Counting hypergraph matchings up to uniqueness threshold

Yitong Yin
Nanjing University, China
yinyt@nju.edu.cn

Jinman Zhao
Nanjing University, China
jinman.zhao@gmail.com

Abstract
We study the problem of approximate counting of weighted matchings in hypergraphs of bounded maximum edge size and maximum degree. The problem is expressed as evaluating the partition function, which is the weighted sum of all matchings in a hypergraph where each matching $M$ is assigned a weight $\lambda^{|M|}$ in terms of a fixed activity parameter $\lambda$. The problem unifies the two most important statistical physics models in approximate counting: the hardcore model for weighted independent set and the monomer-dimer model for weighted matching.

We show that for hypergraph matchings, the critical activity $\lambda_c = \frac{d^d}{k(d-1)^{d+1}}$ is the uniqueness threshold for the Gibbs measure on $(d+1)$-uniform $(k+1)$-regular hypertree. We give an FPTAS for the hypergraphs of maximum edge size at most $d+1$ and maximum degree at most $k+1$, whenever the activity $\lambda < \lambda_c$. This is the first time that a result of this kind is established for a model other than spin systems. We prove this by constructing a hypergraph version of Weitz’s self-avoiding walk tree, and verifying the strong spatial mixing (decay of correlation) of the Gibbs measure.

By a folklore reduction from the hardcore model, there is no FPRAS for the family of hypergraphs as described above if $\lambda > \frac{2k+1+(-1)^k}{k+1} \lambda_c \approx 2\lambda_c$, unless $\text{NP} = \text{RP}$. We also point out a barrier in the existing scheme for proving inapproximability which makes it insufficient to prove the inapproximability approaching the uniqueness threshold for hypergraphs.

1 Introduction
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 partition function for the matchings of $\mathcal{H}$ with activity $\lambda$ is defined as

$$Z_\lambda(\mathcal{H}) = \sum_{M; \text{matching of } \mathcal{H}} w_\lambda(M).$$

Due to a complexity dichotomy [4], the partition functions is $\#P$-hard to compute unless for the trivial settings. We study the problem of approximately computing the partition function for hypergraphs with bounded maximum edge size and bounded maximum degree.

This model represents an interesting subclass of Boolean constraint satisfaction problem (CSP) defined by the matching (packing) constraints. It also unifies the two most important models in

*State Key Laboratory for Novel Software Technology, Nanjing University, China. Supported by NSFC grants 61272081 and 61321491.
approximate counting arising from statistical physics: the hardcore model for graph independent sets and the monomer-dimer model for graph matchings. For instance, consider the family of hypergraphs of maximum edge size bounded by $d + 1$ and maximum degree bounded by $k + 1$:

- When $d = 1$, the model becomes the monomer-dimer model on graphs of maximum degree at most $k + 1$. In a seminal work of Jerrum and Sinclair [12], a fully polynomial-time randomized approximation scheme (FPRAS) is given for approximately computing the partition function. For deterministic algorithms, Bayati et al. [1] gives a deterministic fully polynomial-time approximation scheme (FPTAS) when $k$ is bounded by any constant. As a statistical physics model, it is well-known that the monomer-dimer model has no phase transition. When $d = 2$, the model is called the monomer-trimer model, and has phase transition as shown in [11].

- When $k = 1$, the partition function is actually the sum of weights of independent sets in the dual graph, and the model becomes the hardcore model on graphs of maximum degree at most $d + 1$. Perhaps the most amazing fact for this model is that the transition of computational complexity of approximate counting coincides with the physical phase transition. According to the influential works of Weitz [31], Galanis, Štefankovič and Vigoda [8], and Sly and Sun [28] there is an FPTAS for the partition function when $\lambda < \lambda_c$; and when $\lambda > \lambda_c$ the partition function is inapproximable unless $\text{NP} = \text{RP}$. The critical activity $\lambda_c = \frac{d}{d-1} \frac{d+1}{d} k$, called the uniqueness threshold, is a threshold for a physical phase transition, namely the uniqueness of Gibbs measure on the independent sets of infinite $(d + 1)$-regular tree.

Both the algorithms of Weitz [31] and Bayati et al. [1] utilizes the strong spatial mixing (decay of correlation) property of the Gibbs measure. The Gibbs measure is a probability distribution over configurations proportional to their weights. The weak spatial mixing holds if the boundary-to-variable correlation decays with an exponential rate in the distance between them; and the strong spatial mixing further requires this to hold even conditioning on an arbitrary configuration partially specified on a subset of variables.

For the general model on hypergraphs, the problem has only been studied when $\lambda = 1$, i.e. counting the number of matchings in a hypergraph. In a recent breakthrough in counting CSP, Liu and Lu [18] proves the decay of correlation and the existence of FPTAS for the matchings in 3-uniform hypergraphs of maximum degree at most 4. A similar result is proved for 3-uniform hypergraphs of maximum degree at most 3 in an independent work [5]. Both them rely on case-studies of correlation decay for specific settings of edge sizes and degrees, and hence are insufficient to support a general result with continuous activity $\lambda$.

**Our contributions.** We prove that $\lambda_c(d, k) = \frac{d^k}{(d-1)^{d+1}}$ is the uniqueness threshold for the Gibbs measures on hypergraph matchings in the infinite $(d + 1)$-uniform $(k + 1)$-regular hypertree, which covers as special cases the uniqueness threshold $\lambda_c(d, 1) = \frac{d^d}{(d-1)^{d+1}}$ for the hardcore model on $(d + 1)$-regular tree and lack of phase-transition for the monomer-dimer model on trees (as it corresponds to the case $d = 1$). It also gives an explanation to the bound achieved in [18], as it corresponds to the edge size and degree satisfying the uniqueness threshold when $\lambda = 1$.

We show that the strong spatial mixing and the existence of FPTAS for the hypergraph matchings with activity $\lambda$ up to the uniqueness threshold. As far as we can recall, this is the first time a result of this kind is discovered for a model which is not a spin system.
Theorem 1.1. For any finite \( k \geq 1 \) and \( d \geq 1 \), if \( \lambda < \lambda_c = \frac{d^d}{k(d-1)^{d+1}} \), then for all hypergraphs of maximum edge size at most \( d + 1 \) and maximum degree at most \( k + 1 \), the hypergraph matchings with activity \( \lambda \) exhibit strong spatial mixing with exponential rate, and there exists an FPTAS for the partition function.

The theorem unifies the results of spatial mixing and FPTAS for the hardcore model [31] and the monomer-dimer model [1], and also covers as special cases the results of approximate counting hypergraph matchings in [5,18]. In particular, when \( d = 1 \), the model becomes the monomer-dimer model on graphs of maximum degree at most \( k + 1 \), and the uniqueness condition remains to be satisfiable for unbounded \( k \). In this case, our algorithm outputs an \( \epsilon \)-approximation of the partition function within running time \( (\frac{n}{\epsilon})^{O(\sqrt{\lambda k \log k})} \) even for unbounded \( k \), which achieves the previous best bound by the deterministic algorithm for the monomer-dimer model in [1].

For the hardness of approximation, due to a simple reduction from the hardcore model, the following inapproximability result when \( \lambda > \frac{2k+1+(-1)^{k}}{k+1} \lambda_c \approx 2 \lambda_c \) is folklore.

Theorem 1.2. For any \( k \geq 1 \) and \( d \geq 2 \), if \( \lambda > \frac{2k+1+(-1)^{k}}{k+1} \lambda_c \), where \( \lambda_c = \frac{d^d}{k(d-1)^{d+1}} \), there is no FPRAS for the partition function of hypergraph matchings with activity \( \lambda \) for all hypergraphs of maximum edge size at most \( d + 1 \) and maximum degree at most \( k + 1 \), unless \( \text{NP} = \text{RP} \).

We also explore the possibility of proving the inapproximability when \( \lambda > \lambda_c \). And we show a “barrier” result such that the existing scheme of [6–9, 21, 27, 28] is insufficient to prove the inapproximability approaching the uniqueness threshold.

Remark. For the convenience of analysis, all our results are proved for the independent sets in the dual hypergraphs. Under duality, matchings of a hypergraph of maximum edge size \( d \) and maximum degree \( k \) are equivalent to the independent sets of a hypergraph of maximum edge size \( k \) and maximum degree \( d \). We stress that by our definition, a vertex subset \( I \subseteq V \) is an independent set of a hypergraph \( H = (V,E) \) if \( I \) intersects with every hyperedge in \( H \) for at most one vertex. This should be distinguished with a popular variant of definition for hypergraph independent set (e.g. in [2]), which only requires any independent set \( I \) to contain no hyperedge in \( H \) as subset.

Technical contributions. We define a hypergraph version of Weitz’s self-avoiding walk tree [31], which preserves the marginal probabilities for the hypergraph matchings. The approach has the advantage in showing that for hypergraph matchings, the worst case for the strong spatial mixing is achieved by the infinite \( d \)-uniform \( k \)-regular hypertree, among all hypergraphs of maximum edge size at most \( d \) and maximum degree at most \( k \). Such phenomena were once believed to be true for most statistical physics models, however were found false for even some simple spin systems [16,26].

For the barrier to the inapproximability, we essentially show that the extremal Gibbs measures which realize the uniqueness threshold tightly through the current ways by the existing approaches, cannot be locally-like by any finite hypergraphs. The structure of Gibbs measures on hypergraphs is much more complicated than that for graphs, nevertheless, we characterize the symmetry in the Gibbs measures by a hypergraph version of branching matrix, which is inspired by that of [23], and use it to characterize both the extremal Gibbs measures realizing the uniqueness threshold and the Gibbs measures which have finite hypergraph to be locally alike. This characterization with hypergraph branching matrix is interesting by itself.
Related works. Approximate counting of hypergraph matchings was studied in [13] for hypergraphs with restrictive structures, and in [5, 18] for hypergraphs with bounded edge size and maximum degree. In [2], approximate counting of a variant of hypergraph independent sets was studied, where an independent set only need to not contain any hyperedge.

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 correlation decay was established via the self-avoiding walk tree for the hardcore model [25, 31], monomer-dimer model [1, 24], and two-spin systems [15, 16, 24]. Similar tree-structured recursions were employed to prove correlation decay for multi-spin systems [10, 20, 22], and more general CSPs [17–19].

