FAST COUNTING AND SAMPLING FOR FERROMAGNETIC TWO-SPIN SYSTEMS

WEIMING FENG, HENG GUO, AND YICHUN YANG

Abstract. We introduce two new models equivalent to ferromagnetic two-spin systems: a weighted subgraph model and a random cluster type model. Using these new connections, we obtain an efficient sampling algorithm and a new randomised algorithm that efficiently approximates the partition function of ferromagnetic two-spin systems in certain parameter regimes. No efficient sampling algorithms are known before in this regime, and our new estimation algorithm runs in near-quadratic time for bounded degree graphs and in polynomial time for general graphs, improving upon the previous algorithm of Guo, Liu, and Lu (2020).

1. Introduction

Spin systems model nearest neighbour interactions. They originate from statistical physics, and are now widely studied in other areas including probability theory, machine learning, and theoretical computer science, sometimes under different names such as Markov random fields. The most important computation task in their study is to estimate the so-called partition function, which is a central quantity linking to many of the system's macroscopic properties.

A particularly important case is when the number of spin is 2. These two-state spin systems (or 2-spin systems for short) show drastically different behaviours depending on whether the interaction is repulsive (anti-ferromagnetic) or attractive (ferromagnetic). For anti-ferromagnetic systems, great progress has been made regarding partition function estimation – it is **NP**-hard beyond the tree-uniqueness threshold [Sly10, SS14, GŠV16], and has an efficient algorithm otherwise [Wei06, SST14, LLY13, CLV23, CCYZ25]. The complexity landscape for the anti-ferromagnetic case is completely mapped out.

On the other hand, ferromagnetic systems have demonstrated a more involved complexity landscape. Here the most well-known system is the Ising model, and the pioneer work of Jerrum and Sinclair [JS93] showed that efficient algorithms for the ferromagnetic Ising model exist regardless of the tree-uniqueness threshold. Their approach is direct Markov chain Monte Carlo (MCMC), and this approach was taken further by Goldberg, Jerrum, and Paterson [GJP03] for general ferromagnetic 2-spin systems under certain conditions. Later, Guo and Lu [GL18] applied Weitz's correlation decay method [Wei06] to push the algorithmic threshold even further. In certain regimes, their algorithm is almost optimal, up to a threshold beyond which the problem becomes #**BIS**-hard [LLZ14]. Here #**BIS** stands for the problem of counting independent sets in bipartite graphs, which is conjectured to be intractable to approximate [DGGJ04]. However, Guo and Lu's algorithm does not extend to the case where both spins are attracting. This is not a barrier to algorithms though, as subsequently Guo, Liu, and Lu [GLL20] successfully adapted Barvinok's method [Bar16, PR17] to this setting. These works [JS93, GJP03, GL18, GLL20] constitute the current algorithmic frontier for ferromagnetic 2-spin systems, but there is still a large gap between known algorithmic and hardness thresholds.

A main drawback of the correlation decay method and Barvinok's method is that the running time of the resulting algorithm is often a polynomial with a large exponent. In particular, the exponent for the algorithm in [GLL20] scales logarithmically in the maximum degree, and it runs within polynomial-time only for bounded degree graphs. In this paper, we address this issue by giving a new algorithm

(Weiming Feng) School of Computing and Data Science, The University of Hong Kong, Hong Kong, China

(Heng Guo) School of Informatics, University of Edinburgh, Informatics Forum, Edinburgh, EH8 9AB, United Kingdom

(Yichun Yang) School of Computer Science, Beijing Institute of Technology, Beijing, China

that runs in near-quadratic time for bounded degree graphs, and in polynomial-time when the degree is unbounded. We note that a similar speedup for [GL18] was obtained earlier this year [FGY26], which relies on a new analysis for direct MCMC. In contrast, our result is quite different. We obtain an equivalent (weighted subgraph) formulation of the problem first, and then apply MCMC to the new setting. Our new method applies to the regime in which both spin types are attractive, a setting where the techniques of [GL18, FGY26] encounter some fundamental barriers.

Another drawback of [GLL20] is that due to lack of self-reducibility, no efficient sampling algorithm is known. We also address this via introducing an intermediate random cluster type model so that we can transform a sample from the weighted subgraph model to a sample of spin systems.

Next, we formally define the problems and state our result in more detail.

1.1. **Sampling and counting for ferromagnetic two-spin systems.** Let $G = (V, E)$ be a graph with $n = |V|$ vertices and $m = |E|$ edges. A two-spin system on G is defined by an interaction matrix A and an external field vector B as follows:

$$A = \begin{pmatrix} \beta & 1 \\ 1 & \gamma \end{pmatrix}, \qquad B = \begin{pmatrix} \lambda \\ 1 \end{pmatrix}.$$

For any configuration $\sigma : V \to \{0, 1\}$, its weight is defined by

$$w_G^{\text{spin}}(\sigma; \beta, \gamma, \lambda) = \prod_{(u,v)\in E} A_{\sigma_u,\sigma_v} \prod_{v\in V} B_{\sigma_v} = \beta^{m_0(\sigma)}\gamma^{m_1(\sigma)}\lambda^{n_0(\sigma)},$$

where $m_0(\sigma)$ and $m_1(\sigma)$ are the numbers of edges with both endpoints in state 0 and in state 1, respectively, and $n_0(\sigma)$ is the number of vertices in state 0. The associated Gibbs distribution is given by $\mu^{\text{spin}}(\sigma) \propto w_G^{\text{spin}}(\sigma; \beta, \gamma, \lambda)$. The partition function, namely the normalising factor for the Gibbs distribution, is

$$Z_G^{\text{spin}}(\beta, \gamma, \lambda) := \sum_{\sigma:V\to\{0,1\}} w_G^{\text{spin}}(\sigma; \beta, \gamma, \lambda). \tag{1}$$

By symmetry between the two states, we will assume without loss of generality throughout the paper that $\beta \leqslant \gamma$. The 2-spin system is said to be ferromagnetic if $\beta\gamma > 1$ (and anti-ferromagnetic if $\beta\gamma < 1$).

Consider a ferromagnetic two-spin system with parameters β, γ, and λ such that $\beta \leqslant \gamma$. Goldberg, Jerrum, and Paterson [GJP03] first showed that a fully polynomial-time randomised approximation scheme (FPRAS) for Z_G^{spin} exists if $\lambda \leqslant \frac{\gamma}{\beta}$, by extending the MCMC approach of Jerrum and Sinclair [JS93]. Later, Guo and Lu [GL18] identified a threshold $\lambda_c = \lambda_c(\beta, \gamma) := \left(\frac{\gamma}{\beta}\right)^{\frac{\sqrt{\beta\gamma}}{\sqrt{\beta\gamma}-1}}$, and introduced an fully polynomial-time approximation scheme (FPTAS)[1] if $\lambda < \lambda_c$ with the further assumption of $\beta \leqslant 1$, based on establishing strong spatial mixing (SSM) and the correlation decay method [Wei06]. This threshold is almost optimal as Liu, Lu, and Zhang [LLZ14] showed that if $\lambda > \left(\frac{\gamma}{\beta}\right)^{\left\lfloor \frac{2\sqrt{\beta\gamma}}{\sqrt{\beta\gamma}-1} \right\rfloor/2+1}$, the problem becomes #**BIS**-hard, where approximating #**BIS** is conjectured to be intractable [DGGJ04].

On the other hand, the regime where $\beta > 1$ is still poorly understood. Here SSM does not hold any more for any constant $\lambda > 0$.[2] Thus the approach of [GL18] (as well as the recent improvement of [FGY26]) cannot be extended to this case. However, the lack of correlation decay does not block efficient algorithms, which makes the picture much less clear. Guo, Liu, and Lu [GLL20] used Barvinok's method [Bar16] to identify yet another threshold:

$$\lambda^\star = \lambda^\star(\beta, \gamma) := \left(\frac{\gamma}{\beta}\right)^{\frac{\pi}{2\tan^{-1}\sqrt{\beta\gamma-1}}}, \tag{2}$$

[1] Note that the difference between FPTAS and FPRAS is that FPTAS is deterministic.

[2] Roughly speaking, this is because there is no degree bound. In fact, even the so-called weak spatial mixing does not hold. One can verify this by considering a regular tree of sufficiently large degrees.

and introduced an FPTAS for bounded degree graphs if $\lambda < \lambda^\star$.[3]

Among all ferromagnetic 2-spin systems, the most well-understood special case is the Ising model (namely $\beta = \gamma > 1$). There are rich structures for ferromagnetic Ising models, leading to various equivalent formulations such as the even-subgraph model [vdW41] (also known as high-temperature expansion) and the random cluster model [FK72]. Using these connections, polynomial-time sampling and approximate counting algorithms are obtained [JS93, GJ18]. Furthermore, in the presence of external field $\lambda \neq 1$, there are a near-linear time sampler and a near-quadratic time FPRAS [CZ23].

Part of the reason that the parameter regime $\gamma > \beta > 1$ is not very well-understood is the lack of these equivalent formulations. Our main conceptual contribution is to fill in this gap. We introduced new weighted subgraph models and random cluster type models, and generalise previous equivalences to the whole $\gamma > \beta > 1$ regime. These new connections help us find fast approximate counting and sampling algorithms. We believe they will be useful in fully resolving the approximate counting and sampling complexity of ferromagnetic 2-spin systems in the future.

Our first result is a fast sampling algorithm for $\lambda < \lambda^\star$. In this regime, no efficient sampling algorithm is previously known. The FPTAS of [GLL20] cannot be transformed to a sampler due to lack of self-reducibility (basically, by the same reason why SSM fails).

Theorem 1. *Let $\beta, \gamma, \lambda > 0$ be parameters such that $\gamma > \beta > 1$ and $\lambda < \lambda^\star$. There exists an algorithm that, given any $\epsilon > 0$ and any graph $G = (V, E)$, outputs a random sample $\sigma \in \{0,1\}^V$ in time $O\left(\Delta^C n \log \frac{n}{\epsilon}\right)$ that is at most ϵ away from μ^{spin} in total variation distance, where $n = |V|$ is the number of vertices, Δ is the maximum degree of the graph, and $C = C(\beta, \gamma, \lambda)$ is a constant.*

To achieve Theorem 1, we use Markov chains. However, directly MCMC is not efficient here. To see this, for any constants $\beta > 1, \gamma > 1$, and $\lambda > 0$, consider a sufficiently large complete graph. It is easy to see that the all-1 and all-0 configurations contribute weights exponentially larger than mixed configurations, and therefore the mixing time for any locally defined Markov chains such as the Glauber dynamics is exponentially large. This is also why the recent work of [FGY26], which is a direct MCMC, is restricted to the case of $\beta \leqslant 1 < \gamma$.

Our solution to bypass this issue is via the aforementioned new connections. We consider an equivalent weighted subgraph formulation, which we will explain in Section 1.2 shortly. We show that Glauber dynamics for the weighted subgraphs mixes within the time bound of Theorem 1. However, there appears to be no easy way to transform this subgraph sample to a spin sample. We instead introduce an intermediate random cluster type model. For the three models, we not only establish equivalence among their partition functions, but also show new couplings among their Gibbs distributions. With the couplings, we first transform the subgraph sample to a random cluster sample, and then to a spin sample, to establish Theorem 1. Details can be found in Section 3.2.

Using the sampler above and a simple λ-annealing reduction, we also obtain the following FPRAS.

Theorem 2. *Let $\beta, \gamma, \lambda > 0$ be parameters such that $\gamma > \beta > 1$ and $\lambda < \lambda^\star$. There exists an FPRAS that, given any $\epsilon > 0$ and any graph $G = (V, E)$, outputs a random quantity $\widehat{Z}$ in time $O\left(\Delta^C \cdot \frac{n^2}{\epsilon^2} \cdot \log^2 \frac{n}{\epsilon}\right)$ such that $\widehat{Z} \in [Z(1-\epsilon), Z(1+\epsilon)]$ with probability at least $\frac{2}{3}$, where $Z = Z_G^{spin}(\beta, \gamma, \lambda)$ is the partition function, $n = |V|$ is the number of vertices, Δ is the maximum degree of the graph, and $C = C(\beta, \gamma, \lambda)$ is a constant.*

The algorithm in [GLL20] applies to parameters of the same regime as in Theorem 2, but it runs in time $O\left(\left(\frac{n}{\varepsilon}\right)^{O(\log \Delta)}\right)$, which is no longer a polynomial if the graph has unbounded degrees. This type of running time is typical for FPTASes [Wei06, Bar16, PR17]. Utilising randomness, the algorithm in Theorem 2 runs in polynomial-time for any graph, and in near-quadratic time for bounded degree graphs. We note that near-quadratic time is the best known running time for partition function estimation in many settings

[3] It can be verified that $\lambda^\star < \lambda_c$ [GLL20].

[SVV09, Kol18], although this barrier has been broken very recently when the parameters are lower than the computational transition threshold [AFF+25, CCLZ26]. Their techniques do not apply in our setting, and it is not clear if sub-quadratic time can be achieved here.

Theorem 1 and Theorem 2 cover only the case of $\gamma > \beta > 1$. When $\beta \leqslant 1 \leqslant \gamma$ but $\beta\gamma > 1$, we refer the interested reader to a recent work by Feng, Guo, and Yang [FGY26]. They showed an $\widetilde{O}(n^2)$ mixing time bound for Glauber dynamics in the spin system from the all 1 configuration, when $\lambda < \lambda_c$, and an $\widetilde{O}(n)$ mixing time bound when $\lambda < \sqrt{\gamma/\beta}$. For the Ising model ($\beta = \gamma > 1$), there are the classical sampling algorithm for $\lambda = 1$ by Jerrum and Sinclair [JS93] and the near-linear time sampler for $\lambda \neq 1$ by Chen and Zhang [CZ23]. Approximate counting algorithms in all these settings can be obtained similarly via simulated annealing.

Since the weighted subgraph model is the key to our whole argument, we explain it next.

1.2. **The subgraph world.** We introduce a new weighted subgraph (WSG) model and relate its partition function to that of the two-spin system. This is a generalisation to the classical subgraph world models [JS93]. Let $G = (V, E)$ be a graph. For any subset of edges $S \subseteq E$ and any vertex v, let $\deg_S(v)$ denote the number of edges in S incident to v; in other words, $\deg_S(v)$ is the degree of v in the subgraph (V, S). The model involves a number of parameters. Each edge $e \in E$ is associated with a parameter $\theta_e > 0$, and each vertex $v \in V$ is associated with $\lambda_v > 0$. Denote $\boldsymbol{\theta} = (\theta_e)_{e\in E}$ and $\boldsymbol{\lambda} = (\lambda_v)_{v\in V}$. Let $\rho > 0$ be another parameter. In our applications later, we will have $\rho\lambda_v < 1$ for all $v \in V$. For any subset of edges $S \subseteq E$, the weight of S is given by

$$w_G^{\mathrm{wsg}}(S; \boldsymbol{\theta}, \boldsymbol{\lambda}, \rho) = \prod_{e\in S} \theta_e \prod_{v\in V} (1 + \lambda_v(-\rho)^{\deg_S(v)}). \tag{3}$$

The partition function of the WSG model is then given by

$$Z_G^{\mathrm{wsg}}(\boldsymbol{\theta}, \boldsymbol{\lambda}, \rho) = \sum_{S\subseteq E} w_G^{\mathrm{wsg}}(S; \boldsymbol{\theta}, \boldsymbol{\lambda}, \rho).$$

Note that in the definitions above, we allow non-uniform weights for later convenience.

The following equivalence is shown via the holographic transformation [Val08] in Section 3.1.

Theorem 3. *Consider a two-spin system on a graph $G = (V, E)$ with parameters β, γ, λ such that $\gamma \neq 1$, $\beta + \gamma \neq 2$, and $\beta\gamma \neq 1$. Let $\rho = \frac{\beta-1}{\gamma-1}$ and $\theta = \frac{(\gamma-1)^2}{\beta\gamma-1}$, and let $\mathbf{1}_V = (1)_{v\in V}$ and $\mathbf{1}_E = (1)_{e\in E}$ denote the V-dimensional and E-dimensional all-ones vectors, respectively. Then*

$$Z_G^{\mathit{spin}}(\beta, \gamma, \lambda) = (1 - \rho\theta)^{-|E|} \cdot Z_G^{\mathit{wsg}}(\theta\mathbf{1}_E, \lambda\mathbf{1}_V, \rho).$$

Theorem 3 applies to a wide range of parameters due to its algebraic nature. It is easy to check that the conditions of Theorem 3 are satisfied in our application for $\gamma > \beta > 1$ and $\beta\gamma > 1$.

Remark (Comparison with existing subgraph-world models). It is well known that the ferromagnetic Ising model ($\beta = \gamma$) is equivalent to certain subgraph-world models [JS93, FGW23]. Our result in Theorem 3 generalises this equivalence to non-Ising ferromagnetic two-spin systems with $\beta < \gamma$. The Ising model without fields, namely when $\beta = \gamma$ and $\lambda = 1$, is equivalent to the *even-subgraph* model, where the subgraph (V, S) must satisfy that $\deg_S(v)$ is even for every $v \in V$. In this case, our parameter $\rho = 1$, and only even subgraphs have positive weights in (3). When the external field is not 1, say $\lambda < 1$, the Ising model is equivalent to a subgraph world in which each odd-degree vertex incurs a penalty; this is also captured by (3) with $\rho = 1$. For general ferromagnetic two-spin systems with $\beta < \gamma$, our model imposes a penalty based on the degree of the vertex in the subgraph. We believe this new subgraph-world model is of independent interest.

The main ingredient of our algorithm is an efficient sampling algorithm for the Gibbs distribution of the WSG model. Consider a WSG model on $G = (V, E)$ with parameters $\boldsymbol{\theta}, \boldsymbol{\lambda}, \rho$ that correspond to a ferromagnetic 2-spin system. When $\lambda < \lambda^\star$, the weight function $w_G^{\mathrm{wsg}}(S;\boldsymbol{\theta},\boldsymbol{\lambda},\rho)$ in (3) is positive for all $S \subseteq E$ (see Lemma 11). Let $\mu = \mu^{\mathrm{wsg}}_{G,\boldsymbol{\theta},\boldsymbol{\lambda},\rho}$ over 2^E be the Gibbs distribution induced by the weighted subgraph model. For any subset of edges $S \subseteq E$, the probability of S is given by

$$\mu(S) = \frac{w_G^{\mathrm{wsg}}(S;\boldsymbol{\theta},\boldsymbol{\lambda},\rho)}{Z_G^{\mathrm{wsg}}(\boldsymbol{\theta},\boldsymbol{\lambda},\rho)}. \tag{4}$$

We often view μ as a distribution over E-dimensional Boolean vectors. For any $\sigma \in \{0,1\}^E$, let $S_\sigma \subseteq E$ denote the subset of edges corresponding to σ. Then we write

$$\forall \sigma \in \{0,1\}^E, \quad \mu(\sigma) = \mu(S_\sigma) = \frac{w_G^{\mathrm{wsg}}(S_\sigma;\boldsymbol{\theta},\boldsymbol{\lambda},\rho)}{Z_G^{\mathrm{wsg}}(\boldsymbol{\theta},\boldsymbol{\lambda},\rho)}.$$

To generate a random sample from the distribution μ, we employ the standard Glauber dynamics. Starting from an arbitrary initial configuration $X_0 \in \{0,1\}^E$, in each step $t \geqslant 0$ we update X_t as follows:

- sample an edge $e \in E$ uniformly at random;
- sample $X_e \sim \mu_e^{X_{E-e}}$, where $\mu_e^{X_{E-e}}$ is the conditional distribution on e given X_{E-e}, and update X_t by changing the state of e to X_e to get X_{t+1}.

It is easy to simulate Glauber dynamics: since the weight in (3) is a product of terms, each of which depends only on a single edge or a single vertex, the conditional distribution $\mu_e^{X_{E-e}}$ can be computed in time $O(\Delta)$, where Δ is the maximum degree of the graph.

It is well known that the stationary distribution of the Glauber dynamics is μ. To analyse the efficiency of Glauber dynamics, we study its *mixing time*. Let $(X_t)_{t\geqslant 0}$ be the random configurations generated by Glauber dynamics over time. Given any $\epsilon > 0$, the mixing time is the number of steps required for the chain to reach its stationary distribution up to total variation distance ϵ. Formally,

$$T_{\mathrm{mix}}(\epsilon) = \max_{X_0 \in \{0,1\}^E} \min\{t \geqslant 0 \mid D_{\mathrm{TV}}(X_t,\mu) < \epsilon\},$$

where $D_{\mathrm{TV}}(X_t,\mu) = \frac{1}{2}\sum_{\sigma\in\{0,1\}^E} |\Pr[X_t = \sigma] - \mu(\sigma)|$ is the *total variation distance* between (the law of) X_t and μ. Our main technical contribution is the following mixing time bound for the WSG model. Recall that $\lambda^\star = \lambda^\star(\beta,\gamma) = \left(\frac{\gamma}{\beta}\right)^{\frac{\pi}{2\tan^{-1}\sqrt{\beta\gamma-1}}}$ is the threshold in (2).

