.

We say the $\mathcal{H}_n$ *converge locally* to $\mathbb{T}_{k,d}^{\mathbb{G}}$, and write $\mathcal{H}_n \to_{\text{loc}} \mathbb{T}_{k,d}^{\mathbb{G}}$, if for all $t \geq 0$ and $i \in [\tau_v]$, $B_t(I_{n,i})$ converges to $\mathbb{T}_{k,d}^{\mathbb{G}}(t, i)$ in distribution with respect to the joint law $\mathbb{P}_n$ of $(\mathcal{H}_n, I_{n,i})$: that is,

$$\lim_{n \to \infty} \mathbb{P}_n \left(B_t(I_{n,i}) \cong \mathbb{T}_{k,d}^{\mathbb{G}}(t, i) \right) = 1,$$

where $\cong$ denotes isomorphism which preserves vertex- and hyperedge-types and the incidence relation.

Consider the natural uniform random walk on the incidence graph of $\mathbb{T}_{k,d}^{\mathbb{G}}$, and its projection onto the finitely many disjoint orbits (types) for vertices and hyperedges, which gives a (bipartite) finite Markov chain. It is quite amazing to see that the reversibility of this projected chain determines whether there exists a sequence of finite hypergraphs that converge locally to $\mathbb{T}_{k,d}^{\mathbb{G}}$.

▶ **Theorem 27.** *Let $\mathbb{G}$ be an automorphism group of $\mathbb{T}_{k,d}$ with finitely many orbits for vertices and hyperedges. Let $\boldsymbol{D}$ and $\boldsymbol{K}$ be the branching matrices that corresponds to $\mathbb{G}$ as defined in Section 3. There is a sequence of random finite hypergraphs $\mathcal{H}_n \to_{\mathrm{loc}} \mathbb{T}_{k,d}^{\mathbb{G}}$ if and only if the Markov chain $\boldsymbol{P} = \begin{bmatrix} \boldsymbol{0} & \frac{1}{d+1}\boldsymbol{D} \\ \frac{1}{k+1}\boldsymbol{K} & \boldsymbol{0} \end{bmatrix}$ is time-reversible.*

We say a uniform random walk over a hypergraph $\mathcal{H}$ is a uniform random walk on the incidence graph of $\mathcal{H}$: that is, a random walk moves between vertices and hyperedges. Then the Markov chain $\boldsymbol{P}$ is the projection of the uniform random walk over $\mathbb{T}_{k,d}$ onto the equivalent classes of vertices and hyperedges (i.e. the orbits of the automorphism group $\mathbb{G}$ that corresponds to the $\boldsymbol{D}$ and $\boldsymbol{K}$). Meanwhile, matrix $\begin{bmatrix} \boldsymbol{0} & \boldsymbol{D} \\ \boldsymbol{K} & \boldsymbol{0} \end{bmatrix}$ is the adjacent matrix for a directed bipartite graph that describes the (weighted) incidence relation between vertex- and hyperedge-types in the following way: each directed bipartite edge from vertex-type-i to hyperedge-type-j (or vice versa) is assigned with weight d_{ij} (or k_{ji}). So the Markov chain $\boldsymbol{P}$ is also the random walk on this directed bipartite graph where the transition probability of each directed edge is proportional to its weight.

For the bipartite Markov chain $\boldsymbol{P}$, recall that due to Proposition 11, $\boldsymbol{P}$ must be irreducible. Then the time-reversibility of $\boldsymbol{P}$ is equivalent to the following: There exist positive vectors $\vec{p} = (p_i)_{i \in [\tau_v]}$ and $\vec{q} = (q_j)_{j \in [\tau_e]}$ that satisfy the bipartite detailed balanced equation:

$$p_i d_{ij} = q_j k_{ji}$$

for every $(i,j) \in [\tau_v] \times [\tau_e]$. Without loss of generality, we assume $\sum_i p_i + \sum_j q_j = 1$.