2 Preliminary

For a hypergraph \( H = (V, E) \), the size of a hyperedge \( e \in E \) is its cardinality \(|e|\), and the degree of a vertex \( v \in V \), denoted by \( \deg(v) = \deg_H(v) \), is the number of hyperedges \( e \in E \) incident to \( v \), i.e. satisfying \( v \in e \). A hypergraph \( H \) is \( k \)-uniform if all hyperedges are of the same size \( k \), and is \( d \)-regular if all vertices has the same degree \( d \). An incident graph of a hypergraph \( H = (V, E) \) 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 \( H = (V, E) \) is a set \( M \subseteq E \) of disjoint hyperedges in \( H \). Given an activity parameter \( \lambda > 0 \), the Gibbs measure is a probability distribution over matchings of \( H \) proportional to the weight \( w_\lambda(M) = \lambda^{|M|} \), defined as \( \mu(M) = w_\lambda(M)/Z \), where the normalizing factor \( Z = Z_\lambda(H) = \sum_M w_\lambda(M) \) is the partition function.

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

\[
\mu(I) = \frac{w_\lambda(I)}{Z} = \frac{\lambda^{|I|}}{Z},
\]

where the normalizing factor \( Z = Z_\lambda(H) = \sum_I w_\lambda(I) \) is the partition function for independent sets of \( H \) with activity \( \lambda \).

The independent sets and matchings are equivalent under hypergraph duality. The dual of a hypergraph \( H = (V, E) \), denoted by \( H^* = (E^*, V^*) \), is a 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 \( H \) is one-to-one correspondent to a hyperedge \( v^* \in V^* \) (and a vertex \( e^* \in E^* \)), such that \( e^* \in v^* \) if and only if \( v \in e \). Observe that the duality establishes an isomorphism between the probability space of matchings of a hypergraph \( H \) and that of the independent sets of the dual hypergraph \( H^* \); and 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. With this equivalence, from now on all propositions for hypergraph matchings are stated for the independent sets of the dual graph.

Given the Gibbs measure over independent sets of hypergraph \( H \) and a vertex \( v \), we define the marginal probability \( p_v \) as

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

which is the probability that \( v \) is occupied by an independent set \( I \) sampled from the Gibbs measure.

Given a vertex set \( \Lambda \subseteq V \), a configuration is a \( \sigma_\Lambda \in \{0, 1\}^\Lambda \) which corresponds to an independent
set $I_{\Lambda}$ partially specified over $\Lambda$ 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 by an independent set $I$ sampled from the Gibbs measure conditioning on the configuration of vertices in $\Lambda \subseteq V$ being fixed as $\sigma_\Lambda$.

A hypergraph $\mathcal{T}$ is a hypertree if its incident graph is a tree. Note that in our notion of hypertrees, any two hyperedges can share at most one common vertex, and likewise, any two vertices can be contained in at most one common hyperedge. Two special hypertrees play important roles. We use $T^d_k$ to denote the infinite $(k+1)$-uniform $(d+1)$-regular hypertree, in which every vertex $v$ is incident to $d+1$ distinct hyperedges, and every hyperedge contains $k+1$ distinct vertices. We emphasize that the incident graph of $T^d_k$ is a tree. We also use $\hat{T}^d_k$ to denote the rooted hypertree which is the same as $T^d_k$ except the degree of the root vertex is $d$ instead of $d+1$. We call $\hat{T}^d_k$ the $(k+1)$-uniform $d$-ary hypertree.

A probability measure over independent sets of an infinite hypergraph $\mathcal{H} = (V, E)$ is Gibbs if for any finite region $\Psi \subseteq V$, conditioning on that all the vertices in the outer boundary $\partial \Psi = \{v \in V \setminus \Psi \mid \exists e \in E \text{ s.t. } v \in e \text{ and } e \cap \Psi \neq \emptyset\}$ are unoccupied gives the same distribution as defined by (1). A formal definition of this so-called “DLR compatibility condition” can be found in [30]. We consider only the simple Gibbs measures satisfying the conditional independence property: conditioning on a configuration of a subset $\Lambda$ of vertices, the marginal distributions of vertices separated by $\Lambda$ are independent of each other. There may be more than one infinite volume Gibbs measures. The uniqueness of Gibbs measure is related to the spatial mixing properties.

**Definition 2.1.** 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 $\Lambda$,

$$|p_v^{\sigma,\Lambda} - p_v^{\tau,\Lambda}| \leq \delta(\text{dist}_H(v, \Lambda)),$$

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

**Definition 2.2.** 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 $\Lambda$,

$$|p_v^{\sigma,\Lambda} - p_v^{\tau,\Lambda}| \leq \delta(\text{dist}_H(v, \Delta)),$$

where $\Delta \subseteq \Lambda$ stands for the subset on which $\sigma_\Lambda$ and $\tau_\Lambda$ differ.

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

### 3 Tree recursion and the uniqueness threshold

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 $i = 1, 2, \ldots, d$,
\begin{itemize}
  \item \(|e_i| = k_i + 1|; and
  \item \(e_i = \{v, v_1, v_2, \ldots, v_{k_i}\}\).
\end{itemize}

For \(1 \leq i \leq d\) and \(1 \leq j \leq k_i\), let \(T_{ij}\) be the sub-hypertree rooted by \(v_{ij}\). Recall that all hypertrees considered by us satisfy that any two hyperedges share at most one common vertex, thus all \(v_{ij}\) are distinct and the sub-hypertrees \(T_{ij}\) 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 \(\Lambda\), and for each \(1 \leq i \leq d\) and \(1 \leq j \leq k_i\), let \(\sigma_{\Lambda_{ij}}\) be the restriction of \(\sigma_\Lambda\) on the sub-hypertree \(T_{ij}\). Consider the ratios of marginal probabilities:

\[R^\sigma_T = p^\sigma_{T,v}/\left(1 - p^\sigma_{T,v}\right)\quad \text{and} \quad R^\sigma_{T_{ij}} = p^\sigma_{T_{ij},v_{ij}}/\left(1 - p^\sigma_{T_{ij},v_{ij}}\right).\]

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

\[R^\sigma_T = \lambda \prod_{i=1}^d \frac{1}{1 + \sum_{j=1}^{k_i} R^\sigma_{T_{ij}}}. \tag{2}\]

This is the “tree recursion” for hypergraph independent sets. The tree recursions for the hardcore model [31] and the monomer-dimer model [1] can both be interpreted as its special cases.

**Definition 3.1.** For positive integers \(d\) and \(k\), we define \(\lambda_c(d, k) = \frac{d^d}{k(d-1)^{d+1}}\).

Consider the symmetric version of the tree recursion \(f_{d,k}(x) = \frac{\lambda}{(1+kx)^d}\), which is the tree recursion on the \((k+1)\)-uniform \(d\)-ary hypertree \(\tilde{T}_k^d\). Since \(f_{d,k}(x)\) is monotonically decreasing for \(x \geq 0\), the fixed point equation \(x = f_{d,k}(x)\) has a unique positive solution, which we denote as \(\hat{x}\).

Representing \(\tilde{x} = kx\), the recursion for the new quantity \(\tilde{x}\) is \(\tilde{f}(\tilde{x}) \equiv kf_{d,k}(x) = \frac{\lambda}{(1+\tilde{x})^d}\), which is precisely the tree recursion for the hardcore model on infinite \(d\)-ary tree with activity \(k\lambda\). Then the following proposition is an immediately consequence to the similar well-known results for the hardcore model [14,21].

**Proposition 3.1.** Let \(\lambda_c = \lambda_c(d, k)\). Consider the two-variate system:

\[y = \frac{\lambda}{(1+kx)^d} \quad \text{and} \quad x = \frac{\lambda}{(1+ky)^d}. \tag{3}\]

1. If \(\lambda \leq \lambda_c\), then \(|f'_{d,k}(\hat{x})| \leq 1\) and \(x = y = \hat{x}\) is the unique solution to the above system.

2. If \(\lambda > \lambda_c\), then \(|f'_{d,k}(\hat{x})| > 1\) and the above system has three distinct positive solutions \((\hat{x}, \hat{x}),(x^+, x^-)\) and \((x^-, x^+)\), where \(0 < x^- < \hat{x} < x^+\).

The second part of the above proposition implies that the weak spatial mixing does not hold for independent sets of \(\tilde{T}_k^d\) with activity \(\lambda > \lambda_c(d, k)\). Moreover, we show that \(\lambda_c(d, k) = \frac{k^d}{k(d-1)^{d+1}}\) is the activity threshold for the uniqueness of Gibbs measure on independent sets of \((k+1)\)-uniform \((d+1)\)-regular hypertree \(\tilde{T}_k^d\).

**Theorem 3.2.** For every positive integer \(d, k\), assuming \(\lambda_c = \lambda_c(d, k)\), we have:

1. \(\tilde{T}_k^d\) with activity \(\lambda\) admits multiple Gibbs measures on hypergraph independent sets if \(\lambda > \lambda_c\);
2. $T^d_k$ with activity $\lambda$ admits a unique Gibbs measure on hypergraph independent sets if $\lambda < \lambda_c$.

The non-uniqueness part of the theorem can be proved by a similar analysis to the one in [21] for the hardcore model.

**Proof of the first statement in Theorem 3.2.** Let $\mu$ be a Gibbs measure on independent sets of $T^d_k$ with activity $\lambda$. By the definition of the measure, for any vertex $v$ in $T^d_k$, we have

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

Assume that $v$ is incident to $d+1$ hyperedges $e_1, e_2, \ldots, e_{d+1}$, where for $i = 1, 2, \ldots, d+1$, the $i$-th hyperedge is $e_i = \{v, v_{i1}, v_{i2}, \ldots, v_{ik}\}$, so $v_{ij}$ denotes the $j$-th neighbor in the $i$-th hyperedge incident to $v$. It holds that

$$\mu[\text{all the neighbors of } v \text{ are unoccupied}] = \mu[v \text{ is occupied}] + \mu[v \text{ is unoccupied}] \prod_{i=1}^{d+1} \mu[\forall 1 \leq j \leq k, v_{ij} \text{ is unoccupied} \mid v \text{ is unoccupied}]$$

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

Note that for any two adjacent vertices $v, v_{ij}$ in a hypertree, by the law of total probability, 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

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

Combining everything together we obtain

$$p_v = \lambda (1 - p_v)^{-d} \prod_{i=1}^{d+1} \left(1 - p_v - \sum_{j=1}^{k} p_{v_{ij}}\right),$$  \hspace{1cm} (4)

where $p_v = \mu[v \text{ is occupied}]$ and $p_{v_{ij}} = \mu[v_{ij} \text{ is occupied}]$.

