Improved FPTAS for Multi-Spin Systems

Pinyan Lu[1] and Yitong Yin[2*]

[1] Microsoft Research Asia, China. pinyanl@microsoft.com
[2] State Key Laboratory for Novel Software Technology, Nanjing University, China. yinyt@nju.edu.cn

Abstract. We design deterministic fully polynomial-time approximation scheme (FPTAS) for computing the partition function for a class of multi-spin systems, extending the known approximable regime by an exponential scale. As a consequence, we have an FPTAS for the Potts models with inverse temperature β up to a critical threshold $|\beta| = O(\frac{1}{\Delta})$ where Δ is the maximum degree, confirming a conjecture in [10]. We also give an improved FPTAS for a generalization of counting q-colorings, namely the counting list-colorings. As a consequence we have an FPTAS for counting q-colorings in graphs with maximum degree Δ when $q \geq \alpha\Delta + 1$ for α greater than $\alpha^* \approx 2.58071$. This is so far the best bound achieved by deterministic approximation algorithms for counting q-colorings. All these improvements are obtained by applying a potential analysis to the correlation decay on computation trees for multi-spin systems.

1 Introduction

Spin systems in Statistical Physics are the stochastic models defined by local interactions. In Computer Science, spin systems are used as a theoretical framework for counting or inference problems arising from constraint satisfaction problems, e.g. counting independent sets or q-colorings in graphs, and probability inference in graphical models.

A central problem in this framework is the computation of the *partition function*, which may solve both counting and inference. The problem is #P-hard for almost all nontrivial spin systems [3,4]. A classic approach for approximation of partition function is the Markov Chain Monte Carlo (MCMC) method which relies on the rapid mixing of random walks in the configuration space [5–7, 13–18, 21, 27]. A more contemporary approach is the correlation decay technique introduced by Bandyopadhyay and Gamarnik [1] and Weitz [28], which leads to deterministic fully polynomial-time approximation scheme (FPTAS) for #P-hard counting problems [2, 10, 19, 20, 22, 23, 29].

In these algorithms, the computation of a marginal probability (which is equivalent to the computation of partition function by self-reduction) is reduced

* Supported by the National Science Foundation of China under Grant No. 61272081, 61003023 and 61021062.

to evaluating an exponential-size tree-structured dynamical system. The correlation decay property guarantees that the far-away variables can be disregarded without substantially affecting the marginal probability of interest, thus the true values can be efficiently approximated by evaluating truncated dynamical systems. Two such dynamical systems were proposed: (1) the self-avoiding-walk (SAW) tree [28] for two-state spin systems and (2) the computation tree [10] for all spin systems. For two-state spin systems, FPTAS based on SAW-trees may approach the approximability boundaries, such as [19,20,23,28]. This is because a SAW-tree is a faithful construction of the original spin system on a tree, hence long-range correlations can be used as gadgets in the reduction for the inapproximability [8,25,26]. Very recently, the similar long-range correlations were used to prove inapproximability for multi-spin systems [9]. On the algorithmic side, due to a barrier result of Sly in [24], the original spin systems on trees are no longer capable of simulating all marginal probabilities. The computation tree introduced by Gamarnik and Katz in [10] overcomes this fundamental issue by creating a dynamical system consisting of different instances of spin systems.

1.1 Our results

We design efficient computation trees for multi-spin systems: For a vertex of degree d, our computation tree expends to d branches while the previous one in [10] has $\exp(\Omega(d))$ many branches. We apply a potential analysis to the decay of correlation between variables in computation trees. The potential analysis has been used in [19,20,22,23] for analyzing the correlation decay on the self-avoiding walk trees for two-state spin systems. We show for the first time that this powerful technique can be applied to computation trees for multi-spin systems. Our new construction of efficient computation trees and potential analysis greatly extend the regimes of correlation decay and deterministic FPTAS for these systems.

One of the most well-studied multi-spin systems is the Potts model.

Theorem 1. *For any constant $q \geq 2$, there exists an FPTAS for computing the partition function for q-state Potts models with inverse temperature β and maximum degree Δ satisfying* $3\Delta(\mathrm{e}^{|\beta|} - 1) \leq 1$.

For large Δ, the condition $3\Delta(\mathrm{e}^{|\beta|} - 1) \leq 1$ is translated to that $|\beta| = O(\frac{1}{\Delta})$, which greatly improves the best previous bound $\beta = O\left(\frac{1}{\Delta q^{\Delta}}\right)$ due to Gamarnik and Katz [10] and also confirms a conjecture in [10]. For the anti-ferromagnetic case ($\beta < 0$), our condition is asymptotically tight due to a very recent inapproximability result of Galanis, Štefankovič, and Vigoda [9].

Theorem 1 is a special case of a much more general theorem for the q-state spin systems, also called the Markov random fields. As suggested by [10], the regime of correlation decay for these models is described in terms of $c_{\boldsymbol{A}}$, the maximum ratio between edge parameters. We show that there exists an FPTAS for a family of Markov random fields if $3\Delta(c_{\boldsymbol{A}} - 1) \leq 1$. This exponentially improves the previous best known condition $(c_{\boldsymbol{A}}^{\Delta} - c_{\boldsymbol{A}}^{-\Delta})\Delta q^{\Delta} < 1$ proved in [10]. This general result is formally stated as Theorem 3 in Section 2.

We next study the problem of counting proper q-colorings in an undirected graph. For this problem, the mixing or the tractability condition is usually given in form of $q \geq \alpha\Delta + \beta$ for some constant α and β where Δ is the maximum degree of the graph. The previous best bound for deterministic FPTAS was achieved in [10] for an $\alpha \approx 2.8432$ and some sufficiently large β on triangle-free graphs. Better bounds (with $\alpha < 2$) were known for randomized approximation algorithms [6, 15, 27] or correlation decay only [11, 12]. We prove the following theorem for a constant $\alpha^* \approx 2.58071$ which is formally defined by (2) in Section 2.

Theorem 2. *There exists an FPTAS for counting q-colorings on graphs with maximum degree Δ if q and Δ are constants and $q \geq \alpha\Delta + 1$ for $\alpha > \alpha^*$.*

This is a new record for the deterministic FPTAS for counting q-colorings on general graphs, and we remove the triangle-free requirement in previous correlation-decay based results such as [10, 12]. Theorem 2 is proved as a special case of a theorem for a generalization of q-colorings, called the list-colorings, which is formally stated as Theorem 4 in Section 2.

All the above FPTAS require the degree of the graph and the number of states (colors) to be constant. If we remove this restriction, the algorithms compute a $(1 \pm \epsilon)$-approximation of the true value for any fixed $0 < \epsilon < 1$ in time $n^{O(\log n)}$. This complexity bound was only known previously for simple models like list-colorings but was not known for general multi-spin systems since for such systems the previous computation tree proposed in [10] tries to enumerate all configurations of the local neighborhood at each step. We give a more efficient computation tree which uses exponentially less branches.

2 Definitions and Statements of Results

An instance of a *q-state spin system* or a pair-wise *Markov random field (MRF)* is a tuple $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F})$, where

- $G = (V, E)$ is an undirected graph called the *underlying graph*;
- $\mathcal{X} = [q] = \{1, 2, \ldots, q\}$ is a domain of *spin states*;
- $\boldsymbol{A} = (A_e, e \in E)$ is a tuple where each $A_e : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$ is a symmetric function specifying the *activity* of edge e;
- $\boldsymbol{F} = (F_v, v \in V)$ is a tuple where each $F_v : \mathcal{X} \to \mathbb{R}_{\geq 0}$ specifies the *external field* at vertex v.

The size of an MRF instance is defined as $|\Omega| = \max\{|V|, |\mathcal{X}|\}$. We consider only those MRF instances such that the number of bits used to encode $\boldsymbol{A}$ and $\boldsymbol{F}$ is in polynomial of $n = |V|$ and $q = |\mathcal{X}|$. This does not affect the generality of the problem since we are interested in the approximation algorithms.