In fact, it is easy to check that $\vec{p}\boldsymbol{D} = (k+1)\vec{q}$ and $\vec{q}\boldsymbol{K} = (d+1)\vec{p}$, therefore the $\vec{p}$ and $\vec{q}$ are respectively the left eigenvector of $\boldsymbol{DK}$ and $\boldsymbol{KD}$ both with eigenvalue $(d+1)(k+1)$. Since both $\boldsymbol{DK}$ and $\boldsymbol{KD}$ are irreducible, due to the Perron-Frobenius theorem, the only positive left eigenvectors $\vec{p}$ and $\vec{q}$ are the ones that are associated with the Perron-Frobenius eigenvalue $(d+1)(k+1)$ and are one-dimensional.

Furthermore, it must holds that $\frac{\|\vec{p}\|_1}{\|\vec{q}\|_1} = \frac{k+1}{d+1}$. Denote $\vec{p'} = \frac{\vec{p}}{\|\vec{p}\|_1}$ and $\vec{q'} = \frac{\vec{q}}{\|\vec{q}\|_1}$. We have $\|\vec{p'}\| = \|\vec{q'}\| = 1$ and $p'_i \frac{d_{ij}}{d+1} = q'_j \frac{k_{ji}}{k+1}$ for every $(i,j) \in [\tau_v] \times [\tau_e]$, i.e. $\vec{p'}$ is the *vertex-stationary distribution* and $\vec{q'}$ is the *edge-stationary distribution*. We will mostly use $\vec{p}$ and $\vec{q}$ in our proof of Theorem 27. Recall for the automorphism group $\widehat{\mathbb{G}}$ defined in Section 3 such that $\lambda_c(\mathbb{T}_{k,d}^{\widehat{\mathbb{G}}}) = \lambda_c(\mathbb{T}_{k,d}) = \frac{d^d}{k(d-1)^{d+1}}$, i.e. the uniqueness of $\widehat{\mathbb{G}}$-translation-invariant Gibbs measures on $\mathbb{T}_{k,d}$ represents the uniqueness of all Gibbs measures on $\mathbb{T}_{k,d}$, the branching matrices are given as $\widehat{\boldsymbol{D}} = \begin{bmatrix} 1 & d \\ d & 1 \end{bmatrix}$ and $\widehat{\boldsymbol{K}} = \begin{bmatrix} k & 1 \\ 1 & k \end{bmatrix}$. It is easy to verify that the resulting Markov chain $\widehat{\boldsymbol{P}}$ is not time-reversible. It then follows from Theorem 27 that there does not exist *any* sequence of random finite hypergraphs that converge locally to $\mathbb{T}_{k,d}$ with the symmetry $\widehat{\mathbb{G}}$ assumed by the extremal Gibbs measures μ^+, μ^- whose uniqueness represents the uniqueness of all Gibbs measures.

▶ Remark 28. Given branching matrices $\boldsymbol{D}$ and $\boldsymbol{K}$, instead of considering $\mathcal{H}_n$ that converges locally for every type to the $\mathbb{T}_{k,d}^{\mathbb{G}}$ as in Definition 26, we can alternatively define a sequence $\mathcal{H}_n$ that *converges locally in average* to the $\mathbb{T}_{k,d}^{\mathbb{G}}$: that is, for all $t > 0$, the $B_t(I_n)$ converges to $\mathbb{T}_{k,d}^{\mathbb{G}}(t, I)$ in distribution, where I_n denotes a uniformly random vertex in the finite hypergraph $\mathcal{H}_n$, and I denotes a random vertex-type chosen according to the vertex-stationary distribution $\vec{p'}$. This definition looks more analogous to the local convergence defined in [31] for the anti-ferromagnetic 2-spin system. But we will see the two definitions are equivalent: A sequence $\mathcal{H}_n \to_{\mathrm{loc}} \mathbb{T}_{k,d}^{\mathbb{G}}$ also converges locally to $\mathbb{T}_{k,d}^{\mathbb{G}}$ in average, since by double counting the portion

of vertices of type-i must converge to p_i' as $n \to \infty$; and conversely, a sequence converges locally to $\mathbb{T}_{k,d}^{\mathbb{G}}$ in average must also have $\mathcal{H}_n \to_{\mathrm{loc}} \mathbb{T}_{k,d}^{\mathbb{G}}$, simply because neighborhoods of vertices of different types cannot be isomorphic to each other.

