OPTIMAL MIXING TIME FOR THE ISING MODEL IN THE UNIQUENESS REGIME

XIAOYU CHEN, WEIMING FENG, YITONG YIN, AND XINYUAN ZHANG

Abstract. We prove an optimal $O(n \log n)$ mixing time of the Glauber dynamics for the Ising models with edge activity $\beta \in \left(\frac{\Delta-2}{\Delta}, \frac{\Delta}{\Delta-2}\right)$. This mixing time bound holds even if the maximum degree Δ is unbounded.

We refine the boosting technique developed in [CFYZ21], and prove a new boosting theorem by utilizing the entropic independence defined in [AJK+21]. The theorem relates the modified log-Sobolev (MLS) constant of the Glauber dynamics for a near-critical Ising model to that for an Ising model in a sub-critical regime.

1. Introduction

The Ising model [Isi25] has been extensively studied in Statistic Physics, Probability and Computer Science. Let $G = (V, E)$ be an undirected graph. Let $\beta \in \mathbb{R}_{>0}$ be the *edge activity* and $\boldsymbol{\lambda} = (\lambda_v)_{v \in V} \in \mathbb{R}_{>0}^V$ the *local fields*. A *configuration* $\sigma \in \{-1, +1\}^V$ assigns each vertex $v \in V$ one of the two *spins* in $\{-1, +1\}$. The *Gibbs distribution* μ over $\{-1, +1\}^V$ is defined by:

$$\forall \sigma \in \{-1, +1\}^V, \quad \mu(\sigma) \triangleq \frac{1}{Z} \beta^{m(\sigma)} \prod_{v \in V : \sigma_v = +1} \lambda_v,$$

where $m(\sigma) \triangleq |\{\{u, v\} \in E \mid \sigma_u = \sigma_v\}|$ is the number of monochromatic edges in σ, and the *partition function* $Z = Z_G(\beta)$ is defined by

$$Z \triangleq \sum_{\sigma \in \{-1,+1\}^V} \beta^{m(\sigma)} \prod_{v \in V : \sigma_v = +1} \lambda_v.$$

The Ising model is said to be *ferromagnetic* if $\beta > 1$, and *anti-ferromagnetic* if $\beta < 1$.

Sampling from Gibbs distributions is a fundamental computational task. The *Markov chain Monte Carlo* (MCMC) method is the most extensively studied technique for this task. A canonical Markov chain is the *Glauber dynamics* (a.k.a. *Gibbs sampler*, *heat-bath*). For any distribution μ over $\{-1, +1\}^V$, the chain starts from a *feasible* configuration $X_0 \in \Omega(\mu) \subseteq \{-1, +1\}^V$. In the t-th step, X_{t-1} is updated to X_t as:

- pick a vertex $v \in V$ uniformly at random and let $X_t(u) = X_{t-1}(u)$ for all $u \neq v$.
- sample $X_t(v)$ from the marginal distribution $\mu_v^{X_{t-1}(V \setminus \{v\})}$.

Here, $\mu_v^{X_{t-1}(V \setminus \{v\})}$ denotes the marginal distribution on v projected from μ conditional on the values of all vertices except v being fixed as $X_{t-1}(V \setminus \{v\})$. It is well known the chain is reversible with respect to μ. Its convergence rate is captured by the *mixing time*:

$$\forall 0 < \varepsilon < 1, \quad T_{\mathrm{mix}}(\varepsilon) \triangleq \max_{X_0 \in \Omega(\mu)} \min\{t \mid d_{\mathrm{TV}}(X_t, \mu) \leq \varepsilon\},$$

where $d_{\mathrm{TV}}(X_t, \mu)$ is the *total variational distance* between the distribution of X_t and μ.

The rapid mixing of the Glauber dynamics is intrinsically connected to the *spatial mixing* property of the Gibbs distribution μ. For the Ising model, its spatial mixing is captured by the *uniqueness condition.*

(Xiaoyu Chen, Yitong Yin, Xinyuan Zhang) State Key Laboratory for Novel Software Technology, Nanjing University, 163 Xianlin Avenue, Nanjing, Jiangsu Province, China. E-mails: chenxiaoyu233@smail.nju.edu.cn, yinyt@nju.edu.cn, zhangxy@smail.nju.edu.cn

(Weiming Feng) School of Informatics, University of Edinburgh, Informatics Forum, Edinburgh, EH8 9AB, United Kingdom. E-mail: wfeng@ed.ac.uk

Let $\Delta = \Delta_G$ denote the maximum degree of graph G. There exist two critical thresholds $\beta_c(\Delta)$ and $\bar{\beta}_c(\Delta)$ for the edge activity β, called the *uniqueness thresholds*, defined by

$$\beta_c(\Delta) = \frac{\Delta - 2}{\Delta} \quad \text{and} \quad \bar{\beta}_c(\Delta) = \frac{\Delta}{\Delta - 2}.$$

If $\beta_c(\Delta) < \beta < \bar{\beta}_c(\Delta)$, then for arbitrary local fields $\boldsymbol{\lambda} \in \mathbb{R}_{>0}^V$, the influence of a boundary condition on the marginal distribution at a vertex v in μ decays exponentially in the distance between the vertex and the boundary; and if $\beta < \beta_c(\Delta)$ or $\beta > \bar{\beta}_c(\Delta)$, then there exist such local fields that the boundary-to-vertex correlation persists as the distance between them grows to ∞ [LLY13, SST14, GL18].

The mixing behavior of Glauber dynamics for Ising model undergoes sharp transitions at the uniqueness thresholds. On one hand, if β lies in the non-uniqueness regime, the Glauber dynamics is torpid (slow) mixing. For anti-ferromagnetic Ising model, in the non-uniqueness regime $\beta < \beta_c(\Delta)$, unless $\mathbf{NP} = \mathbf{RP}$, sampling from Ising models is not polynomial-time tractable [SS12, GŠV16]. For ferromagnetic Ising model, in the non-uniqueness regime $\beta > \bar{\beta}_c(\Delta)$, the mixing time of Glauber dynamics is exponential in the size of the graph [GM07], although the sampling problem can be solved otherwise by global Markov chains [JS93, GJ18] or polynomial interpolation [LSS17].

On the other hand, if β lies in the uniqueness regime, the Glauber dynamics mixes in polynomial-time. In a seminal work of Mossel and Sly [MS13], by leveraging the monotonicity of coupling, an $\exp\left(\Delta^{O(1/\delta)}\right) n \log n$ mixing time was proved for the ferromagnetic Ising models with n vertices and $\beta \in (1, \frac{\Delta-\delta}{\Delta-2+\delta}]$, where $\delta \in (0, 1)$ controls the gap to the uniqueness threshold. Recently, based on a powerful "local-to-global" theorem of Alev and Lau [ALO20, AL20] for high-dimensional expander walks, an important concept called *spectral independence* was introduced by Anari, Liu and Oveis Gharan in their seminal work [ALO20], which leads to a series of important progress in studies of mixing times [CLV20, FGYZ21, CGŠV21, Liu21, BCC+21, JPV21, CFYZ21, CLV21b, ALG21, AJK+21]. For the Ising models in the uniqueness regime with $\beta \in [\frac{\Delta-2+\delta}{\Delta-\delta}, \frac{\Delta-\delta}{\Delta-2+\delta}]$, a polynomially-bounded mixing time $n^{O(1/\delta)}$ was first proved by Chen, Liu and Vigoda [CLV20] using the spectral independence; and this mixing time bound was improved to $\Delta^{O(1/\delta)} n \log n$ in their subsequent work [CLV21a] by a uniform block factorization of entropy. On general graphs with unbounded maximum degrees, both these mixing bounds become $n^{O(1/\delta)}$ in the worst case. Very recently, in a work [CFYZ21] by the authors of the current paper, by establishing a boosting theorem for spectral gaps, an $\mathrm{e}^{O(1/\delta)} n^2$ mixing time bound was proved for the Ising models on general n-vertex graphs (with bounded or unbounded maximum degree Δ) in the uniqueness regime with $\beta \in [\frac{\Delta-2+\delta}{\Delta-\delta}, \frac{\Delta-\delta}{\Delta-2+\delta}]$.

An important open problem is then to close the gap to the lower bound of the mixing time [HS07, DP11] and prove an $O(n \log n)$ optimal mixing time for the Ising models in the uniqueness regime. In this paper, we resolve the problem by proving the following theorem.

Theorem 1.1. *For all $\delta \in (0, 1)$, there exists $C_\delta = \exp(O(1/\delta))$ such that for every Ising model on n-vertex graph $G = (V, E)$ with maximum degree $\Delta = \Delta_G \geq 3$ and with edge activity $\beta \in \left[\frac{\Delta-2+\delta}{\Delta-\delta}, \frac{\Delta-\delta}{\Delta-2+\delta}\right]$ and local fields $\boldsymbol{\lambda} \in \mathbb{R}_{>0}^V$, the mixing time of the Glauber dynamics is bounded as*

$$T_{mix}(\varepsilon) \leq C_\delta \frac{\lambda_{\max}}{\lambda_{\min}} n \left(\log \frac{n}{\varepsilon} + \log\log \frac{2\lambda_{\max}}{\lambda_{\min}} \right),$$

where $\lambda_{\max} = \max_{v \in V} \lambda_v$ and $\lambda_{\min} = \min_{v \in V} \lambda_v$.

The theorem gives an optimal $O(n \log n)$ mixing time for both ferromagnetic and anti-ferromagnetic Ising models on general graphs (with bounded or unbounded maximum degrees) in the interior of the uniqueness regime, for constant local fields or the local fields that are in the same order of magnitude.

We further remark that by combining techniques in this paper with a Markov chain called the field dynamics defined in [CFYZ21], we can obtain an $O_\delta(n \log^2 n)$-time sampling algorithm for the Ising models with $\beta \in \left[\frac{\Delta-2+\delta}{\Delta-\delta}, \frac{\Delta-\delta}{\Delta-2+\delta}\right]$, where the constant in $O_\delta(\cdot)$ depends only on δ but not on $\boldsymbol{\lambda}$. This sampling algorithm is postponed to the full version of the paper.

1.1. **Results for general distributions**. We prove a boosting theorem for the *modified log-Sobolev (MLS)* constant for Glauber dynamics on general distributions. Bounds on MLS constants can often give optimal mixing time for Glauber dynamics. Let V be a finite ground set. Let μ be a distribution over $\{-1,+1\}^V$. We use $\Omega(\mu)$ to denote the support of distribution μ. Let $P = P_\mu^{\mathrm{GD}} : \Omega(\mu)\times\Omega(\mu) \to \mathbb{R}_{\ge 0}$ denote the transition matrix of the Glauber dynamics on μ. For any function $f : \Omega(\mu) \to \mathbb{R}_{\ge 0}$, its *Dirichlet form* is defined by:

$$\mathcal{E}_P(f, \log f) \triangleq \langle f, (I-P)\log f\rangle_\mu,$$

where the inner product $\langle f, g\rangle_\mu \triangleq \sum_{\sigma\in\Omega(\mu)} f(\sigma)g(\sigma)\mu(\sigma)$. Define the entropy:

$$\mathbf{Ent}_\mu[f] \triangleq \mathbf{E}_\mu[f\log f] - \mathbf{E}_\mu[f]\log \mathbf{E}_\mu[f],$$

where $\mathbf{E}_\mu[f] \triangleq \sum_{\sigma\in\Omega(\mu)} \mu(\sigma)f(\sigma)$. We use the convention $0\log 0 = 0$ in above definitions.

Our goal is to bound the following *modified log-Sobolev constant* [BT06] for the Glauber dynamics on μ:

$$\rho^{\mathrm{GD}}(\mu) \triangleq \inf\left\{ \frac{\mathcal{E}_P(f,\log f)}{\mathbf{Ent}_\mu[f]} \;\middle|\; f : \Omega(\mu)\to\mathbb{R}_{\ge 0},\ \mathbf{Ent}_\mu[f]\neq 0 \right\}, \tag{1}$$

so that the mixing time of Glauber dynamics can be bounded by,

$$t_{\mathrm{mix}}(\varepsilon) \le \frac{1}{\rho^{\mathrm{GD}}(\mu)}\left(\log\log\frac{1}{\mu_{\min}} + \log\frac{1}{2\varepsilon^2}\right),$$

where $\mu_{\min} \triangleq \min_{\sigma\in\Omega(\mu)}\mu(\sigma)$ is the minimum probability in μ.

For any distribution μ over $\{-1,+1\}^V$, any subset $S\subseteq V$, let μ_S denote the marginal distribution on S induced by μ. We simply use μ_v to denote $\mu_{\{v\}}$ if $S=\{v\}$. For any subset $\Lambda\subseteq V$, any $\sigma\in\Omega(\mu_\Lambda)$, we use $\mu^{\Lambda\leftarrow\sigma}$ (or simply μ^σ if Λ is clear from the context) to denote the distribution over $\{-1,+1\}^V$ induced by μ conditional on the configuration σ on Λ.

Definition 1.2 (signed influence matrix [ALO20]). Let μ be a distribution over $\{-1,+1\}^V$. The *signed influence matrix* $\Psi_\mu^{\mathrm{inf}} : V\times V \to \mathbb{R}$ is defined by

$$\forall u,v\in V,\quad \Psi_\mu^{\mathrm{inf}}(u,v) = \begin{cases} \mu_v^{u\leftarrow+1}(+1) - \mu_v^{u\leftarrow-1}(+1) & \text{if } u\neq v \text{ and } \Omega(\mu_u)=\{-1,+1\};\\ 0 & \text{otherwise.}\end{cases}$$

In [ALO20], a notion of *spectral independence* was defined using the signed influence matrix. Specifically, a distribution μ over $\{-1,+1\}^V$ is spectrally independent if its influence matrix Ψ_μ^{inf} has bounded spectral radius. In this paper, we use the following sufficient condition for the spectral independence (SI).

Definition 1.3 (spectral independence in ∞-norm). Let $\eta>0$. A distribution μ over $\{-1,+1\}^V$ is said to be *η-spectrally independent in ∞-norm* if

$$\left\|\Psi_\mu^{\mathrm{inf}}\right\|_\infty = \max_{u\in V}\sum_{v\in V}\left|\Psi_\mu^{\mathrm{inf}}(u,v)\right| \le \eta.$$

Next, we introduce a notation for the distributions with local fields.

Definition 1.4 (magnetizing a joint distribution with local fields). Let μ be a distribution over $\{-1,+1\}^V$. Let $\boldsymbol{\lambda} = (\lambda_v)_{v\in V} \in \mathbb{R}_{>0}$ be positive local fields. Let $\mu^{(\boldsymbol{\lambda})}$ denote the distribution obtained from imposing the local fields $\boldsymbol{\lambda}$ onto μ. Formally, for any configuration $\sigma\in\{-1,+1\}^V$,

$$\mu^{(\boldsymbol{\lambda})}(\sigma) \propto \mu(\sigma)\prod_{v\in V:\sigma_v=+1}\lambda_v.$$

We simply denote $\mu^{(\boldsymbol{\lambda})}$ by $\mu^{(\lambda)}$ if $\lambda_v=\lambda$ for all $v\in V$.

Definition 1.5 (spectral independence in ∞-norm with all fields). A distribution μ over $\{-1,+1\}^V$ is said to be η-spectrally independent in ∞-norm *with all fields* if $\mu^{(\boldsymbol{\lambda})}$ is η-spectrally independent in ∞-norm for all $\boldsymbol{\lambda} = (\lambda_v)_{v\in V} \in \mathbb{R}_{>0}^V$.

We define the MLS constant under the worst pinning. Let π be a distribution over $\{-1,+1\}^V$. For any $\Lambda \subseteq V$, any $\sigma \in \Omega(\pi_\Lambda)$, let $\rho^{\mathrm{GD}}(\pi^\sigma)$ denote the MLS constant for the Glauber dynamics on π^σ. Let

$$\rho^{\mathrm{GD}}_{\min}(\pi) \triangleq \min_{\Lambda \subset V} \min_{\sigma \in \Omega(\pi_\Lambda)} \rho^{\mathrm{GD}}(\pi^\sigma). \tag{2}$$

We now state our boosting theorem for the MLS constant for general distributions over $\{-1,+1\}^V$.

Theorem 1.6 (boosting theorem for modified log-Sobolev). *Let μ be a distribution over $\{-1,+1\}^V$ and $\eta > 0$. If μ is η-spectrally independent in ∞-norm with all fields, then the following holds for the modified log-Sobolev constants for Glauber dynamics:*

$$\forall \theta \in (0,1), \quad \rho^{\mathrm{GD}}(\mu) \geq \left(\frac{\theta}{\mathrm{e}}\right)^{\eta+3} \rho^{\mathrm{GD}}_{\min}(\mu^{(\theta)}).$$

Theorem 1.6 is a boosting theorem for the MLS constant. When applied on a distribution μ in a near-critical regime, by carefully choosing a parameter θ so that $\mu^{(\theta)}$ enters an easier sub-critical regime, the MLS constant bound for μ is reduced to easier or known MLS constant bounds, by losing a $\theta^{O(\eta)}$ factor.

A similar boosting theorem for the Poincaré constant (spectral gap) was established in [CFYZ21], which requires $\mu^{(\boldsymbol{\lambda})}$ to be spectrally independent for all $\boldsymbol{\lambda} \in (0,1]^V$. Theorem 1.6 asks for a stronger condition, but holds for the MLS constant that can imply optimal mixing time bounds.

1.2. **Open problems**. The key ingredient in this paper is the boosting theorem for MLS constant, which connects the MLS constants of Glauber dynamics on distributions μ and $\mu^{(\theta)}$. However, the distribution μ should be spectrally independent in ∞-norm with *arbitrary* local fields, which restricts its applicability. An open problem is to prove a boosting theorem with weaker assumption, especially the spectral independence with one-sided local fields, so that it can be used to prove optimal mixing times for other 2-spin systems, e.g. the hard-core model.

Another important question is to extend our framework to distributions beyond the boolean domain so that it can be used to prove optimal mixing times for general spin systems, e.g. proper q-colorings of general graphs with bounded or unbounded degrees.

Finally, for the Ising model, our bound for mixing time of Glauber dynamics depends on the ratio $\frac{\lambda_{\max}}{\lambda_{\min}}$. This ratio comes from the analysis of the MLS constant in the easier regime $\mu^{(\theta)}$ (see Section 8 for details). We conjecture that such dependency on the local fields in the mixing time can be removed. This requires a deeper understanding of the MLS constant to the Ising models with heterogenous (and non-constant) fields, which has not been very well studied even in an easy regime.