The *partition function* of an MRF instance $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F})$ is defined as

$$Z(\Omega) \triangleq \sum_{\boldsymbol{x} \in \mathcal{X}^V} \prod_{e=uv \in E} A_e(x_u, x_v) \prod_{v \in V} F_v(x_v).$$

This gives rise to a probability distribution $\mathbb{P}_\Omega$, called the *Gibbs measure*, over all configurations $\boldsymbol{x} \in \mathcal{X}^V$, such that

$$\mathbb{P}_\Omega(\boldsymbol{X} = \boldsymbol{x}) = \frac{\prod_{e=uv\in E} A_e(x_u, x_v) \prod_{v\in V} F_v(x_v)}{Z(\Omega)}.$$

Given an MRF instance $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F})$ with underlying graph $G = (V, E)$, we denote by Δ_G the maximum degree of G and let

$$c_{\boldsymbol{A}} \triangleq \max_{\substack{e\in E \\ w,x,y,z\in\mathcal{X}}} \frac{A_e(x, y)}{A_e(w, z)}.$$

Theorem 3. *Let $\mathbb{M}$ be a family of MRF instances with bounded degree and bounded size of domain. There exists an FPTAS for computing the partition function of MRFs in $\mathbb{M}$ if it holds that*

$$\forall \Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F}) \in \mathbb{M}, \qquad 3\Delta_G(c_{\boldsymbol{A}} - 1) \leq 1. \tag{1}$$

For any family $\mathbb{M}$ of MRFs satisfying (1) *without any restriction on the degree or the size of domain, the algorithm computes a $(1 \pm \epsilon)$-approximation of $Z(\Omega)$ for any fixed $0 < \epsilon < 1$ in time $n^{O(\log n)}$ where $n = |\Omega|$.*

The q-state Potts model is a special class of MRFs $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F})$ with for every $e \in E$, $A_e = A$ such that $A(x, y) = \mathrm{e}^\beta$ if $x = y$ and $A(x, y) = 1$ otherwise. The parameter β is called the *inverse temperature*. It is easy to see that Theorem 1 is a special case of Theorem 3 on Potts models.

Next we consider the proper q-colorings in an undirected graph, which can be easily seen as a special case of MRF. The problem of counting q-colorings is solved by solving its generalization called the list-colorings. A list-coloring instance is a tuple $\Omega = (G, \mathcal{X}, \boldsymbol{L})$ with that

- $G = (V, E)$ is an undirected graph;
- $\mathcal{X} = [q]$ is a domain of q colors;
- $\boldsymbol{L} = (L_v, v \in V)$ such that each $L_v \subseteq \mathcal{X}$ is a list of colors for vertex v.

A proper coloring in a list-coloring instance is a proper q-coloring $\boldsymbol{x} \in \mathcal{X}^V$ of vertices such that $x_v \in L_v$ for every $v \in V$. The list-coloring is a special case of MRFs $(G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F})$ with that for every $e \in E$, $A_e = A$ such that $A(x, y) = 0$ if $x = y$ and $A(x, y) = 1$ if otherwise, and for every $v \in V$, the external field F_v is a Boolean function indicating the color list L_v. For the list-colorings we have $c_{\boldsymbol{A}} = \infty$, thus Theorem 3 does not apply, so we use different algorithm and analysis to prove the following theorem. Let $\alpha^* \approx 2.58071$ be the solution to the equation[3]

$$\frac{\sqrt{2} + \sqrt{2\alpha - 1 - \sqrt{4\alpha - 3}}}{\sqrt{2\alpha(\alpha - 1)}} \exp\left(\frac{3 - 2\alpha + \sqrt{4\alpha - 3}}{4(\alpha - 1)}\right) = 1. \tag{2}$$

[3] The LHS of (2) is in fact monotonously decreasing from $+\infty$ to 0 for $\alpha > 1$, so there is a unique solution α^*.

Theorem 4. *There exists a deterministic FPTAS for counting proper colorings in list-coloring instances $\Omega = (G, \mathcal{X}, \boldsymbol{L})$ with bounded degree Δ_G and bounded number of colors $q = |\mathcal{X}|$ satisfying that there is an $\alpha > \alpha^*$ such that*

$$\forall v \in V, \quad |L_v| \geq \alpha \Delta_G + 1. \tag{3}$$

Obviously Theorem 2 is a special case of Theorem 4 as colorings are just list-clorings with $L_v = \mathcal{X}$ for every $v \in V$.

3 Markov Random Fields

Given an MRF instance defined on the underlying graph $G = (V, E)$, we suppose that for each vertex $v \in V$, the neighbors of v are enumerated as $v_1, v_2, \ldots, v_{\deg(v)}$ where $\deg(v)$ is the degree of v. We define operations called *pinning* and *partial pinning* on MRF instances as follows.

Definition 1. *Given an MRF instance $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F})$, a vertex $v \in V$ and its neighbors $v_1, v_2, \ldots, v_d$ in G, where $d = \deg(v)$, for each spin state $x \in \mathcal{X}$ and each $1 \leq i \leq d+1$, the* partial pinning *of Ω, denoted as $\Omega^i_{v,x}$, is a new MRF instance augmented from Ω as $\Omega^i_{v,x} = (G_v, \mathcal{X}, \widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{F}})$, where $G_v = G \setminus \{v\}$ is the subgraph of G induced by $V \setminus \{v\}$, $\widetilde{\boldsymbol{A}} = (A_e, e \in E \setminus \{vv_1, vv_2, \ldots, vv_d\})$ is the restriction of $\boldsymbol{A}$ on the set of edges in G_v, and $\widetilde{\boldsymbol{F}} = (\widetilde{F}_u, u \in V \setminus \{v\})$ where*

$$\forall y \in \mathcal{X}, \quad \widetilde{F}_u(y) = \begin{cases} A_{uv}(x,y) F_u(y) & \text{if } u \in \{v_1, \ldots, v_{i-1}\}, \\ F_u(y) & \text{otherwise.} \end{cases}$$

The pinning *of Ω is a partial pinning by choosing $i = d+1$, which is denoted as $\Omega_{v,x} = \textsc{Pin}_{v,x}(\Omega) = \Omega^{d+1}_{v,x}$.*

The following identity can be seen as a generalization of the recursion for list-colorings derived in [10]. Compared to the recursion for MRFs in [10], it uses substantially less variables.

Proposition 1. *Let $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F})$ be an MRF instance. For every vertex $v \in V_G$ and its neighbors $v_1, v_2, \ldots, v_d$ where $d = \deg(v)$, and every spin state $x \in \mathcal{X}$, it holds that*

$$\mathbb{P}_\Omega(X_v = x) = \frac{F_v(x) \prod_{i=1}^d \left(A_{vv_i}(x,x) - \sum_{z \neq x} \left(A_{vv_i}(x,x) - A_{vv_i}(x,z) \right) \mathbb{P}_{\Omega^i_{v,x}}(X_{v_i} = z) \right)}{\sum_{y \in \mathcal{X}} F_v(y) \prod_{i=1}^d \left(A_{vv_i}(y,y) - \sum_{z \neq y} \left(A_{vv_i}(y,y) - A_{vv_i}(y,z) \right) \mathbb{P}_{\Omega^i_{v,y}}(X_{v_i} = z) \right)}.$$

Proof. We define that

$$Z_\Omega(X_v = x) \triangleq \sum_{\substack{\boldsymbol{x} \in \mathcal{X}^V \\ x_v = x}} \prod_{uw \in E} A_{uw}(x_u, x_w) \prod_{u \in V} F_u(x_u).$$

It can be verified that $Z_\Omega(X_v = x) = F_v(x)Z(\Omega_{v,x})$ where $\Omega_{v,x} = \textsc{Pin}_{v,x}(\Omega)$ is the pinning of Ω. Then

$$\mathbb{P}_\Omega(X_v = x) = \frac{Z_\Omega(X_v = x)}{\sum_{y\in\mathcal{X}} Z_\Omega(X_v = y)} = \frac{F_v(x)Z(\Omega_{v,x})}{\sum_{y\in\mathcal{X}} F_v(y)Z(\Omega_{v,y})}. \tag{4}$$

By the Definition 1, it holds that $\Omega_{v,x} = \Omega^{d+1}_{v,x}$, and $\Omega^1_{v,x}$ is simply the MRF instance deleting vertex v, which is independent of the choice of x. Therefore,

$$(4) = \frac{F_v(x)Z(\Omega^{d+1}_{v,x})/Z(\Omega^1_{v,x})}{\sum_{y\in\mathcal{X}} F_v(y)Z(\Omega^{d+1}_{v,y})/Z(\Omega^1_{v,y})} = \frac{F_v(x)\prod_{i=1}^d \frac{Z(\Omega^{i+1}_{v,x})}{Z(\Omega^i_{v,x})}}{\sum_{y\in\mathcal{X}} F_v(y)\prod_{i=1}^d \frac{Z(\Omega^{i+1}_{v,y})}{Z(\Omega^i_{v,y})}}. \tag{5}$$