The above equation in fact gives a system of equations which defines the relations between the marginal probability of being occupied $p_v$ for every vertex $v$ in the hypertree $T^d_k$. We then show that when $\lambda > \lambda_c$, there are multiple solutions to the systems.

We classify the vertices in $T^d_k$ into two categories, called the black vertices and the white vertices. This 2-coloring of vertices imposes a classification of hyperedges of $T^d_k$: we call a hyperedge black if it contains $k$ black vertices and 1 white vertex; and a hyperedge white if it contains $k$ white vertices and 1 black vertex. Our 2-colored $T^d_k$ is then constructed as follows:

1. every hyperedge is either black or white;
2. every black vertex is incidents to 1 black hyperedge and $d$ white hyperedges;
3. every white vertex is incidents to 1 white hyperedge and $d$ black hyperedges.

See Figure 1 for an illustration of this 2-coloring of $T^d_k$.

In Section 7, we will see the color classes for this 2-coloring are orbits for a particular automorphism group $G$. We restrict ourselves to the Gibbs measures which are invariant under the translation of automorphisms in $G$, and assume the marginal probability of being occupied for every black vertex is $p_b$ and the marginal probability of being occupied for every white vertex is $p_w$. Substituting the marginal probabilities in (4) for $p_b$ and $p_w$ according to the 2-coloring, we obtain a simplified system consisting of only two equations which describe the constraints that $p_b$ and $p_w$ must satisfy:

$$
\begin{align*}
\begin{cases}
p_b = \lambda(1 - p_b)^{-d}(1 - k p_b - p_w)(1 - p_b - k p_w)^d, \\
p_w = \lambda(1 - p_w)^{-d}(1 - k p_w - p_b)(1 - p_w - k p_b)^d.
\end{cases}
\end{align*}
$$

Define that $x = \frac{p_b}{1 - p_w - k p_b}$ and $y = \frac{p_w}{1 - p_b - k p_w}$. It is easy to verify that under this translation the system (5) is equivalent to (3), which according to proposition 3.1, has more than one solutions when $\lambda > \lambda_c$.

When $\lambda \leq \lambda_c$, as stated in Proposition 3.1, the solution to the tree recursion is unique, and hence immediate from the analysis in the above proof, we have the uniqueness of Gibbs measure if we are restricted to the Gibbs measures which are invariant under the particular automorphism group $G$. In fact, by the spatial mixing, we can prove the uniqueness of Gibbs measures even if generally considering all Gibbs measures. Due to the argument in [30], it holds that weak spatial mixing implies the uniqueness of Gibbs measure.

**Proposition 3.3** (Weitz [30]). If the independent sets of $T^d_k$ with activity $\lambda$ exhibit weak spatial mixing with a rate $\delta(\cdot)$ that goes to zero, then there is a unique Gibbs measure $\mu$ on independent sets of $T^d_k$ with activity $\lambda$.  

8
The second part of Theorem 3.2, the uniqueness of Gibbs measure, can be implied by combining Proposition 3.3 and the strong spatial mixing result we are going to prove (Theorem 5.1).

4 The hypergraph self-avoiding walk tree

We construct a hypergraph version of self-avoiding-walk tree. It allows us to compute the marginal probability in an arbitrary hypergraph with the tree recursion (6), and more importantly it shows that the $(k+1)$-uniform $(d+1)$-regular hypertree $T^d_k$ is the worst case for SSM for hypergraph independent sets among all hypergraphs of maximum edge-size at most $(k+1)$ and maximum degree at most $(d+1)$.

Theorem 4.1. For any positive integers $k, d$ and any positive $\lambda$, if the independent sets of $T^d_k$ with activity $\lambda$ 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 $\lambda$, exhibit strong spatial mixing with the same rate $\delta(\cdot)$.

Under duality, the same holds for the hypergraph matchings.

We then describe 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 incident 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 for some $0 \leq i \leq \ell - 1$, the suffix $(v_i, e_{i+1}, v_{i+1}, \ldots, e_\ell, v_\ell, e', v')$ of the walk forms a simple cycle in the incident 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 to $u$. With this local ordering of hyperedges, given any vertex $v \in V$, a rooted hypertree $\mathcal{T} = \mathcal{T}_{\text{SAW}}(\mathcal{H}, v)$, 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 groups according to which hyperedge they use to extend the original walk $w$, so that self-avoiding walks within the same groups extends $w$ with the same hyperedge $e'$. For every group, 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 group.

3. We further modify the hypertree $\mathcal{T}$ obtained from the above two steps according to how cycles are closed. For any vertex $v$ 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 \( T \) are deleted from \( T \). Any hyperedges whose size becoming 1 caused by this step are also deleted from \( T \).

The construction is illustrated in Figure 2.

We consider the Gibbs measure of a rooted hypertree \( T \) with activity \( \lambda \), and use \( \mathbb{P}_T^{\sigma_\Lambda} \) to denote the marginal probability of the root of \( T \) being occupied conditioning on configuration \( \sigma_\Lambda \). Note that each vertex \( u \) in \( T_{\text{SAW}}(H, v) \) can be naturally identified (many-to-one) to the vertex in \( H = (V, E) \) at which the self-avoiding walk corresponding to \( u \) ends, thus a configuration \( \sigma_\Lambda \) partially specified on a subset \( \Lambda \subset V \) of vertices in \( H \) can be directly translated to a partially specified configuration in \( T_{\text{SAW}}(H, v) \) through the one-to-many association. We abuse the notation and still denote the resulting configuration in \( T = T_{\text{SAW}}(H, v) \) as \( \sigma_\Lambda \), thus \( \mathbb{P}_T^{\sigma_\Lambda} \) is well-defined.

**Theorem 4.2.** Let \( H = (V, E) \) be a hypergraph and \( \lambda > 0 \). For any \( v \in V, \Lambda \subseteq V \) and \( \sigma_\Lambda \in \{0, 1\}^\Lambda \), it holds that \( p_{H,v}^{\sigma_\Lambda} = \mathbb{P}_T^{\sigma_\Lambda} \) where \( T = T_{\text{SAW}}(H, v) \).

**Proof.** The proof follows the same routine as that of Weitz [31], with some extra cares to be taken to avoid the complications caused by hypergraphs.

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

Let \( d \) be the degree of the root of \( T \). Suppose that there are \( k_i \) children contained in \( i \)-th child-edge, where the order is determined during the construction of \( T_{\text{SAW}}(H, v) \). \( T_{ij} \) is the subtree
rooted at the $j$-th child in the $i$-th child-edge. Let $\Lambda_{ij} = \Lambda \cap T_{ij}$ and $\sigma_{\Lambda_{ij}}$ be the restriction of $\sigma_{\Lambda}$ on $\Lambda_{ij}$. Adapting the tree recursion (2) for the self-avoiding walk tree $T$, we have

$$R_{T}^{\sigma} = \lambda \prod_{i=1}^{d} \frac{1}{1 + \sum_{k=1}^{\lambda} R_{T_{ij}}^{\sigma_{\Lambda_{ij}}}},$$

(6)

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

In the following we describe our procedure for calculating $R_{\Lambda,v}^{\sigma}$ 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 $v_1, v_2, \ldots, v_d$, 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 $T_{SAW}(\mathcal{H}, v)$. At the same time, we associated each $v_i$ with an activity of $\lambda^{1/d}$ rather than $\lambda$. 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, $R_{\Lambda,v}^{\sigma}$ equals to the ratio between the probabilities in $\mathcal{H}^v$ with all $v_i$ ($1 \leq i \leq d$) being occupied and unoccupied, conditioning on $\sigma_{\Lambda}$. Let $\tau_i$ 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:

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

where $\sigma_{\Lambda \tau_i}$ means the combination of the two configurations $\sigma_{\Lambda}$ and $\tau_i$.

We can obtain the value of $R_{\mathcal{H}^v,v_i}^{\sigma_{\tau_i}}$ 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

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

where $k_i$ is the number of the vertices other than $v_i$ which is incident to $e_i$ and $\rho_{ij}$ 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 $R_{\mathcal{H},v}^{\sigma}$ in the same manner that equation (6) has:

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

(7)

Notice that the recursion does terminate, since the number of unfixed vertices reduces at least by one in each step because in calculating $R_{\mathcal{H}^v/v_i,u_{ij}}^{\sigma_{\tau_i} \rho_{ij}}$ 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 $R_{H,\Lambda}^\sigma$ results in the same value as using the hypertree procedure for $T_{SAW}(H, v)$ with corresponding condition of $\sigma_\Lambda$ imposed on it. First notice that the calculation carried out by the two procedure is the same, since they share the same function (Equation (6) and (7)) when we view them as recursive calls. Furthermore, we have the same stopping values for the both recursive procedures. During constructing $T_{SAW}(H, v)$, 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 $\tau$. Eliminating a hyperedge with no child also does not affect the final value of $R_{T,\sigma}^\Lambda$.

Thus, what is left to complete the proof is to show that the hypertree $T_{SAW}(H^v/v_i, u_{ij})$ with $(\sigma_\Lambda, \tau_{ij})$’s corresponding condition imposed on it is exactly the same as the subtree of $T_{SAW}(H, v)$ rooted at the $j$-th child vertex of the $i$-th child-edge of the root with $\sigma_\Lambda$’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 $H$ 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 $H$ starting at $u_{ij}$, except that $T_{SAW}(H^v/v_i, u_{ij})$ 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 $H$ is a sub-hypergraph of another hypergraph $G$ if the incident graph of $H$ is a subgraph of that of $G$, and for hypertrees this is the same definition. Note that for hypergraphs, a subgraph is not necessarily a sub-collection of hyperedges, but maybe also by sub-hyperedges. The $T_{SAW}$ of a hypergraph $H$ with maximum edge-size at most $k + 1$ and maximum degree at most $d + 1$ is sub-hypertree of $T_k^d$.

**Proposition 4.3.** Let $T_0 = (V_0, E_0)$ be a rooted hypertree and $T = (V, E)$ its sub-hypertree with the same root. For any $\Lambda \subseteq V$ and any $\sigma_\Lambda \in \{0, 1\}^\Lambda$, there exists a configuration $\sigma_{\Lambda_0} \in \{0, 1\}^\Lambda$ for $\Lambda \subseteq \Lambda_0 \subseteq V_0$, extending the configuration $\sigma_\Lambda$, such that $\mathbb{P}_{T}^{\sigma_\Lambda} = \mathbb{P}_{T_0}^{\sigma_{\Lambda_0}}$.