Theorem 4. *Let $\beta,\gamma,\lambda > 0$ be constants such that $\gamma > \beta > 1$ and $\lambda < \lambda^\star$. Let $\theta = \frac{(\gamma-1)^2}{\beta\gamma-1}$ and $\rho = \frac{\beta-1}{\gamma-1}$. Then, the weight function $w_G^{\mathrm{wsg}}(S;\theta\mathbf{1}_E,\lambda\mathbf{1}_V,\rho) \geqslant 0$ for all $S \subseteq E$. Let μ denote the Gibbs distribution induced by the weighted subgraph model on a graph $G = (V,E)$ with parameters $\theta\mathbf{1}_E$, $\lambda\mathbf{1}_V$, and ρ. Then the mixing time of the Glauber dynamics on μ is at most $O\left(\Delta^C \cdot n \log \frac{n}{\epsilon}\right)$, where $n = |V|$, Δ is the maximum degree of the graph, and $C = C(\beta,\gamma,\lambda)$ is a constant.*

For comparison, in the Ising model, where $\beta = \gamma > 1$, the analogous threshold is $\lambda^\star = 1$. When $\lambda < 1$, Chen and Zhang [CZ23] proved a fast mixing time bound for the WSG dynamics induced by the Ising model. However, our method is based on stability of polynomials, which is completely different from the coupling-based method in [CZ23]. Their coupling crucially relies on the fact that, in the weighted subgraph model induced by the Ising model, every vertex of odd degree receives the same *uniform* penalty. In contrast, in our setting with $\beta < \gamma$, vertices of different degrees receive different penalties in (3). Chen and Gu [CG24] developed a more general coupling-based technique, but our setting does not meet their requirements either.

2. Proof outline of the mixing time

In this section we outline the proof of Theorem 4. Under the parameter assumption in the theorem, the positivity of the weight function is guaranteed can be verified by a straightforward calculation, which is deferred to Section 6.1. Assuming the distribution μ is well-defined, we outline the proof of the mixing time bound.

We use the following notations. For any subset of edges $\Lambda \subseteq E$ and any pinning $\tau \in \{0,1\}^{E\setminus\Lambda}$, we write μ^τ for the conditional distribution of μ given τ. In other words, in μ^τ, the edges in $E \setminus \Lambda$ are pinned by τ and the edges in Λ follow the conditional distribution. Moreover, we write μ_S for the marginal distribution of $S \subseteq E$ under μ. When $S = \{e\}$, we may also write μ_e instead. In particular, μ_e^τ means the marginal distribution of e under the conditional distribution μ^τ.

2.1. **Mixing of Glauber dynamics via spectral independence.** We prove the mixing time bound via the spectral independence technique introduced by Anari, Liu, and Oveis Gharan [ALO24]. Let μ be a distribution over $\{0,1\}^E$. For any two variables $i, j \in E$, the influence from i to j is defined as

$$\Psi_\mu(i,j) = \Pr_{X\sim\mu}[X_j = 1 \mid X_i = 1] - \Pr_{X\sim\mu}[X_j = 1 \mid X_i = 0].$$

Definition 5 (spectral independence [ALO24])**.** The distribution μ is said to be η-spectrally independent if the maximum eigenvalue of the influence matrix Ψ_μ is at most η.

Definition 6 (marginal lower bound [CLV21])**.** A distribution μ over $\{0,1\}^E$ has a marginal lower bound $b > 0$ if, for any subset of edges $\Lambda \subseteq E$, any pinning $\tau \in \{0,1\}^{E\setminus\Lambda}$, and any edge $e \in \Lambda$,

$$\forall c \in \{0,1\}, \quad \mu_e^\tau(c) = \Pr_{X\sim\mu}[X_e = c \mid X_{E\setminus\Lambda} = \tau] \geqslant b.$$

Using the definition of the weighted subgraph model in (3), the following lemma can be established by a straightforward calculation. The proof is deferred to Section 6.2.

Lemma 7. *Let $\theta, \rho, \lambda > 0$ be three constants such that $\rho < 1$ and $\lambda\rho < 1$. The weighted subgraph model on any graph G with parameters $\theta\mathbf{1}_E$, $\lambda\mathbf{1}_V$, and ρ has a marginal lower bound $b = b(\theta, \lambda, \rho) > 0$.*

In particular, Chen, Liu, and Vigoda [CLV21, Theorem 1.12] showed the following. For a Gibbs distribution μ on a graph $G = (V, E)$ with maximum degree Δ, if every conditional distribution induced by μ is η-spectrally independent, and μ has a constant marginal lower bound b, then the mixing time of the Glauber dynamics is bounded by $O\left(\Delta^C \cdot n \cdot (\log\log \frac{1}{\mu_{\min}} + \log\frac{1}{\epsilon})\right)$ for some $C > 0$ depending on η and b, where $\mu_{\min} = \min_{\sigma\in\{0,1\}^E} \mu(\sigma)$.

We summarise and specialise this result for the WSG model in the next lemma. In particular, the constant marginal lower bound is guaranteed by Lemma 7, which is independent from Δ.

Lemma 8. *Let $\beta, \gamma, \lambda > 0$ be constants such that $\gamma > \beta > 1$ and $\lambda < \lambda^\star$. Let $\theta = \frac{(\gamma-1)^2}{\beta\gamma-1}$ and $\rho = \frac{\beta-1}{\gamma-1}$. Let μ denote the Gibbs distribution induced by the WSG model on a graph $G = (V, E)$ with parameters $\theta\mathbf{1}_E$, $\lambda\mathbf{1}_V$, and ρ. Suppose there exists a constant $\eta > 0$ such that for any $\Lambda \subseteq E$ and any $\tau \in \{0,1\}^{E\setminus\Lambda}$, the conditional distribution μ^τ is η-spectrally independent. Then the mixing time of the Glauber dynamics on μ is at most $O\left(\Delta^C \cdot n \log\frac{n}{\epsilon}\right)$, where $C = C(\beta, \gamma, \lambda, \eta)$ is a constant.*

2.2. **Spectral independence via the stability.** By Lemma 8, it remains to show that every conditional distribution μ^τ is η-spectrally independent, for every subset $\Lambda \subseteq E$ and every pinning $\tau \in \{0,1\}^{E\setminus\Lambda}$. We show this by showing stability on the complex plane for a suitable polynomial, and then applying the techniques from [CLV24].

For any configuration $\sigma \in \{0,1\}^E$, let S^σ denote the subset of edges selected by σ: $S^\sigma = \{e \in E \mid \sigma_e = 1\}$. By the definition of the weighted subgraph model in (3), the weight of the configuration σ is $\prod_{e\in S^\sigma} \theta_e \prod_{v\in V}\left(1 + \lambda_v(-\rho)^{\deg_{S^\sigma}(v)}\right)$. We separate out the vertex contribution and define

$$\forall \sigma \in \{0,1\}^E, \quad w(\sigma) := \prod_{v\in V}\left(1 + \lambda_v(-\rho)^{\deg_{S^\sigma}(v)}\right). \tag{5}$$

Given a pinning $\tau \in \{0,1\}^{E\setminus\Lambda}$, we treat the edge activities on the unpinned edges as complex variables z_e and define the conditional partition function polynomial

$$Z^{\tau}(\boldsymbol{z}) = \sum_{\sigma\in\{0,1\}^{E}:\sigma_{E\setminus\Lambda}=\tau} w(\sigma) \prod_{e\in\Lambda:\sigma_e=1} z_e. \tag{6}$$

The product only ranges over edges in Λ, since the edges in $E \setminus \Lambda$ have already been fixed by τ. Thus $Z^{\tau}(\boldsymbol{z})$ is a polynomial in the variables z_e, $e \in \Lambda$. When the pinning is empty, we write $Z(\boldsymbol{z})$ for $Z^{\emptyset}(\boldsymbol{z})$. Substituting $z_e = \theta_e$ recovers the conditional distribution on the unpinned variables:

$$\forall \sigma \in \{0,1\}^{E} \text{ with } \sigma_{E\setminus\Lambda} = \tau, \quad \mu^{\tau}(\sigma) := \frac{w(\sigma)\prod_{e\in\Lambda:\sigma_e=1}\theta_e}{Z^{\tau}(\boldsymbol{\theta})}.$$

Again, when the pinning is empty, $\mu^{\emptyset} = \mu$.

The polynomial in (6) allows us to complexify the edge activities and prove spectral independence through polynomial stability, following [CLV24]. Throughout the paper, $\mathbb{C}$ denotes the complex plane.

Definition 9 (stability of polynomial)**.** For regions $\Gamma_i \subseteq \mathbb{C}$, $i \in [n]$, a multivariate polynomial $p \in \mathbb{C}[x_1,\dots,x_n]$ is $(\prod_{i=1}^{n}\Gamma_i)$-stable if $p(x_1,\dots,x_n) \neq 0$ whenever $x_i \in \Gamma_i$ for all $i \in [n]$.

We specialise [CLV24, Theorem 3.2] to our setting as follows.

Proposition 10. *Let E be a set of variables and $w : \{0,1\}^{E} \to \mathbb{R}$ be a weight function. For any $e \in E$, let $\Gamma_e \subset \mathbb{C}$ be a non-empty open connected region, and let $\boldsymbol{\theta} : E \to \mathbb{R}$ satisfy $\theta_e \in \mathbb{R}_+ \cap \Gamma_e$ for every $e \in E$. Suppose that, for every pinning τ, the conditional partition function $Z^{\tau}(\boldsymbol{z})$ is $\left(\prod_{e\in\Lambda}\Gamma_e\right)$-stable. Then the distribution μ^{τ} is spectrally independent with constant*

$$\eta = \frac{2}{b\delta^2},$$

where b is the marginal lower bound of μ^{τ}, $\delta = \min_{e\in E}\frac{1}{\theta_e}\operatorname{dist}(\theta_e, \partial\Gamma_e)$, $\partial\Gamma_e$ is the boundary of Γ_e, and $\operatorname{dist}(x, \partial\Gamma_e) := \inf_{z\in\partial\Gamma_e}|x - z|$.

We apply Proposition 10 to the WSG model. Once again, recall that our setting is $\gamma > \beta > 1$, $\rho = \frac{\beta-1}{\gamma-1}$, $\theta = \frac{(\gamma-1)^2}{\beta\gamma-1}$, and $\lambda < \lambda^{\star}(\beta,\gamma)$. Let μ be the Gibbs distribution induced by the WSG model on a graph $G = (V, E)$ with parameters $\theta\mathbf{1}_E$, $\lambda\mathbf{1}_V$, and ρ. We will use Proposition 10 to prove spectral independence for every conditional distribution μ^{τ}.

The following self-reduction lets us handle all pinnings uniformly. Fix a pinning $\tau \in \{0,1\}^{E\setminus\Lambda}$. Let $G^{\tau} = (V,\Lambda)$ be the subgraph obtained by deleting the pinned edges in $E \setminus \Lambda$. For each vertex $v \in V$, let k_v be the number of pinned edges incident to v that are fixed to 1 by τ. We absorb the contribution of these pinned edges into the external field by setting

$$\lambda_v^{\tau} = \lambda \cdot (-\rho)^{k_v}\,.$$

The following technical lemma is verified in Section 6.1.

Lemma 11. *For any integer $k \geqslant 0$, it holds that $-1 < -\rho\lambda^{\star} \leqslant \lambda \cdot (-\rho)^{k} \leqslant \lambda$.*

Given the pinning τ, for every $S \subseteq \Lambda$, the conditional weight of the configuration S is, up to the non-zero contribution of the pinned edges, equal to

$$\theta^{|S|}\prod_{v\in V}\left(1 + \lambda_v^{\tau}(-\rho)^{\deg_S(v)}\right).$$

By Lemma 11, the above weight is still non-negative. Hence the conditional distribution μ^{τ} projected onto Λ is the Gibbs distribution of a weighted subgraph model on the subgraph G^{τ} with parameters $\theta\mathbf{1}_{\Lambda}$, $(\lambda_v^{\tau})_{v\in V}$, and ρ. We will use this self-reduction to handle all the conditional distributions.

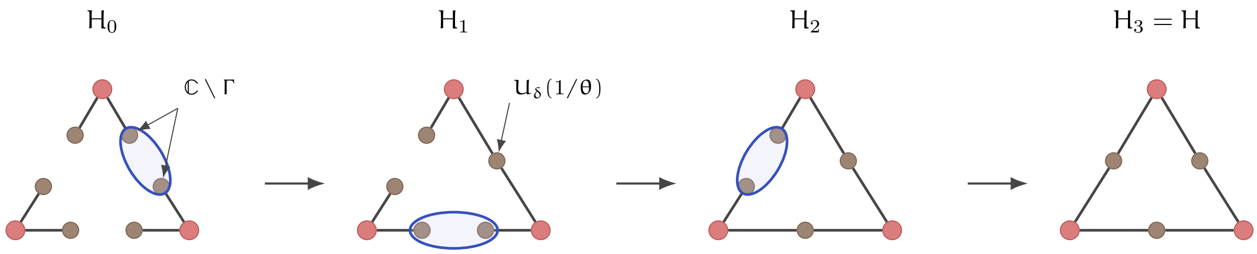


FIGURE 1. Illustration of our contraction procedure for a 3-cycle C_3. Red vertices correspond to original vertices of the graph, and brown vertices correspond to edge variables. We start from H_0, the disjoint union of three disjoint stars. At each step, the blue highlighted pair of brown vertices is contracted into a single vertex. The variables of original brown vertices are *safe* in $\mathbb{C}\setminus\Gamma$, while the contracted brown vertex is *safe* in $\mathcal{U}_\delta(1/\theta)$. After three contractions, we obtain $H_3 = H_{C_3}$.

It remains to establish the required stability condition in Proposition 10. For $x \in \mathbb{C}$ and $\delta > 0$, write $\mathcal{U}_\delta(x) := \{z \in \mathbb{C} \mid |z - x| < \delta\}$ for the open disk of radius δ centered at x.

Theorem 12. *Fix three constants $\gamma > \beta > 1$ and $0 < \lambda < \lambda^*(\beta,\gamma)$. Let $\rho = \frac{\beta-1}{\gamma-1}$ and $\theta = \frac{(\gamma-1)^2}{\beta\gamma-1}$. Let $G = (V, E)$ be a graph and $\boldsymbol{\lambda} = (\lambda_v)_{v\in V}$ be parameters such that $-\rho\lambda \leqslant \lambda_v \leqslant \lambda$ for all $v \in V$. Let $w : \{0,1\}^E \to \mathbb{R}$ be the weight function defined in (5) with respect to graph G and parameters $\boldsymbol{\lambda}$, ρ. There exists a constant $\delta = \delta(\beta,\gamma,\lambda) > 0$ such that the partition function $Z(\boldsymbol{z})$ defined in (6) is $\prod_{e\in E}\mathcal{U}_\delta(\theta)$-stable.*

This theorem is enough for the stability hypothesis in Proposition 10. Indeed, after the self-reduction above, each polynomial $Z^\tau(\boldsymbol{z})$ appearing in Proposition 10 is exactly the partition function in Theorem 12 for the graph G^τ and fields $(\lambda_v^\tau)_{v\in V}$. This is because, in the definition of $Z^\tau(\boldsymbol{z})$ in (6), only the edge activities of the unpinned edges in Λ appear. Thus proving Theorem 12 for arbitrary graphs and all fields in the interval $(-\rho\lambda^\star, \lambda^\star)$ establishes the required stability for all conditional distributions. Theorem 4 is a consequence of combining Lemma 8, Proposition 10, and Theorem 12.

Theorem 12 is our most technically involved ingredient and is shown in Section 4. Here we give a brief overview of the proof. The overall strategy is via the Asano contraction [Asa70, Rue71], similar to previous zero-freeness results for Holant problems [GLLZ21]. Let $G = (V, E)$ be a graph with $|V| = n$ and $|E| = m$. Note that the variables in $Z_w^\tau(\boldsymbol{z})$ are on the edges. Let H_G be the (bipartite) vertex-edge incidence graph of G (see Definition 14). We construct a sequence of graphs $H_0, \ldots, H_m$, where $H_m = H_G$. The initial graph H_0 is a collection of disjoint star graphs. For each $v \in V$, construct a star whose center is v and the leaves correspond to all edges adjacent to v. The union of all these stars is H_0. Thus, in H_0, each edge $e = (u, v) \in E$ corresponds two leaves connecting to u and v, respectively. Call these leaves half-edges, and variables are attached to the half-edges. Order edges in E as $e_1, \ldots, e_m$. Then given H_i, contract the two half edges corresponding to e_{i+1} to obtain H_{i+1}. The two corresponding variables are also identified. After processing all edges, we recover $H_m = H_G$. Figure 1 illustrates this procedure for a 3-cycle.

To prove the desired stability, there are two main ingredients. The first is the stability of the initial graph H_0. Since H_0 is a disjoint union of stars, the stability of the partition function follows from handling each star separately. We carefully choose a region Γ so that the partition function is non-zero whenever no variable lies in Γ. The second ingredient is the stability-preserving property of the contraction step. Namely, assuming that the partition function of H_i is $(\mathbb{C}\setminus\Gamma)\times(\mathbb{C}\setminus\Gamma)$-stable for the pair of contracted half edge variables, we prove that the new partition function for H_{i+1} after contraction remains stable when the contracted variable lies in a small neighborhood $\mathcal{U}_\delta(1/\theta)$ around $1/\theta$. Iterating this argument propagates stability from the initial disjoint union of stars H_0 all the way to the final graph H_m. Consequently, the WSG partition function for G is $\prod_{e\in E}\mathcal{U}_\delta(\theta)$-stable.

The most essential part of this argument is the choice of Γ. Our choice is actually inspired by the zero-freeness result of [GLL20]. Although their result concerns the spin model and we are working in the WSG model, we noticed that after certain variable transformations, the initial stability for H_0 resembles the contraction step in [GLL20]. Then, with the same variable transformation, our contraction step resembles the initial stability in [GLL20]. This is an interesting duality regarding the contraction argument for zero-freeness under holographic transformations, and is also why our Theorem 2 works up to the same λ threshold as in [GLL20]. However, as we need stability for a constant neighbourhood around θ, we also need to strengthen various intermediate results in [GLL20].

In addition, there is one more wrinkle to our proof – the zero-freeness result of [GLL20] is optimised and only works for degrees $\geqslant 2$. The way they handle degree 1 vertices (namely leaves) is to pre-process them first. However, since pinning edges may produce new leaves, this strategy does not work for us. To this end, we identify *bad* leaves, whose external fields are positive and cause trouble to an avoidance lemma (Lemma 34). Observe that any graph can be decomposed into a degree $\geqslant 2$ core and trees appended to the core (see Figure 2 for an illustration). Our strategy is to recursively prune bad leaves and reduce the original graph G to an equivalent bad-leaf reduced graph, whose external fields are updated to some $\boldsymbol{\lambda}'$. Throughout this recursive process, we must ensure that the new fields $\boldsymbol{\lambda}'$ remain in a complex region where the stability argument applies. To control this recursion, we notice that the external fields are updated in a similar fashion to the tree recursion of Weitz [Wei06]. Thus, we use the potential function from [GL18] to design the admissible complex region for the external fields $\boldsymbol{\lambda}'$. Intuitively, we show that the perturbations due to the pruning process do not drift the external fields $\boldsymbol{\lambda}'$ too far away from the real axis, and that they remain inside the region required by the contraction argument. Once we obtain a graph without bad leaves, the zero-freeness argument outlined above applies, which completes the proof of Theorem 12.

3. Equivalent formulations

In this section we show equivalence between the two-spin system and the weighted subgraph model, introduce the random cluster type model, and show its equivalence to the other two models.

3.1. **Weighted subgraphs.** We use the holographic transformation [Val08] to prove the equivalence between the two-spin system and the weighted subgraph model. Let $d \geqslant 1$ be an integer, and let $f : \{0,1\}^d \to \mathbb{C}$ be a function, where $\mathbb{C}$ denotes the complex plane. We write $f = (f_{x_0}, f_{x_1}, \ldots, f_{x_{2^d-1}})$ as either a column or a row vector, where for each $0 \leqslant i < 2^d$, $x_i \in \{0,1\}^d$ is the binary representation of i. In the symmetric case, where f is invariant under permutations of the bits of x_i, we write f succinctly as $f = [f_0, f_1, \ldots, f_d]$, where for each $0 \leqslant j \leqslant d$, f_j is the common value of f_x over all binary strings $x \in \{0,1\}^d$ with Hamming weight $|x| = j$. This is also called the signature of f.

