

AN FPRAS FOR ANTIFERROMAGNETIC ISING MODELS ON RANDOM REGULAR BIPARTITE GRAPHS

ZHIDAN LI, KUAN YANG

ABSTRACT. We design randomized approximation schemes for the partition function of antiferromagnetic Ising models with uniform external field on random regular bipartite graphs. Our algorithm generalizes the approach of Kocurek, Oveis Gharan and Tjowasi (arXiv, 2026) for hardcore models on the same random graph model beyond the uniqueness threshold. We show that, as long as λ is upper bounded by a constant and $\lambda(1 - \beta) \lesssim \Delta^{-1/2}$, an efficient randomized algorithm approximates the partition function with high probability. The algorithm first truncates configurations that are large on either side of the bipartition and then samples from Gibbs distributions conditioned on fixed sizes on one or both sides. To choose an optimal truncation bound, we establish concentration properties of the Gibbs distribution on random regular bipartite graphs. Then we apply high-dimensional expansion and prove trickle-down theorems to obtain fast samplers for the conditioned distributions.

1. INTRODUCTION

The *(Lenz-)Ising model* is one of the most fundamental models in statistical physics, serving as a canonical framework for studying phase transitions, Gibbs measures, and spin correlations. From the algorithmic perspective, approximating the partition function of the Ising model is a central problem in approximate counting, with applications ranging from statistical inference to sampling from Gibbs distributions. We study the Ising model in its subgraph representation, in which configurations are subsets of vertices. This representation is equivalent to the usual spin formulation on regular graphs, and is particularly convenient for bipartite graphs because the configuration can be naturally decomposed according to the two sides of the bipartition. Given a graph $G = (V, E)$ and two parameters $\lambda > 0$ and $\beta \geq 0$, the Ising model with uniform external field λ and edge interaction β on G consists of a finite state space (or configuration space) $\Omega = 2^{V(G)}$ and a weight function $w = w_{G;\lambda,\beta} : \Omega \rightarrow \mathbb{R}$ inducing a partition function $\mathcal{Z}_G(\lambda, \beta)$ and its Gibbs distribution $\mu = \mu_{G;\lambda,\beta}$. A configuration S in Ω is a subset of vertices V , and its weight is defined by

$$(1) \quad w(S) = \lambda^{|S|} \beta^{|E(G[S])|}$$

where $G[S]$ is the induced subgraph of G by S . The *partition function* of the model is the total weight of the system:

$$\mathcal{Z}_G(\lambda, \beta) = \sum_{S \subseteq V(G)} w(S),$$

which is the normalizing factor of the induced Gibbs distribution $\mu_{G;\lambda,\beta}$ on Ω :

$$\forall S \subseteq V(G), \quad \mu(S) = \frac{w(S)}{\mathcal{Z}_G(\lambda, \beta)}.$$

(Zhidan Li) SCHOOL OF COMPUTER SCIENCE, SHANGHAI JIAO TONG UNIVERSITY, SHANGHAI, CHINA. Email: yueEnTeRnAL@sjtu.edu.cn.

(Kuan Yang) JOHN HOPCROFT CENTER FOR COMPUTER SCIENCE, SHANGHAI JIAO TONG UNIVERSITY, SHANGHAI, CHINA. Email: kuan.yang@sjtu.edu.cn.

When $\beta > 1$, the model is *ferromagnetic*; when $\beta < 1$, it is *antiferromagnetic*. When the underlying graph G is regular, this subgraph representation is equivalent, up to a change of parameters, to the usual spin-system formulation; see [Appendix A](#).

In this work, we focus on antiferromagnetic Ising models. It is a significant problem to estimate their partition functions, because the partition function determines the free energy and approximate counting is closely related to sampling from Gibbs distributions and estimating thermodynamic quantities. Unfortunately, exact computation is $\#\mathbf{P}$ -hard even on triangle-free graphs, since counting independent sets ($\#\mathbf{IS}$), which can be viewed as a special case ($\lambda = 1$ and $\beta = 0$) of anti-ferromagnetic Ising models, is $\#\mathbf{P}$ -hard [\[Gre00\]](#). This motivates the search for efficient approximation schemes for $\mathcal{Z}_G(\lambda, \beta)$.

In the study of approximate counting, the uniqueness threshold of the Gibbs measure on infinite Δ -regular trees has been shown to characterize the algorithmic threshold for many important classes of spin systems. On the algorithmic side, seminal works [\[Wei06, LLY13, CCYZ25, CJM⁺26\]](#) show that, in the uniqueness regime or at the threshold, there exists a (randomized) polynomial-time algorithm to approximate $\mathcal{Z}_G(\lambda, \beta)$. In contrast, Sly and Sun [\[SS12\]](#), and Galanis, Štefankovič and Vigoda [\[GŠV16\]](#) show that unless $\mathbf{NP} = \mathbf{RP}$, there is no $\mathbf{FPRAS}$ for approximating the partition functions for general bounded-degree graphs in the non-uniqueness regime. This dichotomy suggests that the tree uniqueness threshold captures an important boundary in the algorithmic complexity of spin systems. However, for bipartite graphs, the situation is different. The problem of (approximately) counting independent sets on a bipartite graph ($\#\mathbf{BIS}$) is widely regarded as an intermediate approximate counting problem between $\mathbf{P}$ and $\mathbf{NP}$; currently no computational hardness results are known, nor is there an approximation algorithm that works for general cases. In addition, Cai et al. [\[CGG⁺16\]](#) shows that approximating the partition function of an antiferromagnetic two-spin system in the non-uniqueness regime is $\#\mathbf{BIS}$ -hard.

Despite the lack of approximation schemes on general graphs in the tree non-uniqueness threshold, recent studies has shown that random bipartite graph models provide an important intermediate setting between trees and worst-case graphs. They retain nontrivial global dependence structure responsible for phase transitions, while their expansion properties allow additional algorithmic techniques. Understanding spin systems on such graphs may reveal whether the tree uniqueness threshold remains an algorithmic barrier in typical instances. For example, approximation schemes beyond the tree uniqueness threshold have recently been obtained for the hard-core model on random regular bipartite graphs: Kocurek, Oveis Gharan and Tjowasi [\[KOGT26\]](#) give an $\mathbf{FPRAS}$ whenever $\lambda = O(\Delta^{-1/2})$. Together with [\[JKP20\]](#), these works give approximation schemes for the hard-core model for all fugacities on random regular bipartite graphs. Meanwhile, for antiferromagnetic Ising models, Geisler et al. [\[GKS⁺26\]](#) show that, despite torpid mixing of Glauber dynamics, an $\mathbf{FPTAS}$ exists for $\mathcal{Z}_G(\lambda, \beta)$ on a family of regular bipartite expanders that includes almost every regular bipartite graph, in a regime far from the uniqueness threshold. These algorithmic results motivate the following question:

Does the antiferromagnetic Ising model on a random regular bipartite graph admit an efficient deterministic or randomized approximation scheme for all uniform external fields and all edge interactions, with high probability?

We partially resolve this question by establishing an $\mathbf{FPRAS}$ in a weak-interaction regime that extends beyond the tree uniqueness threshold. Our main result is the following theorem.

Theorem 1.1. *There exist a positive integer $\Delta_0 \geq 3$ and two positive constants $\lambda_0, C > 0$ such that for every positive integer $\Delta \geq \Delta_0$, real numbers $\lambda > 0$ and $\beta \in [0, 1]$ such that*

$$\lambda \leq \lambda_0, \quad \lambda(1 - \beta) \leq \frac{C}{\sqrt{\Delta}},$$

the following holds with high probability¹ over the random Δ -regular bipartite $2n$ -vertex graph G for all sufficiently large n . There is an **FPRAS** for the partition function of Ising models on G with uniform external field λ at edge interaction β .

We emphasize that the boundedness assumption on λ is essential in our current analysis, since it plays a crucial role to establish the concentration properties of the configuration space. Thus, the theorem concerns a weak-interaction regime rather than the full parameter space.

1.1. Overview of techniques. The main difficulty is that Glauber dynamics may mix slowly beyond the uniqueness threshold. Rather than analyzing local dynamics on the full configuration space directly, we exploit the geometry of suitable conditioned configuration spaces, decompose the Gibbs distribution into a collection of structured slices and prove rapid mixing within each slice; details appear in [Sections 2](#) and [3](#).

Concentration phenomenon. A central geometric property of Ω is *concentration*. We establish two kinds of concentrations: (a) the one-side concentration inequality, and (b) the concentration for *degree profiles*. The one-side concentration inequality gives exponential decay for the number of selected vertices on either side, and the other concentration provides a demonstration of the degree profiles after fixing a vertex subset on one side of size at most half of that side. Together, these estimates allow us to truncate large configurations and reduce the problem to small or balanced ‘slices’ on both sides, while retaining an accurate approximation of $\mathcal{Z}_G(\lambda, \beta)$.

High-dimensional expanders. High-dimensional expanders provide an efficient framework to design and analyze **FPAUSes** for counting problems. Seminal works [[AL20](#), [Opp18](#)] develop tools for deriving rapid mixing of *down-up walks* from ‘local’ properties of the Gibbs distribution.

We apply this framework to analyze the down-up walks for the Gibbs distribution conditional on specific slices. To establish local spectral expansion, we prove trickle-down theorems analogous to those in [[KOGT26](#)], with modifications required for the soft constraints of the antiferromagnetic Ising model. However, extending these techniques is nontrivial. Unlike the hard-core model, which imposes a hard constraint forbidding adjacent occupied vertices, the antiferromagnetic Ising model allows configurations containing adjacent occupied vertices but receiving smaller weights. Consequently, the correlations inside the Gibbs distribution are more subtle, and the techniques developed for hard constraints cannot be directly applied.

1.2. Related works and discussion. Our result gives an **FPRAS** when both λ and $\lambda(1 - \beta)\sqrt{\Delta}$ are upper bounded by constants; this regime includes a portion of the tree non-uniqueness region. In contrast, [[GKSW26](#)] provides an **FPTAS** in a different non-uniqueness regime, when λ is upper bounded and $\lambda(1 - \beta) = \Omega((\ln \Delta)^{3/2} / \Delta^{\min\{1, 1 - \kappa/2\}})$, via abstract polymer models and cluster expansion approach. A gap remains between these parameter regimes. It also remains open whether the upper bound on the vertex fugacity λ is inherent or merely a limitation of the current methods.

To analyze the configuration space, our analysis adapts the *independent-set slices* of [[KOGT26](#)] to this soft-constrained setting. The concentration estimates identify the truncation parameters, while high-dimensional expanders are used to prove rapid mixing of down-up walks on slices of prescribed sizes. Similar local-to-global arguments have been used to sample fixed-size independent sets [[AL20](#), [JMPV23](#)], sample hard-core and antiferromagnetic two-spin systems on random regular graphs [[CCC⁺25](#)], and approximately count the hard-core models on random regular bipartite graphs [[KOGT26](#)].

¹In this paper, when we say that an event occurs ‘with high probability’, we mean that it occurs with probability $1 - o(1)$.

1.3. Organization of this paper. We introduce necessary preliminaries in [Section 2](#). In [Section 3](#), we establish some properties of the Gibbs distribution of antiferromagnetic Ising models on random regular bipartite graphs. Proofs and analysis of trickle-down theorems of high-dimensional expanders are provided in [Section 4](#). Finally, we develop our approximation scheme in [Section 5](#).

Statement on AI use. We apply ChatGPT 5.6 Pro Sol to optimize the choice of parameters in our algorithm. The concentration inequalities are provided by ChatGPT 5.6 Pro Sol after we give our drafts and related materials to it. All calculations in this work are verified by us and we take full responsibility of this paper.