2. Proof outline

We outline the proof of the boosting theorem (Theorem 1.6). The result for Ising model (Theorem 1.1) is an application of the boosting theorem. The proof of Theorem 1.1 is given in Section 8.

To prove Theorem 1.6, the key step is to establish the "magnetized block factorization of entropy". Let $\theta \in (0,1)$ and $\pi = \mu^{(\theta)}$. Formally, for every $\sigma \in \{-1,+1\}^V$,

$$\pi(\sigma) = \frac{\mu(\sigma)\theta^{\|\sigma\|_+}}{Z_\pi}, \quad \text{where } Z_\pi \triangleq \sum_{\sigma \in \{-1,+1\}^V} \mu(\sigma)\theta^{\|\sigma\|_+} \text{ and } \|\sigma\|_+ = |\{v \in V \mid \sigma_v = +1\}|. \tag{3}$$

We use $\mathrm{Bin}(V, 1-\theta)$ to denote the distribution of random subset $R \subseteq V$ generated by including each $v \in V$ into R independently with probability $1-\theta$. Specifically, for every $\Lambda \subseteq V$,

$$\mathop{\mathbf{Pr}}_{R \sim \mathrm{Bin}(V, 1-\theta)}[R = \Lambda] = (1-\theta)^{|\Lambda|}\theta^{|V|-|\Lambda|}.$$

Definition 2.1 (magnetized block factorization of entropy). Let $\theta \in (0,1)$ and $C > 0$. A distribution μ over $\{-1,+1\}^V$ is said to satisfy *θ-magnetized block factorization of entropy with parameter C* if the following holds for $\pi = \mu^{(\theta)}$ and for all $f : \Omega(\mu) \to \mathbb{R}_{\geq 0}$:

$$\mathbf{Ent}_\mu[f] \leq C \cdot \frac{Z_\pi}{\theta^{|V|}} \mathop{\mathbf{E}}_{R \sim \mathrm{Bin}(V, 1-\theta)} \left[\pi_R(\mathbb{1}_R) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}[f]\right], \tag{4}$$

assuming $\pi_R(\mathbb{1}_R) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}[f] = 0$ if $\pi_R(\mathbb{1}_R) = 0$, where $\mathbb{1}_R$ denotes the all-(+1) configuration on $R \subseteq V$.

In the left-hand-side (LHS) of (4), the entropy is evaluated with respect to the original distribution μ; and in the right-hand-side (RHS), the entropy is evaluated with respect to the conditional distribution $\pi^{\mathbb{1}_R}$, where $\pi = \mu^{(\theta)}$ is obtained from μ by changing the local fields by a factor θ. This definition plays a key role to relate μ to $\mu^{(\theta)}$. Theorem 1.6 can be proved by combining the following two lemmas.

Lemma 2.2. *Let $\theta \in (0,1)$ and $C > 0$. For any distribution μ over $\{-1,+1\}^V$ satisfying θ-magnetized block factorization of entropy with parameter C, the Glauber dynamics on μ has the modified log-Sobolev constant*

$$\rho^{\mathrm{GD}}(\mu) \geq \frac{\rho^{\mathrm{GD}}_{\min}(\mu^{(\theta)})}{C}.$$

Lemma 2.3. *Let μ be a distribution over $\{-1,+1\}^V$ and $\eta > 0$. If μ is η-spectrally independent in ∞-norm with all fields, then μ satisfies θ-magnetized block factorization of entropy with parameter*

$$C = \left(\frac{\mathrm{e}}{\theta}\right)^{\eta+3}.$$

The proof of Lemma 2.2 is outlined in Section 2.1, and the proof of Lemma 2.3 is outline in Section 2.2.

2.1. **The MLS constant bounds.** To obtain the MLS constant for the Glauber dynamics $P = P^{\mathrm{GD}}_\mu$ on μ, we need to relate the RHS of (4) with the Dirichlet form $\mathcal{E}_P(f, \log f)$. Formally, we prove the following inequality for any function $f : \{-1,+1\}^V \to \mathbb{R}_{\geq 0}$:

$$\frac{Z_\pi}{\theta^n} \mathop{\mathbf{E}}_{R \sim \mathrm{Bin}(V, 1-\theta)} \left[\pi_R(\mathbb{1}_R) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}[f]\right] \leq \frac{\mathcal{E}_P(f, \log f)}{\rho^{\mathrm{GD}}_{\min}(\pi)}. \tag{5}$$

Combining (4) and (5) implies $\mathbf{Ent}_\mu[f] \leq \frac{C}{\rho^{\mathrm{GD}}_{\min}(\pi)} \mathcal{E}_P(f, \log f)$. This proves Lemma 2.2.

Inequality (5) can be verified by a careful comparison between the MLS constants of the original Glauber dynamics on μ and the Glauber dynamics on $\pi^{\mathbb{1}_R}$. For each $\pi^{\mathbb{1}_R}$, the entropy $\mathbf{Ent}_{\pi^{\mathbb{1}_R}}[f]$ can be bounded by the modified log-Sobolev inequality for Glauber dynamics $P(\pi^{\mathbb{1}_R})$ on $\pi^{\mathbb{1}_R}$:

$$\mathbf{Ent}_{\pi^{\mathbb{1}_R}}[f] \leq \frac{\mathcal{E}_{P(\pi^{\mathbb{1}_R})}(f, \log f)}{\rho^{\mathrm{GD}}(\pi^{\mathbb{1}_R})} \leq \frac{\mathcal{E}_{P(\pi^{\mathbb{1}_R})}(f, \log f)}{\rho^{\mathrm{GD}}_{\min}(\pi)},$$

where $\mathcal{E}_{P(\pi^{\mathbb{1}_R})}(f, \log f)$ is the Dirichlet form for the chain $P(\pi^{\mathbb{1}_R})$. Applying this inequality in the LHS of (5), an upper bound $\mathcal{E}_P(f, \log f)$ with the original chain P emerges from taking an "average" over all Dirichlet forms $\mathcal{E}_{P(\pi^{\mathbb{1}_R})}(f, \log f)$. The detailed calculation is given in Section 7.

2.2. **Magnetized block factorization of entropy**. We outline the proof of Lemma 2.3. To establish the magnetized block factorization of entropy, we use the following transformation of distributions.

Definition 2.4 (k-transformation [CFYZ21]). Let μ be a distribution over $\{-1,+1\}^V$ and $k \geq 1$ an integer. The *k-transformation* of μ, denoted by $\mu_k = \mathsf{Trans}(\mu, k)$, is a distribution over $\{-1,+1\}^{V \times [k]}$ defined as follows. Let $\boldsymbol{X} \sim \mu$. Then $\mu_k = \mathsf{Trans}(\mu, k)$ is the distribution of $\boldsymbol{Y} \in \{-1,+1\}^{V \times [k]}$ constructed as follows:

- if $X_v = -1$, then $Y_{(v,i)} = -1$ for all $i \in [k]$;
- if $X_v = +1$, then $Y_{(v,i^*)} = +1$ and $Y_{(v,i)} = -1$ for all $i \in [k] \setminus \{i^*\}$, where i^* is chosen from $[k]$ uniformly and independently at random.

Lemma 2.3 is proved by combining the following two lemmas that are claimed for arbitrary distribution μ over $\{-1,+1\}^V$ and $\eta > 0$.

Lemma 2.5. *If μ is η-spectrally independent in ∞-norm with all fields, then for every integer $k \geq 1$, the k-transformed distribution $\mu_k = \mathsf{Trans}(\mu, k)$ is $(\eta+1)$-spectrally independent in ∞-norm with all fields.*

Lemma 2.6. *If for every integer $k \geq 1$, $\mu_k = \mathsf{Trans}(\mu, k)$ is η-spectrally independent in ∞-norm with all fields, then for every $\theta \in (0,1)$, μ satisfies θ-magnetized block factorization of entropy with parameter $C = \left(\frac{\mathrm{e}}{\theta}\right)^{\eta+2}$.*

Lemma 2.5 is proved by a coupling between μ and μ_k. The proof is given in Section 5.

We now give an outline of the proof of Lemma 2.6. First, we show that for all sufficiently large k, the distribution μ_k satisfies the *uniform block factorization of entropy* in [CLV21a]. Next, we show that when $k \to \infty$, the uniform block factorization of entropy for μ_k implies the magnetized block factorization of entropy for the original distribution μ.

Definition 2.7 (uniform block factorization of entropy [CP20]). Let V be a set of size $n = |V|$, $1 \leq \ell \leq n$ an integer, and $C > 0$. A distribution μ over $\{-1,+1\}^V$ is said to satisfy the *ℓ-uniform block factorization of entropy with parameter C* if for all $f : \Omega(\mu) \to \mathbb{R}_{\geq 0}$,

$$\mathbf{Ent}_\mu\left[f\right] \leq \frac{C}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \mu[\mathbf{Ent}_S\left[f\right]],$$

where $\mu[\mathbf{Ent}_S\left[f\right]] \triangleq \sum_{\sigma \in \Omega(\mu_{V\setminus S})} \mu_{V\setminus S}(\sigma) \cdot \mathbf{Ent}_{\mu^\sigma}\left[f\right]$.

The following lemma shows that for all sufficiently large integers k, the transformed distribution μ_k satisfies the uniform block factorization of entropy. The lemma can be proved using the approach of *entropic independence* developed in [AASV21, AJK⁺21]. A formal proof is included in Section 4.

Lemma 2.8 ([AASV21, AJK⁺21]). *Let μ be a distribution over $\{-1,+1\}^V$ with $n = |V|$. Let $\eta > 0$ and $\theta \in (0,1)$. For all integers $k > \frac{\eta+2}{\theta n}$, for the k-transformed distribution $\mu_k = \mathsf{Trans}(\mu, k)$, if μ_k is η-spectrally independent in ∞-norm with all fields, then μ_k satisfies $\lceil \theta k n \rceil$-uniform block factorization of entropy with $C = \left(\frac{\mathrm{e}}{\theta}\right)^{\eta+2}$.*

Our next lemma relates the uniform block factorization of entropy for the k-transformed distribution $\mu_k = \mathsf{Trans}(\mu, k)$ to the magnetized block factorization of entropy for μ. A formal proof is given in Section 6.

Lemma 2.9. *Let μ a distribution over $\{-1,+1\}^V$ with $n = |V|$. Let $\theta \in (0,1)$ and $C > 0$. If there is a finite k_0 such that for all integers $k \geq k_0$, the distribution $\mu_k = \mathsf{Trans}(\mu, k)$ satisfies $\lceil \theta k n \rceil$-uniform block factorization with parameter C, then μ satisfies θ-magnetized block factorization of entropy with parameter C.*

Lemma 2.6 is a straightforward consequence of Lemma 2.8 and Lemma 2.9.

Finally, we briefly outline the proof of Lemma 2.9. Fix an integer $k \geq k_0$. We use V_k to denote the $V \times [k]$ and v_i to denote the pair (v, i). Since μ_k satisfies the $\lceil \theta k n \rceil$-uniform block fractorization with parameter C, it holds that for any $g : \Omega(\mu_k) \to \mathbb{R}$,

$$\mathbf{Ent}_{\mu_k}\left[g\right] \leq \frac{C}{\binom{nk}{\lceil \theta n k \rceil}} \sum_{S \in \binom{V_k}{\lceil \theta n k \rceil}} \mu_k\left[\mathbf{Ent}_S\left[g\right]\right] = C \cdot \mathbf{E}_{S \sim \binom{V_k}{\lceil \theta n k \rceil}}\left[\mu_k\left[\mathbf{Ent}_S\left[g\right]\right]\right], \tag{6}$$

where $S \sim \binom{V_k}{\lceil \theta n k \rceil}$ is sampled uniformly at random from $\binom{V_k}{\lceil \theta n k \rceil}$. By the definition of the k-transformation, a mapping from $\Omega(\mu_k)$ to $\Omega(\mu)$ can be naturally constructed as follows: For any $\sigma \in \Omega(\mu_k)$,

$$\forall v \in V, \quad \sigma^\star_v = \begin{cases} +1 & \exists i \in [k] \text{ s.t. } \sigma_{(v,i)} = +1; \\ -1 & \forall i \in [k], \sigma_{(v,i)} = -1. \end{cases}$$

Given any function $f : \Omega(\mu) \to \mathbb{R}_{\geq 0}$, define a function $f^k : \Omega(\mu_k) \to \mathbb{R}_{\geq 0}$ by $f^k(\sigma) = f(\sigma^\star)$ for all $\sigma \in \Omega(\mu_k)$. The following identities can be verified:

$$\mathbf{Ent}_\mu[f] = \mathbf{Ent}_{\mu_k}\left[f^k\right], \tag{7}$$

$$\frac{Z_\pi}{\theta^n} \mathop{\mathbf{E}}_{R\sim\mathrm{Bin}(V,1-\theta)} \left[\pi_R(\mathbb{1}_R) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}[f]\right] = \lim_{k\to\infty} \mathbf{E}_{S\sim\binom{V_k}{\lceil\theta n k\rceil}} \left[\mu_k\left[\mathbf{Ent}_S\left[f^k\right]\right]\right]. \tag{8}$$

Equation (7) follows from the bijection. Equation (8) follows from the concentration property of the random set S. Applying (7) and (8) together with (6) and letting $k \to \infty$ gives:

$$\begin{aligned}
\mathbf{Ent}_\mu[f] &\overset{\text{by (7)}}{=} \lim_{k\to\infty} \mathbf{Ent}_{\mu_k}\left[f^k\right] \\
&\overset{\text{by (6)}}{\leq} \lim_{k\to\infty} \mathbf{E}_{S\sim\binom{V_k}{\lceil\theta n k\rceil}} \left[\mu_k\left[\mathbf{Ent}_S\left[f^k\right]\right]\right] \\
&\overset{\text{by (8)}}{=} \frac{Z_\pi}{\theta^n} \mathop{\mathbf{E}}_{R\sim\mathrm{Bin}(V,1-\theta)} \left[\pi_R(\mathbb{1}_R) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}[f]\right].
\end{aligned}$$

This proves the magnetized block factorization of entropy for μ.

3. Definitions and Preliminaries

3.1. **Markov chain background**. Let Ω be a finite state space and $(X_t)_{t\geq 0}$ is a Markov chain on it. The Markov chain $(X_t)_{t\geq 0}$ can be represented by a transition matrix $P \in \mathbb{R}_{\geq 0}^{\Omega\times\Omega}$. We often use the matrix P to refer to $(X_t)_{t\geq 0}$ when the context is clear. A distribution μ is called a stationary distribution of P if $\mu = \mu P$. A Markov chain is said to be

- *irreducile*, if for any $X, Y \in \Omega$, there is an integer $t \geq 0$ such that $P^t(X,Y) > 0$;
- *aperiodic*, if for any $X \in \Omega$, $\gcd\{t > 0 \mid P^t(X,X) > 0\} = 1$;
- *reversible* with respect to μ, if the following *detailed balanced equation* is satisfied
$$\forall X, Y \in \Omega, \quad \mu(X)P(X,Y) = \mu(Y)P(Y,X),$$
which also implies that μ is a stationary distribution of P.

It is well-known that when a Markov chain is both irreducible and aperiodic, then it has a unique stationary distribution [LPW09].

Let μ be a distribution with support Ω and P be a Markov chain over Ω with the unique stationary distribution μ. To measure the convergence rate of P, we define the *mixing time* of P to be

$$\forall 0 < \varepsilon < 1, \quad T_{\mathrm{mix}}(\varepsilon) \triangleq \max_{X\in\Omega} \min\left\{t \mid d_{\mathrm{TV}}\left(P^t(X,\cdot),\mu\right) \leq \varepsilon\right\},$$

where $d_{\mathrm{TV}}\left(P^t(X,\cdot),\mu\right)$ is the *total variation distance* between $P^t(X,\cdot)$ and μ, which is defined as

$$d_{\mathrm{TV}}\left(P^t(X,\cdot),\mu\right) \triangleq \frac{1}{2}\sum_{Y\in\Omega}\left|P^t(X,Y) - \mu(Y)\right|.$$

There is a sharp connection between the mixing time and the *modified log-Sobolev (MLS) constant*. To introduce it, we define the *Dirichlet form* with respect to P as

$$\mathcal{E}_P(f,g) \triangleq \langle f, (I-P)g\rangle_\mu \triangleq \sum_{X\in\Omega} \mu(X)\, f(X)\, (I-P)g(X),$$

where f, g are two functions over Ω and I denotes the identity matrix. If the Markov chain P is reversible with respect to μ, the Dirichlet form can be written as

$$\mathcal{E}_P(f,g) = \frac{1}{2}\sum_{\sigma,\tau\in\Omega} \mu(\sigma)P(\sigma,\tau)(f(\sigma)-f(\tau))(g(\sigma)-g(\tau)).$$

Moreover, let the *entropy* of a function $f : \Omega \to \mathbb{R}_{\geq 0}$ as

$$\mathbf{Ent}_\mu[f] \triangleq \mathbf{E}_\mu[f \log f] - \mathbf{E}_\mu[f] \log \mathbf{E}_\mu[f],$$

where we use the convention that $0 \log 0 = 0$. Note that when $\mathbf{E}_\mu[f] = 1$, then $\mathbf{Ent}_\mu[f]$ is exactly the *relative entropy* (a.k.a. *KL-divergence*) of $\nu(\cdot) = \mu(\cdot)f(\cdot)$ and $\mu(\cdot)$. Formally

$$\mathbf{Ent}_\mu[f] = D_{\mathrm{KL}}(\nu \parallel \mu) \triangleq \sum_{\sigma \in \Omega} \nu(\sigma) \log \frac{\nu(\sigma)}{\mu(\sigma)}, \quad \text{where } \nu(\sigma) = \mu(\sigma) f(\sigma).$$

For any functions $\hat{f}, \hat{g}$ over $\hat{\Omega}$ such that $\hat{\Omega} \supseteq \Omega$, we simply use $\mathcal{E}_P(\hat{f}, \hat{g})$ and $\mathbf{Ent}_\mu\left[\hat{f}\right]$ to denote $\mathcal{E}_P(\hat{f}_\Omega, \hat{g}_\Omega)$ and $\mathbf{Ent}_\mu\left[\hat{f}_\Omega\right]$, where $\hat{f}_\Omega$ and $\hat{g}_\Omega$ are obtained by restricting $\hat{f}$ and $\hat{g}$ on Ω. The *modified log-Sobolev constant* introduced in [BT06] is defined by:

$$\rho(P) \triangleq \inf \left\{ \frac{\mathcal{E}_P(f, \log f)}{\mathbf{Ent}_\mu[f]} \;\middle|\; f : \Omega(\mu) \to \mathbb{R}_{\geq 0},\ \mathbf{Ent}_\mu[f] \neq 0 \right\}.$$

If P is reversible and all the eigenvalues of P are non-negative, then by [BT06, Corollary 2.8] and [DSC96, Corollary 2.2, (ii)], it holds that

$$t_{\mathrm{mix}}(P, \varepsilon) \leq \frac{1}{\rho(P)} \left(\log \log \frac{1}{\mu_{\min}} + \log \frac{1}{2\varepsilon^2} \right), \tag{9}$$

where $\mu_{\min} = \min_{\sigma \in \Omega} \mu(\sigma)$ denotes the minimum probability in μ. Specifically, [BT06] gives the mixing time for the continuous time variant of P, and [DSC96] shows that when P is reversible and all the eigenvalues of P are non-negative, then the mixing time of P could be bounded by the mixing time of its continous time variant.

