FPTAS for Hardcore and Ising Models on Hypergraphs[*]

Pinyan Lu[1], Kuan Yang[2], and Chihao Zhang[3]

1   School of Information Management and Engineering, Shanghai University of
     Finance and Economics, No. 100 Wudong Road, Yangpu District, Shanghai,
     China
     lu.pinyan@mail.shufe.edu.cn
2   Zhiyuan College, Shanghai Jiao Tong University, No. 800 Dongchuan Road,
     Minhang District, Shanghai, China
     kuan.yang.6@gmail.com
3   Department of Computer Science and Engineering, Shanghai Jiao Tong
     University, No. 800 Dongchuan Road, Minhang District, Shanghai, China
     chihao.zhang@gmail.com

―――― **Abstract** ――――

Hardcore and Ising models are two most important families of two state spin systems in statistic
physics. Partition function of spin systems is the center concept in statistic physics which connects
microscopic particles and their interactions with their macroscopic and statistical properties of
materials such as energy, entropy, ferromagnetism, etc. If each local interaction of the system
involves only two particles, the system can be described by a graph. In this case, fully polynomial-
time approximation scheme (FPTAS) for computing the partition function of both hardcore and
anti-ferromagnetic Ising model was designed up to the uniqueness condition of the system. These
result are the best possible since approximately computing the partition function beyond this
threshold is NP-hard. In this paper, we generalize these results to general physics systems,
where each local interaction may involves multiple particles. Such systems are described by
hypergraphs. For hardcore model, we also provide FPTAS up to the uniqueness condition, and
for anti-ferromagnetic Ising model, we obtain FPTAS under a slightly stronger condition.

1998 ACM Subject Classification F.2.2 Computations on Discrete Structures

Keywords and phrases hard-core model, ising model, hypergraph, spatial mixing, correlation
decay

1   Introduction

In recent couple of years, there are remarkable progress on designing approximate counting
algorithms based on correlation decay approach [26, 1, 7, 22, 13, 23, 14, 20, 15, 19, 18, 17, 16].
Unlike the previous major approximate counting approach that based on random sampling
such as Markov Chain Monte Carlo (MCMC) (see for examples [11, 10, 12, 8, 3, 9, 25, 4, 5, 21]),
correlation decay based approach provides deterministic fully polynomial-time approximation
scheme (FPTAS). New FPTASes were designed for a number of interesting combinatorial
counting problems and computing partition functions for statistic physics systems, where
partition function is a weighted counting function from the computational point of view. One

―――――――――

SYMPOSIUM
ON THEORETICAL
ASPECTS
OF COMPUTER
SCIENCE

most successful example is the algorithm for anti-ferromagnetic two-spin systems [13, 23, 14], including counting independent sets [26]. The correlation decay based FPTAS is beyond the best known MCMC based FPRAS and achieves the boundary of approximability [24, 6].

In this paper, we generalize these results of anti-ferromagnetic two-spin systems to hypergraphs. For physics point of view, this corresponds to spin systems with higher order interactions, where each local interaction involves more than two particles. There are two main ingredients for the original algorithms and analysis on normal graphs (we will use the term normal graph for a graph to emphasize that it is not hypergraphs): (1) the construction of the self-avoiding walk tree by Weitz [26], which transform a general graph to a tree; (2) correlation decay proof for the tree, which enables one to truncate the tree to get a good approximation in polynomial time. However, the construction of the self-avoiding walk tree cannot be extended to hypergraphs, which is the main obstacle for the generalization.

The most related previous work is counting independent sets for hypergraphs by Liu and Lu [17]. They established a computation tree with a two-layers recursive function instead of the self-avoiding walk tree and provided a FPTAS to count the number of independent sets for hypergraphs with maximum degree of 5, extending the algorithm for normal graph with the same degree bound. Their proof was significantly more complicated than the previous one due to the complication of the two-layers recursive function. In particular, the "right" degree bound for the problem is a real number between 5 and 6 if one allow fraction degree in some sense. This integer gap provides some room of flexibility and enables them to do some case-by-case numerical argument to complete the proof. However, the parameters for the anti-ferromagnetic two-spin systems on hypergraphs are real numbers. To get a sharp threshold, we do not have any room for numerical approximation.

1.1 Our Results

We study two most important anti-ferromagnetic two-spin systems on hypergraphs: the hardcore model and the anti-ferromagnetic Ising model. The formal definitions of these two models can be found in Section 2.

Our first result is an FPTAS to compute the partition function of hypergraph hardcore model.

▶ **Theorem 1.** *For hardcore model with a constant activity parameter of λ, there is an FPTAS to compute the partition function for hypergraphs with maximum degree $\Delta \geq 2$ if $\lambda < \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta}$.*

This bound is exactly the uniqueness threshold for the hardcore on normal graphs. Thus, it is tight since normal graphs are special cases of hypergraphs. To approximately compute the partition function beyond this threshold is NP-hard. In particular, The FPTAS in [17] for counting the number of independent sets for hypergraphs with maximum degree of 5 can be viewed as a special case of our result with parameters $\Delta = 5, \lambda = 1$, which satisfies the above uniqueness condition. Another interesting special case is when $\Delta = 2$. This is not an interesting case for normal graphs since a normal graph with maximum degree of 2 is simply a disjoint union of paths and cycles, whose partition function can be computed exactly. However, the problem becomes more complicated on hypergraphs: it can be interpreted as counting weighted edge covers on normal graphs by viewing vertices of degree two as edges and hyperedges as vertices. The exact counting of this problem is known to be #P-complete and an FPTAS was found recently [18]. In our model, the uniqueness bound $\frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta}$ is infinite for $\Delta = 2$ and as a result we give an FPTAS for counting weighted edge covers for any constant edge weight λ. This gives an alternative proof for the main result in [18].

Our second result is on computing the partition function of anti-ferromagnetic Ising model.

▶ **Theorem 2.** *For Ising model with interaction parameter $0 < \beta < 1$ and external field λ, there is an FPTAS to compute the partition function for hypergraphs with maximum degree Δ if $\beta \geq 1 - \frac{2}{2e^{-1/2}\Delta+3}$.*

The tight uniqueness bound for anti-ferromagnetic Ising model on normal graphs is $\beta \geq 1 - \frac{2}{\Delta}$. So, our bound is in the same asymptotic order but a bit worse in the constant coefficient as $2e^{-1/2} \approx 1.213 > 1$. Moreover, our result can apply beyond Ising model to a larger family of anti-ferromagnetic two-spin systems on hypergraphs.

1.2 Our Techniques

We also use the correlation decay approach. Although the framework of this method is standard, in many work along this line of research, new tools and techniques are developed to make this relatively new approach more powerful and widely applicable. This is indeed the case for the current paper as well. We summarize the new techniques we introduced here.

For hardcore model, we replace the numerical case-by-case analysis by a monotone argument with respect to the edge size of the hypergraph which shows that the normal graphs with edge size of 2 is indeed the worst cases. This gives a tight bound for hardcore model.