2. PRELIMINARIES

2.1. Notations. We use e to denote the natural logarithm base and $\ln$ to denote the natural logarithm function $\log_e$. For a natural number n , we use $[n]$ to denote the set $\{1, 2, \dots, n\}$.

For an operator/matrix A , we use $A^\top$ to denote its transpose. The spectral norm of A is defined by $\|A\|_2 := \max_{\|x\|_2=\|y\|_2=1} x^\top A y$. We use $\mathbf{1}_S$ to denote the all-one vector on S and $\mathbf{I}_S$ to denote the identity matrix on S (when context is clear, we simply use $\mathbf{1}$ and $\mathbf{I}$). We use $\lambda_i(A)$ to denote the i -th largest eigenvalue of A and $\lambda_{\max}(A), \lambda_{\min}(A)$ to denote the largest and smallest eigenvalue respectively.

For a random variable X , we denote by $\text{Law}(X)$ its law. The expectation of X is denoted by $\mathbf{E}[X]$ and the variance of X is denoted by $\text{Var}(X)$. We use $M_X(\cdot)$ to denote the moment generating function of X , i.e., $M_X(t) = \mathbf{E}[e^{tX}]$. The entropy function $H(X)$ is defined as $H(X) := -\sum_x \mathbb{P}(X=x) \ln \mathbb{P}(X=x)$ with convention $0 \ln 0 = 0$. With abuse of notations, we use $H(p)$ to denote the relative entropy function of a Bernoulli random variable with rate p .

For two probability distributions P and Q supported a finite state space Ω , the Kullback-Leibler divergence $\mathcal{D}_{\text{KL}}(P \parallel Q)$ is defined by

$$\mathcal{D}_{\text{KL}}(P \parallel Q) := \sum_{x \in \Omega} P(x) \ln \frac{P(x)}{Q(x)},$$

with convention $0 \ln(0/0) = 0$ and $p \ln(p/0) = \infty$ for $p > 0$. For two real numbers $p, q \in [0, 1]$, we simply use $\mathcal{D}_{\text{KL}}(p \parallel q)$ to denote the KL-divergence between two Bernoulli random variables with mean p and q , i.e., $\mathcal{D}_{\text{KL}}(p \parallel q) := \mathcal{D}_{\text{KL}}(\text{Ber}(p) \parallel \text{Ber}(q))$.

The graph $G = (V(G), E(G))$ consists of a vertex set $V(G)$ and an edge set $E(G) \subseteq V(G) \times V(G)$. The bipartite graph is denoted by $G = (V(G) = V_L(G) \cup V_R(G), E(G))$ where $V_L(G) \cap V_R(G) = \emptyset$ and $E(G) \subseteq V_L(G) \times V_R(G)$. When the context is clear, we simply use V, V_L, V_R and E to denote the graph. Given a graph $G = (V, E)$, let $E_G(S, T)$ be the number of edges between S and T . For a vertex subset $S \subseteq V$, let $G[S]$ be the subgraph of G induced by S . Moreover, we define the degree function $d_S(\cdot)$ on V as $d_S(v) := |E_G(\{v\}, S)|$ the number of edges between v and S .

2.2. Useful inequalities. We state some useful inequalities here.

Lemma 2.1 (Markov inequality). *For a non-negative random variable X and $a > 0$, it holds that*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbf{E}[X]}{a}.$$

Moreover, if it always holds that $X \geq 1$, then $\mathbb{P}(X > 0) \leq \mathbf{E}[X]$.

Using Markov inequality with the moment generating function, we can derive the following version of Chernoff bound (for details, refer [\[Ger25\]](#)).

Lemma 2.2 (Chernoff bound). *For a non-negative random variable X and $a, t > 0$, it holds that*

$$\mathbb{P}(X \geq a) \leq M_X(t) e^{-ta}.$$

Moreover, if $X = \sum_{i=1}^n X_i$ where each X_i is an independent Bernoulli random variable with rate p , for every $p < \rho < 1$, it holds that

$$\mathbb{P}(X \geq \rho n) \leq \exp(-n\mathcal{D}_{\text{KL}}(\rho \parallel p)) = \left(\left(\frac{p}{\rho} \right)^\rho \left(\frac{1-p}{1-\rho} \right)^{1-\rho} \right)^n.$$

The Popoviciu's inequality on variances in [Pop35] gives an upper bound on the variance with respect to the range of the random variable.

Proposition 2.3 (Popoviciu's inequality on variances). *For a random variable X , it always holds that*

$$\text{Var}(X) \leq \frac{1}{4}(M - m)^2$$

where M is the possible maximum of X and m is the possible minimum of X .

The following concentration inequality for martingales with bounded increments, which is a special version of Bernstein's inequality and known as *Freedman's inequality*, plays a crucial role in our analysis. For details, we recommend readers to refer [Fre75, DvZ01].

Lemma 2.4 (Freedman's inequality). *Let $\{(Z_i, \mathcal{F}_i)\}_{i=0}^n$ be a martingale with bounded increments ($|Z_i - Z_{i-1}| \leq B$ almost surely) and set*

$$V_n = \sum_{i=1}^n \mathbf{E}[(Z_i - Z_{i-1})^2 \mid \mathcal{F}_{i-1}].$$

Then for every $t, v > 0$, it holds that

$$\mathbb{P}(|Z_n - Z_0| \geq t, V_n \leq v) \leq 2 \exp\left(-\frac{t^2}{2(v + Bt/3)}\right).$$

2.3. Markov chains and mixing time. We state here some basic definitions and lemmas for Markov chains and for more details, please see [LP17].

For a state space Ω , a Markov chain M consists of a transition kernel P on Ω with its stationary distribution π . For brevity, we sometimes use the transition matrix P to denote the Markov chain M . In this work, we only consider *ergodic* and *reversible* Markov chains (see details in [LP17]).

We use standard notation for total variation distance and mixing time. For two probability distributions μ, ν on Ω , the *total variation distance* is defined as

$$\|\mu - \nu\|_{\text{TV}} := \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)| = \max_{\Lambda \subseteq \Omega} |\mu(\Lambda) - \nu(\Lambda)|.$$

For a positive real $\varepsilon \in (0, 1)$, define the *mixing time* of M with error ε as

$$\tau_{\text{mix}}(M, \varepsilon) := \inf \{t \in \mathbb{N} : \forall x \in \Omega, \|P^t(x, \cdot) - \pi\|_{\text{TV}} \leq \varepsilon\}.$$

It is a route way to bound the mixing time by the spectral gap of P . For a random walk P , set $\lambda_*(P) = \max\{\lambda_2(P), |\lambda_{\min}(P)|\}$. The following fact is standard.

Fact 2.5. *For a Markov chain P with respect to μ , it holds that*

$$\tau_{\text{mix}}(P, \varepsilon) \leq O\left(\frac{1}{1 - \lambda_*(P)} \left(\ln \frac{1}{\mu_{\min}} + \ln \varepsilon^{-1}\right)\right)$$

where $\mu_{\min} := \min_{x \in \Omega: \mu(x) > 0} \mu(x)$. Consequently, for its lazy version $\frac{P+I}{2}$, it holds that

$$\tau_{\text{mix}}\left(\frac{P+I}{2}, \varepsilon\right) \leq O\left(\frac{1}{1 - \lambda_2(P)} \left(\ln \frac{1}{\mu_{\min}} + \ln \varepsilon^{-1}\right)\right).$$

2.4. Approximate algorithms. We formally define the concept of approximate algorithms. For a quantity $Z \in \mathbb{R}$ and a tolerance $\varepsilon \in (0, 1)$, we say $\widehat{Z}$ is an ε -approximation to Z if $(1 - \varepsilon)Z \leq \widehat{Z} \leq (1 + \varepsilon)Z$.

Definition 2.6 (FPTAS and FPRAS). For a function Z from a family of instances to reals, a *fully polynomial-time approximation scheme* (**FPTAS**) is a deterministic algorithm such that, for every instance Φ and every tolerance $\varepsilon \in (0, 1)$, it outputs an ε -approximation to $Z(\Phi)$ in time $\text{poly}(|\Phi|, \varepsilon^{-1})$. Similarly, a *fully polynomial-time randomized approximation scheme* (**FPRAS**) is a randomized algorithm such that, for every instance Φ and every tolerance $\varepsilon \in (0, 1)$, it outputs an ε -approximation to $Z(\Phi)$ in time $\text{poly}(|\Phi|, \varepsilon^{-1})$ with probability at least $3/4$.

Definition 2.7 (FPAUS). Fix an input Φ and an induced state space $\Omega = \Omega(\Phi)$. For a non-zero weight function $w : \Omega \rightarrow \mathbb{R}_{\geq 0}$, consider the distribution μ on Ω defined as $\mu(x) \propto w(x)$. A *fully polynomial-time almost-uniform sampler* (**FPAUS**) for μ outputs a sample X from Ω randomly provided a tolerance $\varepsilon \in (0, 1)$ such that $\|\text{Law}(X) - \mu\|_{\text{TV}} \leq \varepsilon$ and the running time is $\text{poly}(|\Phi|, \varepsilon^{-1})$.

For a family of self-reducible problems, [JVV86] shows that an **FPAUS** can induce an **FPRAS**. For more details, refer [JVV86].

2.5. High-dimensional expander. Given a finite ground set U , a *d-dimensional pure simplicial complex* $(\mathcal{X}, \pi)$ is a downward closed collection of subsets of U with all maximal subsets having size d and equipped with a probability distribution π over all maximal subsets. We call a subset in $\mathcal{X}$ a *face* or *facet*, and for every facet $S \in \mathcal{X}$, its dimension $\dim(S)$ is defined as its cardinality $|S|$ and the co-dimension $\text{codim}(S)$ is defined by $\text{codim}(S) = d - \dim(S)$. For every integer $k = 0, 1, \dots, d$, let $\mathcal{X}(k)$ be the collection of facets of dimension k , *i.e.*,

$$\mathcal{X}(k) := \{S \in \mathcal{X} \mid \dim(S) = k\}.$$

For each $k = 0, 1, \dots, d$, we induce a probability distribution π_k according to the marginal distribution of π on $\mathcal{X}(k)$ as

$$(2) \quad \forall S \in \mathcal{X}(k), \quad \pi_k(S) = \frac{1}{\binom{d}{k}} \sum_{T \in \mathcal{X}(d): S \subseteq T} \pi(T).$$

Based on $\{\pi_k\}_{k=0}^d$, we define random walks on $\mathcal{X}$. At first, we consider two kinds of operators $P_k^\uparrow \in \mathbb{R}^{\mathcal{X}(k) \times \mathcal{X}(k+1)}$ for $k = 0, \dots, d-1$ and $P_k^\downarrow \in \mathbb{R}^{\mathcal{X}(k) \times \mathcal{X}(k-1)}$ for $k = 1, \dots, d$ defined as

$$P_k^\uparrow(S, T) = \mathbb{1}[T \subseteq S] \cdot \frac{\pi_{k+1}(T)}{(k+1)\pi_k(S)}, \quad P_k^\downarrow(S, T) = \mathbb{1}[T \subseteq S] \frac{1}{k}.$$

Then there are two natural random walks on the simplicial complexes $\mathcal{X}$:

- (1) the up-down walk $P_k^\Delta = P_k^\uparrow P_{k+1}^\downarrow$ for every $k = 0, 1, \dots, d-1$ and its non-lazy version $P_k^\wedge = \frac{k+2}{k+1} P_k^\Delta - \frac{1}{k+1} \mathbf{I}_{\mathcal{X}(k)}$;
- (2) the down-up walk $P_k^\nabla = P_k^\downarrow P_{k-1}^\uparrow$ for every $k = 1, \dots, d$ and its non-lazy version $P_k^\vee = \frac{k+1}{k} P_k^\nabla - \frac{1}{k} \mathbf{I}_{\mathcal{X}(k)}$.