The partition function of a partial pinning of Ω expands as:

$$Z(\Omega^i_{v,x}) = \sum_{\boldsymbol{x}\in\mathcal{X}^{V\setminus\{v\}}} \prod_{\substack{uw\in E\\ u\neq v\\ w\neq v}} A_{uw}(x_u, x_w) \prod_{u\in V\setminus\{v\}} F_u(x_u) \prod_{j=1}^{i-1} A_{vv_j}(x, x_{v_j}).$$

It can be verified that $Z(\Omega^{i+1}_{v,x}) = \sum_{z\in\mathcal{X}} A_{vv_i}(x,z)\cdot Z_{\Omega^i_{v,x}}(X_{v_i} = z)$. Therefore,

$$\begin{aligned}
(5) &= \frac{F_v(x)\prod_{i=1}^d \sum_{z\in\mathcal{X}} A_{vv_i}(x,z)\cdot \frac{Z_{\Omega^i_{v,x}}(X_{v_i}=z)}{Z(\Omega^i_{v,x})}}{\sum_{y\in\mathcal{X}} F_v(y)\prod_{i=1}^d \sum_{z\in\mathcal{X}} A_{vv_i}(y,z)\cdot \frac{Z_{\Omega^i_{v,y}}(X_{v_i}=z)}{Z(\Omega^i_{v,y})}}\\
&= \frac{F_v(x)\prod_{i=1}^d \sum_{z\in\mathcal{X}} A_{vv_i}(x,z)\cdot \mathbb{P}_{\Omega^i_{v,x}}(X_{v_i}=z)}{\sum_{y\in\mathcal{X}} F_v(y)\prod_{i=1}^d \sum_{z\in\mathcal{X}} A_{vv_i}(y,z)\cdot \mathbb{P}_{\Omega^i_{v,y}}(X_{v_i}=z)}\\
&= \frac{F_v(x)\prod_{i=1}^d\Big(A_{vv_i}(x,x)-\sum_{z\neq x}\big(A_{vv_i}(x,x)-A_{vv_i}(x,z)\big)\mathbb{P}_{\Omega^i_{v,x}}(X_{v_i}=z)\Big)}{\sum_{y\in\mathcal{X}} F_v(y)\prod_{i=1}^d\Big(A_{vv_i}(y,y)-\sum_{z\neq y}\big(A_{vv_i}(y,y)-A_{vv_i}(y,z)\big)\mathbb{P}_{\Omega^i_{v,y}}(X_{v_i}=z)\Big)},
\end{aligned}$$

where the last equation uses the fact that $\sum_{z\in\mathcal{X}} \mathbb{P}_{\Omega^i_{v,y}}(X_{v_i} = z) = 1$.

3.1 Algorithms based on the computation tree recursion

Given an MRF instance $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F})$ on underlying graph $G = (V, E)$, a vertex $v \in V$ with d neighbors $v_1, v_2, \ldots, v_d$ in G and a spin state $x \in \mathcal{X}$, we define the following function:

$$f_{\Omega,v,x}(\boldsymbol{p}) \triangleq \frac{F_v(x)\prod_{i=1}^d\big(A_{vv_i}(x,x)-\sum_{z\neq x}\big(A_{vv_i}(x,x)-A_{vv_i}(x,z)\big)p_{i,x,z}\big)}{\sum_{y\in\mathcal{X}} F_v(y)\prod_{i=1}^d\big(A_{vv_i}(y,y)-\sum_{z\neq y}\big(A_{vv_i}(y,y)-A_{vv_i}(y,z)\big)p_{i,y,z}\big)} \tag{6}$$

over the domain of vectors $\boldsymbol{p} = (p_{i,y,z}, 1 \le i \le d; y, z \in \mathcal{X}; y \neq z) \in [0,1]^{dq(q-1)}$ satisfying that $\sum_{z\neq y} p_{i,y,z} \le 1$ for every $1 \le i \le d$ and $y \in \mathcal{X}$. Due to Proposition 1 we have $\mathbb{P}_\Omega(X_v = x) = f_{\Omega,v,x}(\boldsymbol{p})$ where $p_{i,y,z} = \mathbb{P}_{\Omega^i_{v,y}}(X_{v_i} = z)$ for each $1 \le i \le d$ and $y, z \in \mathcal{X}$ that $y \neq z$. This already gives us a procedure, called the *computation tree recursion*, for computing the exact value of a marginal probability $\mathbb{P}_\Omega(X_v = x)$. Note that it terminates since each partial pinning $\Omega^i_{v,y}$ deletes a vertex v from the current underlying graph.

It is easy to verify the following closure property of the computation tree recursion: If each $p_{i,y,z}$ is replaced by an estimation $\widehat{\mathbb{P}}_{\Omega^i_{v,y}}(X_{v_i} = z)$ of marginal $\mathbb{P}_{\Omega^i_{v,y}}(X_{v_i} = z)$ such that $\widehat{\mathbb{P}}_{\Omega^i_{v,y}}(X_{v_i} = z) \in [0,1]$ and $\sum_{z\neq y} \widehat{\mathbb{P}}_{\Omega^i_{v,y}}(X_{v_i} = z) \le 1$ then the outcome of the recursion $\widehat{\mathbb{P}}_\Omega(X_v = x) = f_{\Omega,v,x}(\boldsymbol{p})$ as an estimation of $\mathbb{P}_\Omega(X_v = x)$ still satisfies that $\widehat{\mathbb{P}}_\Omega(X_v = x) \in [0,1]$ and $\sum_{x\in\mathcal{X}} \widehat{\mathbb{P}}_\Omega(X_v = x) = 1$.

The size of the computation tree can be easily of exponential in the size of the underlying graph. We can run the computation tree recursion up to t levels and use a naive estimation of marginals for the base cases. Formally, for $t \ge 0$, the quantity $\widehat{\mathbb{P}}^{(t)}_\Omega(X_v = x)$ is recursively defined as follows:

- If $t = 0$, let $\widehat{\mathbb{P}}^{(0)}_\Omega(X_v = x) = \frac{F_v(x)}{\sum_{y\in\mathcal{X}} F_v(y)}$.
- If $t > 0$, let $\widehat{\mathbb{P}}^{(t)}_\Omega(X_v = x) = f_{\Omega,v,x}(\widehat{\boldsymbol{p}})$ where $\hat{p}_{i,y,z} = \widehat{\mathbb{P}}^{(t-1)}_{\Omega^i_{v,y}}(X_{v_i} = z)$ for each $1 \le i \le d$ and $y, z \in \mathcal{X}$ that $y \neq z$.

The value of the base case $\widehat{\mathbb{P}}^{(0)}_\Omega(X_v = x)$ is not important due to a correlation decay property we prove later. As shown in [10], on graphs of constant maximum degrees, the quantity $\widehat{\mathbb{P}}^{(t)}_\Omega(X_v = x)$ can be efficiently computed by dynamic programming when $t = O(\log n)$.

The partition function can be approximated from estimations of marginals by the following standard procedure. Enumerate the vertices in V as $v_1, v_2, \ldots, v_n$.

1. Let $\Omega_1 = \Omega$. For $k = 1, 2, \ldots, n$, assuming that the Ω_k is well-defined, use the computation tree recursion to compute $\widehat{\mathbb{P}}^{(t)}_{\Omega_k}(X_{v_k} = x)$ for all $x \in \mathcal{X}$, choose x_k to be the x which maximizes the $\widehat{\mathbb{P}}^{(t)}_{\Omega_k}(X_{v_k} = x)$ and construct $\Omega_{k+1} = \textsc{Pin}_{v_k,x_k}(\Omega_k)$ as a pinning of Ω_k.
2. Compute that $\widehat{Z}(\Omega) = \frac{\prod_{e=uv\in E} A_e(x_u,x_v) \prod_{v\in V} F_v(x_v)}{\prod_{k=1}^n \widehat{\mathbb{P}}^{(t)}_{\Omega_k}(X_{v_k}=x_{v_k})}$ and return $\widehat{Z}(\Omega)$.

This algorithm is the same as the one proposed in [10], except for using a simplified computation tree recursion, thus by the same analysis as in [10], we have the following proposition.

Proposition 2. *Let $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F})$ be an MRF instance such that G has maximum degree Δ and $q = |\mathcal{X}|$. The value of $\widehat{Z}(\Omega)$ can be computed in time* $\mathrm{poly}(|\Omega|) \cdot (q\Delta)^{O(t)}$.