To handle hypergraph with unbounded edge sizes, we need to prove that the decay rate is much smaller for edges of larger size. Such effect is called computationally efficient correlation decay, which has been used in many previous works to obtain FPTASes for systems with unbounded degrees or edge sizes. In all those works, one sets a threshold for the parameter and proves different types of bounds for large and small ones separately. Such artificial separation gets a discontinuous bound which adds some complications in the proof and usually ends with a case-by-case discussion. In particular, this separation is not compatible with the above monotone argument. To overcome this, we propose a new uniform and smooth treatment for this by modifying the decay rate by a polynomial function of the edge size. After this modification, we only need to prove one single bound which automatically provides computationally efficient correlation. We believe that this idea is important and may find applications in other related problems.

For the Ising model, the main difficulty is to get a computation tree as a replacement of the self-avoiding walk tree. We proposed one, which also works for general anti-ferromagnetic two-spin systems on hypergraphs. However, unlike the case of the hardcore model, the computation tree is not of perfect efficiency and this is the main reason that the bound we achieve in Theorem 2 is not tight. To get the computationally efficient correlation decay, we also use the above mentioned uniform and smooth treatment. We also extend our result beyond Ising to a family of anti-ferromagnetic two-spin systems on hypergraphs.

1.3 Discussion and Open Problems

One obvious open question is to close the gap for Ising model, or more generally extend our work to anti-ferromagnetic two-spin systems on hypergraphs with better parameters. However, it seems that it is impossible to obtain a tight result in these models using the computation tree proposed in this paper, due to its imperfectness. How to overcome this is an important open question.

Even for the hardcore model, our result is tight only for the family of all hypergraphs since the normal graphs are special cases. From both physics and combinatorics point of view, it would be very interesting to study the family of w-uniform hypergraphs where each hyperedge is of the same size w. By our monotone argument, it is plausible to conjecture that one can get better bound for larger w. In particular, MCMC based approach does show that larger edge size helps: for hypergraph independent set with maximum degree of Δ and minimum edge size w, an FPRAS for $w \geq 2\Delta + 1$ was shown in [2]. However, their result is not tight. Can we get a tight bounds in terms of Δ and w by correlation decay approach? The high level idea sounds promising, but there is an obstacle to prove such result by our computation tree. To construct the computation tree, we need to construct modified instances. In these modified instances, the size of a hyperedge may decrease to as low as 2. Therefore, even if we start with w-uniform hypergraphs or hypergraphs with minimum edge size of w, we may need to handle the worst case of normal graphs during the analysis. How to avoid this effect is a major open question whose solution may have applications in many other problems.

The fact that larger hyperedge size only makes the problem easier is not universally true for approximation counting. One interesting example is counting hypergraph matchings. FPTAS for counting 3D matchings of hypergraphs with maximum degree 4 is given in [17], and extension to weighted setting are studied in [27]. In particular, a uniqueness condition in this setting is defined in [27], and it is a very interesting open question whether this uniqueness condition is also the transition boundary for approximability.

2   Preliminaries

A hypergraph $G(V, \mathcal{E})$ consists of a vertex set V and a set of hyperedges $\mathcal{E} \subseteq 2^V$. For every hyperedge $e \in \mathcal{E}$ and vertex $v \in V$, we use $e - v$ to denote $e \setminus \{v\}$ and use $e + v$ to denote $e \cup \{v\}$.

2.1   Hypergraph Hardcore Model

The hardcore model is parameterized by the activity parameter $\lambda > 0$. Let $G(V, \mathcal{E})$ be a hypergraph. An independent set of G is a vertex set $I \subseteq V$ such that $e \not\subseteq I$ for every hyperedge $e \in \mathcal{E}$. We use $\mathcal{I}(G)$ to denote the set of independent sets of G. The weight of an independent set I is defined as $w(I) \triangleq \lambda^{|I|}$. We let $Z(G)$ denote the partition function of $G(V, \mathcal{E})$ in the hardcore model, which is defined as

$$Z(G) \triangleq \sum_{I \in \mathcal{I}(G)} w(I).$$

The weight of independent sets induces a Gibbs measure on G. For every $I \in \mathcal{I}$, we use

$$\mathbf{Pr}_G[I] \triangleq \frac{w(I)}{Z(G)}$$

to denote the probability of obtaining I if we sample according to the Gibbs measure. For every $v \in V$, we use

$$\mathbf{Pr}_G[v \in I] \triangleq \sum_{\substack{I \in \mathcal{I}(G) \\ v \in I}} \mathbf{Pr}_G[I]$$

to denote the *marginal probability* of v.

2.2 Hypergraph Two State Spin Model

Now we give a formal definition to hypergraph two state spin systems. This model is parameterized by the external field $\lambda > 0$. An instance of the model is a labeled hypergraph $G(V, \mathcal{E}, (\boldsymbol{\beta}, \boldsymbol{\gamma}))$ where $\boldsymbol{\beta}, \boldsymbol{\gamma} : \mathcal{E} \to \mathbb{R}$ are two labeling functions that assign each edge $e \in \mathcal{E}$ two reals $\boldsymbol{\beta}(e), \boldsymbol{\gamma}(e)$. A configuration on G is an assignment $\sigma : V \to \{0, 1\}$ whose weight $w(\sigma)$ is defined as

$$w(\sigma) \triangleq \prod_{e \in \mathcal{E}} w(e, \sigma) \prod_{v \in V} w(v, \sigma)$$

where for a hyperedge $e = \{v_1, \ldots, v_w\}$

$$w(e, \sigma) \triangleq \begin{cases} \boldsymbol{\beta}(e) & \text{if } \sigma(v_1) = \sigma(v_2) = \cdots = \sigma(v_w) = 0 \\ \boldsymbol{\gamma}(e) & \text{if } \sigma(v_1) = \sigma(v_2) = \cdots = \sigma(v_w) = 1 \\ 1 & \text{otherwise} \end{cases}$$

and for a vertex v,

$$w(v, \sigma) \triangleq \begin{cases} \lambda & \text{if } \sigma(v) = 1 \\ 1 & \text{otherwise.} \end{cases}$$

The partition function of the instance is given by

$$Z(G) = \sum_{\sigma \in \{0,1\}^V} w(\sigma).$$

Similarly, the weight of configurations induces a Gibbs measure on G. For every $\sigma \in \{0, 1\}^V$, we use

$$\mathbf{Pr}_G\left[\sigma\right] \triangleq \frac{w(I)}{Z(G)}$$

to denote the probability of σ in the measure. For every $v \in V$, we use

$$\mathbf{Pr}_G\left[\sigma(v) = 1\right] \triangleq \sum_{\substack{\sigma \in \{0,1\}^V \\ \sigma(v)=1}} \mathbf{Pr}_G\left[\sigma\right]$$

to denote the *marginal probability* of v.