It is direct to observe that these random walks are all reversible with respect to π_k .

An important sub-structure of a simplicial complex plays a crucial role is the *links*. For a facet $S \in \mathcal{X}$, we define the link of $\mathcal{X}$ on S as

$$\mathcal{X}_S := \{T \mid T \cap S = \emptyset \wedge S \cup T \in \mathcal{X}\}.$$

It is similar to define the distributions $\pi_{S,k}$ and random walks $P_{S,k}^\uparrow, P_{S,k}^\downarrow, P_{S,k}^\Delta, P_{S,k}^\nabla, P_{S,k}^\wedge$ and $P_{S,k}^\vee$ on $\mathcal{X}_S$ as notations in $\mathcal{X}$.

The concept of *local-spectral expander* plays a crucial role in the analysis of mixing rate of random walks on a simplicial complex.

Definition 2.8 (Local-spectral expander). For a d -dimensional simplicial complex $(\mathcal{X}, \pi)$, we say it is a $(\gamma_0, \dots, \gamma_{d-2})$ -*local-spectral expander* if for every $k = 0, \dots, d-2$ and $S \in \mathcal{X}(k)$, it holds that $\lambda_2(\mathbf{P}_{S,1}^\wedge) \leq \gamma_k$.

When $\mathcal{X}$ is a local-spectral expander, the spectral gap of the down-up walk $\mathbf{P}_d^\Delta$ is bounded by [AL20].

Theorem 2.9 ([AL20]). For a d -dimensional simplicial complex $(\mathcal{X}, \pi)$ which is a $(\gamma_0, \dots, \gamma_{d-2})$ -local spectral expander, it holds that for every $k = 1, \dots, d$,

$$\lambda_2(\mathbf{P}_k^\nabla) \leq 1 - \frac{1}{k} \prod_{i=0}^{k-2} (1 - \gamma_i).$$

The following trickle-down theorem in [Opp18] is the key ingredient.

Lemma 2.10 (Oppenheim's trickle-down theorem [Opp18]). For a d -dimensional pure simplicial complex $(\mathcal{X}, \pi)$ and a facet $S \in \mathcal{X}$, suppose that for every $v \in \mathcal{X}_S(1)$, $\lambda_2(\mathbf{P}_{S \cup \{v\},1}^\wedge) \leq \gamma < 1$ and $\mathbf{P}_{S,1}^\wedge$ is connected. Then it holds that

$$\lambda_2(\mathbf{P}_{S,1}^\wedge) \leq \frac{\gamma}{1 - \gamma}.$$

2.6. Random regular bipartite graphs. In this part, we introduce the model of random regular bipartite graphs. Let $\text{RBG}(n, \Delta)$ be the probabilistic model generating a graph from all Δ -regular $2n$ -vertex bipartite graphs uniformly at random. For simplicity of analysis, we introduce the *pairing model* to study random regular bipartite graphs.

Definition 2.11 (Pairing model). The pairing model generates a bipartite multi-graph $G = (V = V_L \cup V_R, E)$ in the following way:

- the vertex set is set by $V_L = V_R = [n]$;
- to generate E :
 - define $\mathcal{V}_L = V_L \times [\Delta]$ and $\mathcal{V}_R = V_R \times [\Delta]$, and choose a uniformly random perfect matching between $\mathcal{V}_L$ and $\mathcal{V}_R$;
 - for a pair (u, i) and (v, j) in the perfect matching, generate an edge (u, v) in E .

The following proposition builds a connection between $\text{RBG}(n, \Delta)$ and $\text{PM}(n, \Delta)$.

Proposition 2.12 ([Wor99]). For an event $\mathcal{E}$, if $\mathbb{P}_{\text{PM}(n, \Delta)}(\mathcal{E}) = o(1)$, then $\mathbb{P}_{\text{RBG}(n, \Delta)}(\mathcal{E}) = o(1)$.

For a bipartite graph $G = (V = V_L \cup V_R, E)$, define the matrix $A_G \in \mathbb{R}^{V_L \times V_R}$ as

$$A_G(u, v) = \begin{cases} 1 & (u, v) \in E \\ 0 & (u, v) \notin E \end{cases}.$$

We remark here the matrix A_G plays a significant role in the spectral analysis of G , since the adjacent matrix of G can be denoted by

$$\text{Adj}(G) = \begin{pmatrix} 0 & A_G \\ A_G^\top & 0 \end{pmatrix}.$$

Proposition 2.13 ([KOGT26]). With high probability over $G \sim \text{RBG}(n, \Delta)$, $\lambda_2(A_G) \leq 3\sqrt{\Delta}$.

We also use the following result.

Proposition 2.14 ([KOGT26, Lemma 3.7]). With high probability over $G = (V, E) \sim \text{RBG}(n, \Delta)$, for every $u, v \in V$, $|N_G(u) \cap N_G(v)| \leq 2$.

3. ISING MODELS ON RANDOM REGULAR BIPARTITE GRAPHS

In this section, we capture the geometry of the configuration space of Ising models on random regular bipartite graphs. The main challenge beyond the uniqueness threshold is the presence of strong dependencies among configurations. Instead of analyzing the full Gibbs measure directly, we partition the configuration space according to the number of occupied vertices on each side of the bipartite graph. Each part admits better expansion properties, which will later allow us to construct rapidly mixing Markov chains.

To describe the partition of configurations space Ω , we introduce the concept of *slices* in [KOGT26]. For every natural number $s \geq 0$, we say a vertex subset $A \subseteq V_L$ of size s is an L -side slice of size s , and say a vertex subset $B \subseteq V_R$ of size s is an R -side slice of size s . For two natural numbers $\ell, r \geq 0$, an (ℓ, r) -two-side slice S is a vertex subset of V such that $|S \cap V_L| = \ell$, $|S \cap V_R| = r$ and $w(S) > 0$. For simplicity, let $\Omega^{(\ell, r)}$ be the collection of all (ℓ, r) -two-side slices.

We specify slices with their cardinalities. For every natural number $s = 0, \dots, n$, define the one-side partial partition function $\mathcal{Z}_s^L, \mathcal{Z}_s^R$ as

$$\mathcal{Z}_s^L := \sum_{A \subseteq V_L: |A|=s} \sum_{B \subseteq V_R} w(A \cup B), \quad \mathcal{Z}_s^R := \sum_{B \subseteq V_R: |B|=s} \sum_{A \subseteq V_L} w(A \cup B).$$

Similarly, for natural numbers $\ell = 0, 1, \dots, n$ and $r = 0, 1, \dots, n$, define the two-side partial partition function $\mathcal{Z}_{\ell, r}$ as

$$\mathcal{Z}_{\ell, r} := \sum_{S \in \Omega^{(\ell, r)}} w(S).$$

The corresponding Gibbs distribution is denoted by

$$\begin{aligned} \mu_s^L(S) &:= \mu(S \mid S \subseteq V_L, |S| = s), & \mu_s^R(S) &:= \mu(S \mid S \subseteq V_R, |S| = s), \\ \mu_{\ell, r}(S) &:= \mu(S \mid |S \cap V_L| = \ell, |S \cap V_R| = r). \end{aligned}$$

For a vertex subset $A \subseteq V_L$, define

$$W(A) := \sum_{v \in V_R} \beta^{d_A(v)}.$$

Similarly we can extend the definition of W to every vertex subset of V_R .

Before stating properties for $\mu_{G; \lambda, \beta}$ when G is generated from random regular bipartite graphs, we describe the main idea of the approximating algorithm to make our analysis more understandable. To approximate $\mathcal{Z}_G(\lambda, \beta)$, we pick two proper parameters $\alpha, \rho \in (0, 1)$. Then we define the quantity $\widehat{\mathcal{Z}}$ as

$$\widehat{\mathcal{Z}} := \sum_{\ell \leq \alpha n, r \leq \alpha n} \mathcal{Z}_{\ell, r} + \sum_{\alpha n < s \leq \rho n} \mathcal{Z}_s^L + \sum_{\alpha n < s \leq \rho n} \mathcal{Z}_s^R$$

when $\alpha n < \rho n$ and

$$\widehat{\mathcal{Z}} := \sum_{\ell \leq \rho n, r \leq \rho n} \mathcal{Z}_{\ell, r}$$

otherwise. The purpose of the one-side slices is to capture configurations where one side is moderately large while the other side remains small. **Lemma 3.3** shows that configurations with both sides larger than αn have exponentially small contribution.

Remark 3.1. For a parameter $\theta \in (0, 1)$, we simply use θn instead of $\lfloor \theta n \rfloor$. This takes little impact on our analysis but makes our statements much cleaner. For simplicity when we are saying a fraction is in an interval, it means that we take values θ such that θn is an integer.

3.1. One-side tail bound. The first property we use is the *tail bound* of μ , which can be viewed as a generalization of [KOGT26, Lemma 1.11].

Lemma 3.2. *For a bipartite graph $G = (V = V_L \cup V_R, E)$ with $|V_L| = |V_R| = n$ and two real numbers $\lambda > 0, \beta \in [0, 1]$, the Ising model on G at vertex interaction λ and edge interaction β satisfies that for every $\rho \in (0, 1)$,*

$$\mathbb{P}_{S \sim \mu}(|S \cap V_L| > \rho n) \leq \exp(-n \mathcal{D}_{\text{KL}}(\rho \parallel \lambda/(1+\lambda))).$$

The same tail bound holds for V_R as well.

Proof. For a vertex subset $A \subseteq V_L$, since random variables $\{\mathbb{1}[v \in S]\}_{v \in V_R}$ are independent conditioned on $S \cap V_L = A$, it holds that

$$\mathbb{P}(S \cap V_L = A) \propto \lambda^{|A|} \prod_{v \in V_R} (1 + \lambda \beta^{d_A(v)}).$$

Therefore, for every $A \subseteq V_L$ and every $u \in V_L \setminus A$, since $\beta \leq 1$,

$$\frac{\mathbb{P}(S \cap V_L = A \cup \{u\})}{\mathbb{P}(S \cap V_L = A)} = \lambda \prod_{v \in V_R} \frac{1 + \lambda \beta^{d_{A \cup \{u\}}(v)}}{1 + \lambda \beta^{d_A(v)}} \leq \lambda.$$

Then we obtain that the random variable $X = |S \cap V_L|$ is dominated by n independent Bernoulli random variables $\text{Ber}(\lambda/(1+\lambda))$. By the Chernoff bound Lemma 2.2, we conclude the lemma. $\square$

The following lemma gives a proper approximation to $\mathcal{Z}_G(\lambda, \beta)$.