3.2. **Fractional log-concavity and correlation matrix**. Let V be a finite set. Without loss of generality, we assume that $V = [n] = \{1, 2, \ldots, n\}$. Let μ be a distribution over $\{-1, +1\}^V$. Equivalently, we can view μ as a distribution over the subsets of $[n]$ such that for any *configuration* $\sigma \in \{-1, +1\}^{[n]}$,

$$\mu(\sigma) = \mu(S_\sigma), \quad \text{where } S_\sigma = \{i \in [n] \mid \sigma_i = +1\}. \tag{10}$$

For any $i \in [n]$, we may use i to denote event $i \in S$, and $\bar{i}$ to denote event $i \notin S$. The *signed influence matrix* for μ in Definition 1.2 can be rewritten as follows

$$\forall i, j \in [n], \quad \Psi_\mu^{\mathrm{inf}}(i, j) = \begin{cases} \mathbf{Pr}_\mu[j \mid i] - \mathbf{Pr}_\mu\left[j \mid \bar{i}\right] & \text{if } i \neq j \wedge \mathbf{Pr}_\mu[i] > 0 \wedge \mathbf{Pr}_\mu\left[\bar{i}\right] > 0; \\ 0 & \text{otherwise.} \end{cases}$$

Let $\boldsymbol{\lambda} = (\lambda_i)_{i \in [n]} \in \mathbb{R}_{>0}$ be local fields. Let $\mu^{(\boldsymbol{\lambda})}$ denote the distribution obtained from imposing the local fields $\boldsymbol{\lambda}$ onto μ, which can also be viewed as a distribution over subsets of $[n]$. Formally,

$$\forall S \subseteq [n], \quad \mu^{(\boldsymbol{\lambda})}(S) \propto \mu(S) \prod_{i \in S} \lambda_i.$$

The *generating polynomial* associated to μ is defined by

$$g_\mu(z_1, z_2, \ldots, z_n) \triangleq \sum_{S \subseteq [n]} \mu(S) \prod_{i \in S} z_i.$$

Definition 3.1 (fractional log-concavity[AASV21]). Let $\alpha \in [0, 1]$. A distribution $\mu : 2^{[n]} \to \mathbb{R}_{\geq 0}$ is said to be α-fractionally log-concave (α-FLC) if $\log g_\mu(z_1^\alpha, z_2^\alpha, \ldots, z_n^\alpha)$ is concave, viewed as a function on $\mathbb{R}_{\geq 0}^n$.

The notion of fractional log-concavity is closely related to the following *correlation matrix*.

Definition 3.2 (correlation matrix). Let $\mu : 2^{[n]} \to \mathbb{R}_{\geq 0}$ be a distribution. The correlation matrix $\Psi_\mu^{\mathrm{cor}} : [n] \times [n] \to \mathbb{R}$ is defined by

$$\forall i, j \in [n], \quad \Psi_\mu^{\mathrm{cor}}(i,j) = \begin{cases} 1 - \mathbf{Pr}_\mu[i] & \text{if } i = j; \\ \mathbf{Pr}_\mu[j \mid i] - \mathbf{Pr}_\mu[j] & \text{if } i \neq j \wedge \mathbf{Pr}_\mu[i] > 0; \\ 0 & \text{if } i \neq j \wedge \mathbf{Pr}_\mu[i] = 0. \end{cases}$$

Proposition 3.3 ([AASV21]). *Let $\mu : 2^{[n]} \to \mathbb{R}_{\geq 0}$ be a distribution and $\alpha \in (0,1]$. μ is α-fractionally log-concave if and only if $\lambda_{\max}(\Psi^{\mathrm{cor}}_{\mu^{(\lambda)}}) \leq \frac{1}{\alpha}$ for any positive local fields $\boldsymbol{\lambda} = (\lambda_i)_{i\in[n]} \in \mathbb{R}_{>0}$.*

For the proof of Proposition 3.3, readers can refer to the proof of Lemma 69 and Remark 70 in the full version of [AASV21].

3.3. **Homogeneous distributions and random walks.** Let $\mu : \binom{[n]}{k} \to \mathbb{R}_{\geq 0}$ be a distribution over the size-k subsets of $[n]$. We call such distributions the *homogeneous distributions*.

Let $\Omega \subseteq \binom{[n]}{k}$ denote the support of μ. Let X be the downward closure of Ω. Formally, X is the smallest family such that $\Omega \subseteq X$ and if $\alpha \in X$ then $\beta \in X$ for all $\beta \subseteq \alpha$. In other words, X is the *simplicial complexes* generated by μ. For any *face* $\alpha \in X$, let $|\alpha|$ denote the *dimension* of α. For any integer $0 \leq j \leq k$, let $X(j)$ denote all the faces in X with dimension j.

Definition 3.4 (down/up walk). Let X be the simplicial complexes generated by a homogeneous distribution $\mu : \binom{[n]}{k} \to \mathbb{R}_{\geq 0}$. Let $0 \leq j < k$ be an integer.

- The down walk $D_{k\to j} : X(k) \times X(j) \to \mathbb{R}_{\geq 0}$ is defined by

$$\forall \alpha \in X(k), \beta \in X(j), \quad D_{k\to j}(\alpha, \beta) = \begin{cases} \frac{1}{\binom{k}{j}} & \text{if } \beta \subseteq \alpha; \\ 0 & \text{otherwise.} \end{cases}$$

- The up walk $U_{j\to k} : X(j) \times X(k) \to \mathbb{R}_{\geq 0}$ is defined by

$$\forall \alpha \in X(j), \beta \in X(k), \quad U_{j\to k}(\alpha, \beta) = \begin{cases} \frac{\mu(\beta)}{\sum_{\gamma \in X(k) : \alpha \subseteq \gamma} \mu(\gamma)} & \text{if } \alpha \subseteq \beta; \\ 0 & \text{otherwise.} \end{cases}$$

The following relative entropy decay result is proved in [AJK+21].

Theorem 3.5 ([AJK+21, Theorem 4]). *Suppose $\mu : \binom{[n]}{k} \to \mathbb{R}_{\geq 0}$ is α-fractionally log-concave for some $\alpha \in (0,1]$. Let Ω denote the support of μ. For any integer $0 \leq j \leq k - \lceil 1/\alpha \rceil$, and any distribution ν over Ω,*

$$D_{\mathrm{KL}}\left(\nu D_{k\to j} \,\|\, \mu D_{k\to j}\right) \leq \left(1 - \kappa\left(j, k, \frac{1}{\alpha}\right)\right) D_{\mathrm{KL}}\left(\nu \,\|\, \mu\right),$$

where

$$\kappa\left(j,k,c\right) \triangleq \frac{(k+1-j-c)^{c - \lceil c \rceil} \prod_{i=0}^{\lceil c \rceil - 1} (k-j-i)}{(k+1)^c}. \tag{11}$$

Let $\mu_{(k)} = \mu$. For any $0 \leq j < k$, let $\mu_{(j)} = \mu_{(k)} D_{k\to j}$. Let $f^{(k)} : X(k) \to \mathbb{R}_{\geq 0}$. For any $0 \leq j < k$, define the function $f^{(j)} : X(j) \to \mathbb{R}_{\geq 0}$ by $f^{(j)} = U_{j\to k} f^{(k)}$. We have the following lemma.

Lemma 3.6. *Let $\mu = \mu_{(k)} : \binom{[n]}{k} \to \mathbb{R}_{\geq 0}$, $0 \leq j < k$ and $\kappa \in (0,1)$. Let Ω denote the support of μ. Suppose for any distribution ν over Ω, $D_{\mathrm{KL}}\left(\nu D_{k\to j} \,\|\, \mu D_{k\to j}\right) \leq (1-\kappa) D_{\mathrm{KL}}\left(\nu \,\|\, \mu\right)$. For any function $f^{(k)} : X(k) \to \mathbb{R}_{\geq 0}$,*

$$\mathbf{Ent}_{\mu_{(j)}}\left[f^{(j)}\right] \leq (1-\kappa) \mathbf{Ent}_{\mu_{(k)}}\left[f^{(k)}\right].$$

Proof. First note that if $f^{(k)} = 0$, then the lemma holds trivially. We next prove that we only need to consider the function $f^{(k)}$ with $\mathbf{E}_{\mu_{(k)}}\left[f^{(k)}\right] = 1$. For any $g : X(k) \to \mathbb{R}_{\geq 0}$ and any $c > 0$, it holds that $\mathbf{Ent}_{\mu_{(k)}}[cg] = c\mathbf{Ent}_{\mu_{(k)}}[g]$ and $\mathbf{Ent}_{\mu_{(j)}}\left[(cg)^{(j)}\right] = c\mathbf{Ent}_{\mu_{(j)}}\left[g^{(j)}\right]$ where $g^{(j)} = U_{j\to k}g$ and $(cg)^{(j)} = U_{j\to k}(cg)$ for $j < k$. Suppose $f^{(k)} \neq 0$ and $t = \mathbf{E}_{\mu_{(k)}}\left[f^{(k)}\right] > 0$. The lemma holds for $f^{(k)}$ if and only if the lemma holds for $f' = f^{(k)}/t$ and $\mathbf{E}_{\mu_{(k)}}[f'] = 1$.

Assume $\mathbf{E}_{\mu_{(k)}}\left[f^{(k)}\right] = 1$. Define a distribution ν over $X(k)$ by $\nu(\sigma) = \mu(\sigma)f^{(k)}(\sigma)$ for all $\sigma \in X(k)$, then

$$D_{\mathrm{KL}}(\nu \,\|\, \mu) = \sum_{\sigma\in X(k)} \mu(\sigma)\frac{\nu(\sigma)}{\mu(\sigma)}\log\frac{\nu(\sigma)}{\mu(\sigma)} = \mathbf{E}_{\mu}\left[f^{(k)}\log f^{(k)}\right] \overset{(*)}{=} \mathbf{Ent}_{\mu_{(k)}}\left[f^{(k)}\right],$$

where $(*)$ holds because $\mathbf{E}_{\mu}\left[f^{(k)}\right] = 1$ and $\mu = \mu_{(k)}$. Let $\nu_{(j)} = \nu D_{k\to j}$ and $\mu_{(j)} = \mu D_{k\to j}$. We have

$$D_{\mathrm{KL}}\left(\nu D_{k\to j} \,\|\, \mu D_{k\to j}\right) = \sum_{\sigma\in X(j)} \mu_{(j)}(\sigma)\frac{\nu_{(j)}(\sigma)}{\mu_{(j)}(\sigma)}\log\frac{\nu_{(j)}(\sigma)}{\mu_{(j)}(\sigma)} = \mathbf{Ent}_{\mu_{(j)}}\left[f^{(j)}\right],$$

where the last equation holds because $f^{(j)}(\sigma) = U_{j\to k}f(\sigma) = \sum_{\alpha\in X(k):\sigma\subseteq\alpha} U_{j\to k}(\sigma,\alpha)\frac{\nu(\alpha)}{\mu(\alpha)} = \frac{\nu_{(j)}(\sigma)}{\mu_{(j)}(\sigma)}$ and $\mathbf{E}_{\mu_{(j)}}\left[f^{(j)}\right] = \sum_{\sigma\in X(j)}\mu_{(j)}(\sigma)\frac{\nu_{(j)}(\sigma)}{\mu_{(j)}(\sigma)} = 1$. This proves the lemma. □

3.4. **Multivariate hypergeometric distribution**. Let V be a set of n buckets, each of them has k balls. Suppose we pick ℓ balls from all kn balls uniformly at random, without replacement. For each bucket $v \in V$, let $a_v \in \mathbb{Z}_{\geq 0}$ denote the number of balls picked from the bucket v, then $\boldsymbol{a} = (a_v)_{v\in V}$ follows multivariate hypergeometric distribution.

Formally, given a set V of size n, an integer $k \geq 1$ and an integer $0 \leq \ell \leq kn$, the multivariate hypergeometric distribution $\Pi_{V,k,\ell}$ is defined as follows. The support of $\Pi_{V,k,\ell}$ is defined by

$$\Omega(\Pi_{V,k,\ell}) \triangleq \left\{\boldsymbol{a} = (a_v)_{v\in V} \mid \sum_{v\in V} a_v = \ell \text{ and } \forall v\in V, a_v \in \mathbb{Z}_{\geq 0}\right\}, \tag{12}$$

For any $\boldsymbol{a} \in \Omega(\Pi_{V,k,\ell})$, it holds that

$$\Pi_{V,k,\ell}(\boldsymbol{a}) = \frac{\prod_{v\in V}\binom{k}{a_v}}{\binom{kn}{\ell}}. \tag{13}$$

Lemma 3.7 ([JDP83] and [DR98]). *Let $\boldsymbol{a} \sim \Pi_{V,k,\ell}$. For any $v \in V$ and $\varepsilon \in (0,1)$, it holds that*

$$\mathbf{Pr}\left[\left|\frac{a_v}{k} - \frac{\ell}{kn}\right| \geq \varepsilon\right] \leq 2\exp\left(-2\varepsilon^2 k\right).$$

Proof. For each bucket $v \in V$, let $(v,1),(v,2),\ldots,(v,k)$ denote all balls in v. For each ball (v,i), let $X_{(v,i)}$ indicate whether the ball (v,i) is picked. It holds that $a_v = \sum_{i\in[k]} X_{(v,i)}$. Since $(X_{(v,i)})_{i\in[k]}$ are negative associated [JDP83, Lemma 2.11], the Chernoff-Hoeffding inequality [DR98] can be applied to a_v. □

4. Fractional Log-Concavity and Uniform Block Factorization

In this section, we prove Lemma 2.8. The lemma can be proved by the techniques in [AASV21, AJK+21]. We conclude the proof here for completeness. We prove a general result based on spectral independence.

Definition 4.1 (spectral independence with all fields). Let $\eta > 0$. A distribution μ over $\{-1,+1\}^V$ is said to be *η-spectrally independent with all fields* if $\lambda_{\max}(\Psi^{\mathrm{inf}}_{\mu^{(\boldsymbol{\lambda})}}) \leq \eta$ for any $\boldsymbol{\lambda} = (\lambda_v)_{v\in V} \in \mathbb{R}^V_{>0}$.

The notion of spectral independence was first introduced in [ALO20], then further developed in [CGŠV21, FGYZ21]. Remark that if a distribution is η-spectrally independent in ∞-norm with all fields, then it must be η-spectrally independent with all fields. This is because $\lambda_{\max}(\Psi^{\mathrm{inf}}_{\mu^{(\boldsymbol{\lambda})}}) \leq \left\|\Psi^{\mathrm{inf}}_{\mu^{(\boldsymbol{\lambda})}}\right\|_\infty$.

Proposition 4.2 ([AASV21, AJK⁺21]). *Let $V = [n] = \{1, 2, \ldots, n\}$ and $\eta > 0$. For any distribution μ over $\{-1, +1\}^V$ that is η-spectrally independent with all fields, any integer $\lceil \eta + 1 \rceil < \ell \le n$, μ satisfies ℓ-uniform block factorization of entropy with $C = 1/\kappa(n-\ell, n, \eta+1)$, where $\kappa(\cdot)$ is defined in (11).*

Now, we are ready to prove Lemma 2.8.

Proof of Lemma 2.8. Consider the distribution μ_k. Since μ_k is η-spectrally independent in ∞-norm with all fields, μ_k is $\hat{\eta}$-spectrally independent with all fields, where $\hat{\eta} = \lceil \eta \rceil$. Let $\ell = \lceil \theta n k \rceil$. Since $k > \frac{\eta+2}{\theta n}$, it holds that $\lceil \hat{\eta} + 1 \rceil < \ell \le kn$. By Proposition 4.2, μ_k satisfies ℓ-uniform block factorization of entropy with $C = 1/\kappa(nk - \ell, nk, \hat{\eta}+1)$. Note that $\hat{\eta}+1$ is an integer. As observed in [AJK⁺21], $\kappa(nk-\ell, nk, \hat{\eta}+1)$ defined in (11) equals to $\binom{\ell}{\hat{\eta}+1}/\binom{nk}{\hat{\eta}+1}$. This implies that

$$C = \frac{1}{\kappa(nk-\ell, nk, \hat{\eta}+1)} = \binom{nk}{\hat{\eta}+1}\Big/\binom{\ell}{\hat{\eta}+1} \le \left(\frac{\mathrm{e}nk}{\hat{\eta}+1}\right)^{\hat{\eta}+1} \Big/ \left(\frac{\ell}{\hat{\eta}+1}\right)^{\hat{\eta}+1} \le \left(\frac{\mathrm{e}nk}{\ell}\right)^{\eta+2} \le \left(\frac{\mathrm{e}}{\theta}\right)^{\eta+2}. \qquad \square$$

We now prove Proposition 4.2. We need to introduce some definitions. By (10), we can view μ as a distribution over $2^{[n]}$ such that for any $\sigma \in \{-1,+1\}^{[n]}$, $\mu(S_\sigma) = \mu(\sigma)$, where $S_\sigma = \{i \in [n] \mid \sigma_i = +1\}$. Define the *homogenization* of a distribution μ over $\{-1,+1\}^{[n]}$ is a distribution $\pi = \mu^{\mathrm{hom}}$ over subsets of $[n] \cup \overline{[n]} = \{1, 2, \ldots, n\} \cup \{\bar{1}, \bar{2}, \ldots, \bar{n}\}$. For any $S \subseteq [n]$, define $S^c = \{\bar{i} \mid i \in [n] \setminus S\}$, and let $\pi(S \cup S^c) = \mu(S)$. It is straightforward to verify π is a homogeneous distribution over $\binom{[n]\cup\overline{[n]}}{n}$. Let Ψ_π^{cor} denote the correlation matrix (Definition 3.2) of π and Ψ_μ^{inf} denote the signed influence matrix (Definition 1.2) of μ. The following result is proved in [AASV21] (see the proof of Lemma 71 in the full version of [AASV21]).

