

# OPTIMAL MIXING FOR TWO-STATE ANTI-FERROMAGNETIC SPIN SYSTEMS

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

ABSTRACT. We prove an optimal  $\Omega(n^{-1})$  lower bound for modified log-Sobolev (MLS) constant of the Glauber dynamics for anti-ferromagnetic two-spin systems with  $n$  vertices in the tree uniqueness regime. Specifically, this optimal MLS bound holds for the following classes of two-spin systems in the tree uniqueness regime:

- all *strictly* anti-ferromagnetic two-spin systems (where both edge parameters  $\beta, \gamma \leq 1$ ), which cover the hardcore models and the anti-ferromagnetic Ising models;
- general anti-ferromagnetic two-spin systems on regular graphs.

Consequently, an optimal  $O(n \log n)$  mixing time holds for these anti-ferromagnetic two-spin systems when the uniqueness condition is satisfied. These MLS and mixing time bounds hold for any bounded or unbounded maximum degree, and the constant factors in the bounds depend only on the gap to the uniqueness threshold. We prove this by showing a boosting theorem for MLS constant for distributions satisfying certain spectral independence and marginal stability properties.

## CONTENTS

1. Introduction	2
2. Outline of Proofs	7
3. Preliminaries	10
4. Factorization of Entropy via Product Domination	13
5. Product Domination from Marginally Stable Spectral Independence	17
6. Boosting Modified Log-Sobolev Constant via $k$ -Transformation	20
7. Invariants of $k$ -Transformation	28
8. Applications to Anti-Ferromagnetic Two-Spin Systems	40
9. Acknowledgements	56
References	56
Appendix A. Mixing time from modified log-Sobolev constant	59
Appendix B. Modified log-Sobolev inequality in sub-critical regime	60
Appendix C. Monotonicity of uniqueness condition	63
Appendix D. Boundedness of anti-ferromagnetic two-spin system	64

---

(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](mailto:chenxiaoyu233@smail.nju.edu.cn), [yinyt@nju.edu.cn](mailto:yinyt@nju.edu.cn), [zhangxy@smail.nju.edu.cn](mailto: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](mailto:wfeng@ed.ac.uk)

## 1. INTRODUCTION

Two-state spin systems, or *two-spin systems*, are canonical graphical models arising from pairwise constrained Boolean variables. A two-spin system is specified on an undirected graph  $G = (V, E)$  by three parameters  $\beta, \gamma, \lambda \geq 0$ , where the two edge parameters  $\beta$  and  $\gamma$  specify the *edge activities*, and the vertex parameter  $\lambda$  specifies the *external field*. A *configuration*  $\sigma \in \{-1, +1\}^V$  assigns each vertex  $v \in V$  a  $\pm 1$ -*spin*. This defines a *Gibbs distribution*  $\mu$  over all the configurations  $\sigma \in \{-1, +1\}^V$  by:

$$\forall \sigma \in \{-1, +1\}^V, \quad \mu(\sigma) \triangleq \frac{1}{Z} \beta^{m_+(\sigma)} \gamma^{m_-(\sigma)} \lambda^{n_+(\sigma)},$$

where  $m_{\pm}(\sigma) \triangleq |\{\{u, v\} \in E \mid \sigma_u = \sigma_v = \pm 1\}|$  denotes the number of  $\pm 1$ -monochromatic edges in  $\sigma$ ,  $n_+(\sigma) \triangleq |\{v \in V \mid \sigma_v = +1\}|$  denotes the number of vertices assigned with  $+1$ -spin in  $\sigma$ , and the normalizing factor, known as the *partition function*, is given by:

$$Z \triangleq \sum_{\sigma \in \{-1, +1\}^V} \beta^{m_+(\sigma)} \gamma^{m_-(\sigma)} \lambda^{n_+(\sigma)}.$$

The hardcore models and the Ising models are two classes of extensively studied two-spin systems.

- *Hardcore model with fugacity  $\lambda$* : a two-spin system with  $\beta = 0$  and  $\gamma = 1$ ;
- *Ising model with temperature  $\beta$  and external field  $\lambda$* : a two-spin system with  $\beta = \gamma$ .

A two-spin system is called *ferromagnetic* if  $\beta\gamma > 1$  and *anti-ferromagnetic* if  $\beta\gamma < 1$ . The hardcore models are anti-ferromagnetic. An Ising model is ferromagnetic if  $\beta > 1$  and anti-ferromagnetic if  $\beta < 1$ .

The *Glauber dynamics* (a.k.a. *heat bath*, *Gibbs sampling*) is a canonical Markov chain for sampling from the Gibbs distribution  $\mu$ . Let  $\Omega(\mu)$  denote the support of  $\mu$ . In each transition, the Glauber dynamics updates the configuration  $\sigma \in \Omega(\mu)$  according to the following rules.