3.2 Correlation decay on the computation tree

The above algorithm approximates the marginal probabilities by simulating a tree-structured dynamical system for a limited number of iterations. The accuracy of this approximation relies on the following property of correlation decay.

Definition 2 (Correlation Decay). *Let $\mathbb{M}$ be a family of MRFs. We say that the computation tree recursion exhibits* exponential correlation decay *over $\mathbb{M}$ if there exists a constant $C > 0$ such that given any MRF instance $\Omega \in \mathbb{M}$, for all $t \geq 1$, it holds that*

$$\max_{\substack{v \in V_\Omega \\ x \in \mathcal{X}}} \left| \mathbb{P}_\Omega(X_v = x) - \widehat{\mathbb{P}}^{(t)}_\Omega(X_v = x) \right| \leq \mathrm{poly}(|\Omega|) \cdot \exp(-C \cdot t).$$

A sufficient condition for the exponential correlation decay is that the error of estimation decays by a constant factor in every iteration. However, in general, the systems exhibiting correlation decay may not necessarily decay in every step. This issue has been addressed by a potential-based analysis in [19, 20, 22, 23] for self-avoiding walk trees for 2-spin systems, which is now formalized as the following condition for computation trees for multi-spin systems.

Definition 3 (The Amortized Decay Condition). *Let $\mathbb{M}$ be a family of q-state MRFs. We say that $\mathbb{M}$ satisfies the* Amortized Decay Condition *if there exists a strictly increasing differentiable function $\varphi : [0,1] \to \mathbb{R}$ satisfying the following conditions:*

1. *Let $\Phi(x) = \frac{\mathrm{d}\,\varphi(x)}{\mathrm{d}\,x}$ denote the derivative of function φ. We call $\Phi(\cdot)$ the* potential function*. The values of $\Phi(\cdot)$ and $\frac{1}{\Phi(\cdot)}$ are bounded by* $\mathrm{poly}(q)$ *over domain $[0,1]$.*
2. *Given an MRF instance $\Omega \in \mathbb{M}$, a vertex $v \in V_\Omega$ with $d = \deg(v)$ and a spin state $x \in \mathcal{X}$, let $f = f_{\Omega,v,x}$ be the computation tree recursion defined by* (6)*, and define the* amortized decay rate *as*

$$\kappa(\boldsymbol{p}) \triangleq \sum_{\substack{1 \leq i \leq d \\ y \neq z}} \left| \frac{\partial f(\boldsymbol{p})}{\partial p_{i,y,z}} \right| \frac{\Phi(f(\boldsymbol{p}))}{\Phi(p_{i,y,z})}. \tag{7}$$

There exists a constant $0 < \kappa < 1$, such that for every MRF instance $\Omega \in \mathbb{M}$, vertex $v \in V_\Omega$ and spin state $x \in \mathcal{X}$, it holds that $\kappa(\boldsymbol{p}) \leq \kappa$ for all $\boldsymbol{p} = (p_{i,y,z}, 1 \leq i \leq d \wedge y,z \in \mathcal{X} \wedge y \neq z) \in [0,1]^{dq(q-1)}$ satisfying that $\sum_{z \neq y} p_{i,y,z} \leq 1$ for all i and y.

We may replace the first condition by a more sophisticated bound on the values of $|\Phi(\cdot)|$ and $\frac{1}{|\Phi(\cdot)|}$, which will give us more freedom to choose potential functions, although the current simple bound is sufficient for our analysis.

We say a family $\mathbb{M}$ of MRF instances is *closed under partial pinning* if for every $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F}) \in \mathbb{M}$, every vertex $v \in V_G$ with $d = \deg(v)$, spin state $x \in \mathcal{X}$ and $1 \leq i \leq d$, it holds for the partial pinning $\Omega^i_{v,x}$ of Ω that $\Omega^i_{v,x} \in \mathbb{M}$.

Lemma 1. *Let $\mathbb{M}$ be a family of MRFs which is closed under partial pinning. If $\mathbb{M}$ satisfies the amortized decay condition then the computation tree recursion exhibits exponential correlation decay over $\mathbb{M}$.*

Proof. Pick an MRF instance $\Omega \in \mathbb{M}$, a vertex $v \in V_\Omega$ with d neighbors $v_1, v_2, \ldots, v_d$ and a spin state $x \in \mathcal{X}$. Let $\varphi : [0,1] \to \mathbb{R}$ be the monotone differentiable function and $\Phi(\cdot)$ be its derivative, as required by the amortized decay condition. Consider the corresponding recursion $f = f_{\Omega,v,x}$.

We define the following notations: Let $p = \mathbb{P}_\Omega(X_v = x)$, $\hat{p} = \widehat{\mathbb{P}}^{(t)}_\Omega(X_v = x)$, and for every $1 \le i \le d$ and $y, z \in \mathcal{X}$ that $y \neq z$, let $p_{i,y,z} = \mathbb{P}_{\Omega^i_{v,y}}(X_{v_i} = z)$ and $\hat{p}_{i,y,z} = \widehat{\mathbb{P}}^{(t-1)}_{\Omega^i_{v,y}}(X_{v_i} = z)$. Obviously, we have $p = f(\boldsymbol{p})$ and $\hat{p} = f(\widehat{\boldsymbol{p}})$. We also denote that $\xi = \varphi(p)$, $\hat{\xi} = \varphi(\hat{p})$, $\xi_{i,y,z} = \varphi(p_{i,y,z})$ and $\hat{\xi}_{i,y,z} = \varphi(\hat{p}_{i,y,z})$, respectively. Let $\epsilon = |p - \hat{p}| = |f(\boldsymbol{p}) - f(\widehat{\boldsymbol{p}})|$, $\delta = |\varphi(p) - \varphi(\hat{p})| = |\varphi(f(\boldsymbol{p})) - \varphi(f(\widehat{\boldsymbol{p}}))|$, $\epsilon_{i,y,z} = |p_{i,y,z} - \hat{p}_{i,y,z}|$, and $\delta_{i,y,z} = |\varphi(p_{i,y,z}) - \varphi(\hat{p}_{i,y,z})|$ be the respective errors. We have

$$\delta = \left|\xi - \hat{\xi}\right| = |\varphi(f(\boldsymbol{p})) - \varphi(f(\widehat{\boldsymbol{p}}))| = \left|\varphi\left(f\left(\varphi^{-1}\left(\boldsymbol{q}\right)\right)\right) - \varphi\left(f\left(\varphi^{-1}\left(\widehat{\boldsymbol{q}}\right)\right)\right)\right|.$$

Due to the Mean Value Theorem, there exist $\tilde{\xi}_{i,y,z} \in [0,1]$ and accordingly $\tilde{p}_{i,y,z} = \varphi^{-1}(\tilde{\xi}_{i,y,z})$, $1 \le i \le d$, $y, z \in \mathcal{X}$, $y \neq z$, such that

$$\delta = \sum_{\substack{1 \le i \le d \\ y \neq z}} \left|\frac{\partial f(\widetilde{\boldsymbol{p}})}{\partial \tilde{p}_{i,y,z}}\right| \frac{\Phi(f(\widetilde{\boldsymbol{p}}))}{\Phi(\tilde{p}_{i,y,z})} \cdot \delta_{i,y,z} \le \kappa(\widetilde{\boldsymbol{p}}) \cdot \max_{\substack{1 \le i \le d \\ y \neq z}} \delta_{i,y,z},$$

where $\kappa(\boldsymbol{p})$ is defined by (7). Since $\mathbb{M}$ satisfies the amortized decay condition, there exists a universal constant $\kappa < 1$ such that $\kappa(\widetilde{\boldsymbol{p}}) \le \kappa$. And since $\mathbb{M}$ is closed under partial pinning, every $\Omega^i_{v,y}$ still belongs to $\mathbb{M}$. Therefore, by induction we have that $\delta \le \kappa^t \delta_0$, where $\delta_0 = |\varphi(p_0) - \varphi(\hat{p}_0)|$ such that $p_0 = \mathbb{P}_{\Omega'}(X_u = w)$ and $\hat{p}_0 = \widehat{\mathbb{P}}^{(0)}_{\Omega'}(X_u = w)$ for some $\Omega' \in \mathbb{M}$, $u \in V_{\Omega'}$ and $x \in \mathcal{X}$, where Ω' is an MRF instance resulting from applying t partial pinnings on the original Ω.