Lemma 4.3 ([AASV21]). *The spectrum of Ψ_π^{cor} is the union of $\{\lambda_i + 1\}_{1 \le i \le n}$ and n copies of 0, where $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_n$ are all eigenvalues of Ψ_μ^{inf}.*

To give the next lemma, we view μ as a distribution over $\{-1,+1\}^{[n]}$. We use $\Omega(\mu)$ and $\Omega(\pi)$ to denote the support of μ and π respectively. Fix a function $f : \Omega(\mu) \to \mathbb{R}_{\ge 0}$. Recall $S_\sigma = \{i \in [n] \mid \sigma_i = +1\}$ for any $\sigma \in \{-1,+1\}^{[n]}$. We can construct $f^{(n)} : \Omega(\pi) \to \mathbb{R}_{\ge 0}$ by $f^{(n)}(S_\sigma \cup S_\sigma^c) = f(\sigma)$ for all $\sigma \in \Omega_\mu$, where $S_\sigma^c = \{\bar{i} \mid i \in [n] \setminus S_\sigma\}$. Let X denote the simplicial complexes generated by π. Let $U_\cdot$ and $D_\cdot$ denote the up walk and down walk on X (Definition 3.4). Let $\pi_{(n)} = \pi$ and $\pi_{(j)} = \pi_{(n)} D_{n \to j}$ for all $0 \le j < n$. Let $f^{(j)} = U_{j \to n} f^{(n)}$ for all $0 \le j < n$. The following lemma is proved in [CLV21a] (see the proof of Lemma 2.6 in the full version of [CLV21a]).

Lemma 4.4 ([CLV21a]). *For any $0 \le j \le n$, it holds that*

$$\frac{1}{\binom{n}{j}} \sum_{S \in \binom{[n]}{j}} \mu[\mathbf{Ent}_S[f]] = \mathbf{Ent}_{\pi_{(n)}}\left[f^{(n)}\right] - \mathbf{Ent}_{\pi_{(n-j)}}\left[f^{(n-j)}\right]. \tag{14}$$

Now, we are ready to prove Proposition 4.2.

Proof of Proposition 4.2. We view μ as a distribution over $2^{[n]}$. Let $\pi = \mu^{\mathrm{hom}}$ be the homogenization of μ.

We first show that π is $\frac{1}{\eta+1}$-fractionally log-concave. By Proposition 3.3, we only need to show that $\lambda_{\max}(\Psi^{\mathrm{cor}}_{\pi^{(\lambda)}}) \le \eta + 1$ for any positive local fields $\boldsymbol{\lambda} = (\lambda_i)_{i \in [n] \cup \overline{[n]}} \in \mathbb{R}_{>0}^{[n]\cup\overline{[n]}}$. Fix a $\boldsymbol{\lambda} \in \mathbb{R}_{>0}^{[n]\cup\overline{[n]}}$. Define $\boldsymbol{\phi} = (\phi_i)_{i \in [n]}$ such that for any $i \in [n]$, $\phi_i = \frac{\lambda_i}{\lambda_{\bar{i}}}$. For any subset $S \subseteq [n]$, recall $S^c = \{\bar{i} \mid i \in [n] \setminus S\}$, we have

$$\begin{aligned} \mu^{(\boldsymbol{\phi})}(S) &\propto \mu(S) \prod_{i \in S} \frac{\lambda_i}{\lambda_{\bar{i}}} \propto \left(\prod_{i \in [n]} \lambda_{\bar{i}}\right) \cdot \mu(S) \prod_{i \in S} \frac{\lambda_i}{\lambda_{\bar{i}}} = \mu(S) \prod_{i \in S} \lambda_i \prod_{i \in [n] \setminus S} \lambda_{\bar{i}}, \\ \pi^{(\boldsymbol{\lambda})}(S \cup S^c) &\propto \mu(S) \prod_{i \in S} \lambda_i \prod_{i \in [n] \setminus S} \lambda_{\bar{i}} \propto \mu^{(\boldsymbol{\phi})}(S). \end{aligned}$$

This implies that $\pi^{(\lambda)}$ is the homogenization of $\mu^{(\phi)}$. Since μ is η-spectrally independent with all fields, it holds that $\lambda_{\max}(\Psi^{\mathrm{inf}}_{\mu^{(\phi)}}) \leq \eta$. By Lemma 4.3, we have $\lambda_{\max}(\Psi^{\mathrm{cor}}_{\pi^{(\lambda)}}) \leq \eta + 1$. By Proposition 3.3, π is $\frac{1}{\eta+1}$-fractionally log-concave.

Let X denote the simplicial complexes generated by π. Let $U_\cdot$ and $D_\cdot$ denote the up walk and down walk on X (Definition 3.4). Let $\pi_{(n)} = \pi$ and $\pi_{(j)} = \pi_{(n)} D_{n\to j}$ for all $0 \leq j < n$. Let $\Omega(\pi)$ denote the support of π. Let $f^{(n)} : \Omega(\pi) \to \mathbb{R}_{\geq 0}$ and $f^{(j)} = U_{j\to n} f^{(n)}$ for all $0 \leq j < n$. Combining Theorem 3.5 and Lemma 3.6, we have for any $0 \leq j < n - \lceil \eta + 1 \rceil$, it holds that

$$\mathbf{Ent}_{\pi_{(j)}}\left[f^{(j)}\right] \leq (1 - \kappa(j, n, \eta + 1))\,\mathbf{Ent}_{\pi_{(n)}}\left[f^{(n)}\right], \tag{15}$$

where $\kappa(\cdot)$ is defined in (11).

Now, we view μ as a distribution over $\{-1, +1\}^V$. We use $\Omega(\mu) \subseteq \{-1, +1\}^V$ to denote the support of μ. Fix a function $f : \Omega(\mu) \to \mathbb{R}_{\geq 0}$. Note that for every $\sigma \in \Omega(\mu)$, the set $S_\sigma = \{i \in [n] \mid \sigma_i = +1\}$. We can construct $f^{(n)} : \Omega(\pi) \to \mathbb{R}_{\geq 0}$ by $f^{(n)}(S_\sigma \cup S_\sigma^c) = f(\sigma)$ for all $\sigma \in \Omega(\mu)$. By Lemma 4.4, for all $j \leq n$,

$$\frac{1}{\binom{n}{j}} \sum_{S \in \binom{[n]}{j}} \mu[\mathbf{Ent}_S[f]] = \mathbf{Ent}_{\pi_{(n)}}\left[f^{(n)}\right] - \mathbf{Ent}_{\pi_{(n-j)}}\left[f^{(n-j)}\right]. \tag{16}$$

For any $\lceil \eta + 1 \rceil < \ell \leq n$, we have

$$\begin{aligned}
\mathbf{Ent}_\mu[f] &\overset{(*)}{=} \mathbf{Ent}_{\pi_{(n)}}\left[f^{(n)}\right] = \mathbf{Ent}_{\pi_{(n)}}\left[f^{(n)}\right] - \mathbf{Ent}_{\pi_{(n-\ell)}}\left[f^{(n-\ell)}\right] + \mathbf{Ent}_{\pi_{(n-\ell)}}\left[f^{(n-\ell)}\right] \\
\text{(by (15) and (16))} \quad &\leq \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{[n]}{\ell}} \mu[\mathbf{Ent}_S[f]] + (1 - \kappa(n-\ell, n, \eta+1))\,\mathbf{Ent}_{\pi_{(n)}}\left[f^{(n)}\right] \\
&\overset{(\star)}{=} \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{[n]}{\ell}} \mu[\mathbf{Ent}_S[f]] + (1 - \kappa(n-\ell, n, \eta+1))\,\mathbf{Ent}_\mu[f]
\end{aligned}$$

where $(*)$ and $(\star)$ hold due to the definitions of $\pi_{(n)}$ and $f^{(n)}$. This proves the proposition. □

5. Spectral Independence of Transformed Distributions

In this section, we prove Lemma 2.5. Let $\eta > 0$ be a real number. Recall that μ is a distribution over $\{-1, +1\}^V$ that is η-spectrally independent with all fields. For $k \geq 1$ be an integer, μ_k is the k-transformed distribution of μ, and we want to show that μ_k is $(\eta + 1)$-spectrally independent with all fields. If $k = 1$, the lemma is trivial. We assume $k \geq 2$ in the proof.

For each $(v, i) \in V \times [k]$, we use v_i to denote the pair (v, i). Let $\boldsymbol{\phi} \in \mathbb{R}^{V\times[k]}_{>0}$ be a vector of local fields and $\pi_k \triangleq \mu_k^{(\boldsymbol{\phi})}$ be the magnetized distribution generated by μ_k and the local fields $\boldsymbol{\phi}$. Then, let $\boldsymbol{\varphi} \in \mathbb{R}^V_{>0}$ be another vector of local fields defined as

$$\forall w \in V, \quad \varphi_w \triangleq \frac{1}{k} \sum_{h \in [k]} \phi_{w_h}, \tag{17}$$

and let $\pi \triangleq \mu^{(\boldsymbol{\varphi})}$ be obtained by imposing the local fields $\boldsymbol{\varphi}$ onto μ.

Lemma 5.1. *Let $u_i, v_j \in V \times [k]$, it holds that*

$$\text{if } u \neq v, \quad \left|\Psi^{\mathrm{inf}}_{\pi_k}(u_i, v_j)\right| \leq \frac{\phi_{v_j}}{\sum_{h\in[k]} \phi_{v_h}} \left|\Psi^{\mathrm{inf}}_{\pi}(u, v)\right|; \tag{18}$$

$$\text{if } u = v \text{ and } i \neq j, \quad \left|\Psi^{\mathrm{inf}}_{\pi_k}(u_i, u_j)\right| \leq \frac{\phi_{u_j}}{\sum_{h\in[k]\setminus\{i\}} \phi_{u_h}}. \tag{19}$$

Now, we are ready to prove Lemma 2.5.

Proof of Lemma 2.5. Let u_i be an element of $V \times [k]$, by (18), (19) and the fact that μ is η-spectrally independent with all fields, we have

$$\sum_{v_j \in V\times[k]} \left|\Psi^{\inf}_{\pi_k}(u_i, v_j)\right| \leq \sum_{v\in V} \left|\Psi^{\inf}_{\pi}(u, v)\right| + 1 \leq \eta + 1. \qquad \square$$

The rest of this section is dedicated to the proof of Lemma 5.1. Recall that by the definition of π_k,

$$\pi_k(\sigma) = \frac{\mu_k(\sigma)}{Z_{\boldsymbol{\phi}}} \prod_{\substack{v_i\in V\times[k]\\ \sigma_{v_i}=+1}} \phi_{v_i}, \tag{20}$$

where $Z_{\boldsymbol{\phi}} = \sum_{\sigma\in\Omega(\mu_k)} \mu_k(\sigma) \prod_{v_i\in V\times[k]:\sigma_{v_i}=+1} \phi_{v_i}$ is the normalizing factor.

We prove Lemma 5.1 by verifying (18) and (19).

5.1. **Verification of** (18). Without loss of generality, we may assume that $\Omega(\pi_{k,u_i}) = \{-1, +1\}$, since otherwise (18) holds trivially. In this case, it is equivalent for us to bound

$$\left|\Psi^{\inf}_{\pi_k}(u_i, v_j)\right| = \left|\pi^{u_i\leftarrow+1}_{k,v_j}(+1) - \pi^{u_i\leftarrow-1}_{k,v_j}(+1)\right|.$$

We will calculate $\pi^{u_i\leftarrow+1}_{k,v_j}(+1)$ and $\pi^{u_i\leftarrow-1}_{k,v_j}(+1)$, respectively. For $\pi^{u_i\leftarrow+1}_{k,v_j}(+1)$, by the law of conditional distirbution, it holds that

$$\pi^{u_i\leftarrow+1}_{k,v_j}(+1) = \frac{\mathbf{Pr}_{Y\sim\pi_k}\left[Y_{u_i} = +1 \wedge Y_{v_j} = +1\right]}{\mathbf{Pr}_{Y\sim\pi_k}\left[Y_{u_i} = +1\right]}.$$

By the definition of the distribution π_k, it holds that

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{Y\sim\pi_k}\left[Y_{u_i} = +1 \wedge Y_{v_j} = +1\right] &= \frac{1}{Z_{\boldsymbol{\phi}}} \sum_{\substack{Y\in\Omega(\mu_k):\\ Y_{u_i}=+1\wedge Y_{v_j}=+1}} \mu_k(Y) \prod_{\substack{w_h\in V\times[k]:\\ Y_{w_h}=+1}} \phi_{w_h} \\
&= \frac{\phi_{u_i}\phi_{v_j}}{Z_{\boldsymbol{\phi}}k^2} \sum_{\substack{X\in\Omega(\mu):\\ X_u=+1\wedge X_v=+1}} \mu(X) \prod_{\substack{w\in V\setminus\{u,v\}\\ X_w=+1}} \left(\sum_{h\in[k]} \frac{\phi_{w_h}}{k}\right) = \frac{\phi_{u_i}\phi_{v_j}}{Z_{\boldsymbol{\phi}}k^2} \sum_{\substack{X\in\Omega(\mu):\\ X_u=+1\wedge X_v=+1}} \mu(X) \prod_{\substack{w\in V\setminus\{u,v\}\\ X_w=+1}} \varphi_w,
\end{aligned} \tag{21}$$

where the last equation holds due to the definition of $\boldsymbol{\varphi}$ in (17). Similarly, it holds that

$$\mathop{\mathbf{Pr}}_{Y\sim\pi_k}\left[Y_{u_i} = +1\right] = \frac{\phi_{u_i}}{Z_{\boldsymbol{\phi}}k} \sum_{\substack{X\in\Omega(\mu):\\ X_u=1}} \mu(X) \prod_{\substack{w\in V\setminus\{u\}\\ X_w=+1}} \varphi_w. \tag{22}$$

Combining (21) and (22), we have

$$\begin{aligned}
\pi^{u_i\leftarrow+1}_{k,v_j}(+1) &= \frac{\phi_{v_j}}{k} \cdot \frac{\sum_{X\in\Omega(\mu):X_u=+1\wedge X_v=+1} \mu(X) \prod_{w\in V\setminus\{u,v\}:X_w=+1} \varphi_w}{\sum_{X\in\Omega(\mu):X_u=+1} \mu(X) \prod_{w\in V\setminus\{u\}:X_w=+1} \varphi_w} \\
&= \frac{\phi_{v_j}}{\sum_{h\in[k]} \phi_{v_h}} \cdot \frac{\sum_{X\in\Omega(\mu):X_u=+1\wedge X_v=+1} \mu(X) \prod_{w\in V:X_w=+1} \varphi_w}{\sum_{X\in\Omega(\mu):X_u=+1} \mu(X) \prod_{w\in V:X_w=+1} \varphi_w} = \frac{\phi_{v_j}}{\sum_{h\in[k]} \phi_{v_h}} \cdot \pi^{u\leftarrow+1}_{v}(+1),
\end{aligned} \tag{23}$$

where the last equation holds due to the definition of $\pi = \mu^{(\boldsymbol{\varphi})}$.

Next, we calculate

$$\pi^{u_i\leftarrow-1}_{k,v_j}(+1) = \frac{\mathbf{Pr}_{Y\sim\pi_k}\left[Y_{u_i} = -1 \wedge Y_{v_j} = +1\right]}{\mathbf{Pr}_{Y\sim\pi_k}\left[Y_{u_i} = -1\right]}.$$

Let $\mathbb{1}[\cdot]$ denote the indicator random variable. By the definition of π_k, we have the following equation

$$\begin{aligned}
&\Pr_{Y\sim\pi_k}\left[Y_{u_i}=-1\wedge Y_{v_j}=+1\right]\\
&=\frac{\phi_{v_j}}{Z_{\boldsymbol{\phi}}k}\sum_{\substack{X\in\Omega(\mu):\\X_v=+1}}\mu(X)\left(\mathbb{1}[X_u=+1]\left(\sum_{h\in[k]\setminus\{i\}}\frac{\phi_{u_h}}{k}\right)+\mathbb{1}[X_u=-1]\right)\prod_{\substack{w\in V\setminus\{u,v\}:\\X_w=+1}}\left(\sum_{h\in[k]}\frac{\phi_{w_h}}{k}\right)\\
&=\frac{\phi_{v_j}}{Z_{\boldsymbol{\phi}}k}\sum_{\substack{X\in\Omega(\mu):\\X_v=+1}}\mu(X)\left(\mathbb{1}[X_u=+1]\varphi'_u+\mathbb{1}[X_u=-1]\right)\prod_{\substack{w\in V\setminus\{u,v\}:\\X_w=+1}}\varphi'_w\\
&=\frac{\phi_{v_j}}{Z_{\boldsymbol{\phi}}k}\sum_{\substack{X\in\Omega(\mu):\\X_v=+1}}\mu(X)\prod_{\substack{w\in V\setminus\{v\}:\\X_w=+1}}\varphi'_w
\end{aligned}$$

where the vector $\boldsymbol{\varphi}'\in\mathbb{R}_{>0}$ is the local fields such that

$$\forall w\in V,\quad \varphi'_w=\begin{cases}\frac{1}{k}\sum_{h\in[k]}\phi_{w_h} & \text{if } w\neq u\\ \frac{1}{k}\sum_{h\in[k]\setminus\{i\}}\phi_{w_h} & \text{if } w=u.\end{cases}\tag{24}$$

Similarly, it holds that

$$\Pr_{Y\sim\pi_k}\left[Y_{u_i}=-1\right]=\frac{1}{Z_{\boldsymbol{\phi}}}\sum_{X\in\Omega(\mu)}\mu(X)\left(\mathbb{1}[X_u=+1]\varphi'_u+\mathbb{1}[X_u=-1]\right)\prod_{\substack{w\in V\setminus\{u\}:\\X_w=+1}}\varphi'_w=\frac{1}{Z_{\boldsymbol{\phi}}}\sum_{X\in\Omega(\mu)}\mu(X)\prod_{\substack{w\in V:\\X_w=+1}}\varphi'_w.$$

By a similar calculation to (23), we can verify that

$$\begin{aligned}
\pi_{k,v_j}^{u_i\leftarrow-1}(+1)&=\frac{\phi_{v_j}}{k}\cdot\frac{\sum_{X\in\Omega(\mu):X_v=+1}\mu(X)\prod_{w\in V\setminus\{v\}:X_w=+1}\varphi'_w}{\sum_{X\in\Omega(\mu)}\mu(X)\prod_{w\in V:X_w=+1}\varphi'_w}\\
&=\frac{\phi_{v_j}}{\sum_{h\in[k]}\phi_{v_h}}\cdot\frac{\sum_{X\in\Omega(\mu):X_v=+1}\mu(X)\prod_{w\in V:X_w=+1}\varphi'_w}{\sum_{X\in\Omega(\mu)}\mu(X)\prod_{w\in V:X_w=+1}\varphi'_w}=\frac{\phi_{v_j}}{\sum_{h\in[k]}\phi_{v_h}}\cdot\nu_v(+1),
\end{aligned}\tag{25}$$

where $\nu=\mu^{(\boldsymbol{\varphi}')}$. Combining (23) and (25), we know that

$$\left|\pi_{k,v_j}^{u_i\leftarrow+1}(+1)-\pi_{v_j}^{u_i\leftarrow-1}(+1)\right|=\frac{\phi_{v_i}}{\sum_{h\in[k]}\phi_{v_h}}\cdot\left|\pi_v^{u\leftarrow+1}(+1)-\nu_v(+1)\right|.\tag{26}$$

In order to bound $\left|\pi_v^{u\leftarrow+1}(+1)-\nu_v(+1)\right|$, we construct the following coupling between $\pi_v^{u\leftarrow+1}$ and ν_v:

- sample a random value $c\in\{-1,+1\}$ according to the distribution ν_u;
- sample c_v,c'_v jointly according to the optimal coupling between $\pi_v^{u\leftarrow+1}$ and $\nu_v^{u\leftarrow c}$.