The *anti-ferromagnetic* Ising model is the special case that $\beta \triangleq \boldsymbol{\beta}(e) = \boldsymbol{\gamma}(e) \leq 1$ for all $e \in \mathcal{E}$. In this model, we call β the *interaction parameter* of the model. The hardcore model introduced in previous section is the special case that $\boldsymbol{\beta}(e) = 1$ and $\boldsymbol{\gamma}(e) = 0$ for all $e \in \mathcal{E}$.

The whole proof of Theorem 2 can be found in the full version of the paper. More precisely, we design an FPTAS for the more general two state spin system and establish the following theorem:

▶ **Theorem 3.** *Consider a class of two state spin system with external field λ such that each instance $G(V, \mathcal{E}, (\boldsymbol{\beta}, \boldsymbol{\gamma}))$ in the class satisfies $1 - \frac{2}{2e^{-1/2}\Delta+3} \leq \boldsymbol{\beta}(e), \boldsymbol{\gamma}(e) \leq 1$ where Δ is the maximum degree of G. There exists an FPTAS to compute the partition function for every instance in the class.*

Theorem 2 then follows since it is a special case of Theorem 3.

Actually, the main idea of FPTAS design and proof for this model is similar to the idea we use to solve hypergraph hardcore model. However, the details of recursion function design and techniques for proof of correlation decay property are pretty different from that in hypergraph hardcore model, see the full version of the paper for details.

Hypergraph Hardcore Model

3.1 Recursion for Computing Marginal Probability

We first fix some notations on graph modification specific to hypergraph independent set. Let $G(V, \mathcal{E})$ be a hypergraph.

- For every $v \in V$, we denote $G - v \triangleq (V \setminus \{v\}, \mathcal{E}')$ where $\mathcal{E}' \triangleq \{e \setminus \{v\} \mid e \in \mathcal{E}\}$.
- For every $e \in \mathcal{E}$, we denote $G - e \triangleq (V, \mathcal{E} \setminus \{e\})$.
- Let x be a vertex or an edge and y be a vertex or an edge, we denote $G - x - y \triangleq (G - x) - y$.
- Let $S = \{v_1, \ldots, v_k\} \subseteq V$, we denote $G - S \triangleq G - v_1 - v_2 \cdots - v_k$.
- Let $\mathcal{F} = \{e_1, \ldots, e_k\} \subseteq \mathcal{E}$, we denote $G - \mathcal{F} \triangleq G - e_1 - e_2 \cdots - e_k$.

Let $G(V, \mathcal{E})$ be a hypergraph and $v \in V$ be an arbitrary vertex with degree d. Let $\{e_1, \ldots, e_d\}$ be the set of hyperedges incident to v and for every $i \in [d]$, $e_i = \{v\} \cup \{v_{ij} \mid j \in [w_i]\}$ consists of $w_i + 1$ vertices.

We first define a hypergraph $G'(V', \mathcal{E}')$, which is the graph obtained from G by replacing v by d copies of itself and each e_i contains a distinct copy. Formally, $V' \triangleq (V \setminus \{v\}) \cup \{v_1, \ldots, v_d\}$, $\mathcal{E}' \triangleq \{e \in \mathcal{E} \mid v \notin e\} \cup \{e_i - v + v_i \mid i \in [d]\}$.

For every $i \in [d]$ and $j \in [w_i]$, we define a hypergraph $G_{ij}(V_{ij}, \mathcal{E}_{ij})$:

$$G_{ij} \triangleq G' - \{v_k \mid i \le k \le d\} - \{e_k \mid 1 \le k \le i\} - \{v_{ik} \mid 1 \le k < j\}.$$

Let $R_v = \frac{\mathbf{Pr}_G[v \in I]}{\mathbf{Pr}_G[v \notin I]}$ and $R_{ij} = \frac{\mathbf{Pr}_{G_{ij}}[v_{ij} \in I]}{\mathbf{Pr}_{G_{ij}}[v_{ij} \notin I]}$. We can compute R_v by following recursion:

▶ **Lemma 4.**

$$R_v = \lambda \prod_{i=1}^{d} \left(1 - \prod_{j=1}^{w_i} \frac{R_{ij}}{1 + R_{ij}} \right). \tag{1}$$

Proof. By the definition of R_v, we have

$$R_v = \frac{\mathbf{Pr}_G[v \in I]}{\mathbf{Pr}_G[v \notin I]} = \lambda \cdot \frac{\mathbf{Pr}_{G'}\left[\bigwedge_{i=1}^{d} v_i \in I\right]}{\mathbf{Pr}_{G'}\left[\bigwedge_{i=1}^{d} v_i \notin I\right]}$$

$$= \lambda \cdot \prod_{i=1}^{d} \frac{\mathbf{Pr}_{G'}\left[v_i \in I \wedge \bigwedge_{j=1}^{i-1} v_j \notin I \wedge \bigwedge_{j=i+1}^{d} v_j \in I\right]}{\mathbf{Pr}_{G'}\left[v_i \notin I \wedge \bigwedge_{j=1}^{i-1} v_j \notin I \wedge \bigwedge_{j=i+1}^{d} v_j \in I\right]}$$

For every $i \in [d]$, define $G_i \triangleq G' - \{v_k \mid i < k \le d\} - \{e_k \mid 1 \le k < i\}$, we have

$$\frac{\mathbf{Pr}_{G'}\left[v_i \in I \wedge \bigwedge_{j=1}^{i-1} v_i \notin I \wedge \bigwedge_{j=i+1}^{d} v_j \in I\right]}{\mathbf{Pr}_{G'}\left[v_i \notin I \wedge \bigwedge_{j=1}^{i-1} v_j \notin I \wedge \bigwedge_{j=i+1}^{d} v_j \in I\right]} = \frac{\mathbf{Pr}_{G'}\left[v_i \in I \mid \bigwedge_{j=1}^{i-1} v_j \notin I \wedge \bigwedge_{j=i+1}^{d} v_j \in I\right]}{\mathbf{Pr}_{G'}\left[v_i \notin I \mid \bigwedge_{j=1}^{i-1} v_j \notin I \wedge \bigwedge_{j=i+1}^{d} v_j \in I\right]}$$

$$= \frac{\mathbf{Pr}_{G_i}[v_i \in I]}{\mathbf{Pr}_{G_i}[v_i \notin I]}.$$

This is because fixing $v_j \in I$ is equivalent to removing v_j from the graph and fixing $v_j \notin I$ is equivalent to removing all edges incident to v_j from the graph.