Proof of Theorem 27. We will prove the necessity of the reversibility of the chain by a double counting argument and the sufficiency is proved by explicitly constructing the sequence of the finite hypergraphs.

Double counting

Let $\mathcal{H}_n = (V_n, E_n)$ where $V_n = \biguplus_{s \in \tau_v} V_{n,s}$ and $E_n = \biguplus_{t \in \tau_e} E_{n,t}$. Assume that $\mathcal{H}_n \to_{\mathrm{loc}} \mathbb{T}_{k,d}^{\mathbb{G}}$.

For $\mathbb{T}_{k,d}^{\mathbb{G}}$ such that there is a hypergraph sequence $\mathcal{H}_n = (V_n = \biguplus_{s \in \tau_v} V_{n,s},\ E_n = \biguplus_{t \in \tau_e} E_{n,t})$ converging locally to $\mathbb{T}_{k,d}^{\mathbb{G}}$, we show that the Markov chain $\boldsymbol{P}$ is time reversible. The proof is by a double counting of the number of vertex-hyperedge pairs with specific type combination.

Since the 1-neighborhood of the vertex with each type in $\mathcal{H}_n$ converges in distribution to the 1-neighborhood of the vertex with the same type in $\mathbb{T}_{k,d}^{\mathbb{G}}$, for sufficiently large n, we have all but a $o(1)$-fraction of vertices in $\mathcal{H}_n$ whose local transitions between vertex-types and hyperedge-types within 1-step are given precisely by $\boldsymbol{D}$ and $\boldsymbol{K}$. Thus, for every $(i,j) \in [\tau_v] \times [\tau_e]$, the total number of incident vertex-hyperedge pair (v,e) with $v \in V_{n,i}$ and $e \in E_{n,j}$ (counted from the vertex-side and from the hyperedge-side) is given by

$$d_{ij}(|V_{n,i}| + o(1)) = k_{ji}(|E_{n,j}| + o(1)).$$

As $n \to \infty$, we will have $(d_{ij}|V_{n,i}|)/(k_{ji}|E_{n,j}|) \to 1$, or equivalently

$$\frac{|E_{n,j}|}{|V_{n,i}|} \to \frac{d_{ij}}{k_{ji}}$$

for all $(i,j) \in [\tau_v] \times [\tau_e]$ such that $d_{ij}, k_{ji} \neq 0$. Thus there exists positive p_i, q_j such that $q_j/p_i = d_{ij}/k_{ji}$ for all such (i,j). Since $\boldsymbol{DK}$ and $\boldsymbol{KD}$ are irreducible, we have unique corresponding positive left eigenvectors, which is $(p_i)_{i \in [\tau_v]}, (q_j)_{j \in [\tau_e]}$ here, such that $p_i d_{ij} = q_j k_{ji}$ for all (i,j).

Construction of $\mathcal{H}_n$

Assume the Markov chain $\boldsymbol{P}$ in Theorem 27 to be time-reversible, and let $\vec{p} = (p_i)_{i \in [\tau_v]}$ and $\vec{q} = (q_j)_{j \in [\tau_e]}$ be the unique positive vectors satisfying $p_i d_{ij} = q_j k_{ji}$ for every $(i,j) \in [\tau_v] \times [\tau_e]$ and $\sum_i p_i + \sum_j q_j = 1$. The sequence of finite hypergraph sequence $\mathcal{H}_n$ that converges locally to $\mathbb{T}_{\boldsymbol{K},\boldsymbol{D}}$ is constructed as follows. The number n is approximately the total number of vertices and hyperedges in $\mathcal{H}_n$ (where the approximation is due to rounding).