Given a bipartite graph $H = (V_1 \uplus V_2, E)$, let $\mathcal{F} = (f_v)_{v \in V_1}$ and $\mathcal{G} = (g_u)_{u \in V_2}$ denote the functions on the vertices of H. The Holant is defined as

$$\mathrm{Holant}(H; \mathcal{F} \mid \mathcal{G}) = \sum_{\sigma : E \to \{0,1\}} \prod_{v \in V_1} f_v(\sigma_{E(v)}) \prod_{u \in V_2} g_u(\sigma_{E(u)}),$$

where $E(v)$ and $E(u)$ are the edges incident to v and u, respectively.

Let $\mathbf{T} \in \mathbb{C}^{2\times 2}$ be an invertible matrix, and let $f : \{0,1\}^d \to \mathbb{C}$ be a signature function. We write $\mathbf{T}f = \mathbf{T}^{\otimes d} f$ (resp. $f\mathbf{T} = f\mathbf{T}^{\otimes d}$) for the transformed function when f is viewed as a column (resp. row) vector. Given $\mathrm{Holant}(H; \mathcal{F} \mid \mathcal{G})$ on a bipartite graph $H = (V_1 \uplus V_2, E)$, view all $f_v \in \mathcal{F}$ as row vectors and all $g_u \in \mathcal{G}$ as column vectors. Define the transformed functions as $\mathcal{F}\mathbf{T} = \{f_v \mathbf{T}\}_{v \in V_1}$ and $\mathbf{T}^{-1}\mathcal{G} = \{\mathbf{T}^{-1} g_u\}_{u \in V_2}$, where $\mathbf{T}^{-1}$ is the inverse of $\mathbf{T}$. The following equivalence holds.

Theorem 13 (Valiant's Holant theorem [Val08])**.** $\mathrm{Holant}(H; \mathcal{F}\mathbf{T} \mid \mathbf{T}^{-1}\mathcal{G}) = \mathrm{Holant}(H; \mathcal{F} \mid \mathcal{G})$.

Next, we represent the two-spin system as a Holant instance via the *half-edge* decomposition. Consider a ferromagnetic two-spin system on a graph $G = (V, E)$ defined by parameters $\beta, \gamma, \lambda > 0$. We define the vertex-edge incidence bipartite graph H_G of G by the following.

Definition 14. Given $G = (V, E)$, the vertex-edge incidence graph $H_G = (V_1 \uplus V_2, E_H)$ is a bipartite graph whose left part $V_1 = V$ consists of all vertices of G and whose right part $V_2 = E$ consists of all edges of G. For each edge $e = \{u, v\} \in E$ of G, we add two edges $e_u = \{u, e\}$ and $e_v = \{v, e\}$ to E_H. We call e_u and e_v the half-edges of e.

Note that the degree of each $v \in V_1$ in H_G equals the degree of v in G, which we denote by $d_v = \deg_G(v) = \deg_{H_G}(v)$, while the degree of each $e \in V_2$ in H_G is exactly 2. For each vertex $v \in V_1$, we define the row vector $f_v : \{0,1\}^{d_v} \to \mathbb{C}$ by

$$f_v = \lambda\,(1,0)^{\otimes d_v} + (0,1)^{\otimes d_v} = [\lambda,\ \underbrace{0,\ldots,0}_{(d_v-1)\text{ zeros}},1],$$

In words, if all half-edges incident to v are in state 0, then the value of f_v is λ; if all half-edges incident to v are in state 1, then the value of f_v is 1; otherwise, the value of f_v is 0. For each edge $e = \{u, v\} \in E$ of G, we define the column vector $g_e : \{0,1\}^2 \to \mathbb{C}$ by

$$g_e = (\beta, 1, 1, \gamma)^{\mathsf{T}} = [\beta, 1, \gamma].$$

In words, if the two half-edges of e are both in state 0, then the value of g_e is β; if they are both in state 1, then the value of g_e is γ; otherwise, the value of g_e is 1. We define

$$\mathcal{F} = \{f_v\}_{v\in V_1} \quad \text{and} \quad \mathcal{G} = \{g_e\}_{e\in V_2}. \tag{7}$$

This construction encodes the two-spin system as a Holant instance. Indeed, each spin configuration $\sigma \in \{0,1\}^V$ of G corresponds to the assignment in H_G that gives every half-edge incident to a vertex v the spin value σ_v. Under this correspondence, the vertex function f_v contributes the external field at v, while the edge function g_e contributes the interaction along $e = \{u, v\}$. Comparing the weights term by term, the following proposition is immediate from the definitions of f_v and g_e.

Proposition 15. *The partition function $Z_G^{\mathrm{spin}}(\beta, \gamma, \lambda)$ of the two-spin system on G is given by*

$$Z_G^{\mathrm{spin}}(\beta, \gamma, \lambda) = \mathrm{Holant}(H_G; \mathcal{F} \mid \mathcal{G}),$$

where $\mathcal{F}$ and $\mathcal{G}$ are defined in (7).

We now apply the holographic transformation to the Holant instance of the two-spin system and compute the transformed signatures. Since $\gamma \neq 1$, define

$$\mathbf{T} := \begin{pmatrix} 1 & -\rho \\ 1 & 1 \end{pmatrix}, \qquad \text{where } \rho = \frac{\beta - 1}{\gamma - 1}. \tag{8}$$

Also, as $\beta + \gamma \neq 2$, $\mathbf{T}$ is invertible. For any vertex $v \in V_1$ of H_G, the transformed vertex signature is

$$f_v \mathbf{T}^{\otimes d_v} = \lambda\,(1,-\rho)^{\otimes d_v} + (1,1)^{\otimes d_v} = [1+\lambda, 1-\lambda\rho, \ldots, 1+\lambda(-\rho)^{d_v}]. \tag{9}$$

Thus the transformed vertex signature depends only on the number of incident transformed half-edges in state 1: its value on assignments of Hamming weight k is $1 + \lambda(-\rho)^k$. For the edge signature, recall that $\theta = \frac{(\gamma-1)^2}{\beta\gamma-1}$. Using

$$\mathbf{T}^{-1} = \frac{1}{1+\rho}\begin{pmatrix} 1 & \rho \\ -1 & 1 \end{pmatrix},$$

a direct calculation gives

$$(\mathbf{T}^{-1})^{\otimes 2} g_e = \frac{1}{(1+\rho)^2}\begin{pmatrix} \beta+2\rho+\gamma\rho^2 \\ -(\beta-1)+\rho(\gamma-1) \\ -(\beta-1)+\rho(\gamma-1) \\ \beta+\gamma-2 \end{pmatrix} = \frac{1}{1-\rho\theta}\begin{pmatrix} 1 \\ 0 \\ 0 \\ \theta \end{pmatrix}. \tag{10}$$

Thus the transformed edge signature is an *equality* signature on the two transformed half-edges, with weight 1 on state 00 and weight θ on state 11, multiplied by the common factor $\frac{1}{1-\rho\theta}$. This is the exact reason why we set the parameter ρ in (8). Since every pair of half-edges are forced to be equal, by mapping the half-edges to the edges, this exactly gives the weighted subgraph model on the original graph G. Therefore, by Theorem 13 and Proposition 15,

$$\begin{aligned} Z_G^{\mathrm{spin}}(\beta,\gamma,\lambda) &= \mathrm{Holant}(H_G; \mathcal{F} \mid \mathcal{G}) = \mathrm{Holant}(H_G; \mathcal{F}\mathbf{T} \mid \mathbf{T}^{-1}\mathcal{G}) \\ &= (1-\rho\theta)^{-|E|} \sum_{S\subseteq E} \theta^{|S|} \prod_{v\in V} \left(1+\lambda(-\rho)^{\deg_S(v)}\right) \\ &= (1-\rho\theta)^{-|E|} Z_G^{\mathrm{wsg}}(\theta\mathbf{1}_E, \lambda\mathbf{1}_V, \rho), \end{aligned}$$

where $\deg_S(v)$ is the degree of vertex v in the subgraph $G' = (V,S) \subseteq G$. This proves Theorem 3.

3.2. **Random cluster type model.** Next we introduce a random cluster (RC) type model à la Fortuin and Kasteleyn [FK72], and couple this model with the 2-spin system and the WSG model, respectively. This way, we can transform a sample from the WSG model to a sample from the spin system, and establish Theorem 1.

The parameters are $\gamma > \beta > 1$ and $\lambda > 0$. As before, let $\rho = \frac{\beta-1}{\gamma-1}$ and $\theta = \frac{(\gamma-1)^2}{\beta\gamma-1}$. Given a graph $G=(V,E)$, the RC distribution is defined over subsets S of edges:

$$\mu^{\mathrm{rc}}(S) \propto w^{\mathrm{rc}}(S) := \prod_{C\in\kappa(V,S)} \left(\lambda^{|V(C)|}(\beta-1)^{|E(C)|} + (\gamma-1)^{|E(C)|}\right), \tag{11}$$

where $\kappa(V,S)$ is the set of connected components in the subgraph (V,S), $E(C)$ is the edge set of the component C, and $V(C)$ is the vertex set of the component C. Define the corresponding partition function $Z^{\mathrm{rc}} = Z_G^{\mathrm{rc}}(\beta,\gamma,\lambda) := \sum_{S\subseteq E} w^{\mathrm{rc}}(S)$. It is not hard to verify that for ferromagnetic Ising models ($\beta=\gamma > 1$ and $\lambda = 1$), (11) coincides with the classical Fortuin-Kasteleyn model [FK72]. For ferromagnetic Ising models with external fields ($\beta=\gamma>1$ and $\lambda \neq 1$), (11) coincides with the weighted random cluster model [FGW23].

The next theorem is a generalisation of the equivalence by Fortuin and Kasteleyn [FK72] and the coupling by Edwards and Sokal [ES88] between ferromagnetic Ising models and random cluster models.

Theorem 16. *For any graph G and parameters (β,γ,λ), $Z_G^{rc}(\beta,\gamma,\lambda) = Z_G^{spin}(\beta,\gamma,\lambda)$, where Z^{spin} is defined in (1).*

Moreover, given an RC sample $S\sim\mu^{rc}$, one can obtain a spin sample σ obeying the law of μ^{spin} as follows: independently for each connected component $C\in\kappa(V,S)$, assign spin 0 to all vertices in C with probability $\frac{\lambda^{|V(C)|}(\beta-1)^{|E(C)|}}{\lambda^{|V(C)|}(\beta-1)^{|E(C)|}+(\gamma-1)^{|E(C)|}}$, and spin 1 otherwise.

Proof. Recall that the ferromagnetic 2-spin interaction can be represented as a Holant problem on the vertex-edge incidence graph H_G. A vertex v is assigned the function $f_v = \lambda(1,0)^{\otimes d_v} + (0,1)^{\otimes d_v}$ and the edges assigned $g = (\beta,1,1,\gamma)^{\mathsf{T}}$. Let $g_0 = (1,1,1,1)^{\mathsf{T}}$ and $g_1 = (\beta-1,0,0,\gamma-1)^{\mathsf{T}}$ so that $g = g_0 + g_1$.

We may also view g, g_0, and g_1 as symmetric functions on two variables. Then, we have

$$\begin{aligned}
Z^{\text{spin}} &= \sum_{\sigma:V\to\{0,1\}} \lambda^{n_0(\sigma)} \prod_{(u,v)\in E} g(\sigma_u,\sigma_v) = \sum_{\sigma:V\to\{0,1\}} \lambda^{n_0(\sigma)} \prod_{(u,v)\in E} \left(g_0(\sigma_u,\sigma_v)+g_1(\sigma_u,\sigma_v)\right)\\
&= \sum_{\sigma:V\to\{0,1\}} \lambda^{n_0(\sigma)} \sum_{\tau:E\to\{0,1\}} \prod_{e=(u,v)\in E} g_{\tau(e)}(\sigma_u,\sigma_v)\\
&= \sum_{\tau:E\to\{0,1\}} \sum_{\sigma:V\to\{0,1\}} \lambda^{n_0(\sigma)} \prod_{e=(u,v)\in E} g_{\tau(e)}(\sigma_u,\sigma_v).
\end{aligned}$$

Given an edge assignment τ, let S be the set of edges that are assigned 1. In fact, since g_0 is the all-1 function, its contribution is the same as removing the edge. On the other hand, g_1 is non-zero only if both half edges are 0 or 1. This implies that, the only σ with non-zero weight is to assign the same value for each connected component $C \in \kappa(V,S)$. The weight of assigning 0 is $\lambda^{|V(C)|}(\beta-1)^{|E(C)|}$ and assigning 1 is $(\gamma-1)^{|E(C)|}$. The overall weight is exactly $w^{\text{rc}}(S)$, which proves the first part of the theorem.

For the second part, consider a joint distribution over both edges and vertices. The only allowed configuration is to have the same value in each connected component induced by the edges, and the weight is as given above. It is not hard to see that this distribution projected on the vertices recovers the spin distribution, and on the edges recovers the RC distribution. The second part of the theorem follows. $\square$

The reason we introduce the RC model is because we can also couple it with the WSG model, which is a generalisation of the coupling used in [GJ09, FGW23]. Denote by μ^{wsg} the Gibbs distribution of the WSG model, as defined in (4) with $\boldsymbol{\theta} = \theta\mathbf{1}$ and $\boldsymbol{\lambda} = \lambda\mathbf{1}$. Also recall that $\rho\theta = \frac{(\beta-1)(\gamma-1)}{\beta\gamma-1} < 1$.

Lemma 17. *For a graph $G=(V,E)$ and parameters $\gamma > \beta > 1$ and $\lambda > 0$, let S be a sample from μ^{wsg}. Independently add each remaining edge in $E \setminus S$ with probability $\rho\theta$ to get R. Then R follows the law of μ^{rc}.*

To prove Lemma 17, we will need the following lemma.

Lemma 18. *Let $G=(V,R)$ be a graph. Then,*

$$\sum_{S\subseteq R} \rho^{-|S|} w^{wsg}_{vtx}(S) = \left(\frac{\beta+\gamma-2}{(\beta-1)(\gamma-1)}\right)^{|R|} w^{rc}(R), \tag{12}$$

where $w^{wsg}_{vtx}(S) := \prod_{v\in V}\left(1+\lambda(-\rho)^{\deg_S(v)}\right)$.

Proof. This is a holographic transformation via Theorem 13. Again, let $H = H_G$ be the vertex-edge incidence graph. Then, $w^{\text{rc}}(R)$ is the Holant by giving each $v \in V$ a function $f_v = \lambda(1,0)^{\otimes d_v} + (0,1)^{\otimes d_v}$, and each $e \in E$ a function $g = [(\beta-1), 0, (\gamma-1)]$. This is because with these functions, the only configuration with non-zero weights have the same value in each connected component C, and the weight is $\lambda^{|V(C)|}(\beta-1)^{|E(C)|}$ if the value is 0, and $(\gamma-1)^{|E(C)|}$ if the value is 1.

Given this Holant, we apply the same transformation as in Theorem 3 by $\mathbf{T} := \begin{pmatrix}1 & -\rho\\ 1 & 1\end{pmatrix}$. Recall that $f_v\mathbf{T}^{\otimes d_v}$ is given in (9) already. On the edges, the resulting function is

$$\begin{aligned}
(\mathbf{T}^{-1})^{\otimes 2} g &= \frac{1}{(1+\rho)^2}\begin{pmatrix}\beta-1+(\gamma-1)\rho^2\\ 0\\ 0\\ \beta+\gamma-2\end{pmatrix} = \left(\frac{\gamma-1}{\beta+\gamma-2}\right)^2 \begin{pmatrix}\frac{(\beta-1)(\beta+\gamma-2)}{\gamma-1}\\ 0\\ 0\\ \beta+\gamma-2\end{pmatrix}\\
&= \frac{(\beta-1)(\gamma-1)}{\beta+\gamma-2}\begin{pmatrix}1\\ 0\\ 0\\ \rho^{-1}\end{pmatrix}.
\end{aligned}$$

On the other hand, the left hand side of (12) can be represented as a Holant as well. The vertex function is exactly $f_v \mathbf{T}^{\otimes d_v}$ for any $v \in V$, and the edge function is $[1, 0, \rho^{-1}]$. This proves the lemma. $\square$

Now we can prove Lemma 17.

Proof of Lemma 17. Note that $\mu^{\text{wsg}}(S) = \frac{\theta^{|S|} w_{\text{vtx}}^{\text{wsg}}(S)}{Z^{\text{wsg}}}$. Then, the probability of getting R is

$$\begin{aligned}
&\sum_{S \subseteq R} \frac{\theta^{|S|} w_{\text{vtx}}^{\text{wsg}}(S)}{Z^{\text{wsg}}} (\rho\theta)^{|R|-|S|} (1-\rho\theta)^{|E|-|R|} \\
&= \frac{(\rho\theta)^{|R|} (1-\rho\theta)^{|E|-|R|}}{Z^{\text{wsg}}} \sum_{S \subseteq R} \rho^{-|S|} w_{\text{vtx}}^{\text{wsg}}(S) \\
\text{(by Lemma 18)} \qquad &= \frac{(1-\rho\theta)^{|E|}}{Z^{\text{wsg}}} \left(\frac{\rho\theta}{1-\rho\theta} \right)^{|R|} \left(\frac{\beta+\gamma-2}{(\beta-1)(\gamma-1)} \right)^{|R|} w^{\text{rc}}(R).
\end{aligned}$$

By Theorem 3 and Theorem 16, $Z^{\text{wsg}} = (1-\rho\theta)^{|E|} Z^{\text{spin}} = (1-\rho\theta)^{|E|} Z^{\text{rc}}$. Moreover, $\frac{\rho\theta}{1-\rho\theta} = \frac{(\beta-1)(\gamma-1)}{\beta+\gamma-2}$. Thus, the probability above is $\frac{w^{\text{rc}}(R)}{Z^{\text{rc}}}$, which is exactly $\mu^{\text{rc}}(R)$. $\square$

We remark that one can also strengthen Theorem 16 and Lemma 17 into a grand model as in [FGW23] in a straightforward way. However, it does not seem to benefit our applications so the details are omitted here. In any case, we conclude this section with a proof of Theorem 1.

Proof of Theorem 1. We run Glauber dynamics for the WSG model for $O\left(\Delta^C \cdot n \log \frac{n}{\epsilon}\right)$ steps to get a subgraph S, where C is from Theorem 4. Since each step of Glauber dynamics can be implemented in time $O(\Delta)$, the total running time here is $O\left(\Delta^{C+1} \cdot n \log \frac{n}{\epsilon}\right)$. By Theorem 4 and the coupling inequality, we may couple S with a perfect sample from μ^{wsg} with probability at least $1-\epsilon$. Then, we use Lemma 17 to get an RC sample R, and subsequently use Theorem 16 to get a spin sample σ. This σ can be coupled with a perfect sample from μ^{spin} with probability at least $1-\epsilon$. Thus, the total variation distance of our sample from μ^{spin} is at most ϵ. $\square$

4. Stability of the polynomial

In this section, we prove Theorem 12. We decompose the argument into two steps. First, we reduce an arbitrary graph G to a graph G' by contracting *bad* leaves (the degree-one vertices with positive external fields). Second, we prove the stability result for the reduced graph G'. The latter step relies crucially on the leaf-contraction operation developed in the reduction step, so we present the reduction first.

Before we proceed, we explain why this preliminary reduction is needed. This is due to a technical reason: our proof in the second step cannot directly handle degree-one vertices with positive external fields. (In particular, Lemma 34 does not hold for leaves.) A related obstacle also appears in [GLL20]. One might try to avoid it by preprocessing the original graph G, deleting all degree-one vertices in advance, and then analyzing the weighted subgraph model only on graphs with minimum degree at least two. This shortcut was used in [GLL20], but it is not sufficient here. Our stability result must hold not only for the original graph, but also after arbitrary pinnings of the edge variables. Pinning an edge to 0 is equivalent to deleting that edge. Thus, even if the original graph has no degree-one vertices after preprocessing, the pinning may produce new degree-one vertices that require us to handle.

We next recall some basic settings. Let $\gamma > \beta > 1$ and $0 < \lambda < \lambda^*(\beta, \gamma)$. Let $\rho = \frac{\beta-1}{\gamma-1}$ and $\theta = \frac{(\gamma-1)^2}{\beta\gamma-1}$. Let $G = (V, E)$ be a graph and $\boldsymbol{\lambda} = (\lambda_v)_{v \in V}$ be parameters such that $-\rho\lambda \leqslant \lambda_v \leqslant \lambda$ for all $v \in V$. We will call λ_v the *external field* of vertex v. Let $w : \{0,1\}^E \to \mathbb{R}$ be the weight function defined in (5) with respect to graph G and parameters $\boldsymbol{\lambda}, \rho$. Formally, for any $S \subseteq E$,

$$w(S) = \prod_{v \in V} (1 + \lambda_v (-\rho)^{\deg_S(v)}), \tag{13}$$

where $\deg_S(v)$ is the degree of vertex v in the subgraph $G_S = (V, S)$. The partition function polynomial in (6) (without pinning) is defined as

$$Z(\boldsymbol{z}) = \sum_{S \subseteq E} w(S) \prod_{e \in S} z_e. \tag{14}$$

Note that the weight function w in (13) depends on the graph $G = (V, E)$ and the parameters $\boldsymbol{\lambda}, \rho$. Throughout this section, the parameter $\rho = \frac{\beta-1}{\gamma-1}$ is fixed. However, we may need to vary the graphs and external fields. We will use the following notation to highlight the dependence on the graph and the external fields:

$$Z_G(\boldsymbol{z}; \boldsymbol{\lambda}) = \sum_{S \subseteq E} \prod_{v \in V} (1 + \lambda_v (-\rho)^{\deg_S(v)}) \prod_{e \in S} z_e,$$

where the parameter $\rho = \frac{\beta-1}{\gamma-1}$.