Note that $\left|\pi_v^{u\leftarrow+1}(+1)-\nu_v(+1)\right|$ is the total variation distance between $\pi_v^{u\leftarrow+1}$ and ν_v. By the coupling inequality, we have

$$\left|\pi_v^{u\leftarrow+1}(+1)-\nu_v(+1)\right|\leq\Pr\left[c_v\neq c'_v\right]=\sum_{c\in\Omega(\nu_u)}\nu_u(c)\cdot d_{\mathrm{TV}}\left(\pi_v^{u\leftarrow+1},\nu_v^{u\leftarrow c}\right).$$

Recall $k\geq 2$, thus $\boldsymbol{\phi},\boldsymbol{\varphi}$ and $\boldsymbol{\varphi}'$ are all positive local fields and $\{-1,+1\}=\Omega(\pi_{k,u_i})$. It holds that $\Omega(\mu_u)=\Omega(\pi_u)=\Omega(\nu_u)=\{-1,+1\}$. Note that $\boldsymbol{\varphi}$ and $\boldsymbol{\varphi}'$ only differs at u. It is straightforward to verify that for any $c\in\Omega(\nu_u)$, $\pi_v^{u\leftarrow c}=\nu_v^{u\leftarrow c}$. We have

$$\left|\pi_v^{u\leftarrow+1}(+1)-\nu_v(+1)\right|\leq\sum_{c\in\Omega(\nu_u)}\nu_u(c)\cdot d_{\mathrm{TV}}\left(\pi_v^{u\leftarrow+1},\pi_v^{u\leftarrow c}\right)\leq d_{\mathrm{TV}}\left(\pi_v^{u\leftarrow-1},\pi_v^{u\leftarrow+1}\right)=\left|\Psi_\pi^{\mathrm{inf}}(u,v)\right|.$$

Combining above inequality with (26) proves (18).

5.2. **Verification of** (19). Recall that in this case, we have $u = v$ and $i \neq j$. Without loss of generality, we assume $\Omega(\pi_{k,u_i}) = \{-1, +1\}$, since otherwise (19) holds trivially. In this case, we have

$$\left|\Psi^{\inf}_{\pi_k}(u_i, v_j)\right| = \left|\pi^{u_i \leftarrow +1}_{k,u_j}(+1) - \pi^{u_i \leftarrow -1}_{k,u_j}(+1)\right| = \pi^{u_i \leftarrow -1}_{k,u_j}(+1), \tag{27}$$

where the last equation holds because $\pi^{u_i \leftarrow +1}_{k,u_j}(+1) = 0$. By the definition of the conditional distribution,

$$\pi^{u_i \leftarrow -1}_{k,u_j}(+1) = \frac{\mathbf{Pr}_{Y \sim \pi_k}\left[Y_{u_i} = -1 \wedge Y_{u_j} = +1\right]}{\mathbf{Pr}_{Y \sim \pi_k}\left[Y_{u_i} = -1\right]}.$$

By the definition of π_k, we have

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{Y \sim \pi_k}\left[Y_{u_i} = -1\right] &= \frac{1}{Z_\phi} \sum_{X \in \Omega(\mu)} \mu(X) \left(\mathbb{1}[X_u = +1] \left(\sum_{h \in [k] \setminus \{i\}} \frac{\phi_{u_h}}{k} \right) + \mathbb{1}[X_u = -1] \right) \prod_{\substack{w \in V \setminus \{u\}: \\ X_w = +1}} \left(\sum_{h \in [k]} \frac{\phi_{w_h}}{k} \right) \\
&\geq \frac{1}{Z_\phi} \sum_{X \in \Omega(\mu)} \mu(X) \mathbb{1}[X_u = +1] \left(\sum_{h \in [k] \setminus \{i\}} \frac{\phi_{u_h}}{k} \right) \prod_{\substack{w \in V \setminus \{u\}: \\ X_w = +1}} \left(\sum_{h \in [k]} \frac{\phi_{w_h}}{k} \right) \\
&= \frac{1}{Z_\phi} \left(\sum_{h \in [k] \setminus \{i\}} \frac{\phi_{u_h}}{k} \right) \sum_{\substack{X \in \Omega(\mu): \\ X_u = 1}} \mu(X) \prod_{\substack{w \in V \setminus \{u\}: \\ X_w = +1}} \left(\sum_{h \in [k]} \frac{\phi_{w_h}}{k} \right) \\
&= \frac{1}{Z_\phi} \left(\frac{\sum_{h \in [k] \setminus \{i\}} \phi_{u_h}}{\phi_{u_j}} \right) \sum_{\substack{X \in \Omega(\mu): \\ X_u = 1}} \frac{\mu(X) \phi_{u_j}}{k} \prod_{\substack{w \in V \setminus \{u\}: \\ X_w = +1}} \left(\sum_{h \in [k]} \frac{\phi_{w_h}}{k} \right) \\
&= \frac{\sum_{h \in [k] \setminus \{i\}} \phi_{u_h}}{\phi_{u_j}} \mathop{\mathbf{Pr}}_{Y \sim \pi_k}\left[Y_{u_j} = +1\right] = \frac{\sum_{h \in [k] \setminus \{i\}} \phi_{u_h}}{\phi_{u_j}} \mathop{\mathbf{Pr}}_{Y \sim \pi_k}\left[Y_{u_i} = -1 \wedge Y_{u_j} = +1\right],
\end{aligned}$$

where the last equation holds because if $Y_{u_j} = +1$, then $Y_{u_i} = -1$. This implies that

$$\pi^{u_i \leftarrow -1}_{k,u_j}(+1) = \frac{\mathbf{Pr}_{Y \sim \pi_k}\left[Y_{u_i} = -1 \wedge Y_{u_j} = +1\right]}{\mathbf{Pr}_{Y \sim \pi_k}\left[Y_{u_i} = -1\right]} \leq \frac{\phi_{u_j}}{\sum_{h \in [k] \setminus \{i\}} \phi_{u_h}}. \tag{28}$$

6. Magnetized Block Factorization of Entropy

In this section, we prove Lemma 2.9. Let μ be a distribution over $\{-1, +1\}^V$. Let $\theta \in (0, 1)$ be a real number, recall that $\pi = \mu^{(\theta)}$ is a distribution and Z_π is the normalization factor defined in (3). Recall that we use $\mathrm{Bin}(V, 1 - \theta)$ to denote the distribution of subset $\mathcal{R} \subseteq V$ that is randomly generated by including each $v \in V$ into R independently with probability $1 - \theta$. Let $\eta > 0$. Suppose μ_k satisfies the $\lceil \theta k n \rceil$-uniform block factorization with constant C for all integers $k \geq k_0$. To prove Lemma 2.9, we need to show that for any $f : \{-1, +1\}^V \to \mathbb{R}_{\geq 0}$, it holds that

$$\mathbf{Ent}_\mu\left[f\right] \leq C \cdot \frac{Z_\pi}{\theta^{|V|}} \mathop{\mathbf{E}}_{R \sim \mathrm{Bin}(V, 1-\theta)} \left[\pi_R\left(\mathbb{1}_R\right) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}\left[f\right]\right],$$

where we assume $\pi_R\left(\mathbb{1}_R\right) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}\left[f\right] = 0$ if $\pi_R\left(\mathbb{1}_R\right) = 0$.

Fix an integer $k \geq k_0$ and a function $f : \{-1, +1\}^V \to \mathbb{R}_{\geq 0}$. We define a new function $f^k : \{-1, +1\}^{V_k} \to \mathbb{R}_{\geq 0}$ as

$$\forall \sigma \in \{-1, +1\}^{V_k}, \quad f^k(\sigma) = f(\sigma^\star).$$

Recall that

$$\forall v \in V, \quad \sigma^\star_v = \begin{cases} +1 & \exists i \in [k] \text{ s.t. } \sigma_{(v,i)} = +1; \\ -1 & \forall i \in [k], \sigma_{(v,i)} = -1. \end{cases}$$

Since μ_k satisfies the $\lceil \theta kn \rceil$-uniform block factorization with parameter C, by definition, we have

$$\mathbf{Ent}_{\mu_k}\left[f^k\right] \le \frac{C}{\binom{nk}{\lceil \theta kn \rceil}} \sum_{S \in \binom{V_k}{\lceil \theta kn \rceil}} \mu_k\left[\mathbf{Ent}_S\left[f^k\right]\right].$$

Note that above inequality holds for all $k \ge k_0$. Lemma 2.9 is a straightforward consequence of the following two results

$$\mathbf{Ent}_{\mu_k}\left[f^k\right] = \mathbf{Ent}_\mu\left[f\right]; \tag{29}$$

$$\lim_{k\to\infty} \frac{C}{\binom{nk}{\lceil \theta kn \rceil}} \sum_{S \in \binom{V_k}{\lceil \theta kn \rceil}} \mu_k\left[\mathbf{Ent}_S\left[f^k\right]\right] = C \cdot \frac{Z_\pi}{\theta^{|V|}} \mathop{\mathbf{E}}_{R\sim\mathrm{Bin}(V,1-\theta)}\left[\pi_R\left(\mathbb{1}_R\right) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}\left[f\right]\right]. \tag{30}$$

In the rest of section, we verify (29) and (30).

6.1. **Verification of (29).** By definition, it holds that

$$\mathbf{Ent}_{\mu_k}\left[f^k\right] = \sum_{\sigma\in\Omega(\mu_k)} \mu_k(\sigma) f^k(\sigma)\log f^k(\sigma) - \sum_{\sigma\in\Omega(\mu_k)} \mu_k(\sigma) f^k(\sigma)\log \sum_{\tau\in\Omega(\mu_k)} \mu_k(\tau) f^k(\tau).$$

Note that

$$\sum_{\sigma\in\Omega(\mu_k)} \mu_k(\sigma) f^k(\sigma)\log f^k(\sigma) = \sum_{\alpha\in\Omega(\mu)} \sum_{\sigma\in\Omega(\mu_k):\sigma^\star=\alpha} \mu_k(\sigma) f(\alpha)\log f(\alpha) = \sum_{\alpha\in\Omega(\mu)} \mu(\alpha) f(\alpha)\log f(\alpha);$$

$$\sum_{\sigma\in\Omega(\mu_k)} \mu_k(\sigma) f^k(\sigma) = \sum_{\alpha\in\Omega(\mu)} \sum_{\sigma\in\Omega(\mu_k):\sigma^\star=\alpha} \mu_k(\sigma) f(\alpha) = \sum_{\alpha\in\Omega(\mu)} \mu(\alpha) f(\alpha).$$

This implies the equation $\mathbf{Ent}_{\mu_k}\left[f^k\right] = \mathbf{Ent}_\mu\left[f\right]$.

6.2. **Verification of (30).** We introduce some notations. Let $\boldsymbol{\varphi} \in [0,1]^V$ be a vector. For any subset $\Lambda \subseteq V$, we use $\boldsymbol{\varphi}_\Lambda \in [0,1]^V$ denote the vector induced by restricting $\boldsymbol{\varphi}$ on Λ. Let $\mu^{(\boldsymbol{\varphi}_\Lambda)}$ denote the distribution obtained by imposing local fields $\boldsymbol{\varphi}_\Lambda$ on μ, formally

$$\forall \sigma \in \{-1,+1\}^V, \quad \mu^{(\boldsymbol{\varphi}_\Lambda)}(\sigma) = \frac{\mu(\sigma)}{Z(\boldsymbol{\varphi}_\Lambda)} \prod_{\substack{v\in\Lambda:\\ \sigma_v=+1}} \boldsymbol{\varphi}_\Lambda(v), \quad \text{where } Z(\boldsymbol{\varphi}_\Lambda) \triangleq \sum_{\tau\in\{-1,+1\}^V} \mu(\tau) \prod_{\substack{v\in\Lambda:\\ \tau_v=+1}} \boldsymbol{\varphi}_\Lambda(v).$$

The above definition can be viewed as a generalization of Definition 1.4. The local fields $\boldsymbol{\varphi}_\Lambda$ can only be defined on a subset Λ. It is easy to see $\mu^{(\boldsymbol{\varphi}_\Lambda)} = \mu^{(\boldsymbol{\varphi}_\Lambda \cup \mathbb{1}_{V\setminus\Lambda})}$. Besides, we allow $\boldsymbol{\varphi}_\Lambda(v) = 0$ for some vertices $v \in V$. The distribution $\mu^{(\boldsymbol{\varphi}_\Lambda)}$ is well-defined if and only if $Z(\boldsymbol{\varphi}_\Lambda) > 0$.

Fix an integers $k \ge 1$ and $1 \le \ell \le kn$. We have the following lemma.

Lemma 6.1. *Let $C > 0$ and $k \ge 1$ and $1 \le \ell \le kn$ be two integers. If μ_k satisfies the ℓ-uniform block factorization with constant C, then for any $f : \{-1,+1\}^V \to \mathbb{R}_{\ge 0}$, let*

$$H_f(\ell,k) \triangleq \frac{1}{\binom{nk}{\ell}} \sum_{S\in\binom{V_k}{\ell}} \mu_k\left[\mathbf{Ent}_S\left[f^k\right]\right],$$

it holds that

$$H_f(\ell,k) = \sum_{\boldsymbol{b}:k\boldsymbol{b}\in\Omega(\Pi_{V,k,\ell})} \Pi_{V,k,\ell}(k\boldsymbol{b}) \sum_{\tau\in\Omega(\mu)} \mu(\tau) \sum_{\substack{R\subseteq\tau^{-1}(+1):\\ \forall v\in\tau^{-1}(+1)\setminus R,\ b_v>0}} \prod_{u\in R}(1-b_u) \prod_{v\in\tau^{-1}(+1)\setminus R} b_v \cdot \mathbf{Ent}_{\mu^{(\boldsymbol{b}_{V\setminus R}),\mathbb{1}_R}}\left[f\right].$$

where $\mathbb{1}_R$ is the all-$(+1)$ configuration on R and $\tau^{-1}(+1) = \{v \in V \mid \tau_v = +1\}$, the distribution $\mu^{(\boldsymbol{b}_{V\setminus R}),\mathbb{1}_R}$ is obtained from $\mu^{(\boldsymbol{b}_{V\setminus R})}$ conditional on $\mathbb{1}_R$, and $\Pi_{V,k,\ell}(\cdot)$ is the multivariate hypergeometric distribution in (13).

The above lemma relates $H_f(\ell, k)$ to the multivariate hypergeometric distribution. By the concentration property, we have the following lemma.

Lemma 6.2. *Let $\theta \in (0,1)$ and $\pi = \mu^{(\theta)}$. Let $f : \{-1,+1\}^V \to \mathbb{R}_{\geq 0}$ be a function. It holds that*

$$\lim_{k\to\infty} H_f\left(\lceil \theta n k \rceil, k\right) = \frac{Z_\pi}{\theta^{|V|}} \mathop{\mathbf{E}}_{R\sim \mathrm{Bin}(V,1-\theta)} \left[\pi_R\left(\mathbb{1}_R\right) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}\left[f\right]\right].$$

Equation (30) can be verified by combining Lemma 6.1 and Lemma 6.2. We verify these two lemmas in the rest of this section.

6.3. **Proof of Lemma 6.1.** For any subset $\Lambda \subseteq V_k$, we use $\mu_{k,\Lambda}$ to denote the marginal distribution on Λ induced from μ_k. By definition, we have

$$\begin{aligned} H_f(\ell,k) &= \frac{1}{\binom{kn}{\ell}} \cdot \sum_{S\in\binom{V_k}{\ell}} \sum_{\tau\in\Omega(\mu_{k,V_k\setminus S})} \mu_{k,V_k\setminus S}(\tau)\cdot \mathbf{Ent}_{\mu_k^\tau}\left[f^k\right] \\ &= \frac{1}{\binom{kn}{\ell}} \cdot \sum_{\sigma\in\Omega(\mu_k)} \mu_k(\sigma) \sum_{S\in\binom{V_k}{\ell}} \mathbf{Ent}_{\mu_k^\rho}\left[f^k\right], \quad \text{where } \rho = \rho(\sigma, S) \triangleq \sigma_{V_k\setminus S}. \end{aligned}$$

Recall that $V_k = V \times [k]$. For each $v \in V$, define $C_v = \{v_1, v_2, \ldots, v_k\}$. For configuration $\sigma \in \Omega(\mu_k)$, define

$$F(\sigma) \triangleq \{v \in V \mid \exists\, i \in [k] \text{ s.t. } \sigma_{v_i} = +1\}.$$

Similarly, for partial configuration $\rho = \sigma_{V_k\setminus S}$, this definition could be extended as

$$F(\rho) \triangleq \{v \in V \mid \exists\, i \in [k] \text{ s.t. } v_i \in V_k \setminus S \wedge \sigma_{v_i} = +1\}.$$

Next, we define local fields $\boldsymbol{\phi}_\rho \in [0,1]^{V\setminus F(\rho)}$ such that

$$\forall v \in V \setminus F(\rho), \quad \phi_\rho(v) \triangleq \frac{|S \cap C_v|}{k}.$$

Claim 6.3. *It holds that* $\mathbf{Ent}_{\mu_k^\rho}\left[f^k\right] = \mathbf{Ent}_{\mu^{(\boldsymbol{\phi}_\rho),\mathbb{1}_{F(\rho)}}}[f]$.