Since e_i is the unique hyperedge in G_i that contains v_i, we have

$$\frac{\mathbf{Pr}_{G_i}[v_i \in I]}{\mathbf{Pr}_{G_i}[v_i \notin I]} = 1 - \mathbf{Pr}_{G_i - v_i - e_i}\left[\bigwedge_{j=1}^{w_i} v_{ij} \in I\right] = 1 - \prod_{j=1}^{w_i} \mathbf{Pr}_{G_{ij}}[v_{ij} \in I] = 1 - \prod_{j=1}^{w_i} \frac{R_{ij}}{1 + R_{ij}}.$$

◀

The Uniqueness Condition

Let the underlying graph be an infinite d-ary tree, then the recursion (1) becomes

$$f_{\lambda,d}(x) = \lambda \left(\frac{1}{1+x} \right)^d.$$

Let $\hat{x}$ be the positive fixed-point of $f_{\lambda,d}(x)$, i.e., $\hat{x} > 0$ and $f_{\lambda,d}(\hat{x}) = \hat{x}$. The condition on λ for the uniqueness of the Gibbs measure is that $\left| f'_{\lambda,d}(\hat{x}) \right| < 1$. The following proposition is well-known.

▶ **Proposition 5.** *Let* $\lambda_c = \frac{d^d}{(d-1)^{d+1}}$, *then* $\left| f'_{\lambda_c,d}(\hat{x}) \right| = 1$ *and for every* $0 < \lambda < \lambda_c$, *it holds that* $\left| f'_{\lambda,d}(\hat{x}) \right| < 1$.

3.2 The Algorithm to Compute Marginal Probability

Let $G(V, \mathcal{E})$ be a hypergraph with maximum degree Δ and $v \in V$ be an arbitrary vertex with degree d. Define G_{ij}, R_v, R_{ij} as in Section 3.1. Then the recursion (1) gives a way to compute the marginal probability $\mathbf{Pr}_G [v \in I]$ *exactly*. However, an exact evaluation of the recursion requires a computation tree with exponential size. Thus we introduce the following truncated version of the recursion, with respect to constants $c > 0$ and $0 < \alpha < 1$.

$$R(G, v, L) = \begin{cases} \lambda \prod_{i=1}^{d} \left(1 - \prod_{j=1}^{w_i} \frac{R(G_{ij}, v_{ij}, L)}{1 + R(G_{ij}, v_{ij}, L)} \right) & \text{if } d = \Delta \\ \lambda \prod_{i=1}^{d} \left(1 - \prod_{j=1}^{w_i} \frac{R(G_{ij}, v_{ij}, L - \lfloor 1 + c \log_{1/\alpha} w_i \rfloor)}{1 + R(G_{ij}, v_{ij}, L - \lfloor 1 + c \log_{1/\alpha} w_i \rfloor)} \right) & \text{if } d < \Delta \text{ and } L > 0 \\ \lambda & \text{otherwise.} \end{cases}$$

The recursion can be directly used to compute $R(G, v, L)$ for any given L and it induces a truncated computation tree (with height L in some special metric). It is worth noting that, the case that $d = \Delta$ can only happen at the root of the computation tree, since in each smaller instance, the degree of v_{ij} is decreased by at least one.

We claim that $R(G, v, L)$ is a good estimate of R_v with a suitable choice of c and α, for those (λ, Δ) in the uniqueness region.

▶ **Lemma 6.** *Let* $G(V, \mathcal{E})$ *be a hypergraph with maximum degree* $\Delta \geq 2$. *Let* $v \in V$ *be a vertex with degree* d *and let* $\lambda < \lambda_c = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta}$ *be the activity parameter. There exist constants* $C > 0$ *(more precisely,* $C = 6\lambda\sqrt{1+\lambda}$*) and* $\alpha \in (0, 1)$ *such that*

$$|R(G, v, L) - R_v| \leq C \cdot \alpha^{\max\{0, L\}}$$

for every L.

The whole proof of this lemma is postponed to the next section. Now we can prove Theorem 1 via using this lemma.

Proof of Theorem 1. The input of the FPTAS is an instance $G(V, \mathcal{E})$ and an accuracy parameter $0 < \varepsilon < 1/2$. Assume $V = \{v_1, \ldots, v_n\}$. Note that $I = \varnothing$ is an independent set of G with $w(I) = 1$. Therefore

$$Z(G) = 1/\mathbf{Pr}_G [I] = \left(\mathbf{Pr}_G \left[\bigwedge_{i=1}^{n} v_i \notin I \right] \right)^{-1} = \left(\prod_{i=1}^{n} \mathbf{Pr}_G \left[v_i \notin I \,\middle|\, \bigwedge_{j=1}^{i-1} v_j \notin I \right] \right)^{-1}.$$

For every $1 \leq i \leq n$, we define a graph $G_i(V_i, E_i)$:

- $G_1 \triangleq G$;
- For every $i \geq 2$, $G_i \triangleq G_{i-1} - v_{i-1} - \mathcal{E}'$ where $\mathcal{E}' \triangleq \{e \in \mathcal{E}_{i-1} \mid v_{i-1} \in e\}$ consists of edges in G_{i-1} incident to v_i.

It is straightforward to verify that $\mathbf{Pr}_G\left[v_i \notin I \mid \bigwedge_{j=1}^{i-1} v_j \notin I\right] = \mathbf{Pr}_{G_i}[v_i \notin I]$ for every $1 \leq i \leq n$. Thus,

$$Z(G) = \prod_{i=1}^{n} \left(\mathbf{Pr}_{G_i}[v_i \notin I]\right)^{-1} = \prod_{i=1}^{n} (1 + R_i),$$

where $R_i \triangleq \frac{\mathbf{Pr}_{G_i}[v_i \in I]}{\mathbf{Pr}_{G_i}[v_i \notin I]}$. Let C and α be constants in Lemma 6. We compute $R(G_i, v_i, L)$ with $L = \frac{\log(2Cn\varepsilon^{-1})}{\log \alpha^{-1}}$ for every $1 \leq i \leq n$, then $|R_i - R(G_i, v_i, L)| \leq \frac{\varepsilon}{2n}$. This implies

$$1 - \frac{\varepsilon}{2n} \leq \frac{1 + R_i}{1 + R(G, v_i, L)} \leq 1 + \frac{\varepsilon}{2n}.$$

Let $\hat{Z} = \prod_{i=1}^{n} \left(1 + R(G_i, v_i, L)\right)^{-1}$ be our estimate of the partition function, then it holds that

$$e^{-\varepsilon} \leq \frac{Z(G)}{\hat{Z}} \leq e^{\varepsilon}.$$

It remains to bound the running time of our algorithm. Let $T(L)$ denote the maximum running time of computing $R(G, v, L)$ (over all choices of $d \leq \Delta$ and arbitrary w_i). Then by the definition of $R(G, v, L)$, for every $L > 0$,

$$T(L) \leq \sum_{i=1}^{d} \sum_{j=1}^{w_i} T(L - \lfloor 1 + c \log_{1/\alpha} w_i \rfloor) + O(n).$$

It is easy to verify that $T(L) = n\Delta^{O(L)} = \left(\frac{n}{\varepsilon}\right)^{O(\log \Delta)}$ for our choice of L. Thus our algorithm is an FPTAS for computing $Z(G)$. ◀

3.3 Correlation Decay

In this section, we establish Lemma 6. We first prove some technical lemmas.