Lemma 3.3. *For two parameters $\alpha, \rho \in (0, 1/2]$, assume that $\rho > \frac{\lambda}{1+\lambda}$ and for every vertex subset $A \subseteq V_L$ of size αn ,*

$$(3) \quad \lambda W(A) \leq (1 - \kappa)|A|.$$

Define the quantity

$$\widehat{Z} = \begin{cases} \sum_{\ell \leq \alpha n, r \leq \alpha n} \mathcal{Z}_{\ell, r} + \sum_{\alpha n < s \leq \rho n} \mathcal{Z}_s^L + \sum_{\alpha n < s \leq \rho n} \mathcal{Z}_s^R & \alpha \leq \rho \\ \sum_{\ell \leq \rho n, r \leq \rho n} \mathcal{Z}_{\ell, r} & \text{otherwise} \end{cases}.$$

Then there is a constant $C = C(\kappa) = -\ln(1 - \kappa) - \kappa$ such that

$$\left| \frac{\widehat{Z}}{\mathcal{Z}_G(\lambda, \beta)} - 1 \right| \leq 2e^{-n \mathcal{D}_{\text{KL}}(\rho \parallel \lambda/(1+\lambda))} + e^{-C\alpha n}.$$

Proof. When $\alpha > \rho$, by Lemma 3.2,

$$\left| \frac{\widehat{Z}}{\mathcal{Z}_G(\lambda, \beta)} - 1 \right| \leq \mathbb{P}(|S \cap V_L| > \rho n) + \mathbb{P}(|S \cap V_R| > \rho n) \leq 2e^{-n \mathcal{D}_{\text{KL}}(\rho \parallel \lambda/(1+\lambda))}.$$

When $\alpha \leq \rho$, it holds that by Lemma 3.2,

$$\begin{aligned} \left| \frac{\widehat{Z}}{\mathcal{Z}_G(\lambda, \beta)} - 1 \right| &\leq \mathbb{P}(|S \cap V_L| > \rho n) + \mathbb{P}(|S \cap V_R| > \rho n) \\ &\quad + \mathbb{P}(\alpha n < |S \cap V_L| \leq \rho n, \alpha n < |S \cap V_R| \leq \rho n) \\ &\leq 2e^{-n \mathcal{D}_{\text{KL}}(\rho \parallel \lambda/(1+\lambda))} + \mathbb{P}(\alpha n < |S \cap V_L| \leq \rho n, \alpha n < |S \cap V_R| \leq \rho n). \end{aligned}$$

For a vertex subset $A \subseteq V_L$ of size $> \alpha n$, conditional on $S \cap V_L = A$, it holds that the random variables $\{\mathbb{1}[v \in S]\}_{v \in V_R}$ are independent. Therefore, $\mathbb{1}[v \in S]$ is a Bernoulli random variable with rate

$$p_v := \mathbb{P}(v \in S \mid S \cap V_L = A) = \frac{\lambda \beta^{d_A(v)}}{1 + \lambda \beta^{d_A(v)}}.$$

Then

$$\mathbf{E}[|S \cap V_R| \mid S \cap V_L = A] = \sum_{v \in V_R} p_v \leq \lambda W(A).$$

Pick a subset B of A of size αn and by monotonicity ($\beta \leq 1$), $W(A) \leq W(B)$. Then by Chernoff bound ([Lemma 2.2](#)),

$$\mathbb{P}(|S \cap V_R| > \alpha n \mid S \cap V_L = A) \leq \exp(-\alpha n(-\ln(1-\kappa) - \kappa)).$$

By law of total expectation,

$$\begin{aligned} \mathbb{P}(\alpha n < |S \cap V_L| \leq \rho n, \alpha n < |S \cap V_R| \leq \rho n) \\ &\leq \sum_{A \subseteq V_L: |A| > \alpha n} \mathbb{P}(S \cap V_L = A) \mathbb{P}(|S \cap V_R| > \alpha n \mid S \cap V_L = A) \\ &\leq \sum_{A \subseteq V_L: |A| > \alpha n} \mathbb{P}(S \cap V_L = A) e^{-\alpha n(-\ln(1-\kappa) - \kappa)} \\ &\leq e^{-\alpha n(-\ln(1-\kappa) - \kappa)}. \end{aligned}$$

Substituting it into the upper inequality, we conclude that

$$\left| \frac{\hat{Z}}{\mathcal{Z}_G(\lambda, \beta)} - 1 \right| \leq 2e^{-n\mathcal{D}_{\text{KL}}(\rho \parallel \lambda/(1+\lambda))} + e^{-\alpha(-\ln(1-\kappa) - \kappa)n}$$

which is the desired upper bound. $\square$

3.2. Concentration phenomena. In this part, we reveal some concentration phenomena in the probability distribution $\mu_{G;\lambda,\beta}$. For simplicity of analysis, we consider the pairing model $\text{PM}(n, \Delta)$. Recall the definition of $W(A)$, By symmetry, the normalized expectation of $W(A)$ is defined by

$$\begin{aligned} \text{mean}(\theta) &= \frac{1}{n} \mathbf{E}_{\text{PM}(n, \Delta)}[W(A)] \\ &= \frac{1}{n} \sum_{v \in V_R} \mathbf{E}_{\text{PM}(n, \Delta)}[\beta^{d_A(v)}] \\ &= \frac{1}{\binom{\Delta n}{\Delta}} \sum_{j=0}^{\Delta} \binom{\Delta \theta n}{j} \binom{\Delta(1-\theta)n}{\Delta-j} \beta^j. \end{aligned}$$

Before all, we discuss the following approximation to $\text{mean}(\theta)$ which could give some intuition to our discuss.

Lemma 3.4 (Approximation to normalized expectation). *It holds that $\text{mean}(\theta) = (1 - (1-\beta)\theta)^\Delta + o_n(1)$.*

Proof. Let X be a random variable denoting the number of marked elements in a Δ -size sample drawn from Δn elements containing $\Delta \theta n$ marked ones without replacement. Then by symmetry, $\text{mean}(\theta) = \mathbf{E}[\beta^X]$. Let Y be a random variable denoting the number of marked elements from Δ independent samples in the same population. We construct a coupling between X and Y . For

a process of Y , when the draws are distinct, we use it as X ; otherwise, we resample X from all distinct Δ samples uniformly at random. Then by union bound,

$$\mathbb{P}(X \neq Y) \leq \binom{\Delta}{2} \frac{1}{\Delta n} = \frac{\Delta - 1}{2n}.$$

Since $\beta \in [0, 1]$,

$$\mathbf{E}[\beta^X - \beta^Y] \leq \mathbb{P}(X \neq Y) \leq \frac{\Delta - 1}{2n}.$$

It is clear that $Y \sim \text{Bin}(\Delta, \theta)$. Then we conclude that

$$\text{mean}(\theta) = \mathbf{E}[\beta^Y] + o_n(1) = (1 - \theta + \theta\beta)^\Delta + o_n(1).$$

□

We investigate the following concentration for $W(A)$. For simplicity, we provide the proof of it in [Appendix B](#).

Lemma 3.5. *For a vertex subset $A \subseteq V_L$ with $|A| \leq n/2$, there exists an absolute constant $C > 0$ such that with high probability over the pairing model $\text{PM}(n, \Delta)$, for every $t > 0$,*

$$\mathbb{P}(|W(A) - \mathbf{E}[W(A)]| \geq t) \leq 2 \exp\left(-\frac{t^2}{2(1 - \beta)(2n + 2t/3)}\right).$$

Specially, for the case $\beta = 1$, $W(A)$ is deterministic.

For every positive integer $n > 0$, define the function $h_n : [0, 1] \times \mathbb{R}_{>0} \rightarrow \mathbb{R}$ as

$$h_n(\theta, u) := H(\theta) + \frac{u + 3 \ln(n + 1)}{n}$$

where $H(x) = -x \ln x - (1 - x) \ln(1 - x)$ is the binary entropy function. For convenience, we set the error function as

$$\mathfrak{E}_\beta(\theta, u) = 4\left((1 - \beta)h_n(\theta, u) + \sqrt{(1 - \beta)h_n(\theta, u)}\right).$$

The following corollary, which is a direct result from [Lemma 3.5](#), takes significant impact on our choice of the parameter α .

Corollary 3.6. *Fix a factor $u > 0$. With probability at least $1 - 2e^{-u}$ over $\text{PM}(n, \Delta)$, for every vertex subset $A \subseteq V_L$ or $A \subseteq V_R$ of size $\theta n \leq n/2$,*

$$(4) \quad \left| \frac{W(A)}{n} - \text{mean}(\theta) \right| \leq \mathfrak{E}_\beta(\theta, u).$$

Moreover, taking $u = 2 \ln n$, (1) holds for every vertex subset $A \subseteq V_L$ or $A \subseteq V_R$ of size $\theta n \leq n/2$ with probability $1 - o_n(1)$.

Proof. When $\beta = 1$, $W(A)$ is a deterministic statistic and there remains nothing to do. For $\beta \in [0, 1)$, apply [Lemma 3.2](#) with the choice

$$t = 4n\left((1 - \beta)h_n(\theta, u) + \sqrt{(1 - \beta)h_n(\theta, u)}\right).$$

It holds that for a subset A of size θn , the failure probability is at most $2e^{-2nh_n(\theta, u)}$. By the union bound and $\binom{n}{k} \leq e^{nH(k/n)}$ for $k \leq n/2$, the total failure probability is at most

$$2 \sum_{k=0}^{n/2} \binom{n}{k} 2e^{-2nh_n(k/n, u)} \leq 4(n + 1)^{-6} \sum_{k=0}^{n/2} e^{-nH(k/n) - 2u} \leq 2e^{-u}.$$

Therefore we conclude the lemma. □

Another important concentration phenomenon is the distribution of *degree profiles*. For a subset $A \subseteq V_L$ of size θn , its degree profile $p_A \in \mathbb{R}^{\Delta+1}$ is a normalized vector defined as

$$\forall j = 0, \dots, \Delta, \quad p_A(j) = \frac{1}{n} |\{v \in V_R \mid d_A(v) = j\}|$$

and set its ‘expectation’ profile as

$$\forall j = 0, \dots, \Delta, \quad \Phi_\theta(j) = \binom{\Delta}{j} \theta^j (1 - \theta)^{\Delta-j}.$$

Let $\mathcal{P}(\theta)$ be all admissible degree profiles p for a subset $A \subseteq V_L$ of size θn . Define the residual factor $r_n(\cdot)$ as

$$r_n(u) = \frac{1}{n} ((\Delta + 5) \ln(n + 1) + u + \ln 2(\Delta + 1)).$$

Proposition 3.7. *For every factor $u > 0$, with probability $1 - e^{-u}$ over $\text{PM}(n, \Delta)$, for every vertex subset $A \subseteq V_L$ or $A \subseteq V_R$, it holds that*

$$(5) \quad \mathcal{D}_{\text{KL}}(p_A \parallel \Phi_\theta) \leq H(\theta) + r_n(u)$$

where $\theta = |A|/n$.

The proof of [Proposition 3.7](#) is deferred to [Appendix B](#). Hence for every real number $r \geq 0$, define

$$\mathcal{P}(\theta, r) := \{p \in \mathcal{P}(\theta) : \mathcal{D}_{\text{KL}}(p \parallel \Phi_\theta) \leq H(\theta) + r\},$$

and

$$s_-(\theta, r) := \inf_{p \in \mathcal{P}(\theta, r)} \sum_{j=0}^{\Delta} p(j) \beta^j, \quad s_+(\theta, r) := \sup_{p \in \mathcal{P}(\theta, r)} \sum_{j=0}^{\Delta} p(j) \beta^j.$$

Then consider the following bounds:

$$\underline{s}(\theta) := \max \{\text{mean}(\theta) - \mathfrak{E}_\beta(\theta, u), s_-(\theta, r_n(u))\}, \quad \bar{s}(\theta) := \min \{\text{mean}(\theta) + \mathfrak{E}_\beta(\theta, u), s_+(\theta, r_n(u))\}.$$

Corollary 3.8. *Fix $u > 0$. For the pairing model $G = (V = V_L \cup V_R, E) \sim \text{PM}(n, \Delta)$, with probability at least $1 - 3e^{-u}$, for every vertex subset $A \subseteq V_L$ or $A \subseteq V_R$ of size at most $n/2$,*

$$\underline{s}(|A|/n) \leq \frac{W(A)}{n} \leq \bar{s}(|A|/n).$$

Proof. The corollary is a direct result of [Lemma 3.5](#) and [Proposition 3.7](#). □

The following lemma is also useful in our verification of parameters.