By Claim 6.3, we have

$$H_f(\ell,k) = \frac{1}{\binom{kn}{\ell}} \cdot \sum_{\sigma\in\Omega(\mu_k)} \mu_k(\sigma) \sum_{S\in\binom{V_k}{\ell}} \mathbf{Ent}_{\mu^{(\boldsymbol{\phi}_\rho),\mathbb{1}_{F(\rho)}}}[f], \quad \text{where } \rho = \rho(\sigma, S) \triangleq \sigma_{V_k\setminus S}. \tag{31}$$

Note that S is sampled from $\binom{V_k}{\ell}$ uniformly at random. Let $\boldsymbol{a} = (a_v)_{v\in V}$, where $a_v = |S \cap C_v|$. Vector $\boldsymbol{a}$ follows the distribution multivariate hypergeometric distribution $\Pi_{V,k,\ell}$. We can use the following procedure to generate the random set S: sample $\boldsymbol{a} \sim \Pi_{V,k,\ell}$; sample $S_v \in \binom{C_v}{a_v}$ uniformly at random and let $S = \cup_{v\in V} S_v$. Note that $F(\rho) \subseteq F(\sigma)$. Let $\boldsymbol{b} = \boldsymbol{a}/k$, so that $\phi_\rho(v) = b_v$ for all $v \in V \setminus F(\rho)$. By letting $R = F(\rho)$, we rewrite equation (31) as

$$H_f(\ell,k) = \sum_{\boldsymbol{b}: k\boldsymbol{b}\in\Omega(\Pi_{V,k,\ell})} \Pi_{V,k,\ell}(k\boldsymbol{b}) \sum_{\sigma\in\Omega(\mu_k)} \mu_k(\sigma) \sum_{\substack{R\subseteq F(\sigma): \\ \forall v\in F(\sigma)\setminus R,\ b_v>0}} \prod_{u\in R}(1-b_u) \prod_{v\in F(\sigma)\setminus R} b_v \cdot \mathbf{Ent}_{\mu^{(\boldsymbol{b}_{V\setminus R}),\mathbb{1}_R}}[f].$$

Here, vector $k\boldsymbol{b}$ determines the size of each S_v. Then, for each $R \subseteq F(\sigma)$, the following part of above equation specifies the probability that $F(\rho) = R$:

- when $u \in R$, the $+1$-vertex in C_u should not be selected into S_u, this happens with probability $1 - b_u$;
- when $u \in F(\sigma) \setminus R$, the $+1$-vertex in C_u must be selected into S_u, so it happens with probability b_u.

If there exists $v \in F(\sigma) \setminus R$ such that $b_v = 0$, then by above argument, we know that $\mathbf{Pr}[R = F(\rho)] = 0$. Since it is no need to enumerate this kind of R, we add a constraint for R which says $\forall v \in F(\sigma) \setminus R,\ b_v > 0$. This constraint for enumerating R also ensures that $\mu^{(\boldsymbol{b}_{V\setminus R}),\mathbb{1}_R}$ will always be well defined.

To generate a random configuration $\sigma \sim \mu_k$, one can sample $\tau \sim \mu$, then transform τ to σ by Definition 2.4. It turns out that $F(\sigma)$ can be uniquely determined by τ, because $\tau(v) = +1$ if and only if $v \in F(\sigma)$. To enumerate all configurations $\sigma \in \Omega(\mu_k)$, we first enumerates all configurations $\tau \in \Omega(\mu)$, then enumerate all σ such that $\sigma^\star = \tau$. We have

$$H_f(\ell,k) = \sum_{\boldsymbol{b}:k\boldsymbol{b}\in\Omega(\Pi_{V,k,\ell})} \Pi_{V,k,\ell}(k\boldsymbol{b}) \sum_{\tau\in\Omega(\mu)} \sum_{\substack{R\subseteq\tau^{-1}(+1):\\ \forall v\in\tau^{-1}(+1)\setminus R,\ b_v>0}} \prod_{u\in R}(1-b_u) \prod_{v\in\tau^{-1}(+1)\setminus R} b_v \cdot \mathbf{Ent}_{\mu^{(\boldsymbol{b}_{V\setminus R}),\mathbb{1}_R}}[f] \cdot \sum_{\substack{\sigma\in\Omega(\mu_k):\\ \sigma^\star=\tau}} \mu_k(\sigma)$$

By definition, it is easy to verify that

$$\mu(\tau) = \sum_{\substack{\sigma\in\Omega(\mu_k):\\ \sigma^\star=\tau}} \mu_k(\sigma).$$

This implies that

$$H_f(\ell,k) = \sum_{\boldsymbol{b}:k\boldsymbol{b}\in\Omega(\Pi_{V,k,\ell})} \Pi_{V,k,\ell}(k\boldsymbol{b}) \sum_{\tau\in\Omega(\mu)} \mu(\tau) \sum_{\substack{R\subseteq\tau^{-1}(+1):\\ \forall v\in\tau^{-1}(+1)\setminus R,\ b_v>0}} \prod_{u\in R}(1-b_u) \prod_{v\in\tau^{-1}(+1)\setminus R} b_v \cdot \mathbf{Ent}_{\mu^{(\boldsymbol{b}_{V\setminus R}),\mathbb{1}_R}}[f].$$

This proves Lemma 6.1.

Proof of Claim 6.3. By [CFYZ21, Lemma 6.2], it is straightforward to verify

$$\forall \xi \in \Omega(\mu),\quad \mu^{(\boldsymbol{\phi}_\rho),\mathbb{1}_{F(\rho)}}(\xi) = \sum_{\tau\in\Omega(\mu_k):\tau^\star=\xi} \mu_k^\rho(\tau). \tag{32}$$

By definition, it holds that

$$\mathbf{Ent}_{\mu_k^\rho}\left[f^k\right] = \sum_{\sigma\in\Omega(\mu_k)} \mu_k^\rho(\sigma) f^k(\sigma)\log f^k(\sigma) - \sum_{\sigma\in\Omega(\mu_k)} \mu_k^\rho(\sigma) f^k(\sigma) \log \sum_{\tau\in\Omega(\mu_k)} \mu_k^\rho(\tau) f^k(\tau).$$

In above equation, we enumerate all $\sigma \in \Omega(\mu_k)$. Since $\Omega(\mu_k^\rho) \subseteq \Omega(\mu_k)$, the above equation is correct. By (32), it holds that

$$\begin{aligned}\sum_{\sigma\in\Omega(\mu_k)} \mu_k^\rho(\sigma) f^k(\sigma)\log f^k(\sigma) &= \sum_{\alpha\in\Omega(\mu)} \sum_{\sigma\in\Omega(\mu_k):\sigma^\star=\alpha} \mu_k^\rho(\sigma) f^k(\sigma)\log f^k(\sigma)\\ &= \sum_{\alpha\in\Omega(\mu)} f(\alpha)\log f(\alpha) \sum_{\sigma\in\Omega(\mu_k):\sigma^\star=\alpha} \mu_k^\rho(\sigma) = \sum_{\alpha\in\Omega(\mu)} \mu^{(\boldsymbol{\phi}_\rho),\mathbb{1}_{F(\rho)}}(\alpha) f(\alpha)\log f(\alpha)\end{aligned}$$

Similarly, it holds that

$$\sum_{\sigma\in\Omega(\mu_k)} \mu_k^\rho(\sigma) f^k(\sigma) = \sum_{\alpha\in\Omega(\mu)} \mu^{(\boldsymbol{\phi}_\rho),\mathbb{1}_{F(\rho)}}(\alpha) f(\alpha).$$

Note that $\Omega(\mu^{(\boldsymbol{\phi}_\rho),\mathbb{1}_{F(\rho)}}) \subseteq \Omega(\mu)$. This implies $\mathbf{Ent}_{\mu_k^\rho}\left[f^k\right] = \mathbf{Ent}_{\mu^{(\boldsymbol{\phi}_\rho),\mathbb{1}_{F(\rho)}}}[f]$. ☐

6.4. **Proof of Lemma 6.2**. Fix the distribution μ, parameters $\theta \in (0,1)$, and a function $f : \{-1,+1\}^V \to \mathbb{R}_{\geq 0}$. We define the following function $H : [0,1]^V \to \mathbb{R}$:

$$H(\boldsymbol{b}) = \sum_{\tau \in \Omega(\mu)} \mu(\tau) \sum_{\substack{R \subseteq \tau^{-1}(+1): \\ \forall v \in \tau^{-1}(+1) \setminus R,\ b_v > 0}} \prod_{u \in R} (1 - b_u) \prod_{v \in \tau^{-1}(+1) \setminus R} b_v \cdot \mathbf{Ent}_{\mu^{(\boldsymbol{b}_{V \setminus R}), \mathbb{1}_R}} [f].$$

Let $\ell = \lceil \theta n k \rceil$. We have

$$H_f(\ell, k) = \mathop{\mathbf{E}}_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} [H(\boldsymbol{b})].$$

We define a bad event $\mathcal{B}_{\delta,k}$ for vector $\boldsymbol{b}$ such that

$$\mathcal{B}_{\delta,k} : \exists v \in V, |b_v - \theta| \geq \delta.$$

Claim 6.4. *For any $\varepsilon > 0$, any $\delta > 0$, there exists $K > 0$ such that for all $k \geq K$, $\mathbf{Pr}_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} \left[\mathcal{B}_{\delta,k}\right] \leq \varepsilon$.*

Let $\boldsymbol{\theta}$ denote the constant vector with value θ. Note that $H(\boldsymbol{b})$ is a continuous function in $(0,1)^V$. Given any $\varepsilon > 0$, we can find $0 < \delta = \delta(\varepsilon) < \frac{\theta}{2}$ such that for any $\boldsymbol{b} \in (0,1)^V$ with $\|\boldsymbol{b} - \boldsymbol{\theta}\|_\infty < \delta$,

$$|H(\boldsymbol{b}) - H(\boldsymbol{\theta})| \leq \frac{\varepsilon}{2}.$$

Besides, for any vector $\boldsymbol{b} \in [0,1]^V$, we have

$$H(\boldsymbol{b}) \leq M = M(\mu, f),$$

where M depends only on μ and f. By Claim 6.4, there exists $K > 0$ such that for all $k \geq K$, it holds that $\mathbf{Pr}_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} \left[\mathcal{B}_{\delta,k}\right] \leq \frac{\varepsilon}{2\max\{H(\boldsymbol{\theta}), M\}}$. Hence for all $k \geq K$, we have

$$H_f(\ell,k) = \mathop{\mathbf{E}}_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} [H(\boldsymbol{b})] \geq \left(1 - \mathop{\mathbf{Pr}}_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} \left[\mathcal{B}_{\delta,k}\right]\right) \left(H(\boldsymbol{\theta}) - \frac{\varepsilon}{2}\right) \geq H(\boldsymbol{\theta}) - \varepsilon$$
$$H_f(\ell,k) = \mathop{\mathbf{E}}_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} [H(\boldsymbol{b})] \leq H(\boldsymbol{\theta}) + \frac{\varepsilon}{2} + M \mathop{\mathbf{Pr}}_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} \left[\mathcal{B}_{\delta,k}\right] \leq H(\boldsymbol{\theta}) + \varepsilon.$$

Let $\|\tau\|_+ = |\{v \in V \mid \tau_v = +1\}|$. This implies that

$$\begin{aligned}\lim_{k \to \infty} H_f\left(\lceil \theta n k \rceil, k\right) = H(\boldsymbol{\theta}) &= \sum_{\tau \in \Omega(\mu)} \mu(\tau) \sum_{R \subseteq \tau^{-1}(+1)} (1-\theta)^{|R|} \theta^{\|\tau\|_+ - |R|} \cdot \mathbf{Ent}_{\mu^{(\boldsymbol{\theta}_{V \setminus R}), \mathbb{1}_R}} [f] \\ &= \sum_{R \subseteq V : \mu_R(\mathbb{1}_R) > 0} (1-\theta)^{|R|} \theta^{-|R|} \mathbf{Ent}_{\pi^{\mathbb{1}_R}} [f] \sum_{\substack{\tau \in \Omega(\mu): \\ \tau_R = \mathbb{1}_R}} \mu(\tau) \theta^{\|\tau\|_+},\end{aligned}$$

where the last equation holds due to $\pi = \mu^{(\theta)}$ and $\mu^{(\boldsymbol{\theta}_{V \setminus R}), \mathbb{1}_R} = \pi^{\mathbb{1}_R}$, this is because π and $\mu^{(\boldsymbol{\theta}_{V \setminus R})}$ disagree only at the local fields on R and the configuration on R is fixed by $\mathbb{1}_R$. Furthermore, we have

$$\begin{aligned}\lim_{k \to \infty} H_f\left(\lceil \theta n k \rceil, k\right) &= \frac{Z_\pi}{\theta^{|V|}} \sum_{R \subseteq V : \mu_R(\mathbb{1}_R) > 0} (1-\theta)^R \theta^{|V| - |R|} \mathbf{Ent}_{\pi^{\mathbb{1}_R}} [f] \, \pi_R(\mathbb{1}_R) \\ &= \frac{Z_\pi}{\theta^{|V|}} \mathop{\mathbf{E}}_{R \sim \mathrm{Bin}(V, 1-\theta)} \left[\pi_R\left(\mathbb{1}_R\right) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}} [f]\right],\end{aligned}$$

where we assume $\pi_R\left(\mathbb{1}_R\right) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}} [f] = 0$ if $\pi_R\left(\mathbb{1}_R\right) = 0$.

Proof of Claim 6.4. Recall that $n = |V|$ and $\ell = \lceil \theta k n \rceil$. Observe that when $k \geq \frac{2}{\delta}$,

$$\begin{aligned}
\Pr_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} \left[\mathcal{B}_{\delta,k}\right] &= \Pr_{k\boldsymbol{b} \sim \Pi_{V,k,\ell_k}} \left[\exists v \in V, |b_v - \theta| \geq \delta\right] \\
&\leq \sum_{v \in V} \Pr_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} \left[|b_v - \theta| \geq \delta\right] \\
&\leq \sum_{v \in V} \Pr_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} \left[\left|b_v - \frac{\lceil \theta k n \rceil}{kn}\right| \geq \frac{\delta}{2}\right], \qquad (33)
\end{aligned}$$

where the last inequality holds for $k \geq \frac{2}{\delta}$.

Furthermore, by Lemma 3.7, there exists $K_0 = K_0(\delta, \varepsilon, n)$ such that for $k \geq K_0$,

$$\forall v \in V, \qquad \Pr_{k\boldsymbol{b} \sim \Pi_{V,k,\ell}} \left[\left|b_v - \frac{\lceil \theta k n \rceil}{kn}\right| \geq \frac{\delta}{2}\right] \leq 2\exp\left(\frac{-\delta^2 k}{2}\right) \leq \frac{\varepsilon}{n}, \qquad (34)$$

where the last inequality holds because k is sufficiently large. Combining (33) and (34) proves the claim. $\square$

7. Implication to Modified Log-Sobolev Constant

In this section, we prove Lemma 2.2. Let $P = P_\mu^{\mathrm{GD}}$ denote the Glauber dynamics on μ. We only need to prove the following lemma.

Lemma 7.1. *Let $\theta \in (0,1)$ and μ be a distribution over $\{-1,+1\}^V$. For any function $f : \{-1,+1\}^V \to \mathbb{R}_{\geq 0}$,*

$$\rho_{\min}^{\mathrm{GD}}(\pi) \cdot \frac{Z_\pi}{\theta^n} \mathop{\mathbf{E}}_{R \sim \mathrm{Bin}(V, 1-\theta)} \left[\pi_R\left(\mathbb{1}_R\right) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}\left[f\right]\right] \leq \mathcal{E}_P(f, \log f), \qquad (35)$$

where $n = |V|$, $\pi = \mu^{(\boldsymbol{\theta})}$, $\rho_{\min}^{\mathrm{GD}}(\pi)$ is defined in (2), and Z_π is defined in (3).

Lemma 2.2 is a straightforward consequence of Lemma 7.1. The rest of this section is dedicated to the proof of Lemma 7.1.

For convenience, we define the following notation. Suppose μ is a distribution over $\{-1,+1\}^V$ and $f : \{-1,+1\}^V \to \mathbb{R}_{\geq 0}$ a function, we define the *covariance* between f and $\log f$ with respect to μ as

$$\mathbf{Cov}_\mu\left(f, \log f\right) \triangleq \frac{1}{2} \sum_{\sigma,\tau \in \Omega(\mu)} \mu(\sigma)\mu(\tau)\left(f(\sigma) - f(\tau)\right)\left(\log f(\sigma) - \log f(\tau)\right).$$

Note that for any $f : \{-1,+1\}^V \to \mathbb{R}_{\geq 0}$, it holds that $\mathbf{Cov}_\mu\left(f, \log f\right) \geq 0$.