Theorem 12 is proved in two steps. First, we reduce a general graph to a graph without *bad* leaf vertex. Then we prove stability for such reduced graphs. To state the two lemmas precisely, we introduce some notation. For a real interval I, define its δ-strip by

$$\mathcal{S}_\delta(I) := \{a \in \mathbb{C} : \operatorname{dist}(a, I) < \delta\}.$$

Recall that

$$U_\delta(x) := \{a \in \mathbb{C} : \operatorname{dist}(a, x) < \delta\}$$

is the disk of radius δ centered at x.

Throughout the next two lemmas, fix constants $\gamma > \beta > 1$ and $0 < \lambda < \lambda^\star(\beta, \gamma)$, and set

$$\rho = \frac{\beta - 1}{\gamma - 1}, \qquad \theta = \frac{(\gamma-1)^2}{\beta\gamma - 1}, \qquad I = [-\rho\lambda, \lambda], \quad I_- = [-\rho\lambda, 0].$$

Condition 19. Let $\delta_1 > 0$ be a constant. Let $G = (V, E)$ be a graph and $\boldsymbol{\lambda} = (\lambda_v)_{v \in V} \in \mathbb{C}^V$. It holds that for all vertices $v \in V$, $\lambda_v \in \mathcal{S}_{\delta_1}(I)$, furthermore, if $\deg_G(v) = 1$, then $\lambda_v \in \mathcal{S}_{\delta_1}(I_-)$.

Lemma 20 (Reduction lemma)**.** *For any constant $\delta_1 > 0$, there exists a constant $\delta_2 = \delta_2(\beta, \gamma, \lambda, \delta_1) > 0$ such that the following holds. Let $G = (V, E)$ be a graph and $\boldsymbol{\lambda} = (\lambda_v)_{v \in V} \in \mathbb{R}^V$ be parameters such that $-\rho\lambda \leqslant \lambda_v \leqslant \lambda$ for all $v \in V$. For any vector $\boldsymbol{z} = (z_e)_{e \in E}$ with $z_e \in U_{\delta_2}(\theta)$ for all $e \in E$, there exists a subgraph $G' = (V', E')$ of G, and complex-valued parameters $\boldsymbol{\lambda}' = (\lambda'_v)_{v \in V'} \in \mathbb{C}^{V'}$ such that G' and $\boldsymbol{\lambda}'$ satisfy Condition 19 with constant δ_1 and*

$$Z_{G'}(\boldsymbol{z}_{E'}; \boldsymbol{\lambda}') \neq 0 \text{ if and only if } Z_G(\boldsymbol{z}; \boldsymbol{\lambda}) \neq 0.$$

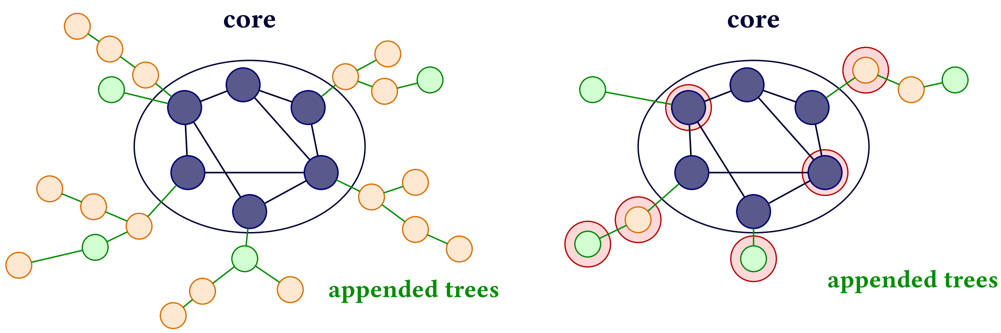
Lemma 21 (Stability of the bad-leaf-reduced core)**.** *There exists $\delta_1 = \delta_1(\beta, \gamma, \lambda) > 0$ such that the following holds. Let $G = (V, E)$ and $\boldsymbol{\lambda} = (\lambda_v)_{v \in V} \in \mathbb{C}^V$ be a graph and a vector such that Condition 19 holds with constant δ_1. Then there exists a constants $\delta_3 = \delta_3(\beta, \gamma, \lambda, \delta_1) > 0$ such that for all $\boldsymbol{z} = (z_e)_{e \in E}$ with $z_e \in U_{\delta_3}(\theta)$ for all $e \in E$,*

$$Z_G(\boldsymbol{z}; \boldsymbol{\lambda}) \neq 0.$$

Assuming the above two lemmas, we prove Theorem 12.

Proof of Theorem 12. Let $\delta_1 = \delta_1(\beta, \gamma, \lambda) > 0$ be the constant from Lemma 21. Since Lemma 20 holds for every $\delta_1 > 0$, we apply it with this value of δ_1. Let $\delta_2 = \delta_2(\beta, \gamma, \lambda, \delta_1)$ be the constant from Lemma 20, and let $\delta_3 = \delta_3(\beta, \gamma, \lambda)$ be the constant from Lemma 21. Define

$$\delta = \min(\delta_2, \delta_3).$$



(a) General graph G.

(b) Graph G' after bad-leaf deletion.

FIGURE 2. An illustration of the leaf-deletion reduction. (a) A general graph G is decomposed into a core of minimum degree at least two and appended trees. In each appended tree, a green vertex u denotes $\lambda_u < 0$, while an orange vertex u denotes $\lambda_u \geqslant 0$. (b) After pruning bad-leaves in G, part of the appended trees is removed. Removed vertices are absorbed into effective external fields on the circled vertices.

Since δ_1 depends only on β, γ, λ, so does δ.

Fix a graph $G = (V, E)$ and a real vector $\boldsymbol{\lambda} = (\lambda_v)_{v \in V} \in \mathbb{R}^V$ such that $-\rho\lambda \leqslant \lambda_v \leqslant \lambda$ for all $v \in V$. Let $\boldsymbol{z} = (z_e)_{e \in E}$ satisfy $z_e \in U_\delta(\theta)$ for all $e \in E$. Since $\delta \leqslant \delta_2$, Lemma 20 produces a reduced graph $G' = (V', E')$ and fields $\boldsymbol{\lambda}'$ with $\lambda'_v \in \mathcal{S}_{\delta_1}(I)$ for all $v \in V'$ and $\lambda'_v \in \mathcal{S}_{\delta_1}(I_-)$ for all $v \in V'$ with $\deg_{G'}(v) = 1$, such that it is enough to prove $Z_{G'}(\boldsymbol{z}_{E'}; \boldsymbol{\lambda}') \neq 0$. Since $\delta \leqslant \delta_3$ and $(G', \boldsymbol{\lambda}')$ satisfies Condition 19 with constant δ_1, this non-vanishing follows from Lemma 21.

Therefore $Z_G(\boldsymbol{z}; \boldsymbol{\lambda}) \neq 0$ for all $\boldsymbol{z} \in \prod_{e \in E} U_\delta(\theta)$, which proves the desired stability. □

Next, we prove Lemma 20 in Section 4.1 and Lemma 21 in Section 4.2.

4.1. **Leaf-deletion reduction.** We give our reduction from general graph G to the leaf-reduced graph G'. We first provide some intuition of the proof idea and then give the formal proof. Let v be a leaf (degree-one vertex) with unique neighbor u and leaf edge $e = \{u, v\}$ in graph G. We call u the *parent* of v. Suppose the current external field at the leaf is λ_v and the current external field at the parent is λ_u. Assume that the external field $\lambda_v \notin \mathcal{S}_{\delta_1}(I_-)$, which violates Condition 19. We will use a process to delete this bad leaf vertex v, and absorb the deleted part into the external field at the parent u.

Fix a configuration $S \subseteq E$ on G such that $e \in S$. Let $S' = S \setminus \{e\}$. We write $t = (-\rho)^{\deg_{S'}(u)}$, where $\deg_{S'}(u)$ is the degree of the vertex u in the subgraph (V, S'). For a configuration S and S', the edge e together with the two endpoints of the leaf edge e contribute the following factor to the partition function polynomial (14):

$$\begin{cases} (1+\lambda_v)(1+\lambda_u t) & \text{for configuration } S', \\ z_e(1-\rho\lambda_v)(1-\rho\lambda_u t) & \text{for configuration } S. \end{cases}$$

The partition function is a sum over all configurations. Both S and S' are valid configurations. If we add the above contributions of u and v in both cases together, we get

$$\begin{aligned} &(1+\lambda_v)(1+\lambda_u t) + z_e(1-\rho\lambda_v)(1-\rho\lambda_u t) \\ &\quad = \big[(1+\lambda_v) + z_e(1-\rho\lambda_v)\big] + \lambda_u t\big[(1+\lambda_v) - \rho z_e(1-\rho\lambda_v)\big]. \end{aligned}$$

Define the following two functions (one may view a as λ_v and z as z_e):

$$C_z(a) = (1+a) + z(1-\rho a), \qquad \Phi_z(a) = \frac{(1+a) - \rho z(1-\rho a)}{(1+a) + z(1-\rho a)}. \tag{15}$$

Then, whenever $C_{z_e}(\lambda_v) \neq 0$, it holds that

$$(1+\lambda_v)(1+\lambda_u t) + z_e(1-\rho\lambda_v)(1-\rho\lambda_u t) = C_{z_e}(\lambda_v)\left(1+\lambda_u t\Phi_{z_e}(\lambda_v)\right). \tag{16}$$

Consider the subgraph $G' = G - \{v, e\}$ obtained from G by removing the leaf vertex v. We can view the factor $(1 + \lambda_u t\Phi_{z_e}(\lambda_v))$ in the RHS of (16) as the new factor contributed at the parent u in the subgraph G'. Therefore, we can write the partition function in the following way:

$$Z_G(\boldsymbol{z};\boldsymbol{\lambda}) = C_{z_e}(\lambda_v) Z_{G'}(\boldsymbol{z};\boldsymbol{\lambda}'), \tag{17}$$

where for all vertices in graph G', the external fields are updated as follows:

$$\begin{cases} \lambda'_u = \lambda_u \Phi_{z_e}(\lambda_v) & \text{for parent } u, \\ \lambda'_w = \lambda_w & \text{for other vertices } w \neq u. \end{cases}$$

Hence, we transfer the stability of $Z_G(\boldsymbol{z};\boldsymbol{\lambda})$ to the stability of $Z_{G'}(\boldsymbol{z};\boldsymbol{\lambda}')$ if $C_{z_e}(\lambda_v) \neq 0$.

Our proof of Lemma 20 is to apply the above deletion process in a systematic way. We introduce the following closure lemma, which plays a crucial role in our proof. Set

$$I_+ = [0, \lambda], \qquad I_- = [-\rho\lambda, 0], \qquad I = I_- \cup I_+.$$

Lemma 22. *For any constant $\delta_1 > 0$, there exists a constant $\delta_2 > 0 = \delta_2(\beta, \gamma, \lambda, \delta_1)$, together with two regions $\mathcal{U}^+, \mathcal{U}^- \subset \mathbb{C}$ such that the following hold. Let $\mathcal{U} = \mathcal{U}^+ \cup \mathcal{U}^-$.*

(i) $I \subset \mathcal{U} \subset \mathcal{S}_{\delta_1}(I)$ *and* $\mathcal{U}^- \subset \mathcal{S}_{\delta_1}(I_-)$.
(ii) *For any* $a \in \mathcal{U}^+$ *and any* $z \in U_{\delta_2}(\theta)$, $C_z(a) \neq 0$.
(iii) *For any integer* $k \geqslant 1$, *any sequence* $a_i \in \mathcal{U}^+$ *and* $z_i \in U_{\delta_2}(\theta)$ *with* $1 \leqslant i \leqslant k$, *it holds that*

$$\forall c \in I, \quad c\prod_{i=1}^{k} \Phi_{z_i}(a_i) \in \mathcal{U}.$$

Lemma 22 is proved in Section 4.3. Assuming the lemma, we define the following leaf-deletion process. Note that $U^- \subset \mathcal{S}_{\delta_1}(I_-)$ is a safe region for Condition 19. We need to keep deleting leaf vertices with external fields in U^+. Formally, consider the following process.

Let $G = (V, E)$ be a graph and $\boldsymbol{\lambda} = (\lambda_v)_{v\in V} \in \mathbb{R}^V$ be a vector of external fields in Lemma 20.

- Set $\lambda'_v \leftarrow \lambda_v$ for all $v \in V$.
- Repeat the following process until G does not have leaf (degree-one) vertex v with $\lambda'_v \in \mathcal{U}^+$:
  - Choose the leaf vertex v with the smallest index such that $\lambda'_v \in \mathcal{U}^+$ and denote its parent by u.
  - Delete the edge $e = (u, v)$ and the leaf v, and update the external field $\lambda'_u \leftarrow \lambda'_u \Phi_{z_e}(\lambda'_v)$.

After the whole process terminates, we denote the partition function of G' by $Z_{G'}(\boldsymbol{z}, \boldsymbol{\lambda}')$. We refer to Figure 2 for an illustration of the process. Now, we are ready to prove Lemma 20.

Proof of Lemma 20. Fix $\delta_1 > 0$, and let $\delta_2 > 0$ and $\mathcal{U} = \mathcal{U}^+ \cup \mathcal{U}^-$ be given by Lemma 22. Also fix $\boldsymbol{z}$ with $z_e \in U_{\delta_2}(\theta)$ for all $e \in E$. We run the deterministic leaf-deletion procedure described above, always deleting the leaf of smallest index among the leaves whose *current* external field $\lambda'_v \in \mathcal{U}^+$. This determines a subgraph $G' = (V', E')$ of G.

We track the external fields and the partition function throughout the process. During the process, let $\lambda_v^{(t)}$ denote the current external field of a vertex v after t leaf deletions. By Lemma 22 (i), $\lambda_v^{(0)} = \lambda_v \in I \subset \mathcal{U}$. We prove by induction on t that the following two properties hold.

- The first property is that for every current vertex u, its current field has the form

$$\lambda_u^{(t)} = \lambda_u \prod_{e=(u,v)\in D_t(u)} \Phi_{z_e}(a_e), \tag{18}$$

where $D_t(u)$ is the set of deleted edges that have been absorbed into u, and each a_e is the current field of the deleted leaf neighbor v at the moment when $e = \{u, v\}$ was deleted. Moreover, for any $e \in D_t(u)$, $a_e \in \mathcal{U}^+$, and $\lambda_u^{(t)} \in \mathcal{U}$.
- The second property is that for any G_t and $\boldsymbol{\lambda}^{(t)}$, the partition function satisfies:

$$Z_G(\boldsymbol{z}; \boldsymbol{\lambda}) = \left(\prod_{e \in E \setminus E(G_t)} C_{z_e}(a_e) \right) Z_{G_t}(\boldsymbol{z}_{E(G_t)}; \boldsymbol{\lambda}^{(t)}). \tag{19}$$

  where $E(G_t)$ is the set of edges in G_t.

Both claims are immediate at $t = 0$. Suppose they hold after t deletions, and let the next deleted leaf be v with parent u and leaf edge $e = (u, v)$. By the choice of the deletion step, the current field $\lambda_v^{(t)}$ lies in $\mathcal{U}^+$. Hence Lemma 22 (ii) gives $C_{z_e}(\lambda_v^{(t)}) \neq 0$. The one-step identity (17) gives

$$Z_{G_t}(\boldsymbol{z}_{E(G_t)}; \boldsymbol{\lambda}^{(t)}) = C_{z_e}(\lambda_v^{(t)}) Z_{G_{t+1}}(\boldsymbol{z}_{E(G_{t+1})}; \boldsymbol{\lambda}^{(t+1)}),$$

where the only changed field is

$$\lambda_u^{(t+1)} = \lambda_u^{(t)} \Phi_{z_e}(\lambda_v^{(t)}).$$

This proves the partition-function identity (19) at time $t + 1$ and updates the field as in (18).

It remains to check that the updated field $\lambda_u^{(t+1)}$ stays in $\mathcal{U}$. By the induction hypothesis, $\lambda_u^{(t)}$ can be written as $\lambda_u^{(t)} = \lambda_u \prod_{i=1}^{k} \Phi_{z_i}(a_i)$, where $\lambda_u \in I$, each $a_i \in \mathcal{U}^+$, and each $z_i \in U_{\delta_2}(\theta)$. The new update appends one more factor with $a_{k+1} = \lambda_v^{(t)} \in \mathcal{U}^+$ and $z_{k+1} = z_e \in U_{\delta_2}(\theta)$. Lemma 22 (iii) implies

$$\lambda_u^{(t+1)} = \lambda_u \prod_{i=1}^{k+1} \Phi_{z_i}(a_i) \in \mathcal{U}, \quad \text{because } \lambda_u \in I.$$

All other current fields are unchanged, so the induction is complete.

At the end of the process, let $G' = G_T = (V', E')$ and let $\boldsymbol{\lambda}' = \boldsymbol{\lambda}^{(T)}$, where T is the total number of leaf deletions. Using Lemma 22 (i), the induction gives $\lambda'_v \in \mathcal{U} \subset \mathcal{S}_{\delta_1}(I)$ for every $v \in V'$. Moreover, by the stopping rule, every degree-one vertex of G' has external field in $\mathcal{U}^- \subset \mathcal{S}_{\delta_1}(I_-)$. It also gives

$$Z_G(\boldsymbol{z}; \boldsymbol{\lambda}) = \prod_{e \in E \setminus E'} C_{z_e}(a_e)\, Z_{G'}(\boldsymbol{z}_{E'}; \boldsymbol{\lambda}'),$$

where each $C_{z_e}(a_e)$ is nonzero by the above analysis. Hence $Z_{G'}(\boldsymbol{z}_{E'}; \boldsymbol{\lambda}') \neq 0$ if and only if $Z_G(\boldsymbol{z}; \boldsymbol{\lambda}) \neq 0$. This proves the lemma. □

4.2. **Stability of the bad-leaf-reduced core.** We now explain the strategy for proving Lemma 21. In the proof, we use the Holant representation of the weighted subgraph model as in Section 3.1. Let $G = (V, E)$ be a graph. Consider the bipartite graph $H_G = (V_1 \uplus V_2, E_H)$ defined in Definition 14, where $V_1 = V$ and $V_2 = E$. For each edge $e = \{u, v\} \in E$, we add two half-edges $e_u = \{u, e\}$ and $e_v = \{v, e\}$ to H_G; we call e_u and e_v the half-edges of e. To simulate the WSG model, the signature function f_v for $v \in V_1$ is defined as

$$f_v = [1 + \lambda_v, 1 + \lambda_v(-\rho), \ldots, 1 + \lambda_v(-1)^{d_v}], \tag{20}$$

where $d_v = \deg_G(v)$ is the degree of v in G. The signature function g_e for $e \in V_2$ is defined as

$$g_e = [1, 0, 0, z_e], \tag{21}$$

where z_e is the variable of the edge e. Let $\mathcal{F} = \{f_v\}_{v \in V_1}$ and $\mathcal{G} = \{g_e\}_{e \in V_2}$. Then, it holds that

$$Z_G(\boldsymbol{z}; \boldsymbol{\lambda}) = \mathrm{Holant}(H_G; \mathcal{F} \mid \mathcal{G}) \tag{22}$$

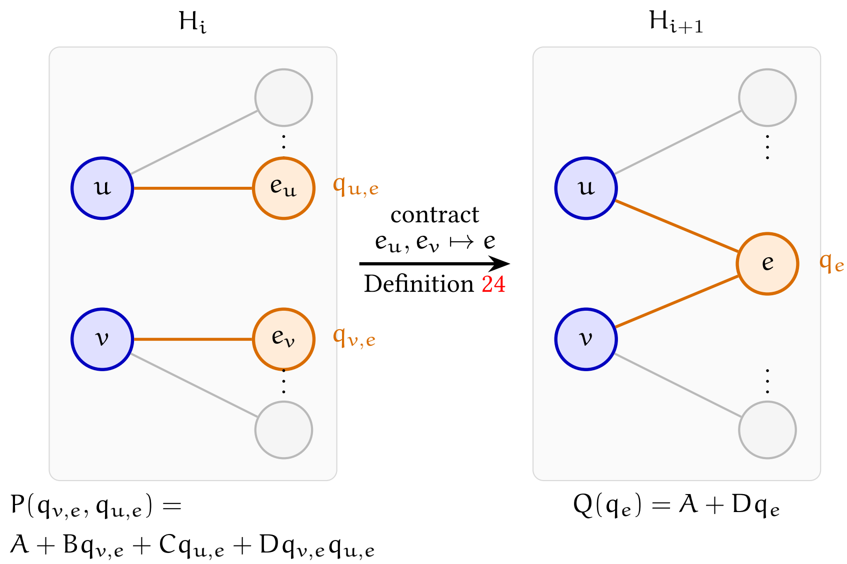


FIGURE 3. One-step contraction from H_i to H_{i+1}, by Definition 24 and Lemma 27.