Lemma 3.9. *On the event of [Corollary 3.8](#), if $(1 - \beta)\theta\Delta \geq 6$, it holds that for all $A \subseteq V_L$ of size θn ,*

$$\frac{W(A)}{n} \leq e^{-(1-\beta)\theta\Delta/2} + \frac{3(H(\theta) + r_n(u))}{(1 - \beta)\theta\Delta/2 - 1}.$$

Proof. Consider the stochastic kernel $\mathcal{K}$ from $\{0, \dots, \Delta\}$ to $\{0, 1\}$ defined by

$$\mathcal{K}(j, 0) = 1 - \beta^j, \quad \mathcal{K}(j, 1) = \beta^j.$$

Then by data processing,

$$\mathcal{D}_{\text{KL}}(W(A)/n \parallel (1 - (1 - \beta)\theta)^\Delta) = \mathcal{D}_{\text{KL}}(p_A \mathcal{K} \parallel \Phi_\theta \mathcal{K}) \leq H(\theta) + r_n(u).$$

If $W(A)/n \leq e^{-(1-\beta)\theta\Delta/2}$, there is nothing to prove. Otherwise,

$$\mathcal{D}_{\text{KL}} \left(W(A)/n \parallel e^{-(1-\beta)\theta\Delta} \right) \leq \mathcal{D}_{\text{KL}} \left(W(A)/n \parallel (1 - (1-\beta)\theta)^\Delta \right) \leq H(\theta) + r_n(u).$$

By the inequality $\mathcal{D}_{\text{KL}} \left(W(A)/n \parallel e^{-(1-\beta)\theta\Delta} \right) \geq W(A)/n((1-\beta)\theta\Delta/2 - 1) - e^{-(1-\beta)\theta\Delta}$, we conclude the lemma. $\square$

4. TRICKLE-DOWN THEOREMS FOR SLICES

The previous section shows that typical slices of the configuration space have a regular geometric structure. In this section, we exploit this structure to prove rapid mixing of local walks.

We state trickle-down theorems for two-side slices and one-side ones respectively. Set parameters $\eta \in (0, 1)$ and $C_{\text{TD}} > 4$. Given $\lambda \leq \lambda_0$ and $\lambda(1-\beta) \leq \frac{C}{\sqrt{\Delta}}$ for some λ_0 such that $(1+\eta)\lambda_0/(1+\lambda_0) \leq 1/2$ and $C > 0$, we take

$$(6) \quad \rho_* := (1+\eta) \frac{\lambda}{1+\lambda}$$

and

$$(7) \quad \alpha_* := \max \left\{ \theta \in [0, \rho_*] : \underline{s}(\theta) - \theta \geq C_{\text{TD}}(1-\beta) \cdot 3\sqrt{\Delta}\theta \right\}.$$

Remark 4.1. Note that α_* is always a valid value and we can further assume that α_*n is a integer.

4.1. Two-side trickle-down theorem. First we discuss the trickle-down theorem for two-side slices. For integers $\ell = 0, \dots, n$ and $r = 0, \dots, n$, the simplicial complex $(\mathcal{X}_{\ell,r}, \mu_{\ell,r})$ induced by $\mu_{\ell,r}$ is constructed as:

- $\mathcal{X}_{\ell,r}$ is the downward closed collection of $\Omega^{(\ell,r)}$;
- the weight function of $\mathcal{X}$ is defined as (2) with $\pi = \mu_{\ell,r}$.

Lemma 4.2. *Suppose that the statements in Proposition 2.13 and Corollary 3.8 holds. Then for every $\ell \leq \alpha_*n$ and $r \leq \alpha_*n$, the facet S in $(\mathcal{X} = \mathcal{X}_{\ell,r}, \pi = \mu_{\ell,r})$ induced by $\mu_{\ell,r}$ of codimension 2 satisfies that*

$$\lambda_2(\mathbf{P}_{S,1}^\wedge) \leq \frac{1}{2(\ell+r)}.$$

for all sufficiently large positive integer n .

Proof. When $\beta = 1$, the result holds trivially. We suppose that $\beta < 1$. For every facet S of codimension 2, let $A = S \cap V_L$ and $B = S \cap V_R$. If $|A| = \ell$ or $|B| = r$, the link of S is a complete graph with positive vertex weights. It is not hard to see $\lambda_2(\mathbf{P}_{S,1}^\wedge) \leq 0$. When $|A| = \ell - 1$ and $|B| = r - 1$,

$$\mathbf{P}_{S,1}^\wedge = \begin{pmatrix} 0 & P \\ Q & 0 \end{pmatrix}$$

where $P \in \mathbb{R}^{(V_L \setminus A) \times (V_R \setminus B)}$, $Q \in \mathbb{R}^{(V_R \setminus B) \times (V_L \setminus A)}$ are defined by

$$P(u, v) = \frac{\beta^{d_A(v)} \beta^{\mathbb{1}[(u,v) \in E]}}{\sum_{z \in V_R \setminus B} \beta^{d_A(z)} \beta^{\mathbb{1}[(u,z) \in E]}},$$

$$Q(v, u) = \frac{\beta^{d_B(u)} \beta^{\mathbb{1}[(u,v) \in E]}}{\sum_{z \in V_L \setminus A} \beta^{d_B(z)} \beta^{\mathbb{1}[(z,v) \in E]}}.$$

For each $u \in V_L \setminus A$ and $v \in V_R \setminus B$, define

$$p_u = \beta^{d_B(u)}, \quad q_v = \beta^{d_A(v)}.$$

Consider the matrix $M \in \mathbb{R}^{(V_L \setminus A) \times (V_R \setminus B)}$

$$M(u, v) := p_u q_v \beta^{\mathbf{1}[(u, v) \in E]} = p_u (1 - (1 - \beta) A_G(u, v)) q_v.$$

Moreover, define diagonal matrices $D_p = \text{diag}(p_u : u \in V_L \setminus A)$, $D_q = \text{diag}(q_v : v \in V_R \setminus B)$ and

$$\begin{aligned} \forall u \in V_L \setminus A, \quad D_L(u, u) &= p_u \sum_{v \in V_R \setminus B} q_v (1 - (1 - \beta) A_G(u, v)), \\ \forall v \in V_R \setminus B, \quad D_R(v, v) &= q_v \sum_{u \in V_L \setminus A} p_u (1 - (1 - \beta) A_G(u, v)). \end{aligned}$$

We represent $\mathbf{P}_{S,1}^\wedge$ in the form of

$$\begin{pmatrix} 0 & D_L^{-1} M \\ D_R^{-1} M^\top & 0 \end{pmatrix}$$

and $M = D_p(\mathbf{1}_{V_L \setminus A} \mathbf{1}_{V_R \setminus B}^\top - (1 - \beta) A_G) D_q$. By monotonicity of $\beta^{d_A(v)}$, the choice of α_* and [Corollary 3.8](#), it holds that

$$\sum_{v \in V_R \setminus B} q_v (1 - (1 - \beta) A_G(u, v)) \geq n \underline{s}(\alpha_*) - \alpha_* n \geq C_{\text{TD}} (1 - \beta) 3\sqrt{\Delta} \alpha_* n - \Delta.$$

Similar bound holds for $\sum_{u \in V_L \setminus A} p_u (1 - (1 - \beta) A_G(u, v))$. Therefore, by [Proposition 2.13](#), it holds that

$$\begin{aligned} \lambda_2(\mathbf{P}_{S,1}^\wedge) &\leq \frac{(1 - \beta) \lambda_2(A_G)}{C_{\text{TD}} (1 - \beta) 3\sqrt{\Delta} (\alpha_* n - \Delta)} \\ &\leq \frac{1 + o_n(1)}{C_{\text{TD}} \alpha_* n} \\ &\leq \frac{1}{2(\ell + r)} \end{aligned}$$

for all sufficiently large positive integer n . □

4.2. One-side trickle-down theorem. Now we discuss the trickle-down theorem on one-side slices. When $\beta = 0$, [\[KOGT26\]](#) proposes a way to derive the trickle-down theorem. We show how to extend their method to adapt the case $\beta > 0$.

By symmetry, we only need to consider the left-side slices μ_s^L . Fix a size $s = 0, \dots, n$. Construct the simplicial complex $(\mathcal{X} = \mathcal{X}_s^L, \pi = \mu_s^L)$ as

- $\mathcal{X}_s^L$ is the downward closed collection of $\binom{V_L}{s}$;
- the weight function of $\mathcal{X}$ is defined as [\(2\)](#) with $\pi = \mu_s^L$.

Recall that for every vertex subset $A \in \binom{V_L}{s}$,

$$\mu_s^L(A) \propto W(A) = \prod_{v \in V_R} (1 + \lambda \beta^{d_A(v)}).$$

Define the sequences $\{\varphi_j\}_{j=0}^\Delta$, $\{\phi_j\}_{j=0}^{\Delta-1}$ and $\{\delta_j\}_{j=0}^{\Delta-2}$ as

$$\varphi_j := 1 + \lambda \beta^j, \quad \phi_j := \ln \frac{\varphi_j}{\varphi_{j+1}}, \quad \delta_j := \frac{\varphi_j \varphi_{j+2}}{\varphi_{j+1}^2} - 1 = \frac{\lambda(1 - \beta)^2 \beta^j}{(1 + \lambda \beta^{j+1})^2}.$$

Now given a facet A of codimension 2, for a vertex $x \in V_L \setminus A$, define

$$q_A(x) = \prod_{v \in N_G(x)} \frac{\varphi_{d_A(v)+1}}{\varphi_{d_A(v)}}.$$

We define the off-diagonal matrix $K \in \mathbb{R}^{(V_L \setminus A) \times (V_L \setminus A)}$ as

$$\forall x \neq y \in V_L \setminus A, \quad K(x, y) = \prod_{v \in N_G(x) \cap N_G(y)} (1 + \delta_{d_A(v)})$$

with convention that the empty product is 1. Then for every $x \in V_L \setminus A$, set

$$Z_A(x) := \sum_{y \in V_L \setminus A} q_A(y) K(x, y).$$

Therefore, a direct calculation gives that for $x \neq y \in V_L \setminus A$,

$$\mathbf{P}_{S,1}^\wedge(x, y) = \frac{q_A(y) K(x, y)}{Z_A(x)}.$$

In addition, the average free energy of p_A is defined as

$$\bar{L}_A = -\frac{1}{n - |A|} \sum_{x \in V_L \setminus A} \ln q_A(x) = \frac{1}{n - |A|} \sum_{j=0}^{\Delta} np_A(j)(\Delta - j)\phi_j.$$

Lemma 4.3. Assume that statements in [Propositions 2.13](#) and [2.14](#) holds. Set

$$U_G(\lambda, \beta) = \lambda(1 - \beta)^2 \lambda_2(A_G)^2 + \lambda^2(1 - \beta)^4 - 1.$$

If $U_G(\lambda, \beta) \leq 0$, it holds that $\lambda_2(\mathbf{P}_{S,1}^\wedge) \leq 0$. Otherwise,

$$\lambda_2(\mathbf{P}_{S,1}^\wedge) \leq \frac{U_G(\lambda, \beta) e^{\bar{L}_A}}{n - |A| - e^{\bar{L}_A}}.$$