The configuration $\sigma_{\Lambda_0}$ just extends $\sigma_\Lambda$ by fixing all the vertices missing in $T$ (actually only those who are closest to the root along each path) to be unoccupied.

**Theorem 4.1** follows immediately from Theorem 4.2 and Proposition 4.3.

**Proof of Theorem 4.1.** Given any hypergraph $H$ of maximum edge-size at most $(k + 1)$ and maximum degree at most $(d + 1)$, by Theorem 4.2 we have $|p_{H,\sigma}^{\sigma_\Lambda} - p_{H,\sigma}^{\tau_\Lambda}| = |\mathbb{P}_{T}^{\sigma_\Lambda} - \mathbb{P}_{T}^{\tau_\Lambda}|$ where $T = T_{SAW}(H, v)$. The distance from the root $v$ to any vertex $u$ in $T$ is no shorter than the distance $H$ between $v$ and the vertex in $H$ to which $u$ is identified. So the SSM with rate $\delta(\cdot)$ on $T$ implies that on the hypergraph $H$.

Since $H$ has maximum edge-size at most $k + 1$ and maximum degree at most $d + 1$, its SAW-tree $T = T_{SAW}(H, v)$ is a sub-hypertree of $T_k^d$. Thus by Proposition 4.3, we have $|\mathbb{P}_{T}^{\sigma_\Lambda} - \mathbb{P}_{T}^{\tau_\Lambda}| = |\mathbb{P}_{T_k^d}^{\sigma_{\Lambda_0}} - \mathbb{P}_{T_k^d}^{\tau_{\Lambda_0}}|$ for some $\sigma_{\Lambda_0}, \tau_{\Lambda_0}$ extending $\sigma_\Lambda, \tau_\Lambda$. The SSM on $T_k^d$ with rate $\delta(\cdot)$ implies that on $T$, which implies the same on the original hypergraph $H$. 

\[ \Box \]
5 Spatial mixing on hypertrees

Recall the definitions in Section 3 that \( f_{d,k}(x) = \frac{\lambda}{(1+kx)^d} \) and \( \hat{x} \) is the unique positive fixed point satisfying \( \hat{x} = f_{d,k}(\hat{x}) \). Let \( f'_{d,k}(\hat{x}) \) be the derivative of \( f_{d,k}(x) \) evaluated at \( x = \hat{x} \).

**Theorem 5.1.** For any positive integers \( d,k \), if \( \lambda < \lambda_c(d,k) \) then the independent sets of the \((k+1)\)-uniform \((d+1)\)-regular hypergraph \( T^d_k \) with activity \( \lambda \) exhibit strong spatial mixing with rate \( \delta(t) \leq C_{d,k,\lambda} \cdot |f'_{d,k}(\hat{x})|^{t/2-1} \), where \( C_{d,k,\lambda} = 2\lambda \sqrt{(d+1)(1+k\lambda)} \sinh^{-1}(\sqrt{k\lambda}) \).

According to Proposition 3.1, \( |f'_{d,k}(\hat{x})| < 1 \) if \( \lambda < \lambda_c(d,k) \). The strong spatial mixing part of the main theorem Theorem 1.1 follows directly by combining Theorem 5.1 and Theorem 4.1, and applying the equivalence between hypergraph independent sets and matchings under duality.

We use a potential analysis to prove the decay of correlation in Theorem 5.1. Let \( \phi : [0,\infty) \to \mathbb{R} \) be a strictly monotone increasing and differentiable function \( \phi : [0,\infty) \to \mathbb{R} \), so the inverse function \( \phi^{-1}(\cdot) \) does exist and is differentiable. The tree recursion

\[
f(x) = \frac{\lambda}{1 + \sum_{j=1}^{k} x_{ij}}
\]

is transformed to a new function \( f^\phi \) defined as follows:

\[
f^\phi(x_{11}, x_{12}, \ldots, x_{dk}) \triangleq \phi(f^{-1}(x_{11}), \phi^{-1}(x_{12}), \ldots, \phi^{-1}(x_{dk})).
\]

Given two input vectors \( x, y \in \phi([0,\infty))^{d\times k} \), denote \( \epsilon = |f^\phi(x) - f^\phi(y)| \), and \( \epsilon_{ij} = |x_{ij} - y_{ij}| \) for every \( 1 \leq i \leq d \) and \( 1 \leq j \leq k \). Due to the mean value theorem, there exists a \( z \in [0,\infty)^{d\times k} \) such that

\[
\epsilon = |f^\phi(x) - f^\phi(y)| \leq \Phi(f(z)) \sum_{i=1}^{d} \sum_{j=1}^{k} \left| \frac{\partial f}{\partial z_{ij}} \right| \frac{\epsilon_{ij}}{\Phi(z_{ij})},
\]

where \( \Phi(x) = \frac{d\phi(x)}{dx} \) denotes the derivative of \( \phi(x) \), and \( \frac{\partial f}{\partial x_{ij}} = \frac{\partial f(x)}{\partial x_{ij}} \big|_{x=z} \) is the partial derivative of \( f(x) \) with respect to \( x_{ij} \) evaluated at \( x = z \).

The **amortized decay rate** \( \alpha(z) \) is defined as

\[
\alpha(z) \triangleq \Phi(f(z)) \sum_{i=1}^{d} \sum_{j=1}^{k} \left| \frac{\partial f}{\partial z_{ij}} \right| \frac{1}{\Phi(z_{ij})} = \Phi(f(z)) f(z) \sum_{i=1}^{d} \frac{1}{1 + \sum_{k_{i=1}^{d} z_{ij}}} \sum_{j=1}^{k} \frac{1}{\Phi(z_{ij})}.
\]

We choose \( \phi(x) = \frac{2}{\sqrt{k}} \sinh^{-1}(\sqrt{kx}) \), so that \( \Phi(x) = \frac{d\phi(x)}{dx} = \frac{1}{\sqrt{x(1+kx)}} \). This choice of potential function is quite similar to the one used in [16, 18, 24, 25], which is basically our potential function without \( k \). We find our choice of potential function is critical for tightly approaching the uniqueness threshold. With this potential function, the \( \alpha(z) \) becomes

\[
\alpha(z) = \sqrt{\frac{f(z)}{1 + kf(z)}} \sum_{i=1}^{d} \frac{\sum_{j=1}^{k} \sqrt{z_{ij}(1+kz_{ij})}}{1 + \sum_{j=1}^{k} z_{ij}}.
\]

We define \( \alpha_{d,k} \triangleq \sup_{z \in [0,\infty)^{d\times k}} \alpha(z) \). By (8), we have \( \epsilon \leq \alpha_{d,k} \max_{i,j} \epsilon_{ij} \).
Lemma 5.2. If \( \lambda \leq \lambda_c(d, k) \) then \( \alpha_{d,k} \leq \sqrt{|f'_{d,k}(\bar{x})|} \). In general, it always holds that \( \alpha_{d,k} \leq \sqrt{kd\lambda} \).

Proof. Given any \( z \in [0, \infty)^{d \times k} \), we first upper bound \( \alpha(z) \) by its symmetrized version where all \( z_{ij} \) are equal. The symmetrization is executed in two steps: we first symmetrize within each \( z_i \), and then symmetrize between \( z_i \)’s.

For every \( 1 \leq i \leq k \), let \( \bar{z}_i = \frac{1}{k} \sum_{j=1}^{k} z_{ij} \). Obviously, \( f(z) = f_k(\bar{z}) \) where

\[
\bar{z} \triangleq (\bar{z}_i)_{i=1,2,...,d} \quad \text{and} \quad f_k(\bar{z}) \triangleq \lambda \prod_{i=1}^{d} \frac{1}{1 + k \bar{z}_i}.
\]

Since function \( g(x) = \sqrt{x(1 + kx)} \) is concave, we have \( \sum_{j=1}^{k} \sqrt{\bar{z}_{ij}(1 + k\bar{z}_{ij})} \leq k\sqrt{\bar{z}_i(1 + k\bar{z}_i)} \) due to Jensen’s inequality. The \( \alpha(z) \) can then be upper bounded by

\[
\alpha(z) \leq \sqrt{\frac{k f_k(\bar{z})}{1 + k f_k(\bar{z})}} \sum_{i=1}^{d} \sqrt{\frac{k \bar{z}_i}{1 + k \bar{z}_i}} \triangleq \alpha(\bar{z}).
\]

Given a \( \bar{z} = (\bar{z}_i)_{i=1,2,...,d} \), let \( \bar{z} \) be defined as \( \bar{z} = \frac{1}{k} \left( \left[ \prod_{i=1}^{d} (1 + k \bar{z}_i)^{1/d} \right] - 1 \right) \). It can be easily verified that \( f_k(\bar{z}) = f_{d,k}(\bar{z}) \).

We further denote that \( w = \ln(1 + k\bar{z}) \) and \( w_i = \ln(1 + k\bar{z}_i) \) for every \( 1 \leq i \leq d \). It holds that \( w = \frac{1}{d} \sum_{i=1}^{d} w_i \). Note that \( \sqrt{\frac{k \bar{z}_i}{1 + k \bar{z}_i}} = \sqrt{1 - e^{-w_i}} \) and the function \( h(x) = \sqrt{1 - e^{-x}} \) is concave. Due to Jensen’s inequality, we have

\[
\sum_{i=1}^{d} \sqrt{\frac{k \bar{z}_i}{1 + k \bar{z}_i}} = \sum_{i=1}^{d} \sqrt{1 - e^{-w_i}} \leq d\sqrt{1 - e^{-w}} = d\sqrt{\frac{k \bar{z}}{1 + k \bar{z}}}
\]

Therefore, \( \alpha(\bar{z}) \) can be upper bounded by its fully symmetrized version

\[
\alpha(\bar{z}) \leq \sqrt{\frac{k df_{d,k}(\bar{z})}{1 + k f_{d,k}(\bar{z})}} \sqrt{\frac{k \bar{z}}{1 + k \bar{z}}} \triangleq \alpha(\bar{z}).
\]

We then upper bound the uni-variate function \( \alpha(z) \). The derivative of function \( \alpha(z) \) is given by

\[
\alpha'(z) = \frac{\alpha(z)}{2} \left[ \frac{f'_{d,k}(z)}{f_{d,k}(z)} + \frac{1}{z} - \frac{k f'_{d,k}(z)}{1 + k f_{d,k}(z)} - \frac{k}{1 + k z} \right] = \frac{\alpha(z)}{2(1 + k z)} \left[ \frac{1}{z} - \frac{kd}{1 + k f_{d,k}(z)} \right].
\]