Let μ be a distribution over $\{-1,+1\}^V$ and let $C > 0$ be a constant. Let P denote the Glauber dynamics for μ. For function $f : \{-1,+1\}^V \to \mathbb{R}_{\geq 0}$, the dirichlet form satisfies

$$\begin{aligned}
\mathcal{E}_P(f, \log f) &= \frac{1}{2} \sum_{\sigma,\tau \in \Omega(\mu)} \mu(\sigma) P(\sigma,\tau)\left(f(\sigma) - f(\tau)\right)\left(\log f(\sigma) - \log f(\tau)\right) \\
&= \frac{1}{2} \sum_{\sigma \in \Omega(\mu)} \mu(\sigma) \frac{1}{n} \sum_{v \in V} \sum_{\tau \in \Omega(\mu^{\sigma_{V \setminus v}})} \mu^{\sigma_{V \setminus v}}(\tau)\left(f(\sigma) - f(\tau)\right)\left(\log f(\sigma) - \log f(\tau)\right) \\
&= \frac{1}{n} \sum_{v \in V} \sum_{\chi \in \Omega(\mu_{V \setminus \{v\}})} \mu_{V \setminus \{v\}}(\chi) \cdot \frac{1}{2} \sum_{\sigma,\tau \in \Omega(\mu^\rho)} \mu^\chi(\sigma)\mu^\chi(\tau)\left(f(\sigma) - f(\tau)\right)\left(\log f(\sigma) - \log f(\tau)\right) \\
&= \frac{1}{n} \sum_{v \in V} \sum_{\chi \in \Omega(\mu_{V \setminus \{v\}})} \mu_{V \setminus \{v\}}(\chi) \cdot \mathbf{Cov}_{\mu^\chi}\left(f, \log f\right) = \frac{1}{n} \sum_{v \in V} \mu[\mathbf{Cov}_v\left(f, \log f\right)], \qquad (36)
\end{aligned}$$

where we use the following notation

$$\mu[\mathbf{Cov}_v\left(f, \log f\right)] \triangleq \sum_{\chi \in \Omega(\mu_{V \setminus \{v\}})} \mu_{V \setminus \{v\}}(\chi) \cdot \mathbf{Cov}_{\mu^\chi}\left(f, \log f\right).$$

Recall that we say P satisfies the *modified log-Sobolev inequality* with constant C if for any $f : \{-1,+1\}^V \to \mathbb{R}_{\geq 0}$, it holds that

$$C \cdot \mathbf{Ent}_\mu [f] \leq \mathcal{E}_P(f, \log f) = \frac{1}{n} \sum_{v \in V} \mu[\mathbf{Cov}_v (f, \log f)]. \tag{37}$$

Now, we are ready to prove Lemma 7.1 Consider the modified log-Sobolev inequality for the Glauber dynamics on $\pi^{\mathbb{1}_R}$. By (37), for all $R \subseteq V$ with $\pi_R(\mathbb{1}_R) > 0$, the Glauber dynamics on $\pi^{\mathbb{1}_R}$ satisfies the modified log-Sobolev inequality with constant $\rho^{\mathrm{GD}}(\pi^{\mathbb{1}_R})$:

$$\forall f \in \mathbb{R}^{\Omega(\mu)}, \quad \rho^{\mathrm{GD}}_{\min}(\pi) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}} [f] \leq \rho^{\mathrm{GD}}(\pi^{\mathbb{1}_R}) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}} [f] \leq \frac{1}{n} \sum_{v \in V} \pi^{\mathbb{1}_R}[\mathbf{Cov}_v (f, \log f)],$$

where $\rho^{\mathrm{GD}}_{\min}(\pi) \leq \rho^{\mathrm{GD}}(\pi^\sigma)$ for all feasible partial configuration σ, as defined in (2). Recall that $\mathrm{Bin}(V, 1-\theta)$ denotes the distribution of subset $R \subseteq V$ that is randomly generated by including each $v \in V$ into R independently with probability $1-\theta$. The LHS of (35) can be upper bounded as follows

$$\begin{aligned} &\rho^{\mathrm{GD}}_{\min}(\pi) \cdot \frac{Z_\pi}{\theta^{|V|}} \mathop{\mathbf{E}}_{R \sim \mathrm{Bin}(V,1-\theta)} \left[\pi_R(\mathbb{1}_R) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}} [f]\right] \\ \leq\ & \frac{Z_\pi}{\theta^{|V|}} \cdot \frac{1}{n} \sum_{v \in V} \mathop{\mathbf{E}}_{R \sim \mathrm{Bin}(V,1-\theta)} \left[\pi_R(\mathbb{1}_R) \cdot \pi^{\mathbb{1}_R}[\mathbf{Cov}_v (f, \log f)]\right]. \end{aligned} \tag{38}$$

Let $I[\cdot]$ denote the indicator variable. The following identity holds for all $f \in \mathbb{R}^{\Omega(\mu)}$:

$$\begin{aligned} \mathop{\mathbf{E}}_{R \sim \mathrm{Bin}(V,1-\theta)} \left[\pi_R(\mathbb{1}_R) \cdot \pi^{\mathbb{1}_R}[\mathbf{Cov}_v (f, \log f)]\right] &\overset{(\star)}{=} \mathop{\mathbf{E}}_{R \sim \mathrm{Bin}(V,1-\theta)} \left[\mathop{\mathbf{E}}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[I\left[R \subseteq \sigma^{-1}(+1)\right] \cdot \mathbf{Cov}_{\pi^\sigma} (f, \log f)\right]\right] \\ &= \mathop{\mathbf{E}}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[\mathop{\mathbf{Pr}}_{R \sim \mathrm{Bin}(V,1-\theta)} \left[R \subseteq \sigma^{-1}(+1)\right] \cdot \mathbf{Cov}_{\pi^\sigma} (f, \log f)\right] \\ &= \mathop{\mathbf{E}}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[\theta^{|V| - \|\sigma\|_+} \cdot \mathbf{Cov}_{\pi^\sigma} (f, \log f)\right], \end{aligned} \tag{39}$$

where the nontrivial equation $(\star)$ holds by verifying for every choice of $v \in V$ and $R \subseteq V$ as follows:

- For the case that $\pi_R(\mathbb{1}_R) > 0$ and $v \notin R$, it holds that
$$\pi_R(\mathbb{1}_R) = \mathop{\mathbf{Pr}}_{\sigma \sim \pi} \left[R \subseteq \sigma^{-1}(+1)\right] = \mathop{\mathbf{Pr}}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[R \subseteq \sigma^{-1}(+1)\right],$$
and $\pi^{\mathbb{1}_R}[\mathbf{Cov}_v (f, \log f)]$ is well-defined, such that
$$\pi^{\mathbb{1}_R}[\mathbf{Cov}_v (f, \log f)] = \mathop{\mathbf{E}}_{\sigma \sim \pi^{\mathbb{1}_R}_{V \setminus \{v\}}} [\mathbf{Cov}_{\pi^\sigma} (f, \log f)] = \mathop{\mathbf{E}}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[\mathbf{Cov}_{\pi^\sigma} (f, \log f) \mid R \subseteq \sigma^{-1}(+1)\right].$$
Therefore,
$$\begin{aligned} \pi_R(\mathbb{1}_R) \cdot \pi^{\mathbb{1}_R}[\mathbf{Cov}_v (f, \log f)] &= \mathop{\mathbf{Pr}}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[R \subseteq \sigma^{-1}(+1)\right] \cdot \mathop{\mathbf{E}}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[\mathbf{Cov}_{\pi^\sigma} (f, \log f) \mid R \subseteq \sigma^{-1}(+1)\right] \\ &= \mathop{\mathbf{E}}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[I\left[R \subseteq \sigma^{-1}(+1)\right] \cdot \mathbf{Cov}_{\pi^\sigma} (f, \log f)\right]. \end{aligned}$$
- For the case that $\pi_R(\mathbb{1}_R) = 0$ or $v \in R$, both sides are 0. On the left-hand-side, if $\pi_R(\mathbb{1}_R) = 0$, then by convention
$$\pi_R(\mathbb{1}_R) \cdot \mathbf{Cov}_{\pi^{\mathbb{1}_R}} (f, \log f) = 0;$$
or else, if $\pi_R(\mathbb{1}_R) > 0$ but $v \in R$, then $\pi^{\mathbb{1}_R}[\mathbf{Cov}_v (f, \log f)]$ is well-defined, but for $\sigma \sim \pi^{\mathbb{1}_R}_{V \setminus \{v\}}$, the $\mathbb{1}_R \uplus \sigma$ gives a configuration fully specified on V and hence $|\Omega(\pi^{\mathbb{1}_R \uplus \sigma})| = 1$. In this case, by definition, the covariance must be 0, i.e.
$$\pi^{\mathbb{1}_R}[\mathbf{Cov}_v (f, \log f)] = \mathop{\mathbf{E}}_{\sigma \sim \pi^{\mathbb{1}_R}_{V \setminus \{v\}}} \left[\mathbf{Cov}_{\pi^{\mathbb{1}_R \uplus \sigma}} (f, \log f)\right] = 0.$$

On the right-hand-side, if $\pi_R(\mathbb{1}_R) = 0$ or $v \in R$, then for $\sigma \sim \pi_{V\setminus\{v\}}$, the event $R \subseteq \sigma^{-1}(+1)$ can never occur, and hence

$$\mathop{\mathbf{E}}_{\sigma\sim\pi_{V\setminus\{v\}}}\left[I\left[R\subseteq\sigma^{-1}(+1)\right]\cdot\mathbf{Cov}_{\pi^\sigma}\left(f,\log f\right)\right]=0.$$

This gives the equation ($\star$) in (39). Meanwhile, the other two equations in (39) follows respectively from linearity of expectation and the fact that $\mathbf{Pr}_{R\subseteq V}\left[R\subseteq\Lambda\right]=\theta^{|V|-|\Lambda|}$ for all $\Lambda\subseteq V$.

Furthermore, it can be verified that

$$\begin{aligned}
&\mathop{\mathbf{E}}_{\sigma\sim\pi_{V\setminus\{v\}}}\left[\frac{1}{\theta^{\|\sigma\|_+}}\mathbf{Cov}_{\pi^\sigma}\left(f,\log f\right)\right]\\
&=\sum_{\sigma\in\Omega(\pi_{V\setminus\{v\}})}\frac{1}{\theta^{\|\sigma\|_+}}\pi_{V\setminus\{v\}}(\sigma)\pi_v^\sigma(-1)\pi_v^\sigma(+1)(f(\sigma_+^v)-f(\sigma_-^v))\left(\log f(\sigma_+^v)-\log f(\sigma_-^v)\right)\\
&=\frac{1}{Z_\pi}\sum_{\sigma\in\Omega(\pi_{V\setminus\{v\}})}\mu_{V\setminus\{v\}}(\sigma)\mu_v^\sigma(-1)\pi_v^\sigma(+1)(f(\sigma_+^v)-f(\sigma_-^v))\left(\log f(\sigma_+^v)-\log f(\sigma_-^v)\right)\\
&\le\frac{1}{Z_\pi}\sum_{\sigma\in\Omega(\pi_{V\setminus\{v\}})}\mu_{V\setminus\{v\}}(\sigma)\mu_v^\sigma(-1)\mu_v^\sigma(+1)(f(\sigma_+^v)-f(\sigma_-^v))\left(\log f(\sigma_+^v)-\log f(\sigma_-^v)\right)\\
&=\frac{1}{Z_\pi}\mu[\mathbf{Cov}_v\left(f,\log f\right)], \qquad (40)
\end{aligned}$$

where $\sigma_\pm^v\in\{-1,+1\}^V$ denote the configurations on V where $\sigma_\pm^v(V\setminus\{v\})=\sigma(V\setminus\{v\})$ and $\sigma_\pm^v(v)=\pm1$, the second equation is due to the chain rule and the relation between π and μ in (3), and the last inequality is due to $\Omega(\pi)=\Omega(\mu)$, $(f(\sigma_+^v)-f(\sigma_-^v))\left(\log f(\sigma_+^v)-\log f(\sigma_-^v)\right)\ge0$ and the relaxation

$$\pi_v^\sigma(+1)\le\mu_v^\sigma(+1), \qquad (41)$$

which holds because $\pi=\mu^{(\theta)}$ is obtained by biasing every variable with a local field $\theta\in(0,1)$. Indeed,

$$\pi_v^\sigma(+1)=\frac{\pi(\sigma_+)}{\pi(\sigma_-)+\pi(\sigma_+)}=\frac{\theta\mu(\sigma_+)}{\mu(\sigma_-)+\theta\mu(\sigma_+)}\le\frac{\mu(\sigma_+)}{\mu(\sigma_-)+\mu(\sigma_+)}=\mu_v^\sigma(+1),$$

where the inequality holds for $\theta\in(0,1)$.

Combining (38), (39), and (40), we have

$$\rho_{\min}^{\mathrm{GD}}(\pi)\cdot\frac{Z_\pi}{\theta^{|V|}}\mathop{\mathbf{E}}_{R\sim\mathrm{Bin}(V,1-\theta)}\left[\pi_R(\mathbb{1}_R)\cdot\mathbf{Ent}_{\pi^{\mathbb{1}_R}}\left[f\right]\right]\le\frac{1}{n}\sum_{v\in V}\mu[\mathbf{Cov}_v\left(f,\log f\right)]=\mathcal{E}_P(f,\log f),$$

where P denotes the Glauber dynamics for μ. This proves Lemma 7.1.

8. Application in Ising Model

In this section, we prove Theorem 1.1. Let $G=(V,E)$ be a graph, $\beta\in\mathbb{R}_{>0}$ and $\boldsymbol{\lambda}=(\lambda_v)_{v\in V}\in\mathbb{R}_{>0}^V$. We use the notation $\mathcal{I}=(G,\beta,\boldsymbol{\lambda})$ to denote an Ising model and $\mu=\mu_{\mathcal{I}}$ to denote the Gibbs distribution defined by $\mathcal{I}$.

In order to apply Theorem 1.6, we introduce the flipping operation in [CFYZ21].

Definition 8.1. Let μ be a distribution over $\{-1,+1\}^V$, $\boldsymbol{\chi}\in\{-1,+1\}^V$ be a direction vector. The flipped distribution $\nu=\mathrm{flip}(\mu,\boldsymbol{\chi})$ over $\{-1,+1\}^V$ is defined as

$$\forall\sigma\in\{-1,+1\}^V,\quad\nu(\sigma)=\mu(\sigma\odot\boldsymbol{\chi}),$$

where $(\sigma\odot\boldsymbol{\chi})_v=\sigma_v\chi_v$ for all $v\in V$.

The following fact is straightforward to verify.

Fact 8.2. *For any distribution μ over $\{-1,+1\}^V$, $\boldsymbol{\chi}\in\{-1,+1\}^V$ and $\nu=\mathrm{flip}(\mu,\boldsymbol{\chi})$, it holds that*

(1) $\rho^{\mathrm{GD}}(\nu) = \rho^{\mathrm{GD}}(\mu)$, $\rho^{\mathrm{GD}}_{\min}(\nu) = \rho^{\mathrm{GD}}_{\min}(\mu)$*, where* $\rho^{\mathrm{GD}}(\cdot)$ *is the modified log-Sobolev constant of Glauber dynamics, and* $\rho^{\mathrm{GD}}_{\min}(\cdot)$ *is defined in* (2)*;*

(2) for any $\boldsymbol{\phi} \in \mathbb{R}^V_{>0}$, $\nu^{(\boldsymbol{\phi})} = \mathrm{flip}(\mu^{(\boldsymbol{\phi}^{\boldsymbol{\chi}})}, \boldsymbol{\chi})$*, where* $\boldsymbol{\phi}^{\boldsymbol{\chi}} \in \mathbb{R}^V_{>0}$ *satisfying* $(\boldsymbol{\phi}^{\boldsymbol{\chi}})_v = \phi_v^{\chi_v}$ *for all* $v \in V$*;*

(3) $\sup_{\boldsymbol{\phi}\in\mathbb{R}^V_{>0}} \left\|\Psi^{\mathrm{Inf}}_{\nu^{(\boldsymbol{\phi})}}\right\|_\infty = \sup_{\boldsymbol{\phi}\in\mathbb{R}^V_{>0}} \left\|\Psi^{\mathrm{Inf}}_{\mu^{(\boldsymbol{\phi})}}\right\|_\infty$.

Given an Ising model $\mathcal{I} = (G, \beta, \boldsymbol{\lambda})$, let $\boldsymbol{\chi}_{\mathcal{I}} \in \{-1, +1\}^V$ be the direction vector defined as follows:

$$\forall v \in V, \quad \chi_v = \begin{cases} +1, & \lambda_v \geq 1 \\ -1, & \lambda_v < 1 \end{cases} \tag{42}$$

We prove that the flipped distribution $\nu = \mathrm{flip}(\mu, \boldsymbol{\chi})$ is spectrally independent with all fields, and with proper local fields, the MLS constant is easy-to-analyze.

Lemma 8.3. *Let* $\delta, C \in (0,1)$*. For any graph* $G = (V, E)$*, any* $\beta \in \mathbb{R}_{>0}$ *and any* $\boldsymbol{\lambda} = (\lambda_v)_{v\in V} \in \mathbb{R}^V_{>0}$*, if* $\beta \in \left[\frac{\Delta-2+\delta}{\Delta-\delta}, \frac{\Delta-\delta}{\Delta-2+\delta}\right]$*, where* $\Delta \geq 3$ *is the maximum degree of* G*, let* $\mu = \mu_{\mathcal{I}}$ *denote the Gibbs distribution defined by the Ising model* $\mathcal{I} = (G, \beta, \boldsymbol{\lambda})$*, the flipped distribution* $\nu = \mathrm{flip}(\mu, \boldsymbol{\chi}_{\mathcal{I}})$ *satisfies*

(1) $\sup_{\boldsymbol{\phi}\in\mathbb{R}^V_{>0}} \left\|\Psi^{\mathrm{Inf}}_{\nu^{(\boldsymbol{\phi})}}\right\|_\infty = \sup_{\boldsymbol{\phi}\in\mathbb{R}^V_{>0}} \left\|\Psi^{\mathrm{Inf}}_{\mu^{(\boldsymbol{\phi})}}\right\|_\infty \leq \frac{2}{\delta}$*;*

(2) If $\min\left(\lambda_v, \frac{1}{\lambda_v}\right) \geq C$ *for all* $v \in V$*, then* $\rho^{\mathrm{GD}}_{\min}(\nu^{(\frac{1}{500})}) \geq \frac{10^{-6}C}{n}$.

We are ready to prove Theorem 1.1.

Proof of Theorem 1.1. Fix an Ising model $\mathcal{I} = (G, \beta, \boldsymbol{\lambda})$ with Gibbs distribution μ. Consider the Glauber dynamics on μ. Note that if $\min\left(\lambda_v, \frac{1}{\lambda_v}\right) \leq \frac{1}{500}$ for all $v \in V$, then the classic path coupling [BD97] yields an $O(n \log n)$ mixing time. Therefore, we may assume that there exists a vertex $v^\star \in V$ satisfying $\min\left(\lambda_{v^\star}, \frac{1}{\lambda_{v^\star}}\right) \geq \frac{1}{500}$, which implies $\min\left(\lambda_v, \frac{1}{\lambda_v}\right) \geq \frac{\lambda_{\min}}{500\lambda_{\max}}$ for all $v \in V$ and $\lambda_{\min} \leq 500$, $\lambda_{\max} \geq \frac{1}{500}$.