Proof. Define a matrix $H \in \mathbb{R}^{(V_L \setminus A) \times V_R}$ as $H(x, v) = \mathbb{1}[(x, v) \in E]$. Set the diagonal matrix $D_\delta = \text{diag}(\delta_{d_A(v)} : v \in V_R)$ and an off-diagonal matrix $M_2 \in \mathbb{R}^{(V_L \setminus A) \times (V_L \setminus A)}$ as

$$\forall x \neq y \in V_L \setminus A, \quad M_2(x, y) = \begin{cases} \delta_{d_A(v_1)} \delta_{d_A(v_2)}, & N_G(x) \cap N_G(y) = \{v_1, v_2\} \\ 0 & \text{otherwise} \end{cases}.$$

By [Proposition 2.14](#), it holds that

$$K = \mathbf{1}_{V_L \setminus A} \mathbf{1}_{V_L \setminus A}^\top + H D_\delta H^\top - \text{diag} \left(1 + \sum_{v \in N_G(x)} \delta_{d_A(v)} : x \in V_L \setminus A \right) + M_2.$$

Recall the definition of $\{\delta_j\}_{j=0}^{\Delta-2}$, it always holds that $\delta_j \leq \lambda(1 - \beta)^2$. Therefore $D_\delta \preceq \lambda(1 - \beta)^2 \mathbf{I}$. For M_2 , note that M_2 is symmetric and of the form

$$\begin{pmatrix} 0 & (m_{xy})_{x,y} \\ (m_{xy})_{x,y}^\top & 0 \end{pmatrix}, \quad m_{xy} \leq \lambda^2(1 - \beta)^4.$$

Therefore $M_2 \preceq \lambda^2(1 - \beta)^4 \mathbf{I}$. Consequently,

$$K \preceq \mathbf{1}_{V_L \setminus A} \mathbf{1}_{V_L \setminus A}^\top + \lambda(1 - \beta)^2 H H^\top + (\lambda^2(1 - \beta)^4 - 1) \mathbf{I}.$$

Note that H is a principal compression of A_G . Therefore by regularity,

$$H H^\top \preceq A_G A_G^\top \preceq \frac{\Delta^2}{n} \mathbf{1}_{V_L} \mathbf{1}_{V_R}^\top + \lambda_2(A_G) \mathbf{I}.$$

Therefore

$$K \preceq \left(1 + \frac{\lambda(1-\beta)^2\Delta^2}{n}\right) \mathbf{1}\mathbf{1}^\top + U_G(\lambda, \beta)\mathbf{I}.$$

A similar argument in [KOGT26] subtracts the rank-one part for $\mathbf{P}_{S,1}^\wedge$. Hence we only need to give bounds to $q_A(x)$ and $Z_A(x)$. For $q_A(x)$, it is clear to see $q_A(x) \leq 1$ since $\beta \leq 1$. For $Z_A(x)$, by Jensen's inequality,

$$\begin{aligned} Z_A(x) &\geq \sum_{y \neq x \in V_L \setminus A} q_A(y) \\ &\geq \sum_{y \in V_L \setminus A} q_A(y) - 1 \\ &\geq (n - |A|)e^{-\bar{L}_A} - 1. \end{aligned}$$

Hence $1/Z_A(x) \leq e^{\bar{L}_A}/(n - |A| - e^{\bar{L}_A})$. Combining all things together, we conclude the lemma. $\square$

5. DESIGN OF OUR APPROXIMATE ALGORITHM

In this section, we give our algorithm to approximate $\mathcal{Z}_G(\lambda, \beta)$ on random regular bipartite graphs and therefore prove [Theorem 1.1](#).

Given $\lambda \leq \lambda_0$ and $\lambda(1-\beta) \leq \frac{C}{\sqrt{\Delta}}$ for some λ_0 and C , we show how to set parameters to apply the above approximation and trickle-down theorems. For some sufficiently large constant $K > 0$, define α_0 be the unique solution of the following equation on $[0, 1/2]$:

$$(1 - (1 - \beta)\alpha)^\Delta = K \left(1 + 3(1 - \beta)\sqrt{\Delta}\right) \alpha.$$

Pick

$$(8) \quad \rho = (1 + \eta) \frac{\lambda}{1 + \lambda}, \quad \alpha = \min \{\alpha_0, \rho\}.$$

Remark 5.1. Given $K > 0$, we only need a largest $\alpha' \leq \alpha_0$ such that $\alpha'n$ is an integer. Therefore it costs polynomial time to determine the value of α' and takes little impact on our analysis. For simplicity, we employ the solution α_0 in following proofs.

Lemma 5.2. *Fix a sufficiently large constant $K > 0$. For $\beta \in (0, 1)$ and sufficiently large factor $\Delta > 0$, it holds that*

$$H(\alpha_0) \leq \frac{(1 + 3(1 - \beta)\sqrt{\Delta})^2 \alpha_0^2}{1 - \beta}.$$

Then on the event of [Corollary 3.8](#), for every $\theta \leq \alpha_0$,

$$\frac{3}{4}(1 - (1 - \beta)\theta)^\Delta \leq \frac{W(A)}{n} \leq \frac{5}{4}(1 - (1 - \beta)\theta)^\Delta.$$

Proof. Take

$$y = (1 - \beta)\Delta\alpha_0, \quad t = (1 - \beta)\sqrt{\Delta}, \quad Q = \frac{(1 + 3(1 - \beta)\sqrt{\Delta})^2 \alpha_0}{1 - \beta}.$$

It is sufficient to show that $\ln(e/\alpha_0) \leq Q$ and by the inequality $H(\alpha_0) \leq \alpha_0 \ln(e/\alpha_0)$ we conclude the lemma.

Recall the choice α_0 and we calculate:

$$\ln(1/\alpha_0) = \ln \left(K \left(1 + 3(1 - \beta)\sqrt{\Delta} \right) \right) - \Delta \ln \left(1 - \frac{y}{\Delta} \right).$$

A standard estimation to $-\ln(1-x)$ on interval $[0, 1)$ and $y/\Delta \leq \alpha_0 \leq 1/2$ gives

$$y \leq -\Delta \ln\left(1 - \frac{y}{\Delta}\right) \leq \frac{y}{1 - y/\Delta} \leq 2y.$$

Therefore

$$\ln\left(K\left(1 + 3(1 - \beta)\sqrt{\Delta}\right)\right) + y \leq \ln(1/\alpha_0) \leq \ln\left(K\left(1 + 3(1 - \beta)\sqrt{\Delta}\right)\right) + 2y.$$

When $3(1 - \beta)\sqrt{\Delta} \geq 1$, it holds that $t \geq 1/3$ and

$$3t \leq 1 + 3(1 - \beta)\sqrt{\Delta} \leq 6t, \quad Q \geq 9y.$$

By the equation determining α_0 , we can also have the following inequalities:

$$\begin{cases} e^{-2y} \leq K(1 + 3(1 - \beta)\sqrt{\Delta}) \frac{y}{(1 - \beta)\Delta} \\ \sqrt{\Delta} \leq 6Kye^{2y} \end{cases}.$$

By a crude inequality $1 + 3(1 - \beta)\sqrt{\Delta} \leq 4\sqrt{\Delta}$, we upper bound the term $\ln(1 + 3(1 - \beta)\sqrt{\Delta})$ as

$$\ln(1 + 3(1 - \beta)\sqrt{\Delta}) \leq \ln 4 + \frac{1}{2} \ln \Delta \leq \ln 24K + 3y$$

and we obtain an additional bound $y \geq (1 + 2 \ln K + \ln 24)$ if

$$\Delta \geq 36K^2(1 + 2 \ln K + \ln 24)^2 e^{4(1 + 2 \ln K + \ln 24)}.$$

Therefore

$$\ln(e/\alpha_0) \leq 1 + 2 \ln K + \ln 24 + 5y \leq 6y \leq Q.$$

When $3(1 - \beta)\sqrt{\Delta} < 1$, then it holds that $t < 1/3$ and thus $1 + 3(1 - \beta)\sqrt{\Delta} < 2$. We bound Q as

$$Q = \left(\frac{1 + 3(1 - \beta)\sqrt{\Delta}}{t}\right)^2 y > 9y.$$

The term $\ln(e/\alpha_0)$ can be upper bounded by $1 + \ln 2K + 2y$. Together with $(1 - \beta)\Delta \leq 2Kye^{2y}$ by the equation determining α_0 , if $(1 - \beta)\Delta \geq L$ where $L = 2K(1 + \ln 2K)e^{2(1 + \ln 2K)}$, it holds that

$$\ln(e/\alpha_0) \leq 3y \leq Q.$$

Otherwise, when $\Delta \geq 2L$, it holds that

$$K(1 + 3(1 - \beta)\sqrt{\Delta})\alpha_0 = (1 - (1 - \beta)\alpha_0)^\Delta \geq (1 - L/\Delta)^\Delta \geq e^{-2L}.$$

Thus $\alpha \geq \alpha_K = e^{-2L}/(K(1 + 3(1 - \beta)\sqrt{\Delta}))$. For sufficiently large Δ depending only on K , it holds that

$$\ln(e/\alpha_0) \leq \ln(e/\alpha_K) \leq \frac{\alpha_K}{1 - \beta} \leq Q.$$

Therefore we prove that $H(\alpha_0) \leq \alpha_0 Q$.

On the event of [Corollary 3.8](#), observe that we only need to show the case $\theta = \alpha_0$. Pick K as a sufficient large positive number and Δ is sufficiently large with respect to K , $H(\alpha_0)$ is at most $(1 - (1 - \beta)\alpha_0)^\Delta/4$ by the equation defining α_0 . Thus we conclude the full lemma. $\square$

The following verification for α and ρ allows us to apply the approximation to the partition function $\mathcal{Z}_G(\lambda, \beta)$ ([Lemma 3.3](#)) and trickle-down theorems ([Lemmas 4.2](#) and [4.3](#)) safely.

Proposition 5.3. For sufficiently large constant $K > 0$, recall that the parameter $\alpha_0 \in [0, 1/2]$ is defined by

$$(1 - (1 - \beta)\alpha_0)^\Delta = K(1 + 3(1 - \beta)\sqrt{\Delta})\alpha_0$$

and $\alpha = \min\{\alpha_0, \rho\}$. There exist constants $\lambda_0 > 0$ and $C > 0$ such that on the event of all results in [Lemma 5.2](#) and the assumption $\lambda \leq \lambda_0$ and $\lambda(1 - \beta) \leq \frac{C}{\sqrt{\Delta}}$, it holds that $\alpha \leq \alpha_*$ and at least one of the following holds:

- (1) $\alpha = \rho$; or
- (2) for all $A \subseteq V_L$ of size αn , it holds that

$$\lambda W(A) \leq \frac{1}{2}|A|.$$

Proof. We show that

$$\lambda \bar{s}(\alpha) \leq \frac{1}{2}\alpha$$

and thus by [Corollary 3.8](#),

$$\lambda W(A) \leq \lambda n \bar{s}(\alpha) \leq \frac{1}{2}|A|.$$

To verify this, note that for all $\theta \leq \alpha_0$, it holds that $(1 - (1 - \beta)\theta)^\Delta \geq K(1 + 3(1 - \beta)\sqrt{\Delta})\theta$. By choosing K large enough, we can make $\alpha \leq \alpha_*$.