By the Mean Value Theorem, there exists a $\tilde{p}_0 \in [0,1]$ such that $\delta_0 = |\varphi(p_0) - \varphi(\hat{p}_0)| \le |\Phi(\tilde{p}_0)|$, which is upper bounded by q^c for some constant c due to the requirement of amortized decay condition, thus $\delta \le \kappa^t \delta_0 \le q^c \kappa^t$. Recall that $\delta = |\varphi(p) - \varphi(\hat{p})|$. Also by the Mean Value Theorem there exists $\tilde{p} \in [0,1]$ such that $\delta = |\varphi(p) - \varphi(\hat{p})| = |\Phi(\tilde{p})||p - \hat{p}| = |\Phi(\tilde{p})|\epsilon$, thus $\epsilon = \frac{\delta}{|\Phi(\tilde{p})|} \le q^c \delta$. Altogether we have that

$$\left|\mathbb{P}_\Omega(X_v = x) - \widehat{\mathbb{P}}^{(t)}_\Omega(X_v = x)\right| = \epsilon \le q^c \delta \le q^c \kappa^t \delta_0 \le q^{2c} \kappa^t.$$

And this holds for every $\Omega \in \mathbb{M}, v \in V_\Omega, x \in \mathcal{X}$ and $t \ge 1$, with the universal constants c and $\kappa < 1$, which implies the exponential correlation decay of computation tree recursion over $\mathbb{M}$.

The following lemma is proved by verifying the amortized decay condition.

Lemma 2. *Let* $\mathbb{M}$ *be a family of MRFs satisfying* (1)*. The computation tree recursion exhibits exponential correlation decay over* $\mathbb{M}$*.*

Proof. Let $\mathbb{M}^*$ be the closure of $\mathbb{M}$ under partial pinning, thus every instance from $\mathbb{M}^*$ is either an instance $\Omega \in \mathbb{M}$ or an outcome of successive partial pinnings of it, and the family $\mathbb{M}^*$ is closed under partial pinning. We show that $\mathbb{M}^*$ satisfies the amortized decay condition. We choose a monotone function $\varphi : [0,1] \to \mathbb{R}$ so that its derivative Φ satisfies that $\Phi(p) = \left(p + \frac{1}{100q}\right)^{-1}$. Thus both $\Phi(\cdot)$ and $\frac{1}{\Phi(\cdot)}$ are bounded by polynomial of q over $[0,1]$.

Let $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F}) \in \mathbb{M}^*$ be an MRF instance on an underlying graph G with maximum degree Δ, $v \in V_G$ a vertex with $d = \deg(v)$, and $x \in \mathcal{X}$ a spin state. Let $f = f_{\Omega,v,x}$ be the recursion defined by (6).

We define some shorthand notations. For each $1 \le i \le d$ and $y, z \in \mathcal{X}$ that $y \neq z$, let $a_{i,y,z} = 1 - \frac{A_{vv_i}(y,z)}{A_{vv_i}(y,y)}$ and $b_y = F_v(y)\prod_{i=1}^d A_{vv_i}(y,y)$, and denote that $s_{i,y} = 1 - \sum_{z\neq y} a_{i,y,z} \cdot p_{i,y,z}$, $s_y = b_y \prod_{i=1}^d s_{i,y}$, and $s = \sum_{y\in\mathcal{X}} s_y$. Then we have

$$f(\boldsymbol{p}) = \frac{b_x \prod_{i=1}^d \left(1 - \sum_{z\neq x} a_{i,x,z} \cdot p_{i,x,z}\right)}{\sum_{y\in\mathcal{X}} b_y \prod_{i=1}^d \left(1 - \sum_{z\neq y} a_{i,y,z} \cdot p_{i,y,z}\right)} = \frac{s_x}{s}.$$

For $\boldsymbol{p} = (p_{i,y,z}, 1 \le i \le d; y, z \in \mathcal{X}; y \neq z) \in [0,1]^{dq(q-1)}$ such that $\sum_{z\neq y} p_{i,y,z} \le 1$ for all i and y, it holds that $s_{i,y} \ge 0$ for any i and y, and $f(\boldsymbol{p}) \in [0,1]$. The partial derivatives satisfy:

$$\left|\frac{\partial f(\boldsymbol{p})}{\partial p_{i,x,z}}\right| = \left|\frac{a_{i,x,z} s_x (s - s_x)}{s^2 \cdot s_{i,x}}\right| = f(\boldsymbol{p})(1 - f(\boldsymbol{p}))\frac{|a_{i,x,z}|}{s_{i,x}},$$

$$\sum_{y\neq x}\left|\frac{\partial f(\boldsymbol{p})}{\partial p_{i,y,z}}\right| = \sum_{y\neq x}\left|\frac{a_{i,y,z} s_x s_y}{s^2 \cdot s_{i,y}}\right| = f(\boldsymbol{p}) \sum_{y\neq x} \frac{s_y}{s} \sum_{i=1}^d \frac{|a_{i,y,z}|}{s_{i,y}}.$$

The amortized decay rate defined by (7) is then bounded as

$$\begin{aligned}
\kappa(\boldsymbol{p}) &= \sum_{\substack{1\le i\le d\\ y\neq z}} \left|\frac{\partial f(\boldsymbol{p})}{\partial p_{i,y,z}}\right| \frac{\Phi(f(\boldsymbol{p}))}{\Phi(p_{i,y,z})} \\
&= f(\boldsymbol{p})(1 - f(\boldsymbol{p}))\Phi(f(\boldsymbol{p})) \sum_{i=1}^d \frac{1}{s_{i,x}} \sum_{z\neq x} \frac{|a_{i,x,z}|}{\Phi(p_{i,x,z})} \\
&\quad + f(\boldsymbol{p})\Phi(f(\boldsymbol{p})) \sum_{y\neq x} \frac{s_y}{s} \sum_{i=1}^d \frac{1}{s_{i,y}} \sum_{z\neq y} \frac{|a_{i,y,z}|}{\Phi(p_{i,y,z})} \\
&\le \sum_{\substack{1\le i\le d\\ z\neq x}} \frac{|a_{i,x,z}|}{s_{i,x}} \cdot \left(p_{i,y,z} + \frac{1}{100q}\right) + \max_{y\neq x} \sum_{\substack{1\le i\le d\\ z\neq y}} \frac{|a_{i,y,z}|}{s_{i,y}} \cdot \left(p_{i,y,z} + \frac{1}{100q}\right) \\
&\le \frac{101}{50}\Delta \cdot \max_{\substack{1\le i\le d\\ z\neq y}} \frac{|a_{i,y,z}|}{s_{i,y}}.
\end{aligned}$$

For every $1 \le i \le d$ and $y, z \in \mathcal{X}$ that $y \neq z$, it can be verified that $s_{i,y} = p_{i,y,y} + \sum_{z\neq y} \frac{A_{vv_i}(y,z)}{A_{vv_i}(y,y)} \cdot p_{i,y,z} \ge \frac{1}{c_{\boldsymbol{A}}}$, and

$$|a_{i,y,z}| = \left|1 - \frac{A_{vv_i}(y,z)}{A_{vv_i}(y,y)}\right| \le \max\left\{\frac{c_{\boldsymbol{A}} - 1}{c_{\boldsymbol{A}}}, c_{\boldsymbol{A}} - 1\right\} \le c_{\boldsymbol{A}} - 1.$$

Note that the partial pinning does not affect the edge activity $\boldsymbol{A}$, thus for $\mathbb{M}$ satisfying (1), for every $\Omega \in \mathbb{M}^*$ the $c_{\boldsymbol{A}}$ still satisfies that $3\Delta(c_{\boldsymbol{A}} - 1) \leq 1$. Therefore,

$$\kappa(\boldsymbol{p}) \leq \frac{101}{50}\Delta c_{\boldsymbol{A}}(c_{\boldsymbol{A}} - 1) \leq \frac{101}{150}(1 + \frac{1}{3\Delta}) \leq \frac{404}{450} < 1.$$

Therefore, the MRF family $\mathbb{M}^*$ satisfies the amortized decay condition. By Lemma 1, the computation tree recursion exhibits exponential correlation decay over $\mathbb{M}^*$ thus also over its subfamily $\mathbb{M}$.