We use the a contraction procedure to show the stability of the partition function for the WSG model. The contraction procedure is as follows:

First, we split every edge into two half-edges, so that each vertex v contributes a local polynomial defined over a star graph. Specifically, for each vertex $v \in V$, let S_v be the star graph rooted at v with one leaf for each edge incident to v. We define H_0 as the disjoint union

$$H_0 = \bigcup_{v \in V} S_v. \tag{23}$$

More formally, the initial bipartite graph H_0 is defined as follows.

Definition 23 (initial bipartite graph)**.** Let $G = (V, E)$ be a graph. The left part of the initial bipartite graph H_0 is the vertex set $V_0 = V$, and the right part is the set of half-edges $E_0 = \{e_u, e_v \mid e = \{u, v\} \in E\}$. For each edge $e \in E$ in graph G, we add two edges $e_u = \{u, e_u\}$ and $e_v = \{v, e_v\}$ to H_0.

By the definition, the left part of H_0 is the vertex set V, and the right part contains $2|E|$ vertices. We then recover the original edge model by *contracting* the half edges. Fix an ordering $e_1, \ldots, e_m$ of the edges of G. Given the graph H_{i-1}, the next graph H_i is defined by applying the Asano contraction [Asa70], defined as follows.

Definition 24 (Asano contraction)**.** Let $e = e_i = \{u, v\}$ be the next edge to be contracted. The contraction operation is defined as follows: find two edges $\{u, e_u\}$ and $\{v, e_v\}$ in the bipartite graph H_{i-1}, and then contract the two vertices e_u and e_v to be one vertex e. The resulting graph is H_i.

Using the above contraction operation, we construct a sequence of bipartite graphs

$$H_0, H_1, \ldots, H_m, \quad \text{for } m = |E|,$$

where H_0 is the disjoint union of the vertex stars and $H_m = H_G$ is the edge-incidence graph. Note that each bipartite graph H_i in this sequence can be written as $H_i = (V_1^{(i)} \uplus V_2^{(i)}, E_H^{(i)})$, where $V_1^{(i)} = V$ and $V_2^{(i)} = \{e_j \mid j \leqslant i\} \cup \{e_u, e_v \mid e = e_j = \{u, v\} \wedge j > i\}$.

Given $H_0, H_1, \ldots, H_m$, we can define a sequence of polynomials $Z_{H_0}, Z_{H_1}, \ldots, Z_{H_m}$. However, due to technical reasons, we define a modified version instead as follows.

Definition 25 (polynomial sequence)**.** Let H_i be a bipartite graph in the sequence. For each $v \in V$ in the left part, let f_v in (20) be the signature function of v. For each $w \in V_2^{(i)}$ in the right part, define

$$\widetilde{g}_w = \begin{cases} [q_w, 1], & \text{if } w \text{ is an uncontracted half-edge vertex,} \\ [q_w, 0, 0, 1], & \text{if } w \text{ is a contracted edge vertex.} \end{cases}$$

Then, the *modified partition function* $\widetilde{Z}_{H_i}$ is defined as follows:

$$\widetilde{Z}_{H_i}(\mathbf{q}; \boldsymbol{\lambda}) = \mathrm{Holant}(H_i; F^{(i)} \mid \widetilde{G}^{(i)}), \tag{24}$$

where $\mathcal{F}^{(i)} = \{f_v\}_{v \in V_1^{(i)}}$ and $\widetilde{G}^{(i)} = \{\widetilde{g}_w\}_{w \in V_2^{(i)}}$.

In the definition above, uncontracted half-edge vertices use the reciprocal unary signature $[q_w, 1]$, while contracted edge vertices use the reciprocal binary equality signature $[q_w, 0, 0, 1]$. Intuitively, this modification is the change of variables $q_w = 1/z_w$ applied to the signatures $[1, z_w]$ and $[1, 0, 0, z_w]$, respectively. This is valid because the variables z_w we consider are bounded away from 0. Note that $\widetilde{Z}_{H_m}(q; \lambda) = \left(\prod_{e \in E} q_e\right) Z_G^{\mathrm{wsg}}(1/q, \lambda)$ as defined in (22). The proof of Lemma 21 has two ingredients: the stability of the initial local star polynomials, and the preservation of stability under each contraction operation. As a result, the final $\widetilde{Z}_{H_m}$ remains stable, as we desired.

Define the local polynomial of vertex v with degree d_v and external field λ_v. Here the variables $z_1, \ldots, z_{d_v}$ represent half-edge variables:

$$\begin{aligned} P_{v,d_v}(z_1, \ldots, z_{d_v}) &:= \sum_{S \subseteq [d_v]} \left(1 + \lambda_v (-\rho)^{|S|}\right) \prod_{i \in S} z_i \\ &= \prod_{i=1}^{d_v} (1 + z_i) + \lambda_v \prod_{i=1}^{d_v} (1 - \rho z_i) \end{aligned} \tag{25}$$

Instead of studying $P_{v,d_v}(z_1, \ldots, z_{d_v})$, we consider a modified local polynomial by setting $q_i = \frac{1}{z_i}$. Define the modified local polynomial:

$$\widetilde{P}_{v,d_v}(q_1, \ldots, q_{d_v}) := \left(\prod_{i=1}^{d_v} q_i\right) P_{v,d_v}(1/q_1, \ldots, 1/q_{d_v}) = \sum_{S \subseteq [d_v]} \left(1 + \lambda_v (-\rho)^{|S|}\right) \prod_{i \notin S} q_i. \tag{26}$$

The modified partition function of H_0 is as given by Definition 25:

$$\widetilde{Z}_{H_0}(\mathbf{q}; \boldsymbol{\lambda}) = \prod_{v \in V} \widetilde{P}_{v,d_v}\left((q_{v,e})_{e \ni v}\right) = \sum_{S \subseteq E_0} \prod_{e \notin S} q_e \prod_{v \in V} \left(1 + \lambda_v (-\rho)^{\deg_S(v)}\right).$$

Here $q_{v,e}$ denotes the half-edge variable at the copy of edge e incident to v. As we discussed before, we first study the stability of $\widetilde{Z}_{H_0}(\mathbf{q}; \boldsymbol{\lambda})$. Since it is a product of local polynomials, it suffices to establish stability of $\widetilde{P}_{v,d_v}(q_{v,1}, \ldots, q_{v,d_v})$ for each vertex v.

We first motivate some definitions that will be essential for our proof. Assume for the moment that $q_{v,i} \neq \rho$ for all $1 \leqslant i \leqslant d_v$. Then the local polynomial can be written as

$$\widetilde{P}_{v,d_v}(q_{v,1}, \ldots, q_{v,d_v}) = \prod_{i=1}^{d_v} (q_{v,i} + 1) + \lambda_v \prod_{i=1}^{d_v} (q_{v,i} - \rho) \tag{27}$$

$$(\text{as we assume } q_{v,i} \neq \rho \text{ for all } i) \qquad = \prod_{i=1}^{d_v} (q_{v,i} - \rho) \left(\prod_{i=1}^{d_v} \left(\frac{1 + q_{v,i}}{q_{v,i} - \rho}\right) + \lambda_v\right). \tag{28}$$

The ratio $\frac{1+q_{v,i}}{q_{v,i}-\rho}$ motivates the following Möbius change of variables:

$$u = \psi(q) := -\frac{1+q}{q-\rho}. \tag{29}$$

Under this assumption, if $\widetilde{P}_{v,d_v}(q_{v,1},\ldots,q_{v,d_v}) = 0$, then $\prod_{i=1}^{d_v} u_{v,i} = (-1)^{d_v+1}\lambda_v$, where $u_{v,i} = \psi(q_{v,i})$. Thus, to prove stability (zero-freeness) of the modified local polynomial, it suffices to rule out this product identity when the variables lie in the desired domain. We will choose a *bad region* for the u-variables so that the identity can hold only when some $u_{v,i}$ lies inside this region. More precisely, the bad region is the open disk

$$D_\epsilon = D(c^\star, r^\star - \epsilon),$$

where $\epsilon > 0$ will be chosen later. The constants $c^\star$ and $r^\star$ are from [GLL20]:

$$c^\star := \frac{\gamma\log(\sqrt{\gamma/\beta})}{\sqrt{\gamma\beta-1}\arctan\sqrt{\gamma\beta-1}-\log(\sqrt{\gamma/\beta})}, \qquad r^\star := \sqrt{\left(c^\star+\frac{1}{\beta}\right)^2+\frac{\beta\gamma-1}{\beta^2}}. \tag{30}$$

Pulling this disk back under the map ψ, we define the corresponding bad region for the q-variables by

$$\Gamma_\epsilon := \{q \in \mathbb{C} : \ \psi(q) \in D(c^\star, r^\star-\epsilon)\}. \tag{31}$$

The following lemma gives the desired stability of each local polynomial.

Lemma 26. *There exist constants $\delta_1, \epsilon > 0$, depending only on γ, β, λ, such that the following holds. Let $G = (V,E)$ and $\boldsymbol{\lambda} \in \mathbb{C}^V$ satisfy Condition 19 with constant δ_1. For any $v \in V$, the polynomial $\widetilde{P}_{v,d_v}(q_1,\ldots,q_{d_v}) \neq 0$ whenever $q_i \in \mathbb{C}\setminus\Gamma_\epsilon$ for all $i \in [d_v]$.*

The proof of Lemma 26 is deferred to Section 4.4. Although the definition of Γ_ϵ was motivated by the case $q_{v,i} \neq \rho$ for all i, the lemma itself also covers the general case in which some $q_{v,i}$ may equal ρ. Since $\widetilde{Z}_{H_0}(\mathbf{q};\boldsymbol{\lambda})$ is a product of local polynomials, Lemma 26 implies the stability of $\widetilde{Z}_{H_0}(\mathbf{q};\boldsymbol{\lambda})$.

We next turn to the contraction step. For each polynomial Z_{H_i} in the sequence, similarly, we define the modified partition function $\widetilde{Z}_{H_i}$ by replacing each variable z with $q = 1/z$. Suppose that at step $i+1$ the edge is $e_{i+1} = \{u,v\}$. The graph H_i still has two half-edge variables $q_{u,e_{i+1}}$ and $q_{v,e_{i+1}}$ corresponding to the two copies of e_{i+1}. The graph H_{i+1} is obtained by replacing these two half-edge variables with a single edge variable $q_{e_{i+1}}$. Equivalently, if $\mathbf{q}'$ denotes all variables not involved in this step, then we define two auxiliary polynomials P and Q as follows:

$$P(q_{u,e_{i+1}}, q_{v,e_{i+1}}) := \widetilde{Z}_{H_i}(\mathbf{q}', q_{u,e_{i+1}}, q_{v,e_{i+1}}; \boldsymbol{\lambda});$$
$$Q(q_{e_{i+1}}) := \widetilde{Z}_{H_{i+1}}(\mathbf{q}', q_{e_{i+1}}; \boldsymbol{\lambda}),$$

where Q is the contracted version of P. The next lemma gives a algebraic description of this one-edge contraction and shows that it preserves the required stability region.

Lemma 27. *Fix $\mathbf{q}'$ and $\boldsymbol{\lambda}$, and define coefficients A, B, C, D depending on $\mathbf{q}'$ and $\boldsymbol{\lambda}$ such that*

$$P(q_{v_e}, q_{u_e}) = \widetilde{Z}_{H_i}(\mathbf{q}', q_{v_e}, q_{u_e}; \boldsymbol{\lambda}) = A + Bq_{v_e} + Cq_{u_e} + Dq_{v_e}q_{u_e}.$$

For every $0 < \epsilon < r^\star/2$, if $P(q_{v_e}, q_{u_e}) \neq 0$ for $q_{v_e}, q_{u_e} \in \mathbb{C}\setminus\Gamma_\epsilon$, then there exists $\delta_3 > 0$, depending only on β, γ, ϵ, such that

$$Q(q_e) = \widetilde{Z}_{H_{i+1}}(\mathbf{q}', q_e; \boldsymbol{\lambda}) = A + Dq_e$$

is zero-free for $q_e \in U_{\delta_3}(1/\theta)$.

We explain why $Q(q_e)$ can be written as $A + Dq_e$ in Lemma 27. In the graph H_i, before contracting the edge $e = \{u,v\}$, the two half-edge variables q_{v_e} and q_{u_e} are independent. Hence, after fixing all other variables $\mathbf{q}'$ and $\boldsymbol{\lambda}$, the polynomial has the form

$$P(q_{v_e}, q_{u_e}) = A + Bq_{v_e} + Cq_{u_e} + Dq_{v_e}q_{u_e}.$$

The four coefficients correspond to the four possible states of the two half-edges in reciprocal coordinates: $Dq_{v_e}q_{u_e}$ is the contribution where neither half-edge is selected, Bq_{v_e} and Cq_{u_e} are the two mixed contributions, and A is the contribution where both half-edges are selected. The combinatorial contraction in Definition 24 merges the two half-edge vertices into one vertex e, whose equality signature allows only the two consistent states: both half-edges are not selected or both are selected. Therefore the two mixed contributions, represented algebraically by Bq_{v_e} and Cq_{u_e}, disappear after contraction. The contracted signature $[q_e, 0, 0, 1]$ assigns weight q_e to the state where neither half-edge is selected and weight 1 to the state where both half-edges are selected. Thus the contracted polynomial is

$$Q(q_e) = A + Dq_e.$$

The proof of the stability of $Q(q_e)$ in Lemma 27 is deferred to Section 4.5. It relies essentially on an argument first given by Ruelle [Rue71].

Given Lemma 26 and Lemma 27, we are now ready to prove Lemma 21.

Proof of Lemma 21. Let ϵ be the constant from Lemma 26. By shrinking ϵ if necessary, we may assume that $0 < \epsilon < r^\star/2$; this only shrinks the stability domain $\mathbb{C} \setminus \Gamma_\epsilon$. By Lemma 26, the modified partition function of the initial graph H_0

$$\widetilde{Z}_{H_0}(\mathbf{q}; \boldsymbol{\lambda}) = \prod_{v \in V} \widetilde{P}_{v,d_v}\big((q_{v,e})_{e \ni v}\big).$$

is stable for each half-edge variable $q_{v,e} \in \mathbb{C} \setminus \Gamma_\epsilon$ for $v \in V$, $e \in E_v$. Now we order the edges of G as $e_1, e_2, \ldots, e_m$. Starting from $\widetilde{Z}_{H_0}$, apply Asano contraction successively from H_i to H_{i+1}. Lemma 27, applied with this ϵ, shows that each contraction preserves stability for the newly created reciprocal edge variable q_e in the disk $U_{\delta_3}(1/\theta)$, where $\delta_3 = \delta_3(\beta, \gamma, \epsilon)$. Since ϵ depends only on β, γ, λ, so does δ_3. Therefore, after all contractions, the resulting polynomial $\widetilde{Z}_{H_m}(\mathbf{q}; \boldsymbol{\lambda})$ is stable whenever $q_e \in U_{\delta_3}(1/\theta)$ for every $e \in E$. Finally, by the construction of the $\mathbf{q}$-variables, the fully contracted polynomial satisfies

$$\widetilde{Z}_{H_m}(\mathbf{q}; \boldsymbol{\lambda}) = \left(\prod_{e \in E} q_e\right) Z_G^{\mathrm{wsg}}(1/\mathbf{q}, \boldsymbol{\lambda}).$$

The factor $\prod_{e \in E} q_e$ is nonzero in a small neighborhood of $1/\theta$ since $\theta \neq 0$. Hence $\widetilde{Z}_{H_m}(\mathbf{q}; \boldsymbol{\lambda}) \neq 0$ implies $Z_G^{\mathrm{wsg}}(\boldsymbol{z}, \boldsymbol{\lambda}) \neq 0$ for $z_e = 1/q_e$. Shrinking the neighborhood if necessary, this gives stability of $Z_G^{\mathrm{wsg}}(\boldsymbol{z}, \boldsymbol{\lambda})$ whenever $z_e \in U_\delta(\theta)$ for every $e \in E$ for some $\delta > 0$ depending only on β, γ, λ. This proves the theorem. □

4.3. **Proof of the closure lemma.** This section is devoted to the proof of Lemma 22. Recall that by our analysis in Section 4.1, we have that

$$C_z(a) = (1 + a) + z(1 - \rho a), \qquad \Phi_z(a) = \frac{(1 + a) - \rho z(1 - \rho a)}{(1 + a) + z(1 - \rho a)}.$$

and the partition function after contracting one leaf edge $e = \{u, v\}$ with leaf vertex v satisfies:

$$Z_G(\boldsymbol{z}; \boldsymbol{\lambda}) = C_{z_e}(\lambda_v) Z_{G \setminus \{v,e\}}(\boldsymbol{z}; \boldsymbol{\lambda}'),$$

where $\lambda'_u = \lambda_u \Phi_{z_e}(\lambda_v)$ and $\lambda'_w = \lambda_w$ for other nodes $w \neq u$.

Fix a constant $\delta_1 > 0$ in Lemma 22. First, we prove Lemma 22 (i). To this end, we will define the region $\mathcal{U}^+$ and $\mathcal{U}^-$ satisfying the conditions in the Lemma. Let $I_+ = [0, \lambda]$ and $I_- = [-\rho\lambda, 0]$, where $0 < \lambda < \lambda^\star$. We use a potential function from [GL18]:

$$h_+(x) = \max\left\{t,\ x \log \frac{\zeta}{x}\right\}, \qquad 0 \leqslant x \leqslant \lambda.$$

The specific choice of ζ, t can be found in [GL18]. In particular, the following holds.

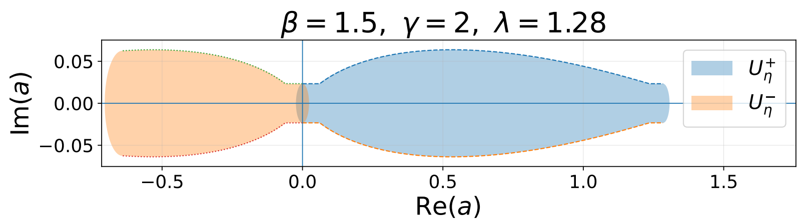


FIGURE 4. An illustration of the region $\mathcal{U}_\eta$ used in the closure lemma.

Lemma 28 (Equation (10) of [GL18][4]). *There exist constants $\zeta \in (\lambda, \lambda^\star)$ and $t > 0$, depending only on β, γ, λ, such that the following holds for a constant $\kappa \in (0, 1)$:*

$$\frac{(\beta\gamma - 1)h_+(x)}{(1+\beta x)(\gamma + x)} \leqslant \kappa \log \frac{\gamma + x}{1 + \beta x} \qquad \forall\, 0 \leqslant x \leqslant \lambda. \tag{32}$$

Next we define the region $\mathcal{U}$ in the Lemma. For $\eta > 0$, define the positive tube and the negative tube by

$$\mathcal{U}_\eta^+ := \bigcup_{x \in [0,\lambda]} D(x, \eta h_+(x)) \quad \text{and} \quad \mathcal{U}_\eta^- := \bigcup_{x \in [0,\rho\lambda]} D(-x, \eta h_+(x)), \tag{33}$$

where $D(x, r)$ is the open disk centered at x with radius r. See Figure 4 for an illustration. We use $\mathcal{U}^+ = \mathcal{U}_\eta^+$ and $\mathcal{U}^- = \mathcal{U}_\eta^-$ for some small enough constant $\eta = \eta(\beta, \gamma, \lambda, \delta_1) > 0$. The value of the constant η will be determined later in the proof.

By choosing η small enough with respect to δ_1, the region $\mathcal{U}_\eta = \mathcal{U}_\eta^+ \cup \mathcal{U}_\eta^-$ satisfies

$$I_+ \subset \mathcal{U}_\eta^+, \qquad I_- \subset \mathcal{U}_\eta^-, \qquad \mathcal{U}_\eta \subset \mathcal{S}_{\delta_1}(I).$$

This verifies the first item of the Lemma.

Next, we prove Lemma 22 (ii). This is guaranteed by continuity of $C_z(x)$ with respect to z and x. When choosing $z = \theta$, we have that for any $0 \leqslant x \leqslant \lambda$,

$$C_\theta(x) = (1 + x) + \theta(1 - \rho x) > C,$$

for some constant $C > 0$. This is because $1 - \rho\theta = \frac{\beta+\gamma-2}{\beta\gamma-1} > 0$. (Recall that $\gamma > \beta > 1$, $\rho = \frac{\beta-1}{\gamma-1}$, and $\theta = \frac{(\gamma-1)^2}{\beta\gamma-1}$.) Then compactness gives, after shrinking η if necessary and taking a sufficiently small constant $\delta_0 > 0$, that

$$C_z(a) \neq 0 \qquad \text{for every } a \in \mathcal{U}_\eta^+,\ z \in U_{\delta_0}(\theta).$$

This verifies the second item of the Lemma.