Suppose $f : D^d \to \mathbb{R}$ is a d-ary function where $D \subseteq \mathbb{R}$ is a convex set, let $\phi : \mathbb{R} \to \mathbb{R}$ be an increasing differentiable function and $\Phi(x) \triangleq \phi'(x)$. The following proposition is a consequence of the mean value theorem:

▶ **Proposition 7.** *For every* $\mathbf{x} = (x_1, \ldots, x_d)$, $\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_d) \in D^d$, *it holds that*
1. $|f(\mathbf{x}) - f(\hat{\mathbf{x}})| = \frac{1}{\Phi(\tilde{x})} |\phi(f(\mathbf{x})) - \phi(f(\hat{\mathbf{x}}))|$ *for some* $\tilde{x} \in D$;
2. $|\phi(f(\mathbf{x})) - \phi(f(\hat{\mathbf{x}}))| \leq \sum_{i=1}^{d} \frac{\Phi(f)}{\Phi(\tilde{x}_i)} \left|\frac{\partial f(\tilde{\mathbf{x}})}{\partial x_i}\right| \cdot |\phi(x_i) - \phi(\hat{x}_i)|$ *for some* $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_d) \in D^d$.

▶ **Lemma 8.** *Let* $\Delta \geq 2$ *be a constant integer and* $\lambda < \lambda_c = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$ *be a constant real. Let* $d < \Delta$ *and* $w_1, \ldots, w_d > 0$ *be integers and* $f = \lambda \prod_{i=1}^{d} \left(1 - \prod_{j=1}^{w_i} \frac{x_{ij}}{1+x_{ij}}\right)$ *be a* $\left(\sum_{i=1}^{d} w_i\right)$-*ary function. Let* $\Phi(x) = \frac{1}{\sqrt{x(1+x)}}$. *Let* $c < \min\left\{\frac{\log(1+\lambda) - \log \lambda}{2 + 4\lambda}, \frac{2\lambda+1}{2} \log\left(\frac{1+\lambda}{\lambda}\right) - 1\right\}$ *be a positive number. There exists a constant* $\alpha < 1$ *depending on* λ *and* d *(but not depending on* w_i *for all* $i \in [d]$*) such that*

$$\sum_{a=1}^{d} w_a^c \sum_{b=1}^{w_a} \frac{\Phi(f)}{\Phi(x_{ab})} \left|\frac{\partial f(\mathbf{x})}{\partial x_{ab}}\right| \leq \alpha < 1$$

for every $\mathbf{x} = (x_{ij})_{i \in [d], j \in [w_i]}$ *where each* $x_{ij} \in [0, \lambda]$

The lemma bounds the amortized decay rate, which is the key to the proof of correlation decay. In previous works, the amortized decay rate is defined as

$$\sum_{a=1}^{d} \sum_{b=1}^{w_a} \frac{\Phi(f)}{\Phi(x_{ab})} \left| \frac{\partial f(\mathbf{x})}{\partial x_{ab}} \right|,$$

without the w_a^c factor. Then one need to give a constant $\alpha < 1$ bound for small w_a and a sub constant bound for large w_a. With this modification, we only need to prove a single bound as above.

Notice that we require c to be a positive constant, so it is necessary to verify that $\frac{2\lambda+1}{2} \log\left(\frac{1+\lambda}{\lambda}\right) - 1 > 0$ for every $\lambda > 0$. To see this, let $h(\lambda) \triangleq \frac{2\lambda+1}{2} \log\left(\frac{1+\lambda}{\lambda}\right) - 1$, then we can compute that

$$h'(\lambda) = \log\left(\frac{1+\lambda}{\lambda}\right) - \frac{1+2\lambda}{2\lambda + 2\lambda^2},$$

$$h''(\lambda) = \frac{1}{2\lambda^2(1+\lambda)^2}.$$

Since $h''(\lambda) > 0$ for every λ, $h'(\lambda)$ is increasing. Along with the fact that $\lim_{\lambda\to\infty} h'(\lambda) = 0$, we have $h'(\lambda) < 0$ for every $\lambda > 0$. This implies that $h(\lambda)$ is decreasing. Also note that

$$\lim_{\lambda\to\infty} h(\lambda) = \lim_{\lambda\to\infty} \log\left(\left(1+\frac{1}{\lambda}\right)^{\lambda}\left(1+\frac{1}{\lambda}\right)^{1/2}\right) - 1 = 0.$$

It holds that $h(\lambda) > 0$ for every $\lambda > 0$. Thus a positive c satisfying $c < h(\lambda)$ exists for every $\lambda > 0$.

Proof of Lemma 8. To simplify the notation, we first let $t_{ij} = \frac{x_{ij}}{1+x_{ij}}$, then for every $i \in [d]$ and $j \in [w_i]$, it holds that $t_{ij} \in \left[0, \frac{\lambda}{1+\lambda}\right]$ and

$$f = \lambda \prod_{i=1}^{d} \left(1 - \prod_{j=1}^{w_i} t_{ij}\right).$$

For every $a \in [d]$ and $b \in [w_i]$, we have

$$\left|\frac{\partial f}{\partial x_{ab}}\right| = \lambda(1-t_{ab})^2 \prod_{\substack{j\in[w_a]\\j\neq b}} t_{aj} \cdot \prod_{\substack{i\in[d]\\i\neq a}} \left(1 - \prod_{j=1}^{w_i} t_{ij}\right) = f \cdot \frac{(1-t_{ab})^2}{t_{ab}} \cdot \frac{\prod_{j=1}^{w_a} t_{aj}}{1 - \prod_{j=1}^{w_a} t_{aj}}.$$

Thus

$$\sum_{a=1}^{d} w_a^c \sum_{b=1}^{w_a} \frac{\Phi(f)}{\Phi(x_{ab})} \left|\frac{\partial f}{\partial x_{ab}}\right| = \sqrt{\frac{f}{1+f}} \sum_{a=1}^{d} \frac{w_a^c \prod_{j=1}^{w_a} t_{aj}}{1 - \prod_{j=1}^{w_a} t_{aj}} \sum_{b=1}^{w_a} \frac{1-t_{ab}}{\sqrt{t_{ab}}}.$$

Let $\mathbf{t} = (t_{ij})_{i\in[d],j\in[w_i]}$, define

$$h(\mathbf{t}) \triangleq \sqrt{\frac{f}{1+f}} \sum_{a=1}^{d} \frac{w_a^c \prod_{j=1}^{w_a} t_{aj}}{1 - \prod_{j}^{w_a} t_{aj}} \sum_{b=1}^{w_a} \frac{1-t_{ab}}{\sqrt{t_{ab}}}$$

$$= \sqrt{\frac{\lambda \prod_{i=1}^{d}\left(1 - \prod_{j=1}^{w_i} t_{ij}\right)}{1 + \lambda \prod_{i=1}^{d}\left(1 - \prod_{j=1}^{w_i} t_{ij}\right)}} \sum_{a=1}^{d} \frac{w_a^c \prod_{j=1}^{w_a} t_{aj}}{1 - \prod_{j}^{w_a} t_{aj}} \sum_{b=1}^{w_a} \frac{1-t_{ab}}{\sqrt{t_{ab}}}.$$