So the critical point of \( \alpha(z) \) is given by \( z = z^* \), where \( z^* \) is the unique positive solution to

\[
1 + k f_{d,k}(z) = kd z.
\]

It also holds that \( \alpha'(z) > 0 \) when \( z < z^* \) and \( \alpha'(z) < 0 \) when \( z > z^* \), thus \( \alpha(z) \) achieves the maximum at \( z = z^* \). So \( \alpha(\bar{z}) \) can be upper bounded by

\[
\alpha(\bar{z}) \leq \alpha(z^*) = \sqrt{\frac{k df_{d,k}(z^*)}{1 + k z^*}}.
\]
where the equality holds by substituting $z^*$ according to (9). We then define a new function $\tilde{\alpha}(x) \triangleq \sqrt{\frac{kd_d(k)x}{1+kx}}$. The function $\tilde{\alpha}(x)$ has the good properties that $\tilde{\alpha}(z^*) = \alpha(z^*)$ and $\tilde{\alpha}(\hat{x}) = \sqrt{|f'_{d,k}(\hat{x})|}$ where $\hat{x} = f_{d,k}(\hat{x})$ is the unique fixed point of $f_{d,k}(x)$. It is also easy to see that $\tilde{\alpha}(x)$ is monotonically decreasing in $x$. Then $\tilde{\alpha}(x) \leq \tilde{\alpha}(0) = \sqrt{kd\lambda}$. Combining with the above argument, we have

\[
\alpha(z) \leq \alpha(\hat{z}) \leq \alpha(z^*) = \tilde{\alpha}(z^*) \leq \sqrt{kd\lambda}.
\]

This proves the generic upper bound of $\alpha_{d,k}$ in the lemma.

We then show the upper bound of $\alpha_{d,k}$ when $\lambda \leq \lambda_c(d,k)$. Assume that $\lambda \leq \lambda_c(d,k)$. Due to proposition 3.1, it holds that $|f_{d,k}(|x|) \leq 1$. We then show that $\hat{x} \leq z^*$. By contradiction, suppose that $\hat{x} > z^*$. Since $z^*$ is the solution to (9) and $f_{d,k}(x)$ is strictly decreasing in $x$, it holds that $1 + k\hat{x} = 1 + kf_{d,k}(\hat{x}) < kd\hat{x}$. On the other hand the uniqueness condition $|f_{d,k}(\hat{x})| = \frac{k\hat{x}}{1+k\hat{x}} \leq 1$ guarantees that $1 + k\hat{x} \geq kd\hat{x}$, a contradiction. Therefore we have $\hat{x} \leq z^*$. Since $\tilde{\alpha}(x)$ is decreasing in $x$, we have $\tilde{\alpha}(z^*) \leq \tilde{\alpha}(\hat{x})$. Combining everything, we have

\[
\alpha(z) \leq \alpha(\hat{z}) \leq \alpha(z^*) = \tilde{\alpha}(z^*) \leq \sqrt{|f'_{d,k}(\hat{x})|}.
\]

Proof of Theorem 5.1. Consider the infinite $(k+1)$-uniform $d$-ary hypertree $\mathcal{T} = \hat{T}_k^d$ rooted by vertex $\rho$. For $1 \leq i \leq d$ and $1 \leq j \leq k$, let $\rho_{ij}$ denote the $j$-th child in the $i$-th hyperedge incident to $\rho$ and $T_{ij}$ the subtree of $\mathcal{T}$ rooted by $\rho_{ij}$. By definition of $\hat{T}_k^d$, each $T_{ij}$ is isomorphic to $\hat{T}_k^d$. Let $\Lambda$ be a set of vertices such that any infinite path from $\rho$ crosses $\Lambda$. Let $\sigma, \tau \in \{0, 1\}^{\Lambda}$ be two independent sets partially specified on $\Lambda$ such that $\sigma$ and $\tau$ disagree over $\Delta \subset \Lambda$. For $1 \leq i \leq d$ and $1 \leq j \leq k$, let $\sigma_{ij}$ and $\tau_{ij}$ be the respective partially specified independent sets obtained from restricting $\sigma$ and $\tau$ on $T_{ij}$. Suppose that the shortest distance from $\rho$ to any vertex in $\Delta$ is $\ell + 1$.

We then prove by induction that

\[
|\phi(R^\sigma_{\ell}) - \phi(R^\tau_{\ell})| \leq \frac{\eta}{\sqrt{k}} \sinh^{-1}(\sqrt{k\lambda}) \cdot \alpha_{d,k}^\ell.
\]

(10)

For all $\ell \geq 0$, if the root $\rho$ is already in $\Lambda$, then $\sigma$ and $\tau$ must assign the same state to $\rho$, which means $R^\sigma_{\ell} = R^\tau_{\ell}$, thus $|\phi(R^\sigma_{\ell}) - \phi(R^\tau_{\ell})| = 0$. Therefore we only need to deal with the cases where $\rho \notin \Lambda$ and hence both $R^\sigma_{\ell}$ and $R^\tau_{\ell}$ are determined by the tree recursion (2), which in our tree $\mathcal{T} = \hat{T}_k^d$, becomes

\[
R^\sigma_{\ell} = f(R^\sigma_{\ell-1}, 1 \leq i \leq d, 1 \leq j \leq k) = \lambda \prod_{i=1}^{d} \frac{1}{1 + \sum_{j=1}^{k} R^\sigma_{\ell-1, ij}},
\]

for $\eta \in \{\sigma, \tau\}$.

When $\ell = 0$, due to the monotonicity of the tree recursion $f(x)$, we have $0 = f(\infty, \ldots, \infty) \leq R^0_{\ell} \leq f(0, \ldots, 0) = \lambda$. Therefore, $|\phi(R^\sigma_{\ell}) - \phi(R^\tau_{\ell})| \leq |\phi(\lambda) - \phi(0)| = \frac{\eta}{\sqrt{k}} \sinh^{-1}(\sqrt{k\lambda})$.

For $\ell \geq 1$, by (8) and the definition of $\alpha_{d,k}$, we have

\[
|\phi(R^\sigma_{\ell}) - \phi(R^\tau_{\ell})| \leq \alpha_{d,k} \max_{i,j} |\phi(R^\sigma_{\ell-1, ij}) - \phi(R^\tau_{\ell-1, ij})|.
\]
Note that the shortest distance from the root $\rho_{ij}$ of each $\mathcal{T}_{ij}$ to the set $\Delta_{ij}$ of vertices on which $\sigma_{ij}$ and $\tau_{ij}$ disagree is at least $\ell$, and due to Lemma 5.2 and Proposition 3.1, $\alpha_{d,k} \leq \sqrt{|f'_{d,k}(\hat{x})|} \leq 1$ when $\lambda < \lambda_c(d,k)$. Then by the induction hypothesis, $|\phi\left(R_{ij}^\sigma\right) - \phi\left(R_{ij}^\tau\right)| \leq \frac{2}{\sqrt{k}} \sinh^{-1}(\sqrt{k\lambda}) \cdot \alpha^{t-1}_{d,k}$. Therefore (10) holds.

Now consider the infinite $(k+1)$-uniform $(d+1)$-regular simple hypertree $\mathcal{T}^* = \mathbb{T}^d_k$. Let $\Psi$ be a finite region in $\mathcal{T}^*$ and $v \in \Psi$. We consider $\mathcal{T}^*$ to be rooted by $v$. For $1 \leq i \leq d+1$ and $1 \leq j \leq k$, let $v_{ij}$ denote the $j$-th child in the $i$-th hyperedge incident to $v$ and $\mathcal{T}_{ij}$ the subtree of $\mathcal{T}^*$ rooted by $v_{ij}$. By definition of $\mathbb{T}^d_k$, each $\mathcal{T}_{ij}^*$ is a $(k+1)$-uniform $d$-ary hypertree $\mathbb{T}^d_k$ rooted at $v_{ij}$.

Let $\sigma, \tau \in \{0,1\}^{\partial \Psi}$ be two independent sets partially specified on the vertex boundary $\partial \Psi$ of $\Psi$. And similarly, for $1 \leq i \leq d+1$ and $1 \leq j \leq k$, let $\sigma_{ij}$ and $\tau_{ij}$ be the respective partially specified independent sets obtained from restricting $\sigma$ and $\tau$ on $\mathcal{T}_{ij}^*$. The tree recursion gives that

$$R_{\mathcal{T}^*}^{\sigma} = \lambda \prod_{i=1}^{d+1} \left( \frac{1}{1 + \sum_{j=1}^{k} R_{\mathcal{T}_{ij}^*}^{\sigma}} \right),$$

for $\eta \in \{\sigma, \tau\}$.

Suppose that $\sigma$ and $\tau$ disagree over $\Delta \subseteq \partial \Psi$ and the shortest distance from $v$ to any vertices in $\Delta$ is $t$. For $t = 1$, by the same monotonicity of the tree recursion and $\alpha_{d+1,k} > \alpha_{d,k}$, we have $|\phi\left(R_{\mathcal{T}^*}^\sigma\right) - \phi\left(R_{\mathcal{T}^*}^\tau\right)| \leq \frac{2}{\sqrt{k}} \sinh^{-1}(\sqrt{k\lambda}) \leq \frac{2}{\sqrt{k}} \sinh^{-1}(\sqrt{k\lambda}) \cdot \alpha_{d+1,k} \cdot \alpha^{-1}_{d,k}$.