Finally, we prove Lemma 22 (iii). We introduce an auxiliary Lemma to prove this item.

Lemma 29. *For all sufficiently small $\eta > 0$, it holds that for any $c \in I$, $k \geqslant 1$, and $a_i \in \mathcal{U}_\eta^+$ with $1 \leqslant i \leqslant k$,*

$$\operatorname{dist}\left(c \prod_{i=1}^{k} \Phi_\theta(a_i), \partial\mathcal{U}_\eta\right) \geqslant \Omega_{\beta,\gamma,\lambda}(\eta),$$

where $\partial\mathcal{U}_\eta$ is the boundary of $\mathcal{U}_\eta$.

We first prove Lemma 22 (iii) via Lemma 29 and then prove Lemma 29. Again, the property (iii) is guaranteed by continuity of $\Phi_z(a)$. For notational convenience, we define a function $M(a) := \Phi_\theta(a)$. We

[4]The results in [GL18] actually work for λ up to the threshold $\lambda_c > \lambda^\star$.

now simplify this expression. Recall that $\rho = \frac{\beta-1}{\gamma-1}$ and $\theta = \frac{(\gamma-1)^2}{\beta\gamma-1}$. Then, the numerator and denominator of $M(a)$ are given by:

$$(1+a) - \rho\theta(1-\rho a) = (1-\rho\theta) + (1+\rho^2\theta)a = \frac{\beta+\gamma-2}{\beta\gamma-1}(1+\beta a),$$
$$(1+a) + \theta(1-\rho a) = (1+\theta) + (1-\rho\theta)a = \frac{\beta+\gamma-2}{\beta\gamma-1}(\gamma+a).$$

The common factor $(\beta+\gamma-2)/(\beta\gamma-1)$ is nonzero under our assumption $\gamma > \beta > 1$. Hence

$$M(a) = \frac{1+\beta a}{\gamma+a}.$$

We remark that $M(a)$ has the same form as the factor on one edge e contributed by the tree recursion in [GL18]. In particular, on the real interval $a \in [-1, 1/\rho]$, the denominator is nonzero since $\gamma + a \geqslant \gamma - 1 > 0$. We have the following properties of $M(a)$:

$$M(-1) = -\rho \geqslant -1, \quad M(0) = \frac{1}{\gamma} > 0, \quad M(1/\rho) = 1,$$

$$\text{and} \quad M'(a) = \frac{\beta\gamma-1}{(\gamma+a)^2} > 0 \quad \text{for } a \in (-1, 1/\rho).$$

By the fact that $\lambda < \lambda^\star < 1/\rho$, there is a constant $q < 1$ such that

$$0 < M(x) \leqslant q < 1 \qquad \text{for every } x \in [0, \lambda].$$

By the definition of $\Phi_z(a)$, we have that $\Phi_z(a)$ is analytic for $a \in \mathcal{U}_\eta^+$ and for z near θ. Thus $\Phi_z(a)$ is uniformly close to $M(a)$:

$$\Phi_z(a) = M(a) + r_z, \quad |r_z| \leqslant C|z-\theta| \text{ for a constant } C > 0.$$

Moreover, for sufficiently small η and δ, there exists $q_1 < 1$ such that

$$|\Phi_z(a)| \leqslant q_1 \quad \text{and} \quad |M(a)| \leqslant q_1 \qquad (a \in \mathcal{U}_\eta^+,\ z \in U_\delta(\theta)).$$

For each i, write

$$\Phi_{z_i}(a_i) = M(a_i) + r_{z_i}, \qquad |r_{z_i}| \leqslant C\delta.$$

Then the difference of the two products can be expanded as

$$\prod_{i=1}^k \Phi_{z_i}(a_i) - \prod_{i=1}^k M(a_i) = \sum_{i=1}^k \left(\prod_{j<i} \Phi_{z_j}(a_j)\right) r_{z_i} \left(\prod_{j>i} M(a_j)\right).$$

Therefore, uniformly in k,

$$\left|\prod_{i=1}^k \Phi_{z_i}(a_i) - \prod_{i=1}^k M(a_i)\right| \leqslant C\delta \sum_{i=1}^k q_1^{k-1} \leqslant C\delta\, k q_1^{k-1} \leqslant C'\delta,$$

Since $\sup_k k q_1^{k-1} < \infty$ and h_+ is lower bounded by constant, by Lemma 29, choosing $\delta \ll \eta$ makes this perturbation smaller than the strict margin obtained above. Thus the conclusions for $z_i = \theta$ persist for all $z_i \in U_\delta(\theta)$. Taking $\mathcal{U}^+ = \mathcal{U}_\eta^+$, $\mathcal{U}^- = \mathcal{U}_\eta^-$, and $\delta_2 = \delta$ proves Lemma 22 (iii).

The remaining part of this section is devoted to the proof of Lemma 29. We write each a_i in $\mathcal{U}_\eta^+$ as

$$a_i = x_i + \xi_i \in \mathcal{U}_\eta^+, \qquad x_i \in [0, \lambda], \qquad |\xi_i| < \eta h_+(x_i). \tag{34}$$

Since M is analytic in a complex neighborhood of $[0, \lambda]$, Taylor expansion gives, uniformly in i and x_i,

$$M(a_i) = M(x_i) + M'(x_i)\xi_i + r_i, \quad \text{where } |r_i| = O_{\beta,\gamma,\lambda}(|\xi_i|^2).$$

Since $0 < M(x_i) < 1$ on $[0,\lambda]$, expanding the product gives

$$\prod_{i=1}^{k} M(a_i) = \prod_{i=1}^{k} M(x_i) + \sum_{i=1}^{k} (M'(x_i)\xi_i) \prod_{j\neq i} M(x_j) + E_k, \tag{35}$$

where E_k is the error term. We need the following lemma to bound the first-order term.

Lemma 30. *There exists $\kappa \in (0,1)$ such that, for every $k \geqslant 1$, every $c \in [0,\lambda]$, and every $x_i \in [0,\lambda]$,*

$$c\sum_{i=1}^{k} M'(x_i)h_+(x_i) \prod_{j\neq i} M(x_j) \leqslant \kappa\, h_+\left(c\prod_{i=1}^{k} M(x_i)\right). \tag{36}$$

Proof. Since $0 < M(x_i) < 1$ and $0 \leqslant c \leqslant \lambda$, we have $c\prod_{i=1}^{k} M(x_i) \in [0,\lambda]$, which is in the domain of h_+. If $c = 0$, then (36) is trivial. Hence assume $c > 0$. Let

$$F = c\prod_{i=1}^{k} M(x_i).$$

We remark that F is exactly the tree recursion function appeared in [GL18]. By $M(x) \in (0,1)$ for $x \in [0,\lambda]$, it holds that $0 < F \leqslant c \leqslant \lambda$ and

$$\frac{M'(x)}{M(x)} = \frac{\beta\gamma - 1}{(1+\beta x)(\gamma + x)}.$$

By Lemma 28,

$$\frac{M'(x_i)}{M(x_i)} h_+(x_i) \leqslant \kappa \log \frac{1}{M(x_i)}.$$

Therefore

$$c\sum_{i} M'(x_i)h_+(x_i) \prod_{j\neq i} M(x_j) = F\sum_{i} \frac{M'(x_i)}{M(x_i)} h_+(x_i) \leqslant \kappa F \sum_{i} \log\frac{1}{M(x_i)} = \kappa F \log\frac{c}{F}.$$

Since $h_+(y) \geqslant y\log(\zeta/y)$ and $0 < F \leqslant c \leqslant \lambda < \zeta$, we have

$$F\log\frac{c}{F} \leqslant h_+(F).$$

□

We next bound the error term in the following lemma.

Lemma 31. *There exists a constant C_1 such that, for any integer $k \geqslant 1$, any $c \in [0,\lambda]$, it holds that*

$$c|E_k| \leqslant C_1\eta^2 h_+\left(c\prod_{i=1}^{k} M(x_i)\right),$$

where E_k is the error term defined in (35).

Proof. To justify this bound, write

$$M(a_i) = m_i + \ell_i + r_i, \qquad m_i = M(x_i), \quad \ell_i = M'(x_i)\xi_i,$$

with $|r_i| \leqslant C|\xi_i|^2$ for some constant $C > 0$. Since h_+ is bounded from above on $[0,\lambda]$, we have $|\ell_i| \leqslant C\eta$ and $|r_i| \leqslant C\eta^2$ by letting constant C be sufficiently large. Also, $0 < m_i \leqslant q < 1$ uniformly on $[0,\lambda]$ for some $q < 1$. Choose η small enough so that $q + C\eta < q_1 < 1$ for some $q_1 < 1$. We view the product $\prod_{i=1}^{k}(m_i+\ell_i+r_i)$ as a polynomial in ℓ_i and r_i with coefficients depending on m_i. Expanding the product and subtracting the constant term and the terms linear in ℓ_i's, the remaining terms are of two types. All the terms with one r_i have total absolute value at most

$$C\eta^2 k q_1^{k-1} \leqslant C'\eta^2, \text{ because } q_1 < 1.$$

Similarly, all the terms with at least two factors from the collection $\{\ell_i, r_i\}$ have total absolute value at most $\sum_{r=2}^k \binom{k}{r}(C\eta)^r q^{k-r}$. Indeed, for $r \geqslant 2$, $\binom{k}{r} = \frac{k(k-1)}{r(r-1)}\binom{k-2}{r-2} \leqslant k^2\binom{k-2}{r-2}$. Therefore, using $q + C\eta < q_1$,

$$\sum_{r=2}^{k}\binom{k}{r}(C\eta)^r q^{k-r} \leqslant C^2\eta^2k^2\sum_{r=2}^{k}\binom{k-2}{r-2}(C\eta)^{r-2}q^{k-r}$$
$$= C^2\eta^2k^2(q+C\eta)^{k-2} \leqslant C^2\eta^2k^2q_1^{k-2} \leqslant C''\eta^2,$$

where the last inequality follows from $q_1 < 1$. Hence $|E_k| \leqslant C_0\eta^2$ uniformly in k. Since $c \leqslant \lambda$ and h_+ is lower bounded by a positive constant on $[0, \lambda]$, this uniform bound implies Lemma 31. $\square$

Finally, we use Lemma 30 and Lemma 31 to prove Lemma 29. First suppose $c \in I_+$. Define

$$F_0 = c\prod_{i=1}^{k} M(x_i) \in I_+,$$

where each $a_i = x_i + \xi_i$ is as in (34). Using Lemma 30 and $|\xi_i| \leqslant \eta h_+(x_i)$, we get

$$\left| c\sum_{i=1}^{k} M'(x_i)\xi_i \prod_{j\neq i} M(x_j)\right| \leqslant \eta\, c\sum_{i=1}^{k} M'(x_i)h_+(x_i)\prod_{j\neq i}M(x_j) \leqslant \kappa\eta h_+(F_0).$$

Together with Lemma 31, this gives

$$\left| c\prod_{i=1}^{k} M(a_i) - F_0\right| \leqslant (\kappa + C_1\eta)\eta h_+(F_0). \tag{37}$$

Note that $\kappa < 1$. Choosing η small enough so that $\kappa + C_1\eta < 1$, we obtain

$$c\prod_{i=1}^{k} M(a_i) \in D(F_0, \eta h_+(F_0)).$$

Since $F_0 \in [0, \lambda]$, this disk is contained in the union definition $\mathcal{U}_\eta^+$ of (33). Hence $c\prod_{i=1}^k M(a_i) \in \mathcal{U}_\eta^+$. In addition, by (37), we have $\mathrm{dist}(c\prod_{i=1}^k M(a_i), \partial\mathcal{U}_\eta^+) \geqslant \Omega(\eta)$ uniformly for $a_i \in \mathcal{U}_\eta^+$, since $h_+(F_0)$ is uniformly lower bounded by constant t. Similarly, for a negative field $-c \in I_-$ with $c \in [0, \rho\lambda] \subset [0, \lambda]$, the same estimate gives

$$\left| -c\prod_{i=1}^{k} M(a_i) - (-F_0)\right| \leqslant (\kappa + C_1\eta)\eta h_+(F_0),$$

where now $F_0 = c\prod_i M(x_i) \in [0, \rho\lambda]$. Thus $-c\prod_{i=1}^k M(a_i) \in \mathcal{U}_\eta^-$. Similarly, by the above inequality, we have $\mathrm{dist}(-c\prod_{i=1}^k M(a_i), \partial\mathcal{U}_\eta^-) \geqslant \Omega(\eta)$. This finishes the proof of Lemma 29.

4.4. **Local stability.** We prove Lemma 26 in this subsection. The following two geometric facts about the disk will be used in our analysis; their proofs are deferred to Section 6.3. We define the closure of a open region $\mathcal{U}$ as $\mathrm{cl}(\mathcal{U})$, which is the smallest closed set containing $\mathcal{U}$.

Claim 32. *When $\gamma > \beta > 1$, it holds that $c^\star > 0$ and $r^\star - c^\star > 1$. Therefore, $D(c^\star, r^\star)$ contains $[-1, c^\star]$.*

Claim 33. *It holds that $\frac{1}{\rho} \notin \mathrm{cl}(D(c^\star, r^\star))$.*

By Claim 33 and $D(c^\star, r^\star)$ being an open set, $\psi(0) = 1/\rho \notin D(c^\star, r^\star - \epsilon)$ for any $\epsilon > 0$. Hence $0 \notin \Gamma_\epsilon$. Let $K_\epsilon := \mathbb{C} \setminus D(c^\star, r^\star - \epsilon)$, and then we have the following technical Lemma. Define the kth-power of a set $S \subseteq \mathbb{C}$ as $S^k := \{\prod_{i=1}^k s_i : s_i \in S\}$.

Lemma 34. *Fix $0 < \lambda < \lambda^\star$, and let $I = [-\rho\lambda, \lambda]$. There exist $\epsilon > 0$ and $\delta' > 0$, depending only on β, γ, λ, such that for every integer $d \geqslant 2$, it holds that*

$$(-1)^{d+1} K_\epsilon^d \cap \mathcal{S}_{\delta'}(I) = \emptyset.$$

Note that Lemma 34 does not hold for $d = 1$ and that is why we contract bad leaves first.

We first use Lemma 34 to prove Lemma 26 and then we prove Lemma 34.

Proof of Lemma 26. Let ϵ_0 and δ_0 be the two constants ϵ and δ' from Lemma 34. We will show Lemma 26 for some constants (ϵ, δ_1) such that $0 < \epsilon < \epsilon_0$ and $0 < \delta_1 < \delta_0$. We remark that since Lemma 34 holds for ϵ_0 and δ_0, it also holds for smaller constants ϵ and δ_1.

Assume first that $q_i \neq \rho$ for every i. Then, by the definition $u_i = \psi(q_i)$, we have

$$\widetilde{P}_{v,d_v}(q_1, \dots, q_{d_v}) = \prod_{i=1}^{d_v} (q_i - \rho) \left((-1)^{d_v} \prod_{i=1}^{d_v} u_i + \lambda_v \right).$$

Thus a zero can occur only if $(-1)^{d_v+1} \prod_{i=1}^{d_v} u_i = \lambda_v$. Since $q_i \notin \Gamma_\epsilon$, we have $u_i \in K_\epsilon$ for every i. For $d_v \geqslant 2$, this contradicts Lemma 34 because $\lambda_v \in \mathcal{S}_{\delta_1}(I)$. For $d_v = 1$, by Condition 19, we have $\lambda_v \in \mathcal{S}_{\delta_1}((-\rho\lambda, 0))$, and if zero can occur, we must have $u_1 \in \mathcal{S}_{\delta_1}((-\rho\lambda, 0))$. We show this cannot happen for sufficiently small ϵ and δ_1. Note that the following two facts hold simultaneously.

- $u_1 \notin D(c^\star, r^\star - \epsilon)$ because $q_1 \notin \Gamma_\epsilon$.
- By Claim 32, $[-1, c^\star]$ is contained in $D(c^\star, r^\star)$, where $c^\star, r^\star > 0$ are constants. Recall that $\lambda < \lambda^\star$. By Lemma 11, $-\rho\lambda > -\rho\lambda^\star > -1$. Hence, both $-\rho\lambda$ and 0 are inner points of $D(c^\star, r^\star)$. For sufficiently small constants ϵ and δ_1, it holds that $S_{\delta_1}((-\rho\lambda, 0)) \subset D(c^\star, r^\star - \epsilon)$.

Combining these two facts, we deduce that $u_1 \in \mathcal{S}_{\delta_1}((-\rho\lambda, 0))$ is impossible.

It remains to handle the case where $q_i = \rho$ for some i. Then the second product in the RHS of (27) vanishes. (Recall that (27) still holds for this case.) The first product can vanish only if $q_j = -1$ for some j. But $q_j = -1$ gives $\psi(q_j) = 0$. Note that by Claim 32, $0 \in D(c^\star, r^\star - \epsilon)$ for sufficiently small constant ϵ, and hence $q_j \in \Gamma_\epsilon$, contradicting the assumption $q_j \notin \Gamma_\epsilon$ in Lemma 26. Therefore no zero occurs in this case either. ☐

Next we prove Lemma 34. We need the following lemma from [GLL20]. Recall that $K_\epsilon = \mathbb{C} \setminus D(c^\star, r^\star - \epsilon)$. Let $K := \mathbb{C} \setminus D(c^\star, r^\star) \subset K_\epsilon$.

Lemma 35 ([GLL20, Lemma 8])**.** *For any $d \geqslant 2$ and any $0 < \lambda < \lambda^\star$, $(-1)^{d+1} K^d \cap [0, \lambda] = \emptyset$.*

It is straightforward to extend Lemma 35 to the following corollary.

Corollary 36. *For any $d \geqslant 2$ and any $0 < \lambda < \lambda^\star$, $(-1)^{d+1} K^d \cap I = \emptyset$ for $I = [-\rho\lambda, \lambda]$.*

Proof. By Lemma 35, we only need to show that $(-1)^{d+1} K^d \cap [-\rho\lambda, 0] = \emptyset$. For $a \in [-\rho\lambda, 0]$, we have $|a| \leqslant \rho\lambda < 1$ due to Lemma 11. By Claim 32, we have $r^\star - c^\star > 1$. Therefore, every $u \in K$ satisfies $|u| > 1$. Every product in K^d has modulus greater than 1, which cannot equal any point of $(-1)^{d+1}[-\rho\lambda, 0]$. ☐

Now we can prove Lemma 34. Comparing to Corollary 36, the difference is that we need to consider K_ϵ, which is slightly larger than K, and we want to avoid a strip around the interval I. This strengthening is necessary for Lemma 27 later.

Proof of Lemma 34. By Claim 32, $r^\star - c^\star > 1$. Set constants

$$M := \max\{\lambda, \rho\lambda\}, \qquad m_0 := \frac{r^\star - c^\star + 1}{2} > 1, \qquad d_0 := \max\{2, \lceil \log_{m_0}(M+1) \rceil\}.$$

Choose $\epsilon_0 > 0$ small enough that $r^\star - c^\star - \epsilon_0 > m_0 > 1$. Then, for every $0 < \epsilon \leqslant \epsilon_0$ and every $u \in K_\epsilon$, $|u| \geqslant |u - c^\star| - c^\star \geqslant r^\star - c^\star - \epsilon \geqslant m_0$. In particular, if $u_1, \dots, u_d \in K_\epsilon$, then $|u_1 \cdots u_d| \geqslant m_0^d$.

If $d > d_0$, then $m_0^d > M+1$. Since $\mathcal{S}_1(I) \subseteq \{a \in \mathbb{C} : |a| < M+1\}$, we have $(-1)^{d+1}K_\epsilon^d \cap \mathcal{S}_1(I) = \emptyset$ for $\delta_0 = 1$, any $d > d_0$, and any $0 < \epsilon \leqslant \epsilon_0$.