When $\alpha = \rho$, there is nothing to do. Otherwise, by [Lemma 5.2](#), we obtain that

$$\frac{\lambda \bar{s}(\alpha)}{\alpha} \leq \left(\frac{5}{4} + o_n(1)\right) \lambda K(1 + 3(1 - \beta)\sqrt{\Delta}) \leq 2K(\lambda + 3\lambda(1 - \beta)\sqrt{\Delta}).$$

Choose λ_0 and C as small constants to make $2K(\lambda_0 + C) \leq 1/2$. Then we conclude the proposition. $\square$

Proposition 5.4. For sufficiently large constant K as in [Lemma 5.2](#), there exist $\lambda_0, C > 0$, a positive integer Δ_0 such that the following holds. Fix $\Delta \geq \Delta_0$ and

$$\lambda \leq \lambda_0, \quad \lambda(1 - \beta) \leq \frac{C}{\sqrt{\Delta}}.$$

For all sufficiently large positive integer n , we take $u = 2 \ln n$ and suppose the events of [Propositions 2.13](#) and [2.14](#), [Corollary 3.8](#) and [Proposition 3.7](#) hold. In addition, assume that $\lfloor \alpha n \rfloor < \lfloor \rho n \rfloor$ (otherwise, there is nothing to do). Then for every $s \in \{\lfloor \alpha n \rfloor + 1, \dots, \lfloor \rho n \rfloor\}$, $s \geq 2$, it holds that for every facet $A \in \binom{V_L}{s-2}$ of codimension 2, by setting $\theta = |A|/n$,

$$\theta \cdot \max\{U_G(\lambda, \beta), 0\} \cdot e^{\bar{L}_A} \leq \begin{cases} O\left(C \frac{\ln \Delta}{\Delta^{3/8}}\right), & (1 - \beta)\theta\Delta \leq \ln \Delta \\ O(C^2), & (1 - \beta)\theta\Delta \geq \ln \Delta \end{cases}.$$

and therefore,

$$1 - \theta - \frac{e^{\bar{L}_A}}{n} > 0, \quad \frac{2(\theta + 2/n)e^{\bar{L}_A} \max\{U_G(\lambda, \beta), 0\}}{1 - \theta - e^{\bar{L}_A}/n} \leq \frac{1}{2}.$$

and thus

$$\lambda_2(\mathbf{P}_{A,1}^\wedge) \leq \frac{1}{2s}.$$

Proof. Since $\beta \in (0, 1)$, it holds that

$$0 \leq \phi_j \leq \lambda(1 - \beta)\beta^j.$$

By definition of $\bar{L}_A$, it holds that

$$\bar{L}_A \leq \frac{\lambda(1-\beta)\Delta}{1-\theta} \frac{W(A)}{n}.$$

To bound $W(A)$, we extend A to A' by adding two extra vertices and thus $|A'| \geq \alpha n$. This operation changes the value of $W(A)$ by at most $2(1-\beta)\Delta$, meaning that

$$W(A) \leq W(A') + 2(1-\beta)\Delta.$$

When $(1-\beta)\theta\Delta \leq \ln \Delta$, by [Lemma 5.2](#),

$$\bar{L}_A \leq 4K\lambda(1-\beta)\Delta(1+3(1-\beta)\sqrt{\Delta})\alpha_0 \leq 12K(\lambda+\lambda(1-\beta)\sqrt{\Delta})\ln \Delta.$$

Pick λ_0 and C as sufficiently small constants such that $12K(\lambda_0+C) \leq 1/8$. Therefore $e^{\bar{L}_A} \leq \Delta^{1/8}$. In addition, by [Proposition 2.13](#), $\max\{U_G(\lambda, \beta), 0\} \leq 100\lambda(1-\beta)^2\Delta$. Therefore,

$$(9) \quad \theta e^{\bar{L}_A} \max\{U_G(\lambda, \beta), 0\} \leq 100\lambda(1-\beta)(1-\beta)\theta\Delta e^{1/8} \leq \frac{100C \ln \Delta}{\Delta^{3/8}}.$$

When $(1-\beta)\theta\Delta > \ln \Delta$, by [Lemma 3.9](#) and $\theta \leq 1/2$,

$$\begin{aligned} \bar{L}_A &\leq \frac{\lambda(1-\beta)\Delta}{1-\theta} \left(e^{-(1-\beta)\theta\Delta/2} + \frac{3(H(\theta) + r_n(u))}{(1-\beta)\theta\Delta/2 - 1} \right) \\ &\leq 2\lambda(1-\beta)\Delta e^{-(1-\beta)\theta\Delta/2} + \frac{6\lambda(1-\beta)\Delta(H(\theta) + r_n(u))}{(1-\beta)\theta\Delta/2 - 1} \\ &\leq 2\lambda(1-\beta)\sqrt{\Delta} + 18\lambda \frac{H(\theta) + r_n(u)}{\theta}. \end{aligned}$$

We take $u = 2\ln n$ and obtain that $r_n(u) = O_\Delta(\ln n/n)$. Then we have

$$\bar{L}_A \leq 2C + 18\lambda \ln(e/\theta) + o_n(1).$$

Take λ_0 and C sufficiently small. Since $\theta \leq \rho = O(\lambda)$, we can obtain that

$$\begin{aligned} \theta e^{-\bar{L}_A} \max\{U_G(\lambda, \beta), 0\} &\leq 10000\lambda(1-\beta)^2\Delta\theta^{1-10000\lambda} \\ &\leq O\left(\lambda^2(1-\beta)^2\Delta\lambda^{-10000\lambda}\right) \\ &\leq O(C^2). \end{aligned}$$

To show (9), we set n_0 sufficient large and for all $n \geq n_0$, the full lemma holds. $\square$

Now we are ready to design our algorithm and thus prove [Theorem 1.1](#).

Proof of Theorem 1.1. When $\beta = 0$, employ the main result of [\[KOGT26\]](#). When $\beta = 1$, it is direct to see

$$\mathcal{Z}_G(\lambda, 1) = (1+\lambda)^{|V|}$$

and we can calculate it precisely in polynomial time.

Assume that $\beta \in (0, 1)$. Take $u = 2\ln n$. Assume that [Propositions 2.13](#) and [2.14](#), [Corollary 3.8](#) and [Proposition 3.7](#) hold. By union bound, these events hold with probability $1 - o_n(1)$. Pick α and ρ as (8). By [Propositions 5.3](#) and [5.4](#), [Lemma 3.3](#), [Lemma 4.2](#) and [Lemma 4.3](#) holds. When $\varepsilon \leq 4(2e^{-n\mathcal{D}_{\text{KL}}(\rho \parallel \lambda/(1+\lambda))} + e^{-\alpha(\ln 2 - 1/2)n})$, we just enumerate all configurations by brute force and calculate the total weights. This terminates in time $\text{poly}(n, \varepsilon^{-1})$.

Otherwise, when $\lfloor \alpha n \rfloor = \lfloor \rho n \rfloor$, we set $\hat{Z} := \sum_{\ell \leq \rho n, r \leq \rho n} \mathcal{Z}_{\ell, r}$ and for a fixed $\ell, r \leq \rho n$, by [Lemmas 4.2](#) and [2.10](#) and thus [Theorem 2.9](#), there is a rapid sampler to sample from $\mu_{\ell, r}$. However, note that the instance is not self-reducible. Instead, by a similar argument in [\[KOGT26\]](#), the random walks on the simplicial complex is self-reducible. Therefore, there is a randomized algorithm to

output an $(\varepsilon/5)$ -approximation to $\mathcal{Z}_{\ell,r}$ in time $\text{poly}(n, \varepsilon^{-1})$, leading to an $(\varepsilon/5)$ -approximation to $\widehat{Z}$. Thus we can design an **FPRAS** for $\mathcal{Z}_G(\lambda, \beta)$.

When $\lfloor \alpha n \rfloor = \lfloor \rho n \rfloor$, we set

$$\widehat{Z} := \sum_{\ell \leq \alpha n, r \leq \alpha n} \mathcal{Z}_{\ell,r} + \sum_{\alpha n < s \leq \rho n} \mathcal{Z}_s^L + \sum_{\alpha n < s \leq \rho n} \mathcal{Z}_s^R$$

By above, we only need to design an efficient sampler for $\mu_{\ell,r}$, μ_s^L and μ_s^R . By [Propositions 5.3](#) and [5.4](#), we apply [Lemmas 4.2](#) and [4.3](#) and thus by [Lemma 2.10](#) and [Theorem 2.9](#), for every $\ell \leq \alpha n, r \leq \alpha n$ and $s \leq \beta$, the down-up walks on all of $\mu_{\ell,r}$, μ_s^L and μ_s^R exhibit polynomial-time mixing. Therefore, we can output an $(\varepsilon/5)$ -approximation to $\widehat{Z}$ in time $\text{poly}(n, \varepsilon^{-1})$ with probability at least $3/4$, which is an ε -approximation to $\mathcal{Z}_G(\lambda, \beta)$. $\square$

REFERENCES

- [AL20] V. L. Alev and L. C. Lau. Improved analysis of higher order random walks and applications. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC 2020)*, STOC 2020, pages 1198–1211, New York, NY, USA, 2020. Association for Computing Machinery. [3](#), [7](#)
- [CCC⁺25] X. Chen, Z. Chen, Z. Chen, Y. Yin, and X. Zhang. Rapid mixing on random regular graphs beyond uniqueness. In *2025 IEEE 66th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 2170–2193, 2025. [3](#)
- [CCYZ25] X. Chen, Z. Chen, Y. Yin, and X. Zhang. Rapid mixing at the uniqueness threshold. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC ’25, pages 879–890, New York, NY, USA, 2025. Association for Computing Machinery. [2](#)
- [CGG⁺16] Jin-Yi Cai, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, Mark Jerrum, Daniel Štefankovič, and Eric Vigoda. #BIS-hardness for 2-spin systems on bipartite bounded degree graphs in the tree non-uniqueness region. *Journal of Computer and System Sciences*, 82(5):690–711, 2016. [2](#)
- [CJM⁺26] X. Chen, Z. Ju, T. Miao, Y. Yin, and X. Zhang. Edge-tilting field dynamics: Rapid mixing at the uniqueness threshold and optimal mixing for swendsen-wang dynamics. *arXiv: 2604.10525*, 2026. [2](#)
- [DvZ01] K. Dzhamaridze and J. H. van Zanten. On Bernstein-type inequalities for martingales. *Stochastic Processes and their Applications*, 93(1):109–117, 2001. [5](#)
- [Fre75] D. A. Freedman. On tail probabilities for martingales. *The Annals of Probability*, 3(1):100–118, 1975. [5](#)
- [Ger25] A. V. Gerbessiotis. A survey of Chernoff and Hoeffding bounds. *arXiv: 2506.15612*, 2025. [4](#)
- [GKSW26] Anna Geisler, Mihyun Kang, Michail Sarantis, and Ronen Wdowinski. Sampling from the antiferromagnetic Ising model on bipartite, regular expander graphs. *arXiv: 2603.02101*, 2026. [2](#), [3](#)
- [Gre00] C. Greenhill. The complexity of counting colourings and independent sets in sparse graphs and hypergraphs. *Computational Complexity*, 9(1):52–72, 2000. [2](#)
- [GŠV16] A. Galanis, D. Štefankovič, and E. Vigoda. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. *Combinatorics, Probability and Computing*, 25(4):500–559, 2016. [2](#)
- [JKP20] M. Jenssen, P. Keevash, and W. Perkins. Algorithms for #BIS-hard problems on expander graphs. *SIAM Journal on Computing*, 49(4):681–710, 2020. [2](#)