For $t \geq 2$, by (8) and the definition of $\alpha_{d,k}$, we have

$$|\phi\left(R_{\mathcal{T}^*}^\sigma\right) - \phi\left(R_{\mathcal{T}^*}^\tau\right)| \leq \alpha_{d+1,k} \max_{i,j} |\phi\left(R_{\mathcal{T}_{ij}^*}^\sigma\right) - \phi\left(R_{\mathcal{T}_{ij}^*}^\tau\right)|.$$

Then the shortest distance from the root $v_{ij}$ of each $\mathcal{T}_{ij}^*$ to the set $\Delta_{ij}$ of vertices on which $\sigma_{ij}$ and $\tau_{ij}$ disagree is at least $t - 1$. Note again that $\alpha_{d,k} \leq 1$ if $\lambda < \lambda_c(d,k)$. Applying (10) with $\mathcal{T} = \mathcal{T}_{ij}^*$ and $\Lambda = \partial \Psi \cap \mathcal{T}_{ij}$, we have

$$|\phi\left(R_{\mathcal{T}^*}^\sigma\right) - \phi\left(R_{\mathcal{T}^*}^\tau\right)| \leq \frac{2}{\sqrt{k}} \sinh^{-1}(\sqrt{k\lambda}) \cdot \alpha_{d+1,k} \cdot \alpha^{-2}_{d,k}. \quad (11)$$

At last, due to the mean value theorem, it holds that

$$|R_{\mathcal{T}^*}^\sigma - R_{\mathcal{T}^*}^\tau| = \frac{1}{\Phi(\xi)} |\phi\left(R_{\mathcal{T}^*}^\sigma\right) - \phi\left(R_{\mathcal{T}^*}^\tau\right)|, \quad (12)$$

where $\xi$ is some value between $R_{\mathcal{T}^*}^\sigma$ and $R_{\mathcal{T}^*}^\tau$. By the same monotonicity of the tree recursion, we still have $R_{\mathcal{T}^*}^\sigma, R_{\mathcal{T}^*}^\tau \in [0, \lambda]$, thus $\frac{1}{\Phi(\xi)} \leq \frac{1}{\Phi(\lambda)} = \sqrt{\lambda(1 + k\lambda)}$. Therefore, (11) and (12) together gives

$$|R_{\mathcal{T}^*}^\sigma - R_{\mathcal{T}^*}^\tau| \leq 2 \sqrt{\frac{\lambda(1 + k\lambda)}{k}} \sinh^{-1}(\sqrt{k\lambda}) \cdot \alpha_{d+1,k} \cdot \alpha^{-2}_{d,k}.$$

Due to Lemma 5.2, $\alpha_{d+1,k} \leq \sqrt{(d+1)k\lambda}$ and if $\lambda < \lambda_c(d,k)$ then $\alpha_{d,k} \leq \sqrt{|f'_{d,k}(\hat{x})|}$. Therefore, if $\lambda < \lambda_c(d,k)$, we have

$$\left| R_{\mathcal{T}_{ij}^*}^\sigma_{ij,v} - R_{\mathcal{T}_{ij}^*}^\tau_{ij,v} \right| \leq C_{d,k,\lambda} \cdot |f'_{d,k}(\hat{x})|^{\frac{1}{2} - 1},$$

16
where \( C_{d,k,\lambda} = 2\lambda \sqrt{(d+1)(1+k\lambda)} \sinh^{-1}(\sqrt{k\lambda}) \). Because of the translation \( p_v = \frac{R_v}{1+R_v} \) between occupation probability and ratio, we have
\[
|p_{\mathcal{T}_{kj}^\varepsilon}^x - p_{\mathcal{T}_{kj}^\tau}^x| \leq |R_{\mathcal{T}_{kj}^\varepsilon} - R_{\mathcal{T}_{kj}^\tau}| \leq C_{d,k,\lambda} \cdot |f'_{d,k}(\hat{x})|^\frac{1}{\varepsilon} - 1,
\]
Which proves the strong spatial mixing of independent sets of \( \mathbb{P}^d_k \) with activity \( \lambda \) with rate \( C_{d,k,\lambda} \cdot |f'_{d,k}(\hat{x})|^\frac{1}{\varepsilon} - 1 \) when \( \lambda < \lambda_c(d,k) \).

\[ \square \]

6 FPTAS for computing the partition function

For \( 0 < \varepsilon < 1 \), a value \( \hat{Z} \) is said to be an \( \varepsilon \)-approximation of \( Z \) if \( (1 - \varepsilon)Z \leq \hat{Z} \leq (1 + \varepsilon)Z \).

**Theorem 6.1.** If \( \lambda < \lambda_c(d,k) \), 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 \( \varepsilon \)-approximation of the partition function for the independent sets of \( \mathcal{H} \) with activity \( \lambda \), within running time \( \left( \frac{n}{\varepsilon} \right)^{O(k \log kd)} \), where \( \kappa = -1/\log |f'_{d,k}(\hat{x})| \).

By duality, when \( \lambda < \lambda_c(d,k) \), 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 \((k+1)\). According to Proposition 3.1, \( |f'_{d,k}(\hat{x})| < 1 \) if \( \lambda < \lambda_c(d,k) \), so the running time is \( \text{Poly}(n, 1/\varepsilon) \) for any bounded \( k \) and \( d \) when \( \lambda < \lambda_c(d,k) \), i.e. the algorithm is an FPTAS. The algorithmic part of the main theorem Theorem 1.1 is proved.

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

**Corollary 6.2.** There exists an algorithm which given any graph \( G \) of maximum degree at most \( \Delta \), and any \( \varepsilon > 0 \), returns an \( \varepsilon \)-approximation of the partition function for the matchings of \( G \) with activity \( \lambda \), within running time \( \left( \frac{n}{\varepsilon} \right)^{O(\sqrt{n \Delta} \log \Delta)} \).

With the construction of hypergraph self-avoiding walk tree and the SSM, the algorithm follows the framework by Weitz [31]. We will describe an algorithm of approximating weighted independent sets in hypergraphs. Under duality this is the same as counting weighted matchings.

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 \( \mu(\emptyset) \) as it is exactly \( 1/Z \). Let \( \emptyset_i \) be the configuration on vertices \( v_1 \) up to \( v_i \) where all of them are unoccupied, and \( p_{\emptyset_i}^{\varepsilon-1} \) the probability of \( v_i \) being occupied conditioning on all vertices \( v_1 \) up to \( v_{i-1} \) being unoccupied. Then we have \( 1/Z = \prod_{i=1}^{n} (1 - p_{\emptyset_i}^{\varepsilon-1}) \). Due to the lower bound \((1 - p_{\emptyset_i}^{\varepsilon-1}) \geq \frac{1}{1+\varepsilon} \) for the probability of vertex unoccupied by an independent set, to get an \( \varepsilon \)-approximation of \( Z \), it suffices to approximate each of \( p_{\emptyset_i}^{\varepsilon-1} \) within an additive error \( 2(1+\varepsilon)^n \).

By Theorem 4.1, we have \( p_{\emptyset_i}^{\varepsilon} = P_{\mathcal{T}_i}^\varepsilon \) 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 \( P_{\mathcal{T}_i}^{\varepsilon} \) can be computed by the tree recursion (2). 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. By the strong spatial mixing provided by Theorem 5.1, if $\lambda < \lambda_c(d, k)$, then the additive error of such estimation of $p^*_v$ is bounded by $C_{d,k,\lambda} \cdot |f'_{d,k}(\hat{x})|^{t/2-1}$. The overall running time of the algorithm is clearly $O(n(kd)^t)$ where $t$ is the depth of the recursion. We shall choose an integer $t$ such that $C_{d,k,\lambda} \cdot |f'_{d,k}(\hat{x})|^{t/2-1} \leq \frac{\epsilon}{2(1+\lambda)n}$, which gives us the time complexity.

7 Inapproximability and a barrier

We restate Theorem 1.2 here under the view of hypergraph independent sets. It is proved by combining a simple reduction in [2] and the inapproximability of the hardcore model [9, 28].

**Theorem 7.1.** For any $k \geq 1$ and $d \geq 2$, if $\lambda > \frac{2k+1+(-1)^k}{k+1} \lambda_c(d, k)$, there is no FPRAS for the partition function of independent sets of hypergraph with activity $\lambda$ for hypergraphs of maximum edge size at most $k + 1$ and maximum degree at most $d + 1$, unless NP=RP.

**Proof.** The reduction is as described in [2], which is reduced from the hardcore model. Given a graph $G(V, E)$ with maximum degree at most $(d + 1)$, we construct a hypergraph $\mathcal{H}(V, E)$ 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 $V_{\mathcal{H}} = \{w_{v, i} \mid w \in V, 1 \leq i \leq t\}$. 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 $2t \leq k + 1$. 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 $G$ 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 [28], for any $d \geq 2$, unless NP=RP, there is no FPRAS for approximately computing $Z_G(\lambda')$ if $\lambda' > \frac{d^d}{(d-1)^{d-1}}$, i.e. when

$$\lambda > \frac{d^d}{(k+1)^2(d-1)^{d-1}} = \frac{2k+1+(-1)^k}{k+1} \lambda_c(d, k).$$

We then explore the possibility of proving the inapproximability when $\lambda > \lambda_c(d, k)$ and find a limitation of the existing approaches towards this goal.

The known gadgets utilizing the non-uniqueness property [6–9, 21, 27, 28] necessarily consist of the following ingredients:

1. The extremal measures achieving the tree non-uniqueness threshold classify the vertices of the regular tree into equivalent classes, e.g. the semi-translation-invariant Gibbs measures which are invariant under parity-preserving automorphisms on the regular tree classify the vertices according to their parities.

2. There is an efficiently constructable finite graph (e.g. random regular bipartite graph) which is “locally like” the infinite regular tree along with the above classification of vertices.
A related sufficient condition for gadgets was formulated as “nearly-independent phase-correlated spins” in a recent work [3].

For hypergraphs, the equivalent classes induced by the symmetry of Gibbs measure can be much complicated. We give a complete characterization of such equivalent classes. We also formalize the notions of “achieving the tree non-uniqueness threshold” for a Gibbs measure and “locally likeness” between hypergraphs with vertex-classifications.

We show such a barrier result: by the existing ways for the Gibbs measures achieving the non-uniqueness threshold, the classification of vertices described in Section 3 (illustrated in Figure 1) is the only one who may tightly achieve the threshold, but there exists no finite hypergraph which is locally like it.

7.1 A branching matrix characterization of Gibbs measures

An automorphism on $\mathcal{H}$ is a one-to-one correspondence from its vertices and hyperedges to themselves so that the incidence relation is preserved. Automorphisms may form groups. Let $G$ be an automorphism group on hypergraph $\mathcal{H}$. A Gibbs measure on independent sets of hypergraph $\mathcal{H}$ is $G$-translation-invariant if it is invariant under all automorphisms from $G$. For example, the semi-translation-invariant measure on $(d+1)$-regular tree in [9,21] is $A$-translation-invariant on the $(d+1)$-regular tree $T^d_1$, where $A$ is the group of all parity-preserving automorphisms.