Proof of Theorem 3: Let $\Omega = (G, \mathcal{X}, \boldsymbol{A}, \boldsymbol{F}) \in \mathbb{M}$ be an MRF instance and $G = (V, E)$. Enumerate the vertices in V as $v_1, v_2, \ldots, v_n$. For each $1 \leq k \leq n$, let $\mathbb{P}^{(t)}_{\Omega_k}(X_{v_k} = x_k)$ be computed by the algorithm in Section 3.1, where $\Omega_1 = \Omega$ and $\Omega_{k+1} = \textsc{Pin}_{v_k, x_k}(\Omega_k)$. It is easy to verify that Ω_k still satisfies the condition (1) for every k since pinning increases neither Δ nor $c_{\boldsymbol{A}}$. Let $\widehat{Z}(\Omega) = \frac{w(\boldsymbol{x})}{\prod_{k=1}^{n} \widehat{\mathbb{P}}^{(t)}_{\Omega_k}(X_{v_k} = x_{v_k})}$ where $w(\boldsymbol{x}) = \prod_{e=uv \in E} A_e(x_u, x_v) \prod_{v \in V} F_v(x_v)$. It holds that

$$\mathbb{P}_\Omega(\boldsymbol{X} = \boldsymbol{x}) = \prod_{k=1}^{n} \mathbb{P}_\Omega(X_{v_k} = x_k \mid \forall 1 \leq i < k, X_{v_i} = x_i).$$

As observed in [10], the marginal probability $\mathbb{P}_\Omega(X_{v_k} = x_k \mid \forall 1 \leq i < k, X_{v_i} = x_i) = \mathbb{P}_{\Omega_k}(X_{v_k} = x_k)$.

Since Ω_k satisfies the condition (1), by Lemma 2, there exists constant $C > 0$ such that

$$\left|\mathbb{P}_{\Omega_k}(X_{v_k} = x_{v_k}) - \widehat{\mathbb{P}}^{(t)}_{\Omega_k}(X_{v_k} = x_{v_k})\right| \leq \mathrm{poly}(|\Omega|) \cdot \exp(-C \cdot t).$$

Thus by choosing appropriate $t = O\left(\log \frac{1}{\epsilon} + \log q + \log n\right)$, it holds for every k that

$$\left|\mathbb{P}_{\Omega_k}(X_{v_k} = x_{v_k}) - \widehat{\mathbb{P}}^{(t)}_{\Omega_k}(X_{v_k} = x_{v_k})\right| \leq \frac{\epsilon}{4qn},$$

and since in the algorithm we always choose the x_{v_k} maximizing the value of $\widehat{\mathbb{P}}^{(t)}_{\Omega_k}(X_{v_k} = x_{v_k})$, we have $\widehat{\mathbb{P}}^{(t)}_{\Omega_k}(X_{v_k} = x_{v_k}) \geq \frac{1}{q}$ thus $\mathbb{P}_{\Omega_k}(X_{v_k} = x_{v_k}) \geq \frac{1}{q} - \frac{\epsilon}{4qn} \geq \frac{1}{2q}$.

By definition we have $\mathbb{P}_\Omega(\boldsymbol{X} = \boldsymbol{x}) = \frac{w(\boldsymbol{x})}{Z(\Omega)}$, thus $Z(\Omega) = \frac{w(\boldsymbol{x})}{\prod_{k=1}^{n} \mathbb{P}_{\Omega_k}(X_{v_k} = x_{v_k})}$. Therefore, we have

$$1 - \epsilon \leq \left(1 - \frac{\epsilon}{2n}\right)^n \leq \frac{Z(\Omega)}{\widehat{Z}(\Omega)} = \prod_{k=1}^{n} \frac{\widehat{\mathbb{P}}^{(t)}_{\Omega_k}(X_{v_k} = x_{v_k})}{\mathbb{P}_{\Omega_k}(X_{v_k} = x_{v_k})} \leq \left(1 + \frac{\epsilon}{2n}\right)^n \leq 1 + \epsilon,$$

which is simplified as that $1 - \epsilon \leq \frac{\widehat{Z}(\Omega)}{Z(\Omega)} \leq 1 + \epsilon$.

By Proposition 2, the total running time is bounded by $\mathrm{poly}(|\Omega|)(q\Delta)^{O(t)}$. Since $t = O\left(\log \frac{1}{\epsilon} + \log q + \log n\right)$, the algorithm is an FPTAS if q and Δ are constants, and in general the running time is bounded by $|\Omega|^{O(\log |\Omega|)}$ for any fixed $0 < \epsilon < 1$. □

4 List-coloring

We consider list-coloring instances $\Omega = (G, \mathcal{X}, \boldsymbol{L})$ satisfying the condition (3). Let $\Delta = \Delta_G$ be the maximum degree of G and define that $\chi(\Delta) = (\alpha - 1)\Delta + 1$. The condition (3) implies the following weaker condition:

$$\forall v \in V, \quad |L_v| - \deg(v) \geq \chi(\Delta). \tag{8}$$

A merit of considering this weaker condition is that it is closed under partial pinning and pinning. The pinning and partial pinning can be defined on list-coloring instances as they are special cases of MRFs. Given a list-coloring instance $\Omega = (G, \mathcal{X}, \boldsymbol{L})$ with underlying graph $G = (V, E)$ and a vertex $v \in V$ with d neighbors $v_1, v_2, \ldots, v_d$, for each color $x \in L_v$, the pinning of Ω is a new list-coloring instance $\Omega_{v,x} = \textsc{Pin}_{v,x}(\Omega) = (G_v, \mathcal{X}, \widehat{\boldsymbol{L}})$, where G_v is the subgraph of G induced by $V \setminus \{v\}$ and $\widehat{\boldsymbol{L}} = (\widehat{L}_u, u \in V \setminus \{v\})$ such that $\widehat{L}_u = L_u \setminus \{x\}$ if u is adjacent to v and $\widehat{L}_u = L_u$ if otherwise; and for each $1 \leq i \leq d+1$, the partial pinning of Ω is a new list-coloring instance $\Omega^i_{v,x} = (G_v, \mathcal{X}, \widetilde{\boldsymbol{L}})$, where $\widetilde{\boldsymbol{L}} = (\widetilde{L}_u, u \in V \setminus \{v\})$ such that $\widetilde{L}_u = L_u \setminus \{x\}$ for $u = v_j$ with $j < i$ and $\widetilde{L}_u = L_u$ for all other u in $V \setminus \{v\}$. The pinning and the partial pinning does not violate the condition (8) since it never increases the maximum degree, and if $|L_v|$ decreases by 1 then also $\deg(v)$ decreases by 1.

The following identity for marginals of list-coloring is proved in [10].

Proposition 3. *Let $\Omega = (G, \mathcal{X}, \boldsymbol{L})$ be a list-coloring instance on graph $G = (V, E)$, $v \in V$ a vertex with d neighbors $v_1, v_2, \ldots, v_d$ where $d = \deg(v)$, and $x \in L_v$ a color. It holds that*

$$\mathbb{P}_\Omega(X_v = x) = \frac{\prod_{i=1}^d \left(1 - \mathbb{P}_{\Omega^i_{v,x}}(X_{v_i} = x)\right)}{\sum_{y \in L_v} \prod_{i=1}^d \left(1 - \mathbb{P}_{\Omega^i_{v,x}}(X_{v_i} = y)\right)}.$$

Some simple lower and upper bounds hold for the marginals, similar to the ones proved in [10].

Lemma 3. *Let $\Omega = (G, \mathcal{X}, \boldsymbol{L})$ be a list-coloring instance with the maximum degree Δ of G, satisfying the condition (8). For any vertex $v \in V_G$ and any color $x \in L_v$, it holds for the marginal probability that $\frac{1}{q \cdot \mathrm{e}^{\frac{1}{\alpha-1}}} \leq \mathbb{P}_\Omega(X_v = x) \leq \frac{1}{\chi(\Delta)}$.*

Proof. The upper bound is easy: conditioning on any coloring of the neighbos of v, the number of remaining colors for v is at least $|L_v| - \deg(v) \geq \chi(\Delta)$, thus marginal probability is at most $\frac{1}{\chi(\Delta)}$. Applying the upper bound $\frac{1}{\chi(\Delta)}$ to the marginals in the numerator of the recursion in Proposition 3 and the trivial upper bound q to the denominator, we have the lower bound $\frac{1}{q \cdot \mathrm{e}^{\frac{1}{\alpha-1}}}$.