For every $\mathbf{t} = (t_{ij})_{i \in [d], j \in [w_i]}$ where each $t_{ij} \in [0, \frac{\lambda}{1+\lambda}]$, define a tuple $\hat{\mathbf{t}} = (\hat{t}_{ij})_{i \in [d], j \in w_i}$ such that for every $i \in [d]$,

$$\hat{t}_{ij} = \begin{cases} \left(\frac{1+\lambda}{\lambda}\right)^{w_i - 1} \prod_{k=1}^{w_i} t_{ik} & \text{if } j = 1 \\ \frac{\lambda}{1+\lambda} & \text{otherwise.} \end{cases}$$

We claim that $h(\mathbf{t}) \le h(\hat{\mathbf{t}})$. To see this, first note that for every $i \in [d]$, $\prod_{j=1}^{w_i} t_{ij} = \prod_{j=1}^{w_i} \hat{t}_{ij}$, it is sufficient to prove that for every $i \in [d]$, $\sum_{j=1}^{w_i} \frac{1 - t_{ij}}{\sqrt{t_{ij}}} \le \sum_{j=1}^{w_i} \frac{1 - \hat{t}_{ij}}{\sqrt{\hat{t}_{ij}}}$. This is a consequence of the Karamata's inequality by noticing that the function $\frac{1 - e^x}{\sqrt{e^x}}$ is convex.

We rename $\hat{t}_{i1}$ to t_i and it is sufficient to upper bound

$$g(\mathbf{t}, \mathbf{w}) \triangleq \sqrt{\frac{\lambda \prod_{i=1}^{d} \left(1 - \left(\frac{\lambda}{1+\lambda}\right)^{w_i - 1} t_i\right)}{1 + \lambda \prod_{i=1}^{d} \left(1 - \left(\frac{\lambda}{1+\lambda}\right)^{w_i - 1} t_i\right)}} \cdot \sum_{i=1}^{d} \frac{w_i^c \left(\frac{\lambda}{1+\lambda}\right)^{w_i - 1} t_i}{1 - \left(\frac{\lambda}{1+\lambda}\right)^{w_i - 1} t_i} \cdot \left(\frac{1 - t_i}{\sqrt{t_i}} + \frac{(w_i - 1)}{\sqrt{\lambda + \lambda^2}}\right) \tag{2}$$

where $t_i \in \left[0, \frac{\lambda}{1+\lambda}\right]$ and $w_i \in \mathbb{Z}^+$ for every $i \in [d]$.

The argument so far is similar to the proof in [17]. In the following, we prove a monotonicity property of each w_i and thus avoid the heavy numerical analysis in [17] and allow us to obtain a tight result.

For every $i \in [d]$, we let $z_i \triangleq 1 - \left(\frac{\lambda}{1+\lambda}\right)^{w_i - 1} t_i$ and thus equivalently $t_i = (1 - z_i) \left(\frac{1+\lambda}{\lambda}\right)^{w_i - 1}$. For every fixed $\mathbf{z} = (z_1, \dots, z_d)$, we can write (2) as

$$g_{\mathbf{z}}(\mathbf{w}) = \sqrt{\frac{\lambda \prod_{i=1}^{d} z_i}{1 + \lambda \prod_{i=1}^{d} z_i}} \sum_{i=1}^{d} \frac{1 - z_i}{z_i} \left(\frac{1 - t_i}{\sqrt{t_i}} + \frac{w_i - 1}{\sqrt{\lambda + \lambda^2}}\right) w_i^c. \tag{3}$$

We show that $g_{\mathbf{z}}(\mathbf{w})$ is monotonically decreasing with w_i for every $i \in [d]$.

Denote $T_i \triangleq \frac{1 - t_i}{\sqrt{t_i}} + \frac{(w_i - 1)}{\sqrt{\lambda + \lambda^2}}$, then

$$\frac{\partial g_{\mathbf{z}}(\mathbf{w})}{\partial w_i} = \sqrt{\frac{\lambda \prod_{i=1}^{d} z_i}{1 + \lambda \prod_{i=1}^{d} z_i}} \cdot \frac{1 - z_i}{z_i} \left(\frac{\partial T_i}{\partial w_i} w_i^c + c w_i^{c-1} T_i\right). \tag{4}$$

The partial derivative (4) is negative for a suitable choice of c:

$$\frac{1 - z_i}{z_i} \cdot \left(\frac{\partial T_i}{\partial z_i} w_i^c + c w_i^{c-1} T_i\right)$$

$$= \frac{1 - z_i}{z_i} \cdot \left(\left(-\frac{1}{2} t_i'(t_i^{-1/2} + t_i^{-3/2}) + \frac{1}{\sqrt{\lambda + \lambda^2}}\right) w_i^c + c w_i^{c-1} \left(\frac{1 - t_i}{\sqrt{t_i}} + \frac{(w_i - 1)}{\sqrt{\lambda + \lambda^2}}\right)\right)$$

$$= \frac{1 - z_i}{z_i} \cdot w_i^{c-1} \left(\left(-\frac{1}{2} \log\left(\frac{1+\lambda}{\lambda}\right)(t_i^{1/2} + t_i^{-1/2}) + \frac{1}{\sqrt{\lambda + \lambda^2}}\right) w_i\right.$$
$$\left. + c\left(t_i^{-1/2} - t_i^{1/2} + \frac{(w_i - 1)}{\sqrt{\lambda + \lambda^2}}\right)\right)$$

$$= \frac{1 - z_i}{z_i} \cdot w_i^{c-1} \left(\frac{(c+1)w_i - c}{\sqrt{\lambda + \lambda^2}} - \right.$$
$$\left. \left(t_i^{1/2}\left(\frac{1}{2} w_i \log\left(\frac{1+\lambda}{\lambda}\right) + c\right) + t_i^{-1/2}\left(\frac{1}{2} w_i \log\left(\frac{1+\lambda}{\lambda}\right) - c\right)\right)\right)$$

Denote

$$p(t,w) \triangleq \frac{(c+1)w-c}{\sqrt{\lambda+\lambda^2}} - \left(t^{1/2}\left(\frac{1}{2}w\log\left(\frac{1+\lambda}{\lambda}\right)+c\right) + t^{-1/2}\left(\frac{1}{2}w\log\left(\frac{1+\lambda}{\lambda}\right)-c\right)\right)$$

Since $c \leq \frac{\log(1+\lambda)-\log\lambda}{2+4\lambda}$, the term

$$t^{1/2}\left(\frac{1}{2}w\log\left(\frac{1+\lambda}{\lambda}\right)+c\right) + t^{-1/2}\left(\frac{1}{2}w\log\left(\frac{1+\lambda}{\lambda}\right)-c\right)$$

achieves its minimum at $t=\frac{\lambda}{1+\lambda}$. Thus

$$p(t,w) \leq p\left(\frac{\lambda}{1+\lambda},w\right) = \left(\frac{\lambda}{1+\lambda}\right)^{1/2}\left(\frac{c+1}{\lambda} - \frac{2\lambda+1}{2\lambda}\log\left(\frac{1+\lambda}{\lambda}\right)\right)w.$$

Moreover, $c < \frac{2\lambda+1}{2}\log\left(\frac{1+\lambda}{\lambda}\right) - 1$ implies that $\frac{c+1}{\lambda} < \frac{2\lambda+1}{2\lambda}\log\left(\frac{1+\lambda}{\lambda}\right)$ holds, which consequently leads to $p\left(\frac{\lambda}{1+\lambda},1\right) < 0$.