Let $\nu = \mathrm{flip}(\mu, \boldsymbol{\chi}_{\mathcal{I}})$, where $\boldsymbol{\chi}_{\mathcal{I}}$ is defined in (42). The first part of Fact 8.2 shows that ν is $\frac{2}{\delta}$-spectrally independent with all fields. The second part shows that $\rho^{\mathrm{GD}}_{\min}(\nu^{(\frac{1}{500})}) \geq \frac{10^{-6}C}{n}$ where $C = \frac{\lambda_{\min}}{500\lambda_{\max}}$. Using Theorem 1.6 with $\theta = \frac{1}{500}$, the Glauber dynamics on ν has MLS constant

$$\rho^{\mathrm{GD}}(\nu) \geq \left(\frac{10^{-\frac{12}{\delta}-30}\lambda_{\min}}{\lambda_{\max}}\right)\frac{1}{n}.$$

Furthermore, since $\lambda_{\min} \leq 500$ and $\lambda_{\max} \geq \frac{1}{500}$, we have

$$\nu_{\min} = \mu_{\min} \geq \left(\min_{v\in V}\min_{0\leq s\leq\Delta}\min\left(\frac{\lambda_v\beta^s}{\lambda_v\beta^s + \beta^{\deg(v)-s}}, \frac{\beta^{\deg(v)-s}}{\lambda_v\beta^s + \beta^{\deg(v)-s}}\right)\right)^n \geq \left(\frac{\lambda_{\min}}{14000\lambda_{\max}}\right)^n.$$

By Fact 8.2, $\rho^{\mathrm{GD}}(\nu) = \rho^{\mathrm{GD}}(\mu)$. Note that the transition matrix of Glauber dynamics has non-negative eigenvalues [DGU14, AL20]. By (9), the mixing time of the Glauber dynamics can be bounded by

$$\begin{aligned} T_{\mathrm{mix}}(\varepsilon) &\leq \frac{1}{\rho^{\mathrm{GD}}(\mu)}\left(\log\log\frac{1}{\mu_{\min}} + \log\frac{1}{2\varepsilon^2}\right) \leq \frac{\lambda_{\max}}{\lambda_{\min}}10^{30+\frac{12}{\delta}}n\left(\log n + \log\log\frac{14000\lambda_{\max}}{\lambda_{\min}} + \log\frac{1}{2\varepsilon^2}\right) \\ &= \exp\left(O(1/\delta)\right)\frac{\lambda_{\max}}{\lambda_{\min}}n\left(\log\frac{n}{\varepsilon} + \log\frac{2\lambda_{\max}}{\lambda_{\min}}\right). \end{aligned}$$

This proves the theorem. □

We need following two lemmas to prove Lemma 8.3.

Lemma 8.4 ([CLV21a]). *Let μ be the Gibbs distribution of Ising model on graph $G = (V, E)$ with maximum degree $\Delta \geq 3$, edge activity β and local fields $\boldsymbol{\lambda} = (\lambda_v)_{v\in V}$. If there exists $\alpha \in (0, 1)$ such that*

$$h(y) \triangleq \left|\frac{(1-\beta^2)e^y}{(\beta e^y + 1)(\beta + e^y)}\right| \leq \frac{1-\alpha}{\Delta - 1}, \tag{43}$$

then it holds that $\|\Psi_\mu^{\inf}\|_\infty \leq \frac{2}{\alpha}$.

Lemma 8.5 ([Mar19, SS20]). *Let μ be a distribution with support $\{-1,+1\}^V$ and P_μ^{GD} be the Glauber dynamics on μ. Let $\alpha = \min_{v\in V}\min_{\sigma\in\{-1,+1\}^V}\mu_v^{\sigma_{V\setminus\{v\}}}(\sigma_v)$ denote the marginal lower bound. Furthermore, denote the Dobrushin's influence matrix of μ by $A = (A_{u,v})_{u,v\in V}$, which is defined as*

$$A_{u,v} = \begin{cases} \max_{(\sigma,\tau)\in S(u,v)} d_{\mathrm{TV}}\left(\mu_v^\sigma, \mu_v^\tau\right) & \text{if } u \neq v \\ 0 & \text{if } u = v, \end{cases}$$

where $S(u, v)$ contains all pairs of configurations $\sigma, \tau \in \{-1,+1\}^{V\setminus\{v\}}$ that disagree only at vertex u. Let $|V| = n$. If $\|A\|_2 < 1$, then the Glauber dynamics on μ has modified log-Sobolev constant

$$\rho^{\mathrm{GD}}(\mu) \geq \frac{\alpha(1-\|A\|_2)^2}{2n}.$$

We are now ready to present the proof of Lemma 8.3

Proof of Lemma 8.3. For any $\boldsymbol{\lambda} \in \mathbb{R}^V_{>0}$, note that when $\beta \in \left[\frac{\Delta-2+\delta}{\Delta-\delta}, \frac{\Delta-\delta}{\Delta-2+\delta}\right]$,

$$\left|\frac{(1-\beta^2)e^y}{(\beta e^y + 1)(\beta + e^y)}\right| \leq \frac{|1-\beta|}{1+\beta} \leq \frac{1-\delta}{\Delta-1}.$$

Together with Lemma 8.4 and Fact 8.2, we have $\sup_{\boldsymbol{\phi}\in\mathbb{R}^V_{>0}}\|\Psi^{\inf}_{\nu(\boldsymbol{\phi})}\|_\infty = \sup_{\boldsymbol{\phi}\in\mathbb{R}^V_{>0}}\|\Psi^{\inf}_{\mu(\boldsymbol{\phi})}\|_\infty \leq \frac{2}{\delta}$.

Let $\boldsymbol{\chi} = \boldsymbol{\chi}_{\mathcal{I}}$. Recall that $\nu^{\left(\frac{1}{500}\right)} = \mathrm{flip}\left(\mu^{\left(\frac{1}{500}\right)^{\boldsymbol{\chi}}}, \boldsymbol{\chi}\right)$ and $\rho^{\mathrm{GD}}_{\min}\left(\nu^{\left(\frac{1}{500}\right)}\right) = \rho^{\mathrm{GD}}_{\min}\left(\mu^{\left(\frac{1}{500}\right)^{\boldsymbol{\chi}}}\right)$. Let $\pi \triangleq \mu^{\left(\frac{1}{500}\right)^{\boldsymbol{\chi}}}$. Note that π is the Gibbs distribution of the Ising model with external field $\boldsymbol{\lambda}^\star \in \mathbb{R}^V_{>0}$, where

$$\lambda_v^\star = \begin{cases} 500\lambda_v & \lambda_v \geq 1; \\ \frac{\lambda_v}{500}, & \lambda_v < 1. \end{cases}$$

It suffices to prove that

$$\rho^{\mathrm{GD}}_{\min}(\pi) \geq \frac{10^{-6}C}{n}.$$

Fix a subset $\Lambda \subseteq V$ of size $m = |V|$. Fix a partial configuration $\tau \in \{-1,+1\}^{V\setminus\Lambda}$. We analyze the MLS constant for the Glauber dynamics $P^{\mathrm{GD}}_{\pi^\tau}$ on π^τ. Note that π^τ is a distribution over $\{-1,+1\}^V$, where the configuration on $V \setminus \Lambda$ is fixed to τ. We consider the distribution π^τ_Λ and the Glauber dynamics $P^{\mathrm{GD}}_{\pi^\tau_\Lambda}$. The following relation of MLS constants is straightforward to verify

$$\rho^{\mathrm{GD}}(\pi^\tau) = \frac{m}{n}\rho^{\mathrm{GD}}(\pi^\tau_\Lambda). \tag{44}$$

We use Lemma 8.5 to analyze the MLS constant for Glauber dynamics on π^τ_Λ. Note that $\frac{C}{500} \leq \min\left(\lambda^\star_v, \frac{1}{\lambda^\star_v}\right) \leq \frac{1}{500}$ holds for all $v \in V$. Let α be the marginal lower bound for π^τ_Λ. It holds that

$$\begin{aligned}
\alpha &\geq \min_{v\in V} \min_{\sigma\in\{-1,+1\}^V} \pi_v^{\sigma_{V\setminus\{v\}}}(\sigma_v)\\
&\geq \min_{v\in V}\min_{0\leq s\leq d\leq\Delta}\min\left(\frac{\lambda^\star_v\beta^s}{\lambda^\star_v\beta^s+\beta^{d-s}}, \frac{\beta^{d-s}}{\lambda^\star_v\beta^s+\beta^{d-s}}\right)\\
&\overset{(*)}{\geq} \min\left(\frac{\lambda^\star_v}{\lambda^\star_v+27}, \frac{1}{27\lambda^\star_v+1}\right)\\
&\geq \frac{C}{2\times 10^4},
\end{aligned}$$

where $(*)$ follows from $\max_{-\Delta\leq s\leq\Delta}\beta^s \leq \left(1+\frac{2}{\Delta-2}\right)^\Delta \leq 27$ when $\Delta \geq 3$. Moreover, let A denote the Dobrushin's influence matrix for π^τ_Λ. It holds that

$$\begin{aligned}
\max\left(\|A\|_1, \|A\|_\infty\right) &\leq \Delta \sup_{\substack{\lambda\in\mathbb{R}_{>0}\\ \min(\lambda,\frac{1}{\lambda})\leq\frac{1}{500}}} \max_{0\leq s<d\leq\Delta}\left|\frac{\lambda\beta^s}{\lambda\beta^s+\beta^{d-s}} - \frac{\lambda\beta^{s+1}}{\lambda\beta^{s+1}+\beta^{d-s-1}}\right|\\
&= \Delta \sup_{\substack{\lambda\in\mathbb{R}_{>0}\\ \min(\lambda,\frac{1}{\lambda})\leq\frac{1}{500}}} \max_{0\leq s<d\leq\Delta}\left|\frac{\lambda\beta^{d-1}(1-\beta^2)}{\lambda\beta^{d-1}(\beta^2+1)+\lambda^2\beta^{2s+1}+\beta^{2d-2s-1}}\right|\\
&\leq \sup_{\substack{\lambda\in\mathbb{R}_{>0}\\ \min(\lambda,\frac{1}{\lambda})\leq\frac{1}{500}}} \max_{0\leq s<d\leq\Delta}\min\left(\frac{\lambda\Delta\left|1-\beta^2\right|}{\beta^{d-2s}}, \frac{\Delta\left|1-\beta^2\right|}{\lambda\beta^{2s-d+2}}\right)\\
&\overset{(*)}{\leq} \frac{3}{5},
\end{aligned}$$

where $(*)$ follows from the fact that, for any $\lambda\in\mathbb{R}_{>0}$ satisfying $\min\left(\lambda,\frac{1}{\lambda}\right)\leq\frac{1}{500}$,

$$\max_{0\leq s<d\leq\Delta}\min\left(\frac{\lambda\Delta\left|1-\beta^2\right|}{\beta^{d-2s}}, \frac{\Delta\left|1-\beta^2\right|}{\lambda\beta^{2s-d+2}}\right) \leq \begin{cases}\min\left(\lambda,\frac{1}{\lambda}\right)\Delta\left(1-\beta^2\right)\beta^{-\Delta} \leq 300\min\left(\lambda,\frac{1}{\lambda}\right) < \frac{3}{5} & ,\beta\in\left(\frac{\Delta-2}{\Delta},1\right]\\ \min\left(\lambda,\frac{1}{\lambda}\right)\Delta\left(\beta^2-1\right)\beta^{\Delta-2} \leq 300\min\left(\lambda,\frac{1}{\lambda}\right) < \frac{3}{5} & ,\beta\in\left[1,\frac{\Delta}{\Delta-2}\right).\end{cases}$$

Hence, $\|A\|_2 \leq \sqrt{\|A\|_1\|A\|_\infty} \leq \frac{3}{5}$. Combining Lemma 8.5 and (44), we have

$$\rho^{\mathrm{GD}}(\pi^\tau) = \frac{m}{n}\rho^{\mathrm{GD}}(\pi^\tau_\Lambda) \geq \frac{m}{n}\cdot\frac{C\times(2/5)^2}{2\times 10^4\times 2m} \geq \frac{C}{10^6 n}.$$

The above inequality holds for any τ, which implies $\rho^{\mathrm{GD}}_{\min}\left(\nu^{\left(\frac{1}{500}\right)}\right) = \rho^{\mathrm{GD}}_{\min}(\pi) \geq \frac{10^{-6}C}{n}$. □

9. Acknowledgements

Weiming Feng is supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 947778).

We would like to thank Heng Guo for helpful discussions.

References

[AASV21] Yeganeh Alimohammadi, Nima Anari, Kirankumar Shiragur, and Thuy-Duong Vuong. Fractionally log-concave and sector-stable polynomials: counting planar matchings and more. In *STOC*, pages 433–446. ACM, 2021.

[AJK+21] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence in high-dimensional expanders: Modified log-sobolev inequalities for fractionally log-concave polynomials and the ising model. *arXiv preprint arXiv:2106.04105*, 2021.

[AL20] Vedat Levi Alev and Lap Chi Lau. Improved analysis of higher order random walks and applications. In *STOC*, pages 1198–1211, 2020.

[ALG21] Dorna Abdolazimi, Kuikui Liu, and Shayan Oveis Gharan. A matrix trickle-down theorem on simplicial complexes and applications to sampling colorings. *arXiv preprint arXiv:2106.03845*, 2021. (To appear in *FOCS*'21).

[ALO20] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *FOCS*, pages 1319–1330, 2020.

[BCC+21] Antonio Blanca, Pietro Caputo, Zongchen Chen, Daniel Parisi, Daniel Štefankovič, and Eric Vigoda. On mixing of markov chains: Coupling, spectral independence, and entropy factorization. *arXiv preprint arXiv:2103.07459*, 2021.

[BD97] Russ Bubley and Martin Dyer. Path coupling: A technique for proving rapid mixing in markov chains. In *FOCS*, pages 223–231, 1997.

[BT06] Sergey G. Bobkov and Prasad Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.*, 19(2):289–336, 2006.

[CFYZ21] Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Rapid mixing of glauber dynamics via spectral independence for all degrees. *arXiv preprint arXiv:2105.15005*, 2021. (To appear in *FOCS*'21).

[CGŠV21] Zongchen Chen, Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Rapid mixing for colorings via spectral independence. In *SODA*, pages 1548–1557, 2021.

[CLV20] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Rapid mixing of Glauber dynamics up to uniqueness via contraction. In *FOCS*, pages 1307–1318, 2020. arXiv:2004.09083.

[CLV21a] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of Glauber dynamics: Entropy factorization via high-dimensional expansion. In *STOC*, 2021. arXiv:2011.02075.

[CLV21b] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Spectral independence via stability and applications to holant-type problems. *arXiv preprint arXiv:2106.03366*, 2021. (To appear in *FOCS*'21).

[CP20] Pietro Caputo and Daniel Parisi. Block factorization of the relative entropy via spatial mixing. *arXiv preprint arXiv:2004.10574*, 2020.

[DGU14] Martin Dyer, Catherine Greenhill, and Mario Ullrich. Structure and eigenvalues of heat-bath Markov chains. *Linear Algebra Appl.*, 454:57–71, 2014.

[DP11] Jian Ding and Yuval Peres. Mixing time for the Ising model: a uniform lower bound for all graphs. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(4):1020–1028, 2011.

[DR98] Devdatt Dubhashi and Desh Ranjan. Balls and bins: A study in negative dependence. *Random Structures and Algorithms*, 13(2):99–124, 1998.

[DSC96] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 1996.

[FGYZ21] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Rapid mixing from spectral independence beyond the boolean domain. In *SODA*, pages 1558–1577, 2021.

[GJ18] Heng Guo and Mark Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. *The Annals of Applied Probability*, 28(2):1292–1313, 2018.

[GL18] Heng Guo and Pinyan Lu. Uniqueness, spatial mixing, and approximation for ferromagnetic 2-spin systems. *ACM Trans. Comput. Theory*, 10(4):Art. 17, 25, 2018. (conference version in *RANDOM*'16).

[GM07] Antoine Gerschenfeld and Andrea Montanari. Reconstruction for models on random graphs. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007), October 20-23, 2007, Providence, RI, USA, Proceedings*, pages 194–204. IEEE Computer Society, 2007.

[GŠV16] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. *Combinatorics, Probability and Computing*, 25(04):500–559, 2016.

[HS07] Thomas P. Hayes and Alistair Sinclair. A general lower bound for mixing of single-site dynamics on graphs. *The Annals of Applied Probability*, pages 931–952, 2007.

[Isi25] Ernst Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, 1925.

[JDP83] Kumar Joag-Dev and Frank Proschan. Negative association of random variables, with applications. *Ann. Statist.*, 11(1):286–295, 1983.

[JPV21] Vishesh Jain, Huy Tuan Pham, and Thuy Duong Vuong. Spectral independence, coupling with the stationary distribution, and the spectral gap of the Glauber dynamics. *arXiv preprint arXiv:2105.01201*, 2021.

[JS93] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on Computing*, 22(5):1087–1116, 1993.

[Liu21] Kuikui Liu. From coupling to spectral independence and blackbox comparison with the down-up walk. *arXiv preprint arXiv:2103.11609*, 2021.

[LLY13] Liang Li, Pinyan Lu, and Yitong Yin. Correlation decay up to uniqueness in spin systems. In *SODA*, pages 67–84. SIAM, 2013. arXiv:1111.7064.

[LPW09] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times.* American Mathematical Soc., 2009.

[LSS17] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. The Ising partition function: Zeros and deterministic approximation. In *FOCS*, pages 986–997. IEEE Computer Society, 2017.

[Mar19] Katalin Marton. Logarithmic Sobolev inequalities in discrete product spaces. *Combin. Probab. Comput.*, 28(6):919–935, 2019.

[MS13] Elchanan Mossel and Allan Sly. Exact thresholds for Ising-Gibbs samplers on general graphs. *Ann. Probab.*, 41(1):294–328, 2013.

[SS12] Allan Sly and Nike Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *FOCS*, pages 361–369, 2012.

[SS20] Holger Sambale and Arthur Sinulis. Logarithmic Sobolev inequalities for finite spin systems and applications. *Bernoulli*, 26(3):1863–1890, 2020.

[SST14] Alistair Sinclair, Piyush Srivastava, and Marc Thurley. Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. *Journal of Statistical Physics*, 155(4):666–686, 2014. (conference version in *SODA*'12).