It remains to handle degrees d between 2 and d_0. Let

$$R := \max\{r^\star - c^\star, M+1\}, \qquad L_\epsilon := K_\epsilon \cap \mathrm{cl}\left(D(0,R)\right).$$

For $0 < \delta \leqslant 1$, avoiding $\mathcal{S}_\delta(I)$ for products from L_ϵ is enough. Indeed, if $u_1, \ldots, u_d \in K_\epsilon$ and some $u_i \notin L_\epsilon$, then $|u_i| > R$, while $|u_j| \geqslant m_0 > 1$ for $j \neq i$. Hence $|\prod_{j=1}^d u_j| > Rm_0^{d-1} \geqslant M+1$. Thus $(-1)^{d+1}\prod_j u_j \notin \mathcal{S}_\delta(I)$, because every point of $\mathcal{S}_\delta(I)$ has modulus less than $M+1$. Consequently,

$$(-1)^{d+1}L_\epsilon^d \cap \mathcal{S}_\delta(I) = \emptyset \quad \Longrightarrow \quad (-1)^{d+1}K_\epsilon^d \cap \mathcal{S}_\delta(I) = \emptyset. \tag{38}$$

We now prove the left-hand side of (38) for each fixed $2 \leqslant d \leqslant d_0$ and some δ_d depending on d. Note that the set L_ϵ is compact. We need some standard notations. For a point $a \in \mathbb{C}$ and a nonempty set $S \subseteq \mathbb{C}$, write $\mathrm{dist}(a, S) := \inf_{z \in S} |a - z|$. For nonempty compact sets $A, B \subseteq \mathbb{C}$, the Hausdorff distance is defined by

$$\mathrm{dist}_H(A, B) := \max\left\{\sup_{a \in A} \mathrm{dist}(a, B),\ \sup_{b \in B} \mathrm{dist}(b, A)\right\}.$$

The compact sets L_ϵ converge, as $\epsilon \to 0$, in Hausdorff distance to $L_0 := K \cap \mathrm{cl}\left(D(0,R)\right)$. Since multiplication is uniformly continuous on the fixed compact set $\mathrm{cl}\left(D(0,R)\right)^d$, the product sets $L_\epsilon^d = \{\prod_{i=1}^d u_i : u_i \in L_\epsilon\}$ also converge in Hausdorff distance to L_0^d. Moreover, $L_0 \subseteq K$, so Corollary 36 gives $(-1)^{d+1}L_0^d \cap I = \emptyset$. Both sets are compact, and thus their distance is positive. Write

$$\eta_d := \mathrm{dist}\left(I, (-1)^{d+1}L_0^d\right) > 0,$$

where $\eta_d > 0$ is a positive constant depending on $\beta, \gamma, \lambda, d$. Since L_ϵ^d converges in Hausdorff distance to L_0^d, there exists $\epsilon_d > 0$ such that the following holds:

$$\mathrm{dist}_H\left((-1)^{d+1}L_\epsilon^d, (-1)^{d+1}L_0^d\right) < \eta_d/3 \qquad \text{whenever } 0 < \epsilon \leqslant \epsilon_d.$$

Taking $\delta_d := \eta_d/3$, we get

$$(-1)^{d+1}L_\epsilon^d \cap \mathcal{S}_{\delta_d}(I) = \emptyset \qquad \text{for all } 0 < \epsilon \leqslant \epsilon_d.$$

Finally choose constants ϵ and δ' as follows: $\epsilon := \min\{\epsilon_0, \epsilon_2, \ldots, \epsilon_{d_0}\}$ and $\delta' := \min\{\delta_0, \delta_2, \ldots, \delta_{d_0}\}$. The large-degree argument covers all $d > d_0$, while the finite-degree argument and (38) cover all $2 \leqslant d \leqslant d_0$. Therefore $(-1)^{d+1}K_\epsilon^d \cap \mathcal{S}_{\delta'}(I) = \emptyset$ for every $d \geqslant 2$, as required. □

4.5. **Contraction preserves stability.** Next we prove Lemma 27. We first state several auxiliary lemmas needed for the proof. Consider the single-variable polynomial $P(x) = A + 2Bx + Cx^2$ with parameters $A, B, C \in \mathbb{C}$. Its corresponding polar form in two variables x_1, x_2 is defined as

$$\widehat{P}(x_1, x_2) = A + B(x_1 + x_2) + Cx_1x_2.$$

Let Γ be an open connected region in $\mathbb{C}$. We say that Γ is a circular region if it is a disk, or the complement of a disk in $\mathbb{C}$. The classical Grace–Szegő–Walsh coincidence theorem has the following immediate consequence.

Proposition 37 (Grace–Szegő–Walsh coincidence theorem [BB09, Theorem 11])**.** *Let Γ be a circular region and $C > 0$ in $\widehat{P}(x_1, x_2)$. The polynomial $\widehat{P}(x_1, x_2)$ is $\Gamma \times \Gamma$-stable if and only if $P(x)$ is Γ-stable.*

The next ingredient is the classical Asano contraction lemma [Asa70, Rue71]. Here we use a simplified version for quadratic polynomials.

Lemma 38 (Asano contraction). *Let $\Gamma \subset \mathbb{C}$ with $0 \notin \Gamma$. If the multivariate polynomial*

$$P(x_1, x_2) = A + Bx_1 + Cx_2 + Dx_1x_2$$

is stable for $(x_1, x_2) \in (\mathbb{C} \setminus \Gamma) \times (\mathbb{C} \setminus \Gamma)$, then

$$Q(x) := A + Dx$$

is stable for $x \in \mathbb{C} \setminus \left(-\Gamma^2\right)$, where $\Gamma^2 := \{x \cdot y : x, y \in \Gamma\}$.

Proof. Since $0 \notin \Gamma$, we have $A = P(0,0) \neq 0$. Without loss of generality, we assume $D \neq 0$, since otherwise the polynomial Q is a constant and thus always stable. By the assumption on P, we have $\widetilde{P}(x) = A + (B+C)x + Dx^2$ is stable for $x \in \mathbb{C} \setminus \Gamma$. Thus $\widetilde{P}(x) = 0$ can occur only when $x \in \Gamma$. Suppose the two roots of $\widetilde{P}(x) = 0$ are $\zeta_1, \zeta_2 \in \Gamma$. By Vieta's formula, we have $\zeta_1\zeta_2 = A/D$. On the other hand, $Q(x) = 0$ can occur only for $x = -A/D \in -\Gamma^2$. Therefore, $Q(x)$ is stable for $x \in \mathbb{C} \setminus \left(-\Gamma^2\right)$. □

For our analysis, we need an auxiliary property of the disk $D(c^\star, r^\star - \epsilon)$, as well as an inequality for the parameters. The proofs of the following two results are deferred to Section 6.3.

Lemma 39. *Let ζ_1, ζ_2 be the two roots of $\beta x^2 + 2x + \gamma = 0$. For any $0 < \epsilon < r^\star/2$, it holds that*

$$\mathrm{cl}\left(D(c^\star, r^\star - \epsilon)\right) \cap \{\zeta_1, \zeta_2\} = \emptyset.$$

Claim 40. *It holds that $\rho\theta < 1$.*

Finally, we are ready to prove Lemma 27.

Proof of Lemma 27. Fix $0 < \epsilon < r^\star/2$. Applying Lemma 38 with the region Γ_ϵ, it suffices to show that a sufficiently small neighborhood of $1/\theta$ is contained in $\mathbb{C} \setminus (-\Gamma_\epsilon^2)$.

By Claim 33 and the continuity of ψ at 0, we have $0 \notin \mathrm{cl}\left(\Gamma_\epsilon\right)$, where $\mathrm{cl}\left(\cdot\right)$ denotes the closure of a set. Hence there is a constant $a_0 > 0$ such that $\Gamma_\epsilon \cap D(0, a_0) = \emptyset$. Set

$$B_0 := |1/\theta| + 1, \qquad R_0 := \frac{B_0}{a_0}, \qquad \Gamma_{\epsilon,R} := \Gamma_\epsilon \cap \mathrm{cl}\left(D(0, R_0)\right).$$

If $q_1, q_2 \in \Gamma_\epsilon$ and $-q_1q_2 \in U_1(1/\theta)$, then $|q_1q_2| < B_0$, while $|q_i| \geqslant a_0$ for $i = 1, 2$. Thus $|q_i| < R_0$ for $i = 1, 2$, and hence $q_1, q_2 \in \Gamma_{\epsilon,R}$. Therefore, to find a neighborhood of $1/\theta$ avoiding $-\Gamma_\epsilon^2$, it suffices to show that $\frac{1}{\theta} \notin \mathrm{cl}\left(-\Gamma_{\epsilon,R}^2\right)$. Now, we are working with a bounded set $\Gamma_{\epsilon,R}$.

Suppose, for contradiction, that $1/\theta \in \mathrm{cl}\left(-\Gamma_{\epsilon,R}^2\right)$. Since $\Gamma_{\epsilon,R}$ is bounded, the continuity of multiplication gives $q_1, q_2 \in \mathrm{cl}\left(\Gamma_{\epsilon,R}\right) \subseteq \mathrm{cl}\left(\Gamma_\epsilon\right)$ such that $1/\theta = -q_1q_2$. Since $q_i \in \mathrm{cl}\left(\Gamma_\epsilon\right)$, we have $q_i \neq \rho$, and the Möbius variables $u_i := \psi(q_i)$ lie in $\mathrm{cl}\left(D(c^\star, r^\star - \epsilon)\right)$. Moreover $u_i \neq -1$, since $\psi(q) = -1$ has no finite solution. Solving $u_i = -(1+q_i)/(q_i - \rho)$, we obtain for $i = 1, 2$ that $q_i = \frac{\rho u_i - 1}{1 + u_i}$. Using $\frac{1}{\theta} = -q_1q_2$, we get the following equation:

$$\frac{1}{\theta} = -\frac{(1 - \rho u_1)(1 - \rho u_2)}{(1 + u_1)(1 + u_2)}.$$

Equivalently,

$$(1 + u_1)(1 + u_2) + \theta(1 - \rho u_1)(1 - \rho u_2) = 0.$$

Expanding and using the definitions of $\theta = (\gamma - 1)^2/(\beta\gamma - 1)$ and $\rho = (\beta - 1)/(\gamma - 1)$, this becomes

$$(1 - \rho\theta)(\beta u_1 u_2 + u_1 + u_2 + \gamma) = 0.$$

By Claim 40, $1 - \rho\theta \neq 0$, and therefore $\beta u_1 u_2 + u_1 + u_2 + \gamma = 0$. Since $u_1, u_2 \in \mathrm{cl}\left(D(c^\star, r^\star - \epsilon)\right) \subset D(c^\star, r^\star - \epsilon/2)$, Proposition 37 applied to the circular region $D(c^\star, r^\star - \epsilon/2)$ implies that $\beta x^2 + 2x + \gamma = 0$ for some $x \in D(c^\star, r^\star - \epsilon/2)$. This contradicts Lemma 39, applied with $\epsilon/2$.

Thus $1/\theta \notin \mathrm{cl}\left(-\Gamma^2_{\epsilon,R}\right)$. Hence there exists $0 < \delta_3 \leqslant 1$ such that $U_{\delta_3}(1/\theta) \subseteq \mathbb{C} \setminus \mathrm{cl}\left(-\Gamma^2_{\epsilon,R}\right)$. To finish the proof, we still need to go back to Γ_ϵ . Take any $z \in U_{\delta_3}(1/\theta)$. If $z = -q_1q_2$ for some $q_1, q_2 \in \Gamma_\epsilon$, then $z \in U_1(1/\theta)$, since $\delta_3 \leqslant 1$. By the choice of R_0, this forces $q_1, q_2 \in \Gamma_{\epsilon,R}$, and hence $z \in -\Gamma^2_{\epsilon,R}$. This contradicts the choice of δ_3. Therefore, $U_{\delta_3}(1/\theta) \subseteq \mathbb{C} \setminus (-\Gamma^2_\epsilon)$. The construction of δ_3 depends only on β, γ, ϵ. This proves the lemma. □

5. Simulated annealing reductions

The sampling algorithm in Theorem 1 implies an FPRAS for $Z^{\mathrm{spin}}_G(\beta, \gamma, \lambda)$ by annealing. We could apply the adaptive procedure of [SVV09], but here a simple non-adaptive annealing for the external field λ in the original spin system will do. We include a proof for completeness.

Fix constants $\beta, \gamma, \lambda > 0$ such that $\gamma > \beta > 1$ and $\lambda < \lambda^\star$. Let $G = (V, E)$ and $m = |E|$. Assume $0 < \epsilon \leqslant 1$. Set

$$q = 1 - \frac{1}{5n}.$$

Choose $\ell = O_\lambda(n \log(n/\epsilon))$ large enough so that, for the sequence

$$\lambda_\ell = \lambda, \qquad \lambda_{i-1} = q\lambda_i \quad \text{for } i = 1, \dots, \ell,$$

we have $\lambda_0 \leqslant \epsilon/(20n)$. For notational convenience, set $\lambda_{\ell+1} = \lambda_\ell/q$. For $0 \leqslant i \leqslant \ell + 1$, write

$$Z_i = Z^{\mathrm{spin}}_G(\beta, \gamma, \lambda_i), \qquad \mu_i = \mu^{\mathrm{spin}}_G(\beta, \gamma, \lambda_i) \quad (0 \leqslant i \leqslant \ell).$$

The target partition function is Z_ℓ, and

$$Z_\ell = \frac{Z_\ell}{Z_{\ell-1}} \cdot \frac{Z_{\ell-1}}{Z_{\ell-2}} \cdots \frac{Z_1}{Z_0} \cdot Z_0.$$

Lemma 41. *For the above choice of ℓ,*

$$\gamma^m \leqslant Z_0 \leqslant e^{\epsilon/10}\gamma^m.$$

Proof. The all-1 configuration contributes exactly γ^m, so $Z_0 \geqslant \gamma^m$. For any configuration σ with $k = n_0(\sigma)$ vertices in spin 0, the edge contribution is at most γ^m, because $\gamma > \beta > 1$. Therefore

$$Z_0 \leqslant \gamma^m \sum_{k=0}^{n} \binom{n}{k} \lambda_0^k = \gamma^m (1 + \lambda_0)^n \leqslant e^{\epsilon/10}\gamma^m,$$

where the last inequality follows from $\lambda_0 \leqslant \epsilon/(20n)$. □

To estimate the telescoping product, consider the following estimator:

$$W_i(\sigma) = \frac{w^{\mathrm{spin}}_G(\sigma; \beta, \gamma, \lambda_i)}{w^{\mathrm{spin}}_G(\sigma; \beta, \gamma, \lambda_{i-1})} = \left(\frac{\lambda_i}{\lambda_{i-1}}\right)^{n_0(\sigma)} \qquad \text{where } \sigma \sim \mu_{i-1}.$$

Let $W_1, W_2, \dots, W_\ell$ be independent random variables, with W_i sampled using μ_{i-1}, and define $W = \prod_{i=1}^{\ell} W_i$. A simple calculation shows that

$$\mathbb{E}[W_i] = \frac{Z_i}{Z_{i-1}} \quad \text{and} \quad \mathbb{E}[W] = \frac{Z_\ell}{Z_0}. \tag{39}$$

The following lemma bounds the variance of W_i and W.

Lemma 42. *For every $1 \leqslant i \leqslant \ell$,*

$$\mathrm{Var}\,(W_i) \leqslant \mathbb{E}[W_i^2] = \frac{Z_{i+1}}{Z_{i-1}}.$$

As a consequence,

$$\mathrm{Var}\,(W) \leqslant \mathbb{E}[W^2] \leqslant O_{\beta,\gamma,\lambda}(1)\left(\frac{Z_\ell}{Z_0}\right)^2.$$

Proof. Since the dependence on λ in $w_G^{\text{spin}}(\sigma;\beta,\gamma,\lambda)$ is exactly $\lambda^{n_0(\sigma)}$, and since $\lambda_{i-1}\lambda_{i+1} = \lambda_i^2$, we have

$$\frac{w_G^{\text{spin}}(\sigma;\beta,\gamma,\lambda_i)^2}{w_G^{\text{spin}}(\sigma;\beta,\gamma,\lambda_{i-1})} = w_G^{\text{spin}}(\sigma;\beta,\gamma,\lambda_{i+1}).$$

Hence

$$\mathbb{E}[W_i^2] = \frac{1}{Z_{i-1}} \sum_{\sigma:V\to\{0,1\}} w_G^{\text{spin}}(\sigma;\beta,\gamma,\lambda_{i+1}) = \frac{Z_{i+1}}{Z_{i-1}}.$$

Independence gives

$$\mathbb{E}[W^2] = \prod_{i=1}^{\ell} \frac{Z_{i+1}}{Z_{i-1}} = \frac{Z_{\ell+1}Z_\ell}{Z_1 Z_0}.$$

For every adjacent pair $\lambda_j = q^{-1}\lambda_{j-1}$,

$$1 \leqslant \frac{Z_j}{Z_{j-1}} = \mathbb{E}_{\mu_{j-1}}\left[\left(\frac{\lambda_j}{\lambda_{j-1}}\right)^{n_0(\sigma)}\right] \leqslant q^{-n} = O(1).$$

Thus $Z_0/Z_1 \leqslant 1$ and $Z_{\ell+1}/Z_\ell = O(1)$, which implies

$$\mathbb{E}[W^2] \leqslant O(1)\left(\frac{Z_\ell}{Z_0}\right)^2.$$

The variance bounds follow from $\mathrm{Var}\,(X) \leqslant \mathbb{E}[X^2]$. ☐

Lemma 43. *Assume that we can draw independent samples from $\mu_0, \mu_1, \ldots, \mu_{\ell-1}$. There is an estimator $\widehat{Z}_\ell$ such that, using $N = O_{\beta,\gamma,\lambda}(\epsilon^{-2})$ independent samples of the random variable W,*

$$\Pr\left[\widehat{Z}_\ell \in (1\pm\epsilon)Z_\ell\right] \geqslant \frac{9}{10}.$$

Proof. For each $r = 1,\ldots,N$, independently sample $\sigma_i^{(r)} \sim \mu_{i-1}$ for every $i = 1,\ldots,\ell$, and define

$$W_i^{(r)} = \left(\frac{\lambda_i}{\lambda_{i-1}}\right)^{n_0(\sigma_i^{(r)})}, \qquad W^{(r)} = \prod_{i=1}^{\ell} W_i^{(r)}.$$

Let

$$\overline{W} = \frac{1}{N}\sum_{r=1}^{N} W^{(r)} \qquad \text{and} \qquad \widehat{Z}_\ell = \gamma^m \overline{W}.$$

By (39), $\mathbb{E}[W^{(r)}] = Z_\ell/Z_0$. By Lemma 42,

$$\mathrm{Var}\left(W^{(r)}\right) \leqslant O_{\beta,\gamma,\lambda}(1)\left(\frac{Z_\ell}{Z_0}\right)^2.$$

Since $(W^{(r)})_r$ are independent,

$$\mathrm{Var}\left(\overline{W}\right) \leqslant \frac{O_{\beta,\gamma,\lambda}(1)}{N}\left(\frac{Z_\ell}{Z_0}\right)^2.$$

Taking $N = O_{\beta,\gamma,\lambda}(\epsilon^{-2})$ sufficiently large and applying Chebyshev's inequality gives

$$\Pr\left[\overline{W} \in \left(1 \pm \frac{\epsilon}{3}\right)\frac{Z_\ell}{Z_0}\right] \geqslant \frac{9}{10}.$$

On this event, $\frac{\widehat{Z}_\ell}{Z_\ell} = \frac{\gamma^m}{Z_0} \cdot \frac{\overline{W}}{Z_\ell/Z_0}$. By the estimate for Z_0 in Lemma 41, $e^{-\epsilon/10} \leqslant \frac{\gamma^m}{Z_0} \leqslant e^{\epsilon/10}$. Thus, for $0 < \epsilon \leqslant 1$, we have the following bound

$$\frac{\widehat{Z}_\ell}{Z_\ell} \in \left[e^{-\epsilon/10}\left(1 - \frac{\epsilon}{3}\right), e^{\epsilon/10}\left(1 + \frac{\epsilon}{3}\right)\right] \subseteq [1-\epsilon, 1+\epsilon].$$

This proves the claim. □

Proof of Theorem 2. We implement the estimator in Lemma 43 using the sampler from Theorem 1. Since all annealing fields satisfy $\lambda_i \leqslant \lambda < \lambda^\star$ for all $i = 0, \ldots, \ell - 1$, the constant in the sampling bound can be chosen uniformly as $C = C(\beta, \gamma, \lambda)$.

Let $N = O_{\beta,\gamma,\lambda}(\epsilon^{-2})$ be the number of independent copies of W used by the estimator. Each copy of W requires one sample from each of $\mu_0, \mu_1, \ldots, \mu_{\ell-1}$, so the total number of samples needed is $N\ell$. For each required sample, run the sampler of Theorem 1 with total variation error at most $1/(100N\ell)$. Its running time is

$$O\left(\Delta^C n \log(100N\ell n)\right).$$

By the union bound and the standard coupling interpretation of total variation distances, all approximate samples can be coupled with exact independent samples with probability at least $99/100$. Conditional on the success of the coupling, Lemma 43 gives a $(1 \pm \epsilon)$ approximation to $Z_\ell = Z_G^{\text{spin}}(\beta, \gamma, \lambda)$ with probability at least $9/10$. Hence the overall success probability is at least $9/10 - 1/100 > 2/3$.