4.1 The computation tree recursion with adjustment

Given a list-coloring instance $\Omega = (G, \mathcal{X}, \boldsymbol{L})$ on graph $G = (V, E)$, a vertex $v \in V$ with d neighbors $v_1, v_2, \ldots, v_d$ and a color $x \in L_v$, the computation tree recursion $f_{\Omega,v,x}$ can be defined on the domain of all $\boldsymbol{p} = (p_{i,y}, 1 \le i \le d \wedge y \in L_v) \in [0,1]^{d|L_v|}$:

$$f_{\Omega,v,x}(\boldsymbol{p}) \triangleq \frac{\prod_{i=1}^{d}(1-p_{i,x})}{\sum_{y\in L_v}\prod_{i=1}^{d}(1-p_{i,y})}. \tag{9}$$

For $t \ge 0$, the quantity $\widehat{\mathbb{P}}^{(t)}_{\Omega}(X_v = x)$ is recursively defined as follows:

- If $t = 0$, let $\widehat{\mathbb{P}}^{(0)}_{\Omega}(X_v = x) = \frac{1}{|L_v|}$.
- If $t > 0$, let $\widehat{\mathbb{P}}^{(t)}_{\Omega}(X_v = x) = \min\left\{\frac{1}{|L_v|-d}, f_{\Omega,v,x}(\widehat{\boldsymbol{p}})\right\}$, where the $\widehat{\boldsymbol{p}}$ is taken as that $\hat{p}_{i,y} = \widehat{\mathbb{P}}^{(t-1)}_{\Omega^i_{v,x}}(X_{v_i} = y)$ for each $1 \le i \le d$ and $y \in L_v$.

Note that the only difference from the MRF case is the truncation of the value of $f(\widehat{\boldsymbol{p}})$ so that $\mathbb{P}_{\Omega}(X_v = x)$ never goes beyond the naive upper bound $\frac{1}{|L_v|-d}$. We call this procedure the *computation tree recursion with adjustment*. It is the same as the procedure proposed in [10] except with a more simplified value truncation.

The estimation $\widehat{Z}(\Omega)$ of the partition function is computed from these estimations $\widehat{\mathbb{P}}^{(t)}_{\Omega}(X_v = x)$ of marginal probabilities by the same algorithm as in Section 3.1. The same complexity bound as in Proposition 2 still holds.

4.2 Correlation Decay

The correlation decay of the computation tree recursion with adjustment can be defined in the same way as in Definition 2.

Lemma 4. *The computation tree recursion with adjustment exhibits exponential correlation decay on list-coloring instances satisfying the condition* (8) *with* $\alpha > \alpha^*$ *where* $\alpha^* \approx 2.58071$ *is defined by* (2) *in Section 2.*

Proof. Let $\Omega = (G, \mathcal{X}, \boldsymbol{L})$ be a list-coloring instance on the underlying graph $G = (V, E)$ with the maximum degree $\Delta = \Delta(G)$ satisfying the condition (8). It can be verified that all the list-coloring instances generated by recursively applying partial pinnings on Ω still satisfy the condition (8).

Let $v \in V$ be a vertex with d neighbors $v_1, v_2, \ldots, v_d$, $x \in L_v$ a color, and $f = f_{\Omega,v,x}$ the recursion defined by (9). It holds that $\mathbb{P}_{\Omega}(X_v = x) = f(\boldsymbol{p})$ where $\boldsymbol{p} = (p_{i,y}, 1 \le i \le d \wedge y \in L_v)$ and each $p_{i,y} = \mathbb{P}_{\Omega^i_{v,x}(X_{v_i}=y)}$. We choose the monotone differentiable function $\varphi : [0,1] \to \mathbb{R}$ so that its derivative is $\Phi(p) = \frac{\mathrm{d}\,\varphi(p)}{\mathrm{d}\,p} = \frac{1}{(1-p)\sqrt{p}}$. We define the amortized decay rate in the same way as (7) by:

$$\kappa(\boldsymbol{p}) \triangleq \sum_{\substack{1\le i\le d\\ y\in L_v}} \left|\frac{\partial f(\boldsymbol{p})}{\partial p_{i,y}}\right| \frac{\Phi(f(\boldsymbol{p}))}{\Phi(p_{i,y})}.$$

By the same analysis as in Lemma 1, due to the mean value theorem, we have

$$\begin{aligned}
&\left|\varphi\left(\mathbb{P}_{\Omega}(X_v = x)\right) - \varphi\left(\widehat{\mathbb{P}}^{(t)}_{\Omega}(X_v = x)\right)\right| \\
\le & \kappa(\boldsymbol{p}) \cdot \max_{\substack{1\le i\le d\\ y\in L_v}} \left|\varphi\left(\mathbb{P}_{\Omega^i_{v,x}}(X_{v_i} = y)\right) - \varphi\left(\widehat{\mathbb{P}}^{(t-1)}_{\Omega^i_{v,x}}(X_{v_i} = y)\right)\right|,
\end{aligned}$$

for some $\boldsymbol{p} = (p_{i,y}, 1 \le i \le d \wedge y \in L_v)$ such that the value of each $p_{i,y}$ is between $\mathbb{P}_{\Omega^i_{v,x}}(X_{v_i} = y)$ and $\widehat{\mathbb{P}}^{(t-1)}_{\Omega^i_{v,x}}(X_{v_i} = y)$. By Lemma 3, we have $\mathbb{P}_{\Omega}(X_{v_i} = y) \le \frac{1}{\chi(\Delta)}$ and due to the definition of the algorithm, $\widehat{\mathbb{P}}^{(t)}_{\Omega}(X_{v_i} = y) \le \frac{1}{|L_v|-d} \le \frac{1}{\chi(\Delta)}$, thus $p_{i,y} \le \frac{1}{\chi(\Delta)}$ for any $1 \le i \le d$ and $y \in L_v$.

By our choice of $\Phi(\cdot)$, it can be verified that

$$\begin{aligned}
\kappa(\boldsymbol{p}) &= \sum_{i=1}^{d} \left|\frac{\partial f(\boldsymbol{p})}{\partial p_{i,x}}\right| \frac{\Phi(f(\boldsymbol{p}))}{\Phi(p_{i,x})} + \sum_{\substack{1\le i\le d\\ y\in L_v\setminus\{x\}}} \left|\frac{\partial f(\boldsymbol{p})}{\partial p_{i,y}}\right| \frac{\Phi(f(\boldsymbol{p}))}{\Phi(p_{i,y})} \\
&\le \sqrt{f(\boldsymbol{p})}\left(\sum_{i=1}^{d}\sqrt{p_{i,x}} + \sum_{i=1}^{d}\max_{y\ne x}\sqrt{p_{i,y}}\right) \\
&\le \sqrt{\frac{\prod_{i=1}^{d}(1-p_{i,x})}{(d+\chi(\Delta))\left(1-\frac{1}{\chi(\Delta)}\right)^{d}}}\left(\sum_{i=1}^{d}\sqrt{p_{i,x}} + \frac{d}{\sqrt{\chi(\Delta)}}\right),
\end{aligned} \tag{10}$$

where the last inequality is due to that $p_{i,x} \le \frac{1}{\chi(\Delta)}$ and $|L_v| \le d + \chi(\Delta)$. Let $\bar{p} = 1 - \left(\prod_{i=1}^{d}(1-p_{i,x})\right)^{\frac{1}{d}}$. Then $\bar{p} \le \frac{1}{\chi(\Delta)}$ since all $p_{i,x}$ satisfy so, and $\prod_{i=1}^{d}(1-p_{i,x}) = (1-\bar{p})^d$. Let $\ell_i = \ln(1-p_{i,x})$, thus $\sum_{i=1}^{d}\ell_i = d\ln(1-\bar{p})$. The function $g(x) = \sqrt{1-\mathrm{e}^x}$ is concave over $x \le 0$, thus by Jensen's inequality,

$$\sum_{i=1}^{d}\sqrt{p_{i,x}} = \sum_{i=1}^{d} g(\ell_i) \le d\cdot g\left(\frac{1}{d}\sum_{i=1}^{d}\ell_i\right) = d\sqrt{\bar{p}}.$$

Therefore, (10) can be bounded by its symmetrized form as follows:

$$\begin{aligned}
\kappa(\boldsymbol{p}) \le \kappa(\bar{p}) &\triangleq \frac{d}{\sqrt{d+\chi(\Delta)}}\left(\frac{1-\bar{p}}{1-\frac{1}{\chi(\Delta)}}\right)^{\frac{d}{2}}\left(\sqrt{\bar{p}} + \frac{1}{\sqrt{\chi(\Delta)}}\right) \\
&\le \frac{\Delta}{\sqrt{\Delta+\chi(\Delta)}}\left(\frac{1-\bar{p}}{1-\frac{1}{\chi(\Delta)}}\right)^{\frac{\Delta}{2}}\left(\sqrt{\bar{p}} + \frac{1}{\sqrt{\chi(\Delta)}}\right).
\end{aligned}$$

where the last inequality is due to that $\bar{p} \le \frac{1}{\chi(\Delta)}$ and $d \le \Delta$.