In all, we choose a positive constant $c < \min\left\{\frac{\log(1+\lambda)-\log\lambda}{2+4\lambda}, \frac{2\lambda+1}{2}\log\left(\frac{1+\lambda}{\lambda}\right)-1\right\}$, and this results in $p\left(\frac{\lambda}{1+\lambda},w\right) \leq p\left(\frac{\lambda}{1+\lambda},1\right) < 0$.

In light of the monotonicity of w_i's, for every fixed $\mathbf{z}$, $g_{\mathbf{z}}(\mathbf{w})$ achieves its maximum when $\mathbf{w}=\mathbf{1}$. Thus

$$\max_{\mathbf{t}\in\left[0,\frac{\lambda}{1+\lambda}\right]^d} g(\mathbf{t},\mathbf{w}) = \max_{\substack{\mathbf{z}=(z_1,\ldots,z_d) \\ \forall i\in[d], z_i\in\left[1-\left(\frac{\lambda}{1+\lambda}\right)^{w_i-1},1\right]}} g_{\mathbf{z}}(\mathbf{w})$$

$$\leq \max_{\substack{\mathbf{z}=(z_1,\ldots,z_d) \\ \forall i\in[d], z_i\in\left[1-\left(\frac{\lambda}{1+\lambda}\right)^{w_i-1},1\right]}} g_{\mathbf{z}}(\mathbf{1}) \leq \max_{\mathbf{z}\in[0,1]^d} g_{\mathbf{z}}(\mathbf{1}).$$

Actually, the case that all w_i's are 1 corresponds to counting weighted independent sets on normal graphs and arguments to bound $g_{\mathbf{z}}(\mathbf{1})$ can be found in [14]. For the sake of completeness, we give a proof of $g_{\mathbf{z}}(\mathbf{1}) \leq \alpha < 1$ (Lemma 9) below.   ◄

▶ **Lemma 9.** *Let $\Delta > 1$, be a constant. Assume $\lambda < \lambda_c = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta}$ be a constant and $d < \Delta$. Then for some constant $\alpha < 1$, $g_{\mathbf{z}}(\mathbf{1}) \leq \alpha < 1$ where $\mathbf{z} = (z_1,\ldots,z_d) \in [0,1]^d$.*

Proof. Let $\lambda'_c \triangleq \frac{d^d}{(d-1)^{d+1}}$ be the uniqueness threshold for the d-ary tree. Then $\lambda < \lambda_c \leq \lambda'_c$.

Plugging $\mathbf{w}=\mathbf{1}$ into (3), we have

$$g_{\mathbf{z}}(\mathbf{1}) = \sqrt{\frac{\lambda\prod_{i=1}^d z_i}{1+\lambda\prod_{i=1}^d z_i}} \cdot \sum_{i=1}^d \sqrt{1-z_i}.$$

Let $z = \left(\prod_{i=1}^d z_i\right)^{\frac{1}{d}}$, it follows from Jensen's inequality that

$$g(\mathbf{t},\mathbf{1}) \leq d\sqrt{\frac{\lambda z^d(1-z)}{1+\lambda z^d}} < d\sqrt{\frac{\lambda'_c z^d(1-z)}{1+\lambda'_c z^d}} \tag{5}$$

Recall that $f_{\lambda,d}(x) = \lambda\left(\frac{1}{1+x}\right)^d$. Let $\hat{x}$ be the positive fixed-point of $f_{\lambda'_c,d}(x)$ and $\hat{z} = \frac{1}{1+\hat{x}}$. We show that $d\sqrt{\frac{\lambda'_c z^d(1-z)}{1+\lambda'_c z^d}}$ achieves its maximum when $z = \hat{z}$. The derivative of $\frac{\lambda'_c z^d(1-z)}{1+\lambda'_c z^d}$ with respect to z is

$$\left(\frac{\lambda'_c z^d(1-z)}{1+\lambda'_c z^d}\right)' = -\frac{\lambda'_c z^{d-1}}{(1+\lambda'_c z^d)^2}\left(z + \lambda'_c z^{d+1} - d(1-z)\right).$$

Since $\lambda'_c = \frac{d^d}{(d-1)^{d+1}}$, the above achieves maximum at $\tilde{z} = \frac{d-1}{d}$. If we let $\tilde{x} = \frac{1-\tilde{z}}{\tilde{z}}$, then it is easy to verify that $f_{\lambda'_c, d}(\tilde{x}) = \tilde{x}$, which implies $\hat{z} = \tilde{z}$ because of the uniqueness of the positive fixed-point.

Therefore, we have for some $\alpha < 1$,

$$g_{\mathbf{z}}(\mathbf{1}) \leq \alpha < d\sqrt{\frac{\lambda'_c \hat{z}^d(1-\hat{z})}{1 + \lambda'_c \hat{z}^d}} = \left| f'_{\lambda'_c, d}(\hat{x}) \right| = 1. \qquad \blacktriangleleft$$

We are now ready to prove the main lemma.

Proof of Lemma 6. Let $\Phi(x) = \frac{1}{\sqrt{x(1+x)}}$ and $\phi(x) = \int \Phi(x)\,\mathrm{d}x = 2\sinh^{-1}(\sqrt{x})$. We first apply induction on $\ell \triangleq \max\{0, L\}$ to show that, if $d < \Delta$, then $|\phi(R_v) - \phi(R(G, v, L))| \leq 2\sqrt{\lambda}\alpha^L$ for some constant $\alpha < 1$.