- [JMPV23] V. Jain, M. Michelen, H. T. Pham, and T. Vuong. Optimal mixing of the down-up walk on independent sets of a given size . In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1665–1681, Los Alamitos, CA, USA, November 2023. IEEE Computer Society. [3](#)
- [JVV86] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, 1986. [6](#)
- [KOGT26] N. Kocurek, S. Oveis Gharan, and D. Tjowasi. Sampling from the hardcore model on random regular bipartite graphs above the uniqueness threshold. *arXiv: 2604.21847*, 2026. [2](#), [3](#), [7](#), [8](#), [9](#), [14](#), [16](#), [19](#)
- [LLY13] L. Li, P. Lu, and Y. Yin. Correlation decay up to uniqueness in spin systems. In *Proceedings of the 2013 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 67–84, 2013. [2](#)
- [LP17] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times: Second Edition*, volume 107. American Mathematical Society, 2017. [5](#)
- [Opp18] I. Oppenheim. Local spectral expansion approach to high dimensional expanders part I: Descent of spectral gaps. *Discrete & Computational Geometry*, 59(2):293–330, 2018. [3](#), [7](#)
- [Pop35] T. Popoviciu. Sur les équations algébriques ayant toutes leurs racines réelles. *Mathematica*, 9(129-145):20, 1935. [5](#)
- [SS12] A. Sly and N. Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science, FOCS '12*, pages 361–369, USA, 2012. IEEE Computer Society. [2](#)
- [Wei06] D. Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing, STOC '06*, pages 140–149, New York, NY, USA, 2006. Association for Computing Machinery. [2](#)
- [Wor99] N. C. Wormald. *Models of Random Regular Graphs*, pages 239–298. London Mathematical Society Lecture Note Series. Cambridge University Press, 1999. [7](#)

APPENDIX A. EQUIVALENCE OF MODELS

We discuss here the equivalence between subgraph-world models and spin systems on regular graphs. Fix a Δ -regular graph $G = (V, E)$. For the physical Ising model on G with inverse temperature J and external field h , define the partition function

$$\mathcal{Z}_G^{\text{sp}}(h, J) = \sum_{\sigma \in \{-, +\}^V} \exp \left(-J \sum_{(u, v) \in E} \sigma_u \sigma_v + h \sum_{v \in V} \sigma_v \right).$$

Since G is a Δ -regular graph, it holds that

$$\mathcal{Z}_G^{\text{sp}}(h, J) = e^{-J\Delta n - 2hn} \mathcal{Z}_G(\lambda, \beta)$$

where $\beta = e^{-4J}$ and $\lambda = e^{2J+2h}$. Therefore it suffice to consider the Ising model defined by the subgraph world when the underlying graph is regular.

APPENDIX B. PROOFS OF CONCENTRATION INEQUALITIES

We state deferred proofs for concentration inequalities in [Section 3](#) here.

Lemma 3.5. *For a vertex subset $A \subseteq V_L$ with $|A| \leq n/2$, there exists an absolute constant $C > 0$ such that with high probability over the pairing model $\text{PM}(n, \Delta)$, for every $t > 0$,*

$$\mathbb{P}(|W(A) - \mathbf{E}[W(A)]| \geq t) \leq 2 \exp\left(-\frac{t^2}{2(1-\beta)(2n+2t/3)}\right).$$

Specially, for the case $\beta = 1$, $W(A)$ is deterministic.

Proof. When $\beta = 1$, $W(A) = n$ is deterministic. So we assume that $\beta \in [0, 1)$. We apply the following inequality: assume that there are N elements with K marked. Define the random variable X as the number of marked elements in a picked uniform subset of size M . Then

$$(10) \quad \mathbf{E}[\beta^X] \leq e^{-\frac{(1-\beta)KM}{N}}.$$

To show (10), for every element i , define $w_i = \beta$ if i is marked and otherwise $w_i = 1$. Then by the Maclaurin's inequality,

$$\mathbf{E}[\beta^X] = \frac{\sum_{1 \leq i_1 < \dots < i_M \leq N} w_{i_1} \cdots w_{i_M}}{\binom{N}{M}} \leq \left(\frac{1}{N} \sum_{i=1}^N w_i\right)^M \leq \left(1 - \frac{(1-\beta)K}{N}\right)^M \leq e^{-\frac{(1-\beta)KM}{N}}.$$

Now we turn to the lemma. Define X be the $\Delta|A|$ selected right half-edges and we uniformly give a random order to $X_1, \dots, X_{\Delta|A|}$. Define the Doob martingale

$$\forall i = 1, \dots, \Delta|A|, \quad Z_i = \mathbf{E}[W(A) \mid X_1, \dots, X_{i-1}].$$

Now we fix a step $i = 1, \dots, \Delta|A|$ and a history $\mathcal{F}_{i-1} = \sigma(X_1, \dots, X_{i-1})$. Let U be the remaining half-edges. For a possible next half-edge $y \in U$, define $f(y) := \mathbf{E}[W(A) \mid \mathcal{F}_{i-1}, X_i = y]$. For two distinct possible half-edges $x, y \in U$, we bound $|f(x) - f(y)|$ by construct a coupling of their remaining selected half-edges. When $X_i = x$, generate F_x uniformly from the $(|A| - i)$ -size subset of $U \setminus \{x\}$. We construct F_y from F_x as

$$F_y := \begin{cases} F_x & y \notin F_x \\ (F_x \setminus \{y\}) \cup \{x\} & y \in F_x \end{cases}.$$

It is trivial to see (F_x, F_y) is a valid coupling of the respective uniform distributions of possible remaining selected half-edges conditional on $X_i = x$ and $X_i = y$. When $y \in F_x$, it holds that $F_x \cup \{x\} = F_y \cup \{y\}$, meaning that the final selected subsets coincide. Therefore, they share the same value of $W(A)$. For $y \notin F_x$, by coupling, $F_x = F_y$ and thus the final selected subsets differ only at x and y . In this case, the subset $F := F_x = F_y$ is an uniform $(\Delta|A| - i)$ -size subset of $U \setminus \{x, y\}$.

When x, y are incident to a same vertex, the two cases show a same value of $W(A)$. Otherwise, suppose that x is incident to u and y is incident to v . Let z_u, z_v be the number of half-edges incident to u and v respectively hitting A in the history profile $X_1, \dots, X_{i-1}$ and Z_u, Z_v be the number of half-edges incident to u and v respectively hitting A in the future profile F . The difference of two future values of $W(A)$ is at most

$$\begin{aligned} |(\beta^{z_u+1+Z_u} + \beta^{z_v+Z_v}) - (\beta^{z_u+Z_u} + \beta^{z_v+1+Z_v})| &= (1-\beta)|\beta^{z_u+Z_u} - \beta^{z_v+Z_v}| \\ &\leq (1-\beta)(\beta^{z_u+Z_u} + \beta^{z_v+Z_v}). \end{aligned}$$

By (10) with $N = \Delta n - i - 1$, $K = \Delta - z_u - 1$ and $M = \Delta|A| - i$,

$$\begin{aligned}\mathbf{E}[\beta^{z_u + Z_u}] &= \beta^{z_u} \mathbf{E}[\beta^{Z_u}] \\ &\leq \beta^{z_u} \exp\left(-(1-\beta) \frac{(\Delta - z_u - 1)(\Delta|A| - i)}{\Delta n - i - 1}\right) \\ &\leq \exp\left(-(1-\beta) \left(z_u + \frac{\Delta - z_u - 1}{\Delta n - i - 1}(\Delta|A| - i)\right)\right) \\ &\leq \exp\left(-(1-\beta) \frac{\Delta|A| - i}{2n}\right).\end{aligned}$$

The same bound also holds for v . Therefore,

$$|f(x) - f(y)| \leq 2(1-\beta)e^{-(1-\beta)(\Delta|A|-i)/(2n)}.$$

Then by Popoviciu's inequality on variances (Proposition 2.3) and $1 - e^{-x} \geq x/2$ on $[0, 1]$,

$$\begin{aligned}V_{\Delta|A|} &:= \sum_{i=1}^{\Delta|A|} \mathbf{E}[(Z_i - Z_{i-1})^2 \mid \mathcal{F}_{i-1}] \\ &\leq \frac{1}{4} \sum_{i=0}^{\Delta|A|-1} 4(1-\beta)^2 e^{-(1-\beta)i/n} \\ &\leq \frac{(1-\beta)^2}{1 - e^{-(1-\beta)/n}} \\ &\leq 2(1-\beta)n.\end{aligned}$$

Trivially, $|Z_i - Z_{i-1}| \leq 2(1-\beta)$. Using Lemma 2.4 with $v = 2(1-\beta)n$ and $B = 2(1-\beta)$, it holds that

$$\begin{aligned}\mathbb{P}(|W(A) - \mathbf{E}[W(A)]| \geq t) &= \mathbb{P}(|W(A) - \mathbf{E}[W(A)]| \geq t, V_{\Delta|A|} \leq 2(1-\beta)n) \\ &\leq 2 \exp\left(-\frac{t^2}{2(1-\beta)(2n + 2t/3)}\right).\end{aligned}$$

Thus we conclude the lemma. $\square$

Proposition 3.7. *For every factor $u > 0$, with probability $1 - e^{-u}$ over $\text{PM}(n, \Delta)$, for every vertex subset $A \subseteq V_L$ or $A \subseteq V_R$, it holds that*

$$(5) \quad \mathcal{D}_{\text{KL}}(p_A \parallel \Phi_\theta) \leq H(\theta) + r_n(u)$$

where $\theta = |A|/n$.

Proof. For a subset $A \subseteq V_L$ of size θn , the probability of an admissible degree profile p is

$$P := \frac{\binom{n}{np(0), \dots, np(\Delta)} \prod_{j=0}^{\Delta} \binom{\Delta}{j}^{np(j)}}{\binom{\Delta n}{\Delta \theta n}}.$$

We use the following crude bound

$$\binom{n}{np_0, \dots, np_\Delta} \leq e^{nH(p)}, \quad \binom{\Delta n}{\Delta \theta n} \geq \frac{e^{\Delta n H(\theta)}}{\Delta n + 1}$$

and obtain that

$$P \leq (\Delta n + 1) e^{nH(p) - \Delta n H(\theta)} \prod_{j=0}^{\Delta} \binom{\Delta}{j}^{np(j)}.$$

On the other side, for every admissible p , it holds that $\sum_j p(j) = 1$ and $\sum_j jp(j) = \Delta\theta$. Therefore,

$$\begin{aligned}\mathcal{D}_{\text{KL}}(p \parallel \Phi_\theta) &= \sum_{j=0}^{\Delta} p(j) \ln \frac{p(j)}{\binom{\Delta}{j} \theta^j (1-\theta)^{\Delta-j}} \\ &= \Delta H(\theta) - H(p) - \sum_{j=0}^{\Delta} p(j) \ln \binom{\Delta}{j}.\end{aligned}$$

Substituting the identity into the upper bound, it holds that

$$\mathbb{P}(p_A = p) \leq (\Delta n + 1) e^{-n \mathcal{D}_{\text{KL}}(p \parallel \Phi_\theta)}.$$

By union bound and the definition of $r_n(u)$,

$$\begin{aligned}\mathbb{P}(\exists A \text{ violates (5)}) &\leq 2 \sum_{k=0}^n \binom{n}{k} (n+1)^{\Delta+1} (\Delta n + 1) \exp(-n(H(k/n) + r_n(u))) \\ &\leq 2n(\Delta n + 1)(n+1)^{\Delta+1} \exp(-nr_n(u)) \\ &\leq e^{-u}.\end{aligned}$$

Thus we conclude the lemma. □