Let $\bar{p} = \frac{\rho}{\chi(\Delta)}$ for $\rho \in [0,1]$. It holds that $\kappa(\bar{p}) \le \frac{(\sqrt{\rho}+1)}{\sqrt{\alpha(\alpha-1)}}\exp\left(-\frac{\rho-1}{2(\alpha-1)}\right)$, whose maximum is achieved when $\rho = \frac{1}{2}(2\alpha-1-\sqrt{4\alpha-3})$, such that

$$\kappa(\bar{p}) \le \kappa_\alpha \triangleq \frac{\sqrt{2}+\sqrt{2\alpha-1-\sqrt{4\alpha-3}}}{\sqrt{2\alpha(\alpha-1)}}\exp\left(\frac{3-2\alpha+\sqrt{4\alpha-3}}{4(\alpha-1)}\right).$$

It can be verified that κ_α is is monotonously decreasing from $+\infty$ to 0 for $\alpha > 1$, so $\kappa_\alpha < 1$ if $\alpha > \alpha^*$ where α^* is the unique solution to $\kappa_\alpha = 1$, as defined by (2).

Since the condition (8) is closed under partial pinning, by induction we have

$$\begin{aligned}&\left|\varphi\left(\mathbb{P}_\Omega(X_v = x)\right) - \varphi\left(\widehat{\mathbb{P}}^{(t)}_\Omega(X_v = x)\right)\right| \\ \leq &\kappa^t \left|\varphi\left(\mathbb{P}_{\Omega'}(X_u = z)\right) - \varphi\left(\widehat{\mathbb{P}}^{(0)}_{\Omega'}(X_u = z)\right)\right|,\end{aligned}$$

where $\Omega' = (G', \mathcal{X}, \boldsymbol{L}')$ is a list-coloring instance resulting from recursively applying t partial pinnings on the original Ω. By the same mean value theorem argument as in Lemma 1, we have $\left|\mathbb{P}_\Omega(X_v = x) - \widehat{\mathbb{P}}^{(t)}_\Omega(X_v = x)\right| \leq \frac{\Phi(\tilde{p}_0)}{\Phi(\tilde{p})}\kappa_\alpha^t$, for some $\tilde{p} \in [0,1]$ and some $\tilde{p}_0$ between $\mathbb{P}_{\Omega'}(X_u = w)$ and $\widehat{\mathbb{P}}^{(0)}_{\Omega'}(X_u = w) = \frac{1}{|L'_u|}$. Recall that the condition (8) is closed under partial pinning. It holds that $\frac{1}{q} \leq \frac{1}{|L'_v|} \leq \frac{1}{\chi(\Delta(G'))}$, and by Lemma 3 it also holds that $\frac{1}{q\cdot \mathrm{e}^{1/(\alpha-1)}} \leq \mathbb{P}_{\Omega'}(X_u = w) \leq \frac{1}{\chi(\Delta(G'))}$. Therefore, $\tilde{p}_0 \in \left[\frac{1}{q\cdot \mathrm{e}^{1/(\alpha-1)}}, \frac{1}{\chi(\Delta(G'))}\right]$. By our choice of $\Phi(p)$, we have $\frac{\Phi(\tilde{p}_0)}{\Phi(\tilde{p})} \leq \frac{\sqrt{q}\cdot \mathrm{e}^{\frac{1}{2(\alpha-1)}}}{1-\frac{1}{\chi(\Delta(G'))}} = O(\sqrt{q})$.

In conclusion, if the condition (8) is satisfied with $\alpha > \alpha^* \approx 2.58071$, there exists a constant $\kappa < 1$ such that $\left|\mathbb{P}_\Omega(X_v = x) - \widehat{\mathbb{P}}^{(t)}_\Omega(X_v = x)\right| \leq O(\sqrt{q})\kappa^t$.

Proof of Theorem 4: We first prove the theorem under the weaker condition (8), which is closed under pinning and partial pinning. The proof is the same as the proof of Theorem 3. The theorem with the stronger condition (3) follows as a consequence. □

References

1. A. Bandyopadhyay and D. Gamarnik. Counting without sampling: Asymptotics of the log-partition function for certain statistical physics models. *Random Structures & Algorithms*, 33(4):452–479, 2008.
2. M. Bayati, D. Gamarnik, D. Katz, C. Nair, and P. Tetali. Simple deterministic approximation algorithms for counting matchings. In *Proceedings of STOC*, pages 122–127, 2007.
3. J.-Y. Cai and X. Chen. A decidable dichotomy theorem on directed graph homomorphisms with non-negative weights. In *Proceedings FOCS*, pages 437–446, 2010.
4. J.-Y. Cai, X. Chen, and P. Lu. Graph homomorphisms with complex values: A dichotomy theorem. In *Proceedings of ICALP*, pages 275–286, 2010.
5. M. Dyer, M. Jerrum, and E. Vigoda. Rapidly mixing markov chains for dismantleable constraint graphs. In *RANDOM*, pages 68–77. 2002.
6. M. E. Dyer, A. M. Frieze, T. P. Hayes, and E. Vigoda. Randomly coloring constant degree graphs. In *Proceedings of FOCS*, pages 582–589, 2004.
7. M. E. Dyer and C. S. Greenhill. On markov chains for independent sets. *Journal of Algorithms*, 35(1):17–49, 2000.

8. A. Galanis, D. Štefankovič, 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. *arXiv preprint arXiv:1305.2902*, 2013.
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. D. Gamarnik, D. Katz, and S. Misra. Strong spatial mixing for list coloring of graphs. *arXiv preprint arXiv:1207.1223*, 2012.
12. L. Goldberg, R. Martin, and M. Paterson. Strong spatial mixing with fewer colors for lattice graphs. *SIAM Journal on Computing*, 35(2):486, 2005.
13. L. A. Goldberg and M. Jerrum. A polynomial-time algorithm for estimating the partition function of the ferromagnetic ising model on a regular matroid. In *Proceedings of ICALP*, pages 521–532, 2011.
14. L. A. Goldberg, M. Jerrum, and M. Paterson. The computational complexity of two-state spin systems. *Random Structures & Algorithms*, 23(2):133–154, 2003.
15. T. P. Hayes. Randomly coloring graphs of girth at least five. In *Proceedings of STOC*, pages 269–278, 2003.
16. M. Jerrum. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Random Structures & Algorithms*, 7(2):157–166, 1995.
17. M. Jerrum and A. Sinclair. Polynomial-time approximation algorithms for the ising model. *SIAM Journal on Computing*, 22(5):1087–1116, 1993.
18. M. Jerrum, A. Sinclair, and E. Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *Journal of the ACM*, 51:671–697, July 2004.
19. L. Li, P. Lu, and Y. Yin. Approximate counting via correlation decay in spin systems. In *Proceedings of SODA*, pages 922–940, 2012.
20. L. Li, P. Lu, and Y. Yin. Correlation decay up to uniqueness in spin systems. In *Proceedings of SODA*, pages 67–84, 2013.
21. M. Luby and E. Vigoda. Approximately counting up to four (extended abstract). In *Proceedings of STOC*, pages 682–687, 1997.
22. R. Restrepo, J. Shin, P. Tetali, E. Vigoda, and L. Yang. Improved mixing condition on the grid for counting and sampling independent sets. In *Proceedings of FOCS*, pages 140–149, 2011.
23. A. Sinclair, P. Srivastava, and M. Thurley. Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. In *Proceedings of SODA*, pages 941–953, 2012.
24. A. Sly. Uniqueness thresholds on trees versus graphs. *The Annals of Applied Probability*, 18(5):1897–1909, 2008.
25. A. Sly. Computational transition at the uniqueness threshold. In *Proceedings of FOCS*, pages 287–296, 2010.
26. A. Sly and N. Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *Proceedings of FOCS*, pages 361–369, 2012.
27. E. Vigoda. Improved bounds for sampling coloring. In *Proceedings of FOCS*, pages 51–59, 1999.
28. D. Weitz. Counting independent sets up to the tree threshold. In *Proceedings of STOC*, pages 140–149, 2006.
29. Y. Yin and C. Zhang. Approximate counting via correlation decay on planar graphs. In *Proceedings of SODA*, pages 47–66, 2013.