- For each $s \in [\tau_v]$ and $t \in [\tau_e]$, let $V_{n,s}$ be the set of $\lceil p_s n \rceil$ vertices of type s, and $E_{n,t}$ be the set of $\lceil q_t n \rceil$ hyperedges of type t. We then describe hypergraphs $\mathcal{H}_n = (V_n, E_n)$ where $V_n = \biguplus_{s \in \tau_v} V_{n,s}$ and $E_n = \biguplus_{t \in \tau_e} E_{n,t}$.
- For each $s \in [\tau_v]$ and $t \in [\tau_e]$, let $N_{s,t} \triangleq \lceil d_{st} p_s n \rceil = \lceil k_{ts} q_t n \rceil$. Sample a uniformly random permutation $f : [N_{s,t}] \to [N_{s,t}]$, and create an incidence between the i-th vertex in $V_{n,s}$ and the j-th hyperedge in $E_{n,t}$ for every $(a, b = f(a))$ with $a \in i + \lceil p_s n \rceil \mathbb{Z}$ and $b \in j + \lceil q_t n \rceil \mathbb{Z}$.

Note that as normalized Perron eigenvectors for irreducible integer matrices, the $\vec{p}$ and $\vec{q}$ must be rational. Then there are infinitely many n such that $N_{s,t}/|V_{n,s}| = d_{st}$ and

$N_{s,t}/|E_{n,t}| = k_{ts}$. Without loss of generality, we can consider only these n, since otherwise it will contribute at most $o(1)$-fractions of bad neighborhoods.

Viewing multi-edges in the incidence graph of $\mathcal{H}_n$ as different edges, it holds that each vertex of type-s is incident to exactly d_{st} hyperedges of type-t and each hyperedge of type-t is incident by exactly k_{ts} vertices of type-s. Therefore it is sufficient to show that for any finite $r > 0$ the probability that the r-neighborhood of a vertex in $\mathcal{H}_n$ has no circle is 1 as $n \to \infty$, i.e. almost surely the r-neighborhood of a vertex in $\mathcal{H}_n$ is a hypertree. This can be proved easily by a standard routine of Galton-Watson branching process (see e.g. Ch. 9 in [15]) since the neighborhood is of constant size and the probability of reencountering a vertex or an edge from a population whose size goes to ∞ goes to 0. ◀

References

1. Mohsen Bayati, David Gamarnik, Dimitriy Katz, Chandra Nair, and Prasad Tetali. Simple deterministic approximation algorithms for counting matchings. In *Proceedings of the 39th ACM Symposium on Theory of Computing (STOC)*, pages 122–127, 2007.

2. Ivona Bezakova, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, and Daniel Stefankovic. Counting independent sets in hypergraphs when strong spatial mixing fails. *arXiv preprint arXiv:1510.09193*, 2015.

3. Magnus Bordewich, Martin Dyer, and Marek Karpinski. Path coupling using stopping times and counting independent sets and colorings in hypergraphs. *Random Structures & Algorithms*, 32(3):375–399, 2008.

4. Amir Dembo, Andrea Montanari, et al. Ising models on locally tree-like graphs. *The Annals of Applied Probability*, 20(2):565–592, 2010.

5. Andrzej Dudek, Marek Karpinski, Andrzej Ruciński, and Edyta Szymańska. Approximate counting of matchings in (3, 3)-hypergraphs. In *SWAT*, pages 380–391, 2014.

6. Martin Dyer, Leslie A Goldberg, and Mark Jerrum. Counting and sampling H-colourings. In *RANDOM*, pages 51–67, 2002.

7. Andreas Galanis, Qi Ge, Daniel Štefankovič, Eric Vigoda, and Linji Yang. Improved inapproximability results for counting independent sets in the hard-core model. *Random Structures & Algorithms*, 45(1):78–110, 2014.

8. Andreas Galanis and Leslie Ann Goldberg. The complexity of approximately counting in 2-spin systems on k-uniform bounded-degree hypergraphs. *arXiv preprint arXiv:1505.06146*, 2015.

9. Andreas Galanis, Daniel Stefankovic, and Eric Vigoda. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. arXiv preprint arXiv:1203.2226, 2012.

10. Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability for antiferromagnetic spin systems in the tree non-uniqueness region. In *Proceedings of the 46th ACM Symposium on Theory of Computing (STOC)*, pages 823–831, 2014.

11. David Gamarnik and Dmitriy Katz. Correlation decay and deterministic FPTAS for counting colorings of a graph. *Journal of Discrete Algorithms*, 12:29–47, 2012.

12. David Gamarnik, Dmitriy Katz, and Sidhant Misra. Strong spatial mixing of list coloring of graphs. *Random Structures & Algorithms*, 2013.

13. Ole J Heilmann. Existence of phase transitions in certain lattice gases with repulsive potential. *Lettere Al Nuovo Cimento (1971–1985)*, 3(3):95–98, 1972.

14. Ole J Heilmann and Elliott H Lieb. Theory of monomer-dimer systems. *Communications in Mathematical Physics*, 25(3):190–232, 1972.

15. Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random graphs*, volume 45. John Wiley & Sons, 2011.

16 Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM Journal on Computing*, 18(6):1149–1178, 1989.

17 Marek Karpinski, Andrzej Rucinski, and Edyta Szymanska. Approximate counting of matchings in sparse uniform hypergraphs. In *Proceedings of the Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 72–79, 2013.

18 Frank P Kelly. Stochastic models of computer communication systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 47(3):379–395, 1985.

19 Liang Li, Pinyan Lu, and Yitong Yin. Approximate counting via correlation decay in spin systems. In *Proceedings of the 23rd ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 922–940, 2012.

20 Liang Li, Pinyan Lu, and Yitong Yin. Correlation decay up to uniqueness in spin systems. In *Proceedings of the 24th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 67–84, 2013.

21 Chengyu Lin, Jingcheng Liu, and Pinyan Lu. A simple FPTAS for counting edge covers. In *Proceedings of the 25th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 341–348, 2014.

22 Jingcheng Liu and Pinyan Lu. FPTAS for counting monotone CNF. In *Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1531–1548, 2015.

23 Pinyan Lu, Menghui Wang, and Chihao Zhang. FPTAS for weighted fibonacci gates and its applications. In *Proceedings of the 41st International Colloquium on Automata, Languages and Programming (ICALP)*, pages 787–799, 2014.

24 Pinyan Lu, Kuan Yang, and Chihao Zhang. Fptas for hardcore and ising models on hypergraphs. *arXiv preprint arXiv:1509.05494*, 2015.

25 Pinyan Lu and Yitong Yin. Improved FPTAS for multi-spin systems. In *RANDOM*, pages 639–654, 2013.

26 Elchanan Mossel, Dror Weitz, and Nicholas Wormald. On the hardness of sampling independent sets beyond the tree threshold. *Probability Theory and Related Fields*, 143(3-4):401–439, 2009.

27 Ricardo Restrepo, Jinwoo Shin, Prasad Tetali, Eric Vigoda, and Linji Yang. Improved mixing condition on the grid for counting and sampling independent sets. *Probability Theory and Related Fields*, 156(1-2):75–99, 2013.

28 Alistair Sinclair, Piyush Srivastava, and Marc Thurley. Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. *Journal of Statistical Physics*, 155(4):666–686, 2014.

29 Alistair Sinclair, Piyush Srivastava, and Yitong Yin. Spatial mixing and approximation algorithms for graphs with bounded connective constant. In *Proceedings of the 54th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 300–309, 2013.

30 Allan Sly. Computational transition at the uniqueness threshold. In *Proceedings of the 51st IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 287–296, 2010.

31 Allan Sly, Nike Sun, et al. Counting in two-spin models on d-regular graphs. *The Annals of Probability*, 42(6):2383–2416, 2014.

32 Frank Spitzer. Markov random fields on an infinite tree. *The Annals of Probability*, pages 387–398, 1975.

33 Dror Weitz. Combinatorial criteria for uniqueness of Gibbs measures. *Random Structures & Algorithms*, 27(4):445–475, 2005.

34 Dror Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the 38th ACM Symposium on Theory of Computing (STOC)*, pages 140–149, 2006.