Given an automorphism group $G$ on a hypergraph $\mathcal{H} = (V,E)$, the orbits induced by the natural group action by the automorphisms on vertices form a classification $C$ of vertices, which partitions $V$ into equivalent classes, such that for a $G$-translation-invariant Gibbs measure, the vertices from the same orbits have the same marginal probability. We call such $C$ an orbit classification induced by automorphism group $G$. Not all classifications of vertices are orbits. We then characterize the finite orbit classifications by a notion of branching matrix for hypertrees.

We extend the idea of branching matrix introduced in [23,29] for generating infinite trees with finite types of vertices, to the hypertrees. For a $(k+1)$-uniform hypergraph, supposed that we classify all vertices into $t$ types, every hyperedge can be identified to an edge-profile, which is a $t$-tuple $(k_1,k_2,...,k_t)$ of nonnegative integers satisfying $\sum_{i=1}^{t} k_i = k + 1$, such that the hyperedge contains precisely $k_i$ vertices of each type $i$. We denote $\tau(k,t) = \binom{t+k}{t-1}$, which gives the total number of possible edge-profiles for the $(k + 1)$-uniform hyperedges composed from $t$ types of vertices, because $\tau(k,t)$ is the total number of multisets of size $k + 1$ defined from $t$ distinct types of elements. We enumerate the edge-profiles in the lexicographic order of $j = (k_1,k_2,...,k_t)$.

**Definition 7.1.** A hypergraph branching matrix for generating $(k + 1)$-uniform hypertree, is a $t \times \tau$ matrix $M$ whose entries are nonnegative integers, and satisfying that $\tau = \tau(k,t)$ and $M_{ij} = 0$ if and only if the edge-profile $j = (k_1,k_2,...,k_t)$ has $k_i = 0$.

Given a hypergraph branching matrix $M$ of size $t \times \tau(k,t)$, a pair $(T,C)$ of an infinite $(k + 1)$-uniform hypertree $T$ and a classification $C$ of vertices of $T$ into $t$ types are said to be generated by $M$ if every vertex in $T$ of type $i$ is incident to exactly $M_{ij}$ hyperedges of edge-profile $j$.

For instance, the hypertree $T^2_3$ with two types of vertices illustrated in Figure 1 is generated by the hypergraph branching matrix $M = (0 0 1 0 0 0)$.

**Proposition 7.2.** A finite classification $C$ of vertices of an infinite hypertree $T$ is an orbit classification if and only if there is a hypergraph branching matrix $M$ generating $(T,C)$. 

19
Proof. For any finite orbit classification \( \mathcal{C} \) of vertices of \( T = (V,E) \) induced by an automorphisms group \( G \). By definition of orbits, for any vertex \( u, v \in V \), \( u, v \) are equivalent under \( \mathcal{C} \) if and only if there exists an automorphism \( \phi \in G \) that maps \( u \) to \( v \), i.e. \( \phi(u) = v \); and every automorphism \( \phi \in G \) preserves classes in \( \mathcal{C} \). Therefore, it can be verified that any two vertices \( u \) and \( v \) of the same class under \( \mathcal{C} \) have precisely the same number of hyperedges for each edge-profile, which means there is a hypergraph branching matrix generating \( (T, \mathcal{C}) \).

Let \( (T, \mathcal{C}) \) be generated by a hypergraph branching matrix \( M \). For every pair of vertices \( u, v \) with the same type, by generating the hypertree according to \( M \) starting from \( u \) and \( v \) respectively, we obtain an automorphism \( \phi_{u \rightarrow v} \) which maps \( u \) to \( v \) and preserves the types of all vertices. Let \( G = \langle \{ \phi_{u \rightarrow v} \mid \forall u, v \text{ with the same type} \} \rangle \) be the group generated from all such automorphisms, \( \mathcal{C} \) is the orbits classification induced by it. \( \square \)

7.2 Gibbs measures supporting the tree recursion

Recall that by Proposition 3.1, the uniqueness threshold \( \lambda_c(d, k) \) is the threshold for \( \lambda \) to make the following two-variate system:

\[
\frac{q^+}{1 - q^+} = \lambda \left( 1 + \frac{kq^-}{1 - q^-} \right)^{-d} \quad \text{and} \quad \frac{q^-}{1 - q^-} = \lambda \left( 1 + \frac{kq^+}{1 - q^+} \right)^{-d} \quad (13)
\]

have unique solution, where \( q^+ \) and \( q^- \) are the marginal probabilities of the root being occupied in the extremal measures on the \((k + 1)\)-uniform \( d \)-ary hypertree \( \tilde{T}_d^k \).

As noted in [21] for the hardcore model, a natural way of translating the marginal probability of a vertex \( v \) in the \((k + 1)\)-uniform \((d + 1)\)-regular hypertree \( \tilde{T}_d^k \) to the marginal probability of the root of \( \tilde{T}_d^k \) is to pick an arbitrary hyperedge \( e \) incident to \( v \) and conditioning on all vertices in \( e \) except \( v \) being unoccupied.

To conclude, the framework consists of the following two steps:

- (edge pinning) For every vertex \( v \) in \( T_d^k \), we may pick a hyperedge \( e \) incident to \( v \) according to certain rules and conditioning on all vertices in \( e \) except \( v \) being unoccupied.

- (tree simulation) The goal is that there are only two kinds of resulting marginal probabilities \( q^+, q^- \) obtained from the above procedure and they satisfy (13).

In Section 3, a family of \( G \)-translation-invariant Gibbs measures is constructed according to the framework, whose orbits on \( T_d^k \) (Figure 1) is characterized by the following hypergraph branching matrix:

\[
B^{k,d}_d = \begin{pmatrix}
0 & d & 0 & \cdots & 0 & 1 & 0 \\
0 & 1 & 0 & \cdots & 0 & d & 0 \\
\end{pmatrix}_{k+2}
\quad (14)
\]

And for \( k = 1 \), \( B^{1,d}_d = \begin{pmatrix}
0 & d+1 & 0 \\
0 & d+1 & 0 \\
\end{pmatrix} \) gives the parity-preserving orbits on \( d \)-regular tree \( T_1^d \). We then show that this is the only possible classification of vertices in \( T_d^k \) to achieve the above goal.

**Proposition 7.3.** For any \( G \)-translation-invariant Gibbs measures on \( T_d^k \) constructed according to the above framework, the orbits of \( G \) must be generated by \( B^{k,d}_d \).
Proof. Since in the tree simulation step it is required to result in exactly two kinds of marginal probabilities $q^\pm$, we consider only the automorphism groups $G$ on $T^d_k$ with two orbits. Denote the two orbits as $V^+$ and $V^-$. Since the Gibbs measure is $G$-translation-invariant, we may assume the marginal probabilities of being occupied for the vertices from $V^+$ and $V^-$ to be $p^+$ and $p^-$ respectively. The hypergraph branching matrix $M$ generating $(T^d_k, C)$ at this circumstance is a $2 \times (k + 2)$ non-negative integer matrix:

$$M = \begin{pmatrix} d_0^+ & d_1^+ & \cdots & d_{k+1}^+ \\ d_0^- & d_1^- & \cdots & d_{k+1}^- \end{pmatrix},$$

which obeys that $\sum_{i=0}^{k+1} d_i^\pm = d + 1$, since $T^d_k$ is $(d + 1)$-regular.

We recall some generic relations between marginal probabilities of vertices in $T^d_k$. Let $v$ be a vertex in $T^d_k$. Assume $e_v \ni v$ to be the incident hyperedge chosen in the edge pinning step of the framework. By the definition of Gibbs measure, we have

$$\mu[v \text{ is occupied}] = \mu[v \text{ is occupied} \mid \text{all the neighbors incident to } e_v \text{ are unoccupied}]$$
$$\mu[\text{all the neighbors incident to } e_v \text{ are unoccupied}].$$

Let $v_j$ denote the $j$-th neighbour of $v$ in hyperedge $e_v$. It holds that

$$\mu[\text{all the neighbors incident to } e_v \text{ are unoccupied}]$$
$$= \mu[v \text{ is occupied}] + \mu[v \text{ is unoccupied}] \left( 1 - \sum_{j=1}^{k} \mu[v_j \text{ is occupied} \mid v \text{ is unoccupied}] \right)$$
$$= \mu[v \text{ is occupied}] + \mu[v \text{ is unoccupied}] \left( 1 - \sum_{j=1}^{k} \frac{\mu[v_j \text{ is occupied}]}{\mu[v \text{ is unoccupied}]} \right).$$

We further denote that

$$p_v = \mu[v \text{ is occupied}]$$
$$q_v = \mu[v \text{ is occupied} \mid \text{all the neighbors incident to } e_v \text{ are unoccupied}].$$

Then combining the above equations, we have

$$q_v = \frac{p_v}{1 - \sum_{j=1}^{k} p_{v_j}}. \tag{15}$$

According to the tree simulation step in the framework, only two possible marginal probabilities $q^\pm$ can be produced by this procedure, which requires that the hyperedge $e_v$ selected in the edge pinning step must have the same edge-profile for all vertices $v$ from the same orbit.

Suppose that for all $v^+ \in V^+$, the selected hyperedge $e_{v^+}$ has edge-profile $(r, k + 1 - r)$; and for all $v^- \in V^-$, the selected hyperedge has edge-profile $(s, k + 1 - s)$. Then Equation (15) is translated to

$$q^+= \frac{p^+}{1 - (r-1)p^+ - (k+1-r)p^-}, \tag{16}$$
$$q^-= \frac{p^-}{1 - (s-1)p^+ - (k+1-s)p^-}.$$
The tree simulation step further requires that $q^\pm$ satisfy (13). Substituting $q^\pm$ with $p^\pm$ according to (16) gives that

$$\begin{align*}
p^+(1 - s p^+ - (1 - s) p^-)^d &= \lambda(1 - s p^+ - (k + 1 - s) p^-)^d(1 - r p^+ - (k + 1 - r) p^-), \\
p^-(1 - (r - k) p^+ - (k + 1 - r) p^-)^d &= \lambda(1 - r p^+ - (k + 1 - r) p^-)^d(1 - s p^+ - (k + 1 - s) p^-),
\end{align*}$$

(17)

Yet according to (4) in Section 3 and Proposition 7.2, we have the relationship between $p^+$ and $p^-$ given by

$$p^+(1 - p^\pm)^d = \lambda \prod_{i=0}^{k+1} (1 - i p^+ - (k + 1 - i) p^-)^{d_i^\pm}$$

(18)

for the two types of vertices.