If $\ell = 0$, note that $R_v \in [0, \lambda]$, thus $|\phi(R(G, v, L)) - \phi(R_v)| \leq 2\sqrt{\lambda}$. We now assume $L = \ell > 0$ and the lemma holds for smaller ℓ. For every $i \in [d]$ and $j \in [w_i]$, we denote $x_{ij} = R_{ij}$ and $\hat{x}_{ij} = R(G_{ij}, v_{ij}, L - \lfloor 1 + c\log_{1/\alpha} w_i\rfloor)$. Let $\mathbf{x} = (x_{ij})_{i\in[d], j\in[w_i]}$, $\hat{\mathbf{x}} = (\hat{x}_{ij})_{i\in[d], j\in[w_i]}$. Let $f = \lambda \prod_{i=1}^{d}\left(1 - \prod_{j=1}^{w_i} \frac{x_{ij}}{1-x_{ij}}\right)$, then it follows from Proposition 7 that for some $\tilde{\mathbf{x}} = (\tilde{x}_{ij})_{i\in[d], j\in[w_i]}$ with each $\tilde{x}_{ij} \in [0, \lambda]$

$$\begin{aligned}
|\phi(R_v) - \phi(R(G, v, L))| &\leq \sum_{i=1}^{d}\sum_{j=1}^{w_i} \frac{\Phi(f)}{\Phi(\tilde{x}_{ij})}\left|\frac{\partial f(\tilde{\mathbf{x}})}{\partial x_{ij}}\right| \cdot |\phi(x_{ij}) - \phi(\hat{x}_{ij})| \\
&\overset{(\spadesuit)}{\leq} 2\sqrt{\lambda}\sum_{i=1}^{d}\sum_{j=1}^{w_i} \frac{\Phi(f)}{\Phi(\tilde{x}_{ij})}\left|\frac{\partial f(\tilde{\mathbf{x}})}{\partial x_{ij}}\right|\alpha^{L - \lfloor 1 + c\log_{1/\alpha} w_i\rfloor} \\
&\overset{(\heartsuit)}{\leq} 2\sqrt{\lambda}\alpha^L.
\end{aligned}$$

($\spadesuit$) follows from the induction hypothesis and ($\heartsuit$) is due to Lemma 8.

The case that $d = \Delta$ can only happen at the root of our computational tree. Following the arguments in the proofs of 8, 9 and the bound in (5), it is easy to see that a universal constant upper bound for the error contraction exists, i.e.,

$$\sum_{i=1}^{\Delta}\sum_{j=1}^{w_i} \frac{\Phi(f)}{\Phi(\tilde{x}_{ij})}\left|\frac{\partial f(\tilde{\mathbf{x}})}{\partial x_{ij}}\right| < \max_{z\in[0,1]}\sqrt{\frac{\Delta^2(\Delta-1)^{\Delta-1}z^\Delta(1-z)}{(\Delta-2)^\Delta + (\Delta-1)^{\Delta-1}z^\Delta}} < 3.$$

Thus $|\phi(R_v) - \phi(R(G, v, L))| \leq 6\sqrt{\lambda}\alpha^L$ for every v.

Then the lemma follows from Proposition 7, since

$$\begin{aligned}
|R_v - R(G, v, L)| &= \frac{1}{\Phi(\tilde{x})} \cdot |\phi(R_v) - \phi(R(G, v, L))| &&\text{for some } \tilde{x} \in [0, \lambda] \\
&\leq 6\lambda\sqrt{1+\lambda} \cdot \alpha^L &&\blacktriangleleft
\end{aligned}$$

References

1 Antar Bandyopadhyay and David Gamarnik. Counting without sampling: Asymptotics of the log-partition function for certain statistical physics models. *Random Structures & Algorithms*, 33(4):452–479, 2008.

2 Magnus Bordewich, Martin Dyer, and Marek Karpinski. Path coupling using stopping times and counting independent sets and colorings in hypergraphs. *Random Structures & Algorithms*, 32(3):375–399, 2008. `doi:10.1002/rsa.20204`.

3 Martin Dyer, Mark Jerrum, and Eric Vigoda. Rapidly mixing Markov chains for dismantleable constraint graphs. In *Randomization and Approximation Techniques in Computer Science*, pages 68–77. Springer, 2002.

4 Martin E. Dyer, Alan M. Frieze, and Mark Jerrum. On counting independent sets in sparse graphs. *SIAM Jounal on Computing*, 31(5):1527–1541, 2002. URL: `http://epubs.siam.org/sam-bin/dbq/article/38384`.

5 Martin E. Dyer and Catherine S. Greenhill. On Markov chains for independent sets. *Journal of Algorithms*, 35(1):17–49, 2000.

6 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.

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

8 Leslie Ann Goldberg and Mark Jerrum. A polynomial-time algorithm for estimating the partition function of the ferromagnetic Ising model on a regular matroid. *SIAM Journal on Computing*, 42(3):1132–1157, 2013.

9 Mark 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.

10 Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on Computing*, 22(5):1087–1116, 1993.

11 Mark Jerrum and Alistair Sinclair. The Markov chain monte carlo method: an approach to approximate counting and integration. *Approximation algorithms for NP-hard problems*, pages 482–520, 1996.

12 Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *Journal of the ACM*, 51:671–697, July 2004. `doi:10.1145/1008731.1008738`.

13 Liang Li, Pinyan Lu, and Yitong Yin. Approximate counting via correlation decay in spin systems. In *Proceedings of the 23th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'12)*, pages 922–940. SIAM, 2012.

14 Liang Li, Pinyan Lu, and Yitong Yin. Correlation decay up to uniqueness in spin systems. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'13)*, pages 67–84. SIAM, 2013.

15 Chengyu Lin, Jingcheng Liu, and Pinyan Lu. A simple FPTAS for counting edge covers. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'14)*, pages 341–348, 2014. `doi:10.1137/1.9781611973402.25`.

16 Jingcheng Liu and Pinyan Lu. FPTAS for #BIS with degree bounds on one side. In *Proceedings of the 47th Annual ACM Symposium on Theory of Computing (STOC'15)*, 2015.

17 Jingcheng Liu and Pinyan Lu. FPTAS for counting monotone CNF. In *Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'15)*, pages 1531–1548, 2015.

18 Jingcheng Liu, Pinyan Lu, and Chihao Zhang. FPTAS for counting weighted edge covers. In *In Proceedings of the 22nd European Symposium on Algorithms (ESA'14)*, pages 654–665, 2014.

19 Pinyan Lu, Menghui Wang, and Chihao Zhang. FPTAS for weighted Fibonacci gates and its applications. *Proceedings of the 41st International Colloquium on Automata, Languages and Programming (ICALP'14)*, pages 787–799, 2014.

20 Pinyan Lu and Yitong Yin. Improved FPTAS for multi-spin systems. In *Proceedings of APPROX-RANDOM*, pages 639–654. Springer, 2013.

21 Michael Luby and Eric Vigoda. Approximately counting up to four. In *Proceedings of the 29th Annual ACM Symposium on Theory of computing (STOC'97)*, pages 682–687. ACM, 1997.

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

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

24 Allan Sly and Nike Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS'12)*, pages 361–369. IEEE, 2012.

25 Eric Vigoda. Improved bounds for sampling colorings. *Journal of Mathematical Physics*, 41(3):1555–1569, 2000.

26 Dror Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC'06)*, pages 140–149. ACM, 2006.

27 Yitong Yin and Jinman Zhao. Counting hypergraph matchings up to uniqueness threshold. *arXiv preprint arXiv:1503.05812*, 2015.