The number of annealing stages is $\ell = O_{\beta,\gamma,\lambda}(n \log(n/\epsilon))$, and each stage requires $N = O_{\beta,\gamma,\lambda}(\epsilon^{-2})$ samples. Also $\log(100N\ell n) = O_{\beta,\gamma,\lambda}(\log(n/\epsilon))$. Therefore the overall running time is

$$O\left(\Delta^C \cdot \frac{n^2}{\epsilon^2} \cdot \log^2 \frac{n}{\epsilon}\right),$$

after adjusting the constant $C = C(\beta, \gamma, \lambda)$. This proves the claimed FPRAS. □

6. Missing proofs

In this section we collate deferred proofs.

6.1. Validity of the Gibbs distribution.

The positivity result in Theorem 4 is a consequence of Lemma 11.

Proof of Lemma 11. By $\rho < 1$, we only need to prove $\rho\lambda^\star < 1$. Set parameters

$$t = \sqrt{\beta\gamma}, \qquad h = \sqrt{\gamma/\beta}, \qquad \alpha = \arctan\sqrt{\beta\gamma - 1}.$$

Then $t = \sec\alpha$ for $1 < h < t$, where $h > 1$ follows from $\gamma > \beta$, and $h < t$ follows from $\beta = t/h > 1$. Moreover, by definition of $\lambda^\star$, we have

$$\lambda^\star = \left(\frac{\gamma}{\beta}\right)^{\pi/(2\alpha)} = h^{\pi/\alpha},$$

and

$$\frac{1}{\rho} = \frac{\gamma - 1}{\beta - 1} = \frac{th - 1}{t/h - 1}.$$

Thus it is enough to prove $h^{\pi/\alpha} < \frac{th-1}{t/h-1}$. Let $y = \log h$, so $0 < y < \log t$, and define

$$F(y) = \log\frac{te^y - 1}{te^{-y} - 1} - \frac{\pi}{\alpha}y.$$

Then $F(0) = 0$. We prove $F'(y) > 0$ for $0 < y < \log t$. Then it holds that $F(y) > 0$. In particular, for $y = \log h$, we have $\log \frac{th-1}{t/h-1} > \frac{\pi}{\alpha} \log h$. So $\frac{1}{\rho} = \frac{th-1}{t/h-1} > h^{\pi/\alpha} = \lambda^\star$. Indeed,

$$F'(y) = \frac{te^y}{te^y - 1} + \frac{te^{-y}}{te^{-y} - 1} - \frac{\pi}{\alpha}.$$

To prove $F'(y) > 0$, we let

$$S(y) = \frac{te^y}{te^y - 1} + \frac{te^{-y}}{te^{-y} - 1}.$$

Since the function $r \mapsto r/(r-1)^2$ is decreasing for $r > 1$, and since $te^{-y} \leqslant te^y$, we have

$$S'(y) = -\frac{te^y}{(te^y - 1)^2} + \frac{te^{-y}}{(te^{-y} - 1)^2} \geqslant 0.$$

Therefore $S(y) \geqslant S(0) = \frac{2t}{t-1}$. Using $t = \sec \alpha$, this gives $S(y) \geqslant \frac{2}{1-\cos\alpha}$. Since $0 < \alpha < \pi/2$, the concavity of cos on $[0, \pi/2]$ implies

$$\cos \alpha > 1 - \frac{2\alpha}{\pi} \quad \Longrightarrow \quad \frac{2}{1 - \cos \alpha} > \frac{\pi}{\alpha}.$$

Thus

$$F'(y) = S(y) - \frac{\pi}{\alpha} > 0.$$

Since $F(0) = 0$, we obtain $F(y) > 0$ for every $y > 0$. Equivalently, $\rho\lambda^\star < 1$. □

6.2. Proof of the marginal lower bound.

Proof of Lemma 7. Fix an edge $e = \{u, v\} \in E$. To prove the marginal lower bound, it suffices to consider the worst pinning $\tau \in \{0,1\}^{E \setminus \{e\}}$ for all edges except e. Suppose d_u edges incident to u are pinned to 1 and d_v edges incident to v are pinned to 1. Then, the conditional distribution μ_e^τ satisfies

$$\begin{aligned}
\mu_e^\tau(1) &= \frac{\theta(1+\lambda(-\rho)^{d_u+1})(1+\lambda(-\rho)^{d_v+1})}{\theta(1+\lambda(-\rho)^{d_u+1})(1+\lambda(-\rho)^{d_v+1}) + (1+\lambda(-\rho)^{d_u})(1+\lambda(-\rho)^{d_v})} \\
&= \frac{1}{1 + \frac{(1+\lambda(-\rho)^{d_u})(1+\lambda(-\rho)^{d_v})}{\theta(1+\lambda(-\rho)^{d_u+1})(1+\lambda(-\rho)^{d_v+1})}} \geqslant \frac{1}{1 + \frac{(1+\lambda)^2}{\theta(1-\lambda\rho)^2}} = \Omega_{\theta,\lambda,\rho}(1).
\end{aligned}$$

The last inequality holds because $\lambda\rho < 1$ by Lemma 11 and $\rho < 1$ so that $(1+\lambda(-\rho)^{d_u})(1+\lambda(-\rho)^{d_v}) \leqslant (1+\lambda)^2$ and $\theta(1+\lambda(-\rho)^{d_u+1})(1+\lambda(-\rho)^{d_v+1}) \geqslant \theta(1-\lambda\rho)^2$. Similarly,

$$\begin{aligned}
\mu_e^\tau(0) &= \frac{(1+\lambda(-\rho)^{d_u})(1+\lambda(-\rho)^{d_v})}{\theta(1+\lambda(-\rho)^{d_u+1})(1+\lambda(-\rho)^{d_v+1}) + (1+\lambda(-\rho)^{d_u})(1+\lambda(-\rho)^{d_v})} \\
&= \frac{1}{1 + \frac{\theta(1+\lambda(-\rho)^{d_u+1})(1+\lambda(-\rho)^{d_v+1})}{(1+\lambda(-\rho)^{d_u})(1+\lambda(-\rho)^{d_v})}} \geqslant \frac{1}{1 + \frac{\theta(1+\lambda)^2}{(1-\lambda\rho)^2}} = \Omega_{\theta,\lambda,\rho}(1).
\end{aligned}$$

By taking the minimum of the two values, we obtain the marginal lower bound $b = \Omega_{\theta,\lambda,\rho}(1)$. □

6.3. Other missing proofs in Section 4.

Proof of Claim 32. First we prove $c^\star > 0$. To see this, by Equation (30), we write $c^\star = \gamma \log(\sqrt{\gamma/\beta})/\Phi$ with $\Phi = \sqrt{\gamma\beta - 1} \arctan \sqrt{\gamma\beta - 1} - \log(\sqrt{\gamma/\beta})$. We prove $\Phi > 0$. We let

$$s = \sqrt{\gamma\beta - 1} \quad \Longrightarrow \quad 1 + s^2 = \gamma\beta$$

We consider

$$h(t) = t \arctan t - \frac{1}{2}\log(1+t^2) \quad \text{for } t > 0.$$

$$h'(t) = \arctan t + \frac{t}{1+t^2} - \frac{t}{1+t^2} = \arctan t > 0 \quad \text{for } t > 0.$$

We note that $h(0) = 0$, so $h(s) > 0$ for $s > 0$. Therefore, $s \arctan s > \frac{1}{2}\log(1+s^2) = \frac{1}{2}\log(\gamma\beta)$ for $s > 0$. By $\beta > 1$, it holds that $\gamma\beta > \frac{\gamma}{\beta}$. Therefore,

$$\frac{1}{2}\log(\gamma\beta) > \frac{1}{2}\log\left(\frac{\gamma}{\beta}\right) = \log\left(\sqrt{\frac{\gamma}{\beta}}\right).$$

Therefore, $\Phi = \sqrt{\gamma\beta - 1}\arctan\sqrt{\gamma\beta-1} - \log(\sqrt{\gamma/\beta}) > 0$, which implies $c^\star > 0$.

Next we prove $r^\star - c^\star > 1$. By Equation (30),

$$(r^\star)^2 = (c^\star)^2 + \frac{2c^\star}{\beta} + \frac{\gamma}{\beta}.$$

Since $c^\star > 0$, proving $r^\star - c^\star > 1$ is equivalent to prove $(r^\star)^2 > (c^\star+1)^2$, which is equivalent to show

$$\frac{2c^\star}{\beta} + \frac{\gamma}{\beta} > 2c^\star + 1.$$

Rearranging gives

$$c^\star < \frac{\gamma - \beta}{2(\beta - 1)}.$$

For notational simplicity, we substitute $c^\star = \gamma L/(s\alpha - L)$ with parameters $L = \log\sqrt{\frac{\gamma}{\beta}}$, $s = \sqrt{\gamma\beta - 1}$, $\alpha = \arctan s$, we reduce this to show

$$(\gamma - \beta)s\alpha > (2\beta\gamma - \beta - \gamma)L. \tag{40}$$

To verify (40), we write $q = \gamma/\beta > 1$ and define

$$H_\beta(q) = (q-1)s\arctan s - \frac{(2\beta-1)q - 1}{2}\log q, \qquad s = \sqrt{\beta^2 q - 1}.$$

Then (40) is equivalent to $H_\beta(q) > 0$. At $\beta = 1$,

$$H_1(q) = (q-1)\left(s\arctan s - \frac{1}{2}\log(1+s^2)\right), \qquad s = \sqrt{q-1},$$

By former computation, we already prove $H_1(q) = h(q) > 0$. For $\beta > 1$, we prove $H_\beta(q)$ increases as β increases. We prove $\partial H_\beta(q)/\partial\beta > 0$. By

$$s = \sqrt{\beta^2 q - 1} \quad \Longrightarrow \quad \frac{\partial s}{\partial \beta} = \frac{q\beta}{s}.$$

Then it holds that

$$\frac{\partial H_\beta(q)}{\partial \beta} = (q-1)\left(\frac{q\beta}{s}\arctan s + \frac{1}{\beta}\right) - q\log q.$$

Using $\arctan s > \frac{s}{\sqrt{1+s^2}}$, we have

$$\frac{q\beta}{s}\arctan s > \frac{q\beta}{\sqrt{1+s^2}} = \frac{q\beta}{\sqrt{q}\beta} = \sqrt{q}.$$

Then it holds that

$$\begin{aligned}\frac{\partial H_\beta(q)}{\partial \beta} &> (q-1)\sqrt{q} - q\log q \\ &= (t^2-1)t - 2t^2\log t, \quad \text{where } t = \sqrt{q} \\ &= t^2\left(t - \frac{1}{t} - 2\log t\right).\end{aligned}$$

Let $f(t) = t^2(t - \frac{1}{t} - 2\log t)$. By direct computation, we have $f(1) = 0$ and $f'(t) > 0$ for $t > 1$. Therefore, $f(t) > 0$ for $t > 1$. Therefore, $\partial H_\beta(q)/\partial\beta > 0$. Therefore, $H_\beta(q) > H_1(q) > 0$. So we have $c^\star < \frac{\gamma-\beta}{2(\beta-1)}$, which is equavalent to $r^\star - c^\star > 1$. ☐

Proof of Claim 33. We prove that $1/\rho \notin \mathrm{cl}(D(c^\star, r^\star))$. Since $1/\rho$ and $c^\star$ are real, it is enough to show

$$\left(\frac{1}{\rho} - c^\star\right)^2 > (r^\star)^2.$$

Recall that

$$\rho = \frac{\beta-1}{\gamma-1}, \qquad \frac{1}{\rho} = \frac{\gamma-1}{\beta-1},$$

and

$$(r^\star)^2 = \left(c^\star + \frac{1}{\beta}\right)^2 + \frac{\beta\gamma-1}{\beta^2}.$$

Hence

$$\left(\frac{1}{\rho} - c^\star\right)^2 - (r^\star)^2 = \left(\frac{1}{\rho}\right)^2 - \frac{\gamma}{\beta} - 2c^\star\left(\frac{1}{\rho} + \frac{1}{\beta}\right).$$

A direct simplification gives

$$\left(\frac{1}{\rho}\right)^2 - \frac{\gamma}{\beta} = \frac{(\gamma-\beta)(\beta\gamma-1)}{\beta(\beta-1)^2},$$

and

$$\frac{1}{\rho} + \frac{1}{\beta} = \frac{\beta\gamma-1}{\beta(\beta-1)}.$$

Therefore

$$\left(\frac{1}{\rho} - c^\star\right)^2 - (r^\star)^2 = \frac{\beta\gamma-1}{\beta(\beta-1)}\left(\frac{\gamma-\beta}{\beta-1} - 2c^\star\right).$$

Thus it suffices to prove

$$2c^\star < \frac{\gamma-\beta}{\beta-1}.$$

which is already justified in the proof of Claim 32. ☐

Proof of Lemma 39. The two roots of $\beta x^2 + 2x + \gamma = 0$ are $\zeta_\pm = \frac{-1\pm i\sqrt{\beta\gamma-1}}{\beta}$. By the definition of $r^\star$, we have

$$(r^\star)^2 = \left(c^\star + \frac{1}{\beta}\right)^2 + \frac{\beta\gamma-1}{\beta^2} = |\zeta_\pm - c^\star|^2.$$

Therefore both roots lie on the boundary of the disk $D(c^\star, r^\star)$. Since $\epsilon > 0$, we have

$$|\zeta_\pm - c^\star| = r^\star > r^\star - \epsilon.$$

Thus neither root belongs to the closed disk $\mathrm{cl}\left(D(c^\star, r^\star - \epsilon)\right)$. Equivalently, $\mathrm{cl}\left(D(c^\star, r^\star - \epsilon)\right) \cap \{\zeta_1, \zeta_2\} = \emptyset$. ☐

Proof of Claim 40. Recall that

$$\rho = \frac{\beta-1}{\gamma-1}, \qquad \theta = \frac{(\gamma-1)^2}{\beta\gamma-1}.$$

Therefore

$$\rho\theta = \frac{\beta-1}{\gamma-1} \cdot \frac{(\gamma-1)^2}{\beta\gamma-1} = \frac{(\beta-1)(\gamma-1)}{\beta\gamma-1}.$$

Since $\beta > 1$ and $\gamma > 1$, we have

$$\beta\gamma - 1 - (\beta-1)(\gamma-1) = \beta + \gamma - 2 > 0.$$

Hence

$$(\beta - 1)(\gamma - 1) < \beta\gamma - 1,$$

and consequently $\rho\theta < 1$. $\square$

References

[AFF+25] Konrad Anand, Weiming Feng, Graham Freifeld, Heng Guo, and Jiaheng Wang. Approximate counting for spin systems in sub-quadratic time. *TheoretiCS*, 4, 2025. 4

[ALO24] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *SIAM J. Comput.*, 53(6):S20–1, 2024. 6

[Asa70] Taro Asano. Theorems on the partition functions of the Heisenberg ferromagnets. *J. Phys. Soc. Japan*, 29:350–359, 1970. 8, 18, 27

[Bar16] Alexander I. Barvinok. *Combinatorics and Complexity of Partition Functions*, volume 30 of *Algorithms and combinatorics*. Springer, 2016. 1, 2, 3

[BB09] Julius Borcea and Petter Brändén. Pólya–Schur master theorems for circular domains and their boundaries. *Ann. of Math.*, 170(1):465–492, 2009. 27

[CCLZ26] Xiaoyu Chen, Zongchen Chen, Kuikui Liu, and Xinyuan Zhang. Subquadratic counting via perfect marginal sampling. *arXiv*, abs/2604.02235, 2026. 4

[CCYZ25] Xiaoyu Chen, Zongchen Chen, Yitong Yin, and Xinyuan Zhang. Rapid mixing at the uniqueness threshold. In *STOC*, pages 879–890. ACM, 2025. 1

[CG24] Zongchen Chen and Yuzhou Gu. Fast sampling of b-matchings and b-edge covers. In *SODA*, pages 4972–4987. SIAM, 2024. 5

[CLV21] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of Glauber dynamics: entropy factorization via high-dimensional expansion. In *STOC*, pages 1537–1550. ACM, 2021. 6

[CLV23] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Rapid mixing of Glauber dynamics up to uniqueness via contraction. *SIAM J. Comput.*, 52(1):196–237, 2023. 1

[CLV24] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Spectral independence via stability and applications to Holant-type problems. *TheoretiCS*, 3, 2024. 6, 7

[CZ23] Xiaoyu Chen and Xinyuan Zhang. A near-linear time sampler for the Ising model with external field. In *SODA*, pages 4478–4503. SIAM, 2023. 3, 4, 5

[DGGJ04] Martin E. Dyer, Leslie Ann Goldberg, Catherine S. Greenhill, and Mark Jerrum. The relative complexity of approximate counting problems. *Algorithmica*, 38(3):471–500, 2004. 1, 2

[ES88] Robert G. Edwards and Alan D. Sokal. Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm. *Phys. Rev. D (3)*, 38(6):2009–2012, 1988. 11

[FGW23] Weiming Feng, Heng Guo, and Jiaheng Wang. Swendsen-Wang dynamics for the ferromagnetic Ising model with external fields. *Inf. Comput.*, 294:105066, 2023. 4, 11, 12, 13

[FGY26] Weiming Feng, Heng Guo, and Minji Yang. Rapid mixing in positively weighted restricted Boltzmann machines. In *FOCS*, 2026. to appear, available at `abs/2604.00963`. 2, 3, 4

[FK72] Cees M. Fortuin and Piet W. Kasteleyn. On the random-cluster model. I. Introduction and relation to other models. *Physica*, 57:536–564, 1972. 3, 11

[GJ09] Geoffrey R. Grimmett and Svante Janson. Random even graphs. *Electron. J. Comb.*, 16(1), 2009. 12

[GJ18] Heng Guo and Mark Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. *Ann. Appl. Probab.*, 28(2):1292–1313, 2018. 3

[GJP03] Leslie Ann Goldberg, Mark Jerrum, and Mike Paterson. The computational complexity of two-state spin systems. *Random Struct. Algorithms*, 23(2):133–154, 2003. 1, 2

[GL18] Heng Guo and Pinyan Lu. Uniqueness, spatial mixing, and approximation for ferromagnetic 2-spin systems. *ACM Trans. Comput. Theory*, 10(4):17:1–17:25, 2018. 1, 2, 9, 21, 22, 23, 24

[GLL20] Heng Guo, Jingcheng Liu, and Pinyan Lu. Zeros of ferromagnetic 2-spin systems. In *SODA*, pages 181–192. SIAM, 2020. 1, 2, 3, 9, 13, 20, 26

[GLLZ21] Heng Guo, Chao Liao, Pinyan Lu, and Chihao Zhang. Zeros of Holant problems: Locations and algorithms. *ACM Trans. Algorithms*, 17(1):4:1–4:25, 2021. 8

[GŠV16] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. *Combin. Probab. Comput.*, 25(4):500–559, 2016. 1

[JS93] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.*, 22(5):1087–1116, 1993. 1, 2, 3, 4

[Kol18] Vladimir Kolmogorov. A faster approximation algorithm for the Gibbs partition function. In *COLT*, volume 75 of *Proceedings of Machine Learning Research*, pages 228–249. PMLR, 2018. 4

[LLY13] Liang Li, Pinyan Lu, and Yitong Yin. Correlation decay up to uniqueness in spin systems. In *SODA*, pages 67–84. SIAM, 2013. 1

[LLZ14] Jingcheng Liu, Pinyan Lu, and Chihao Zhang. The complexity of ferromagnetic two-spin systems with external fields. In *RANDOM*, LIPIcs, pages 843–856. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2014. 1, 2

[PR17] Viresh Patel and Guus Regts. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM J. Comput.*, 46(6):1893–1919, 2017. 1, 3

[Rue71] David Ruelle. Extension of the Lee-Yang circle theorem. *Phys. Rev. Lett.*, 26:303–304, 1971. 8, 21, 27

[Sly10] Allan Sly. Computational transition at the uniqueness threshold. In *FOCS*, pages 287–296. IEEE Computer Society, 2010. 1

[SS14] Allan Sly and Nike Sun. Counting in two-spin models on d-regular graphs. *Ann. Probab.*, 42(6):2383–2416, 2014. 1

[SST14] Alistair Sinclair, Piyush Srivastava, and Marc Thurley. Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. *J. Stat. Phys.*, 155(4):666–686, 2014. 1

[SVV09] Daniel Stefankovic, Santosh S. Vempala, and Eric Vigoda. Adaptive simulated annealing: A near-optimal connection between sampling and counting. *J. ACM*, 56(3):18:1–18:36, 2009. 4, 29

[Val08] Leslie G. Valiant. Holographic algorithms. *SIAM J. Comput.*, 37(5):1565–1594, 2008. 4, 9

[vdW41] Bartel L. van der Waerden. Die lange Reichweite der regelmäßigen Atomanordnung in Mischkristallen. *Zeitschrift für Physik*, 118(7):473–488, 1941. 3

[Wei06] Dror Weitz. Counting independent sets up to the tree threshold. In *STOC*, pages 140–149. ACM, 2006. 1, 2, 3, 9