Comparing (17) and (18), the only feasible value for $(r, s)$ is $(k, 1)$ for all positive integer $k$. It means that $d_1^+ = d_k^- = d, d_1^- = d_k^+ = 1$ and all other $d_i^\pm = 0$ for $k \geq 2$; and $d_1^+ = d_1^- = d + 1$ and all other $d_j^\pm = 0$ for $k = 1$. Thus $B^{k, d}$ is the only possible hypergraph branching matrix satisfying the requirement.

7.3 Locally tree-like finite hypergraphs

We characterize the orbits for infinite hypertree $\mathbb{T}_k^d$ which can be locally like by finite hypergraphs.

Given a hypergraph $\mathcal{H} = (V, E)$, for every vertex $v \in V$ we denote by $N^r_\mathcal{H}(v)$ the $r$-neighborhood of $v$ in $\mathcal{H}$, which is a hypergraph formed by vertices and hyperedges within distance $r$ from $v$.

**Definition 7.2.** A hypergraph $\mathcal{G}$ is said to be $r$-locally like a classified hypergraph $(\mathcal{H}, \mathcal{C})$, if there is an epimorphism (a homomorphism which is surjective) $\phi$ from $\mathcal{H}$ to $\mathcal{G}$ satisfying:

- $\phi$ is also an isomorphism from $N^r_\mathcal{H}(v)$ to $N^r_\mathcal{G}(\phi(v))$ for every vertex $v$ in $\mathcal{H}$;
- no two vertices in $\mathcal{H}$ classified differently by $\mathcal{C}$ are mapped to the same vertex in $\mathcal{G}$ by $\phi$.

In particular, $\mathcal{G}$ is locally like $(\mathcal{H}, \mathcal{C})$ if it is 1-locally like $(\mathcal{H}, \mathcal{C})$.

**Proposition 7.4.** Let $\mathcal{C}$ be a classification of vertices of an infinite hypertree $\mathcal{T}$ into two types, where $(\mathcal{T}, \mathcal{C})$ is generated by the hypergraph branching matrix

$$M = \begin{pmatrix}
d_0^+ & d_1^+ & \cdots & d_{k+1}^+
d_0^- & d_1^- & \cdots & d_{k+1}^-
\end{pmatrix}.$$

If there is a finite hypergraph locally like $(\mathcal{T}, \mathcal{C})$, then for all $j \in \{0, 1, \ldots, k + 1\}$, it holds that

$$d_j^+ d_{k+1-j}^- (k + 1 - j)^2 = d_{k+1-j}^+ d_j^- j^2.$$

**Proof.** The proof is by double counting. We denote the two types for the vertices in the classification $\mathcal{C}$ as $\{+, -\}$. Assume that a finite hypergraph $\mathcal{H} = (U, F)$ is locally like $(\mathcal{T}, \mathcal{C})$, through an epimorphism $\phi$ from $\mathcal{T}$ to $\mathcal{H}$. By definition of being locally like, every vertex $v$ of $\mathcal{H}$ is assigned a type in $\{+, -\}$, which is the unique type of all vertices in the preimage $\phi^{-1}(v)$ according to $\mathcal{C}$.
For every vertex \( v \) in \( \mathcal{H} \) and any vertex \( u \) from the preimage \( \phi^{-1}(v) \), the neighborhood \( N^1_T(u) \) is isomorphic to \( N^1_H(v) \). Thus if \( v \) is classified as \( \pm \), then \( v \) is incident to \( d^\pm_j \) hyperedges of edge-profile \((j, k+1 - j)\) for each \( 0 \leq j \leq k+1 \).

We count the number of vertex-edge pairs \( (v, e) \in U \times F \) with particular vertex-type and edge-profile. Let \( U^\pm \) be the set of vertices in \( \mathcal{H} \) of type \( \pm \), and \( F^j \) the set of hyperedges in \( \mathcal{H} \) with edge-profile \((j, k+1 - j)\). Let \( P_{(\pm, j)} \) denote the number of pairs that \( v \in U^\pm \) and \( e \in F^j \).

Every hyperedge with edge-profile \((j, k+1 - j)\) contains exactly \( j \) vertices of type \( + \) and \( k+1 - j \) vertices of type \( - \); and on the other hand, for every vertex \( v \in U^\pm \), there are exactly \( d^\pm_j \) hyperedges \( e \) with edge-profile \((j, k+1 - j)\) containing \( v \). Thus by double counting, for every \( j \in \{0, 1, \ldots, k+1\} \), we have

\[
P_{(+, j)} = d^+_j |U^+| = j |F^j|;
\]
\[
P_{(-, j)} = d^-_j |U^-| = (k+1-j) |F^j|,
\]
which gives us the following equation for every \( j \in \{0, 1, \ldots, k+1\} \):

\[
d^+_j (k+1-j) |U^+| = d^-_j j |U^-|.
\]

Follow the same routine by considering hyperedges in \( F^{k+1-j} \) with edge-profile \((k+1-j, j)\), we have another equation

\[
d^+_{k+1-j} j |U^+| = d^-_{k+1-j} (k+1-j) |U^-|.
\]

Combining with the previous equation, \(|U^\pm|\) are canceled and we get

\[
d^+_j d^-_{k+1-j} (k+1-j)^2 = d^+_j d^-_{k+1-j} j^2.
\]

It can be easily verified that there is no finite hypergraph locally like the \( B^{k,d} \) defined in (14).

**Corollary 7.5.** For any positive integer \( k, d \), there is no finite hypergraph that is locally like \((\mathbb{T}^d_k, \mathcal{C})\) which is generated by \( B^{k,d} \), unless \( k = 1 \).

For \( k = 1 \), the \((\mathbb{T}^d_1, \mathcal{C})\) generated by \( B^{1,d} \) degenerates to the \((d+1)\)-regular infinite tree with parity coloring, which is locally like by the complete bipartite graph \( K_{d+1,d+1} \), and can also be \( r\)-locally like by random \((d+1)\)-regular bipartite graphs for larger \( r \).

**Acknowledgment.** We are very grateful to Heng Guo for many inspirational discussions. We also want to thank Pinyan Lu for his helpful comments and suggestions.

**References**

[1] M. Bayati, D. Gamarnik, D. Katz, C. Nair, and P. 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] M. Bordewich, M. Dyer, and M. Karpinski. Path coupling using stopping times and counting independent sets and colorings in hypergraphs. *Random Structures & Algorithms*, 32(3):375–399, 2008.
[3] J.-Y. Cai, A. Galanis, L. A. Goldberg, H. Guo, M. Jerrum, D. Štefankovič, and E. Vigoda. Bis-hardness for 2-spin systems on bipartite bounded degree graphs in the tree non-uniqueness region. In RANDOM, volume 28, pages 582–595, 2014.

[4] J.-Y. Cai, H. Guo, and T. Williams. A complete dichotomy rises from the capture of vanishing signatures. In Proceedings of the 45th ACM Symposium on Theory of Computing (STOC), pages 635–644, 2013.

[5] A. Dudek, M. Karpinski, A. Ruciński, and E. Szymańska. Approximate counting of matchings in (3, 3)-hypergraphs. In SWAT, pages 380–391, 2014.

[6] M. Dyer, L. A. Goldberg, and M. Jerrum. Counting and sampling $h$-colourings. In RANDOM, pages 51–67, 2002.

[7] A. Galanis, Q. Ge, D. Štefankovič, E. Vigoda, and L. Yang. Improved inapproximability results for counting independent sets in the hard-core model. Random Structures & Algorithms, 45(1):78–110, 2014.

[8] A. Galanis, D. Stefankovic, and E. Vigoda. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. arXiv preprint arXiv:1203.2226, 2012.

[9] A. Galanis, D. Štefankovič, and E. 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.

[10] D. Gamarnik and D. Katz. Correlation decay and deterministic fptas for counting colorings of a graph. Journal of Discrete Algorithms, 12:29–47, 2012.

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

[12] M. Jerrum and A. Sinclair. Approximating the permanent. SIAM Journal on Computing, 18(6):1149–1178, 1989.

[13] M. Karpinski, A. Rucinski, and E. Szymanska. Approximate counting of matchings in sparse uniform hypergraphs. In Proceedings of the Workshop on Analytic Algorithmics and Combinatorics (ANALCO), pages 72–79, 2012.

[14] F. P. Kelly. Stochastic models of computer communication systems. Journal of the Royal Statistical Society. Series B (Methodological), 47(3):379–395, 1985.

[15] L. Li, P. Lu, and Y. 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.

[16] L. Li, P. Lu, and Y. 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.

[17] C. Lin, J. Liu, and P. Lu. A simple fptas for counting edge covers. In Proceedings of the 25th ACM-SIAM Symposium on Discrete Algorithms (SODA), pages 341–348, 2014.
[18] J. Liu and P. Lu. Fptas for counting monotone cnf. In Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pages 1531–1548, 2015.

[19] P. Lu, M. Wang, and C. 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.

[20] P. Lu and Y. Yin. Improved fptas for multi-spin systems. In RANDOM, pages 639–654, 2013.

[21] E. Mossel, D. Weitz, and N. Wormald. On the hardness of sampling independent sets beyond the tree threshold. Probability Theory and Related Fields, 143(3-4):401–439, 2009.

[22] C. Nair and P. Tetali. The correlation decay (cd) tree and strong spatial mixing in multi-spin systems. arXiv preprint math/0701494, 2007.

[23] R. Restrepo, J. Shin, P. Tetali, E. Vigoda, and L. Yang. Improved mixing condition on the grid for counting and sampling independent sets. Probability Theory and Related Fields, 156(1-2):75–99, 2013.

[24] A. Sinclair, P. Srivastava, and M. Thurley. Approximation algorithms for two-state antiferromagnetic spin systems on bounded degree graphs. Journal of Statistical Physics, 155(4):666–686, 2014.

[25] A. Sinclair, P. Srivastava, and Y. 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.

[26] A. Sly. Uniqueness thresholds on trees versus graphs. The Annals of Applied Probability, 18(5):1897–1909, 2008.

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

[28] A. Sly and N. Sun. The computational hardness of counting in two-spin models on d-regular graphs. In Proceedings of the 53rd IEEE Symposium on Foundations of Computer Science (FOCS), pages 361–369, 2012.

[29] J. C. Vera, E. Vigoda, and L. Yang. Improved bounds on the phase transition for the hard-core model in 2-dimensions. In RANDOM, pages 699–713, 2013.

[30] D. Weitz. Combinatorial criteria for uniqueness of gibbs measures. Random Structures & Algorithms, 27(4):445–475, 2005.

[31] D. 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.