- pick a vertex  $v \in V$  uniformly at random;
- replace the spin  $\sigma_v$  according to the marginal distribution  $\mu_v^{\sigma_{V \setminus \{v\}}}$ .

It is well-known that Glauber dynamics is reversible with respect to the distribution  $\mu$  [LPW17]. Let  $P$  denote the transition matrix of the Glauber dynamics defined as above. The *lazy Glauber dynamics* is defined by in each transition, it is lazy with probability  $1/2$  and does the same transition as the Glauber dynamics otherwise. The transition matrix of the lazy Glauber dynamics is  $P_{zz} = \frac{1}{2}(I + P)$ , where  $I$  is the identity matrix. The *mixing time* of a Markov chain  $(X_t)_{t \geq 0}$  is defined by:

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

where  $d_{\text{TV}}(X_t, \mu)$  denotes the total variation distance between the distribution of  $X_t$  and  $\mu$ .

The *modified log-Sobolev (MLS) constant* [BT06] plays an important role in tight analysis of mixing times. Let  $P : \Omega(\mu) \times \Omega(\mu) \rightarrow \mathbb{R}_{\geq 0}$  denote the transition matrix of the Glauber dynamics on  $\mu$ . For any function  $f, g : \Omega(\mu) \rightarrow \mathbb{R}_{\geq 0}$ , the *Dirichlet form* is defined by:

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

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

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

where  $\mathbf{E}_{\mu}[f] \triangleq \sum_{\sigma \in \Omega(\mu)} \mu(\sigma)f(\sigma)$ . In above definitions, we assume  $0 \log 0 = 0$ .

The modified log-Sobolev constant for the Glauber dynamics on  $\mu$  is given by:

$$(1) \quad \rho^{\text{GD}}(\mu) \triangleq \inf \left\{ \frac{\mathcal{E}_P(f, \log f)}{\text{Ent}_{\mu}[f]} \mid f : \Omega(\mu) \rightarrow \mathbb{R}_{\geq 0}, \text{Ent}_{\mu}[f] \neq 0 \right\}.$$

It bounds the mixing time of lazy Glauber dynamics as follows: Denoted  $\mu_{\min} \triangleq \min_{\sigma \in \Omega(\mu)} \mu(\sigma)$ , we have

$$T_{\text{mix}}(\varepsilon) \leq \frac{4}{\rho^{\text{GD}}(\mu)} \left( \log \log \frac{1}{\mu_{\min}} + 200 \right) \log \frac{1}{\varepsilon},$$

Proving mixing time upper bound is reduced to establishing the modified log-Sobolev inequality (MLSI) that lower bounds the MLS constant. However, this task used to be notoriously difficult, especially when the maximum degree of the model was unbounded and no marginal probability lower bound was assumed.

In a breakthrough of Chen, Liu and Vigoda [CLV21, CLV20], a modified log-Sobolev inequality was established for Glauber dynamics on two-spin systems, which gave optimal lower bounds on the modified log-Sobolev constants for anti-ferromagnetic two-spin systems up to the critical phase-transition threshold known as the uniqueness condition, assuming constant maximum degree and marginal lower bound. To get rid of this dependency on degree or marginal bounds, recently Anari *et al.* [AJK<sup>+</sup>21a, AJK<sup>+</sup>21b] introduced a notion called “entropic independence”, which combined together with the “field dynamics” introduced in [CFYZ21b], could imply modified log-Sobolev inequalities without assuming degree or marginal bounds, in particular, for the two-spin systems where the uniqueness condition holds for all external fields, e.g. the Ising models with arbitrary local fields. A major technical challenge throughout is to establish optimal modified log-Sobolev inequalities in field-dependent uniqueness regimes without assuming degree or marginal bounds, e.g. for hardcore models or Ising models with uniform external fields.

**1.1. Results for two-spin systems.** We prove an  $e^{-O(1/\delta)}n^{-1}$  lower bound for the MLS constant for Glauber dynamics for the anti-ferromagnetic two-spin systems with  $n$  vertices in the tree uniqueness regime with a slack  $\delta \in (0, 1)$ . This MLS bound is asymptotically optimal in  $n$  and implies an optimal  $O(n \log n)$  mixing time for the Glauber dynamics when the uniqueness condition is satisfied with a constant gap  $\delta$ .

Consider two-spin systems on graph  $G = (V, E)$  with parameters  $(\beta, \gamma, \lambda)$ . By symmetry, we can assume:

$$(2) \quad 0 \leq \beta \leq \gamma, \quad \gamma > 0 \quad \text{and} \quad \lambda > 0.$$

A tuple  $(\beta, \gamma, \lambda)$  is called anti-ferromagnetic if it further satisfies  $\beta\gamma < 1$  in addition to this.

The following uniqueness condition for anti-ferromagnetic two-spin system was characterized in [LLY13].

**Definition 1.1.** Let  $d \geq 1$  be an integer. An anti-ferromagnetic  $(\beta, \gamma, \lambda)$  is  $d$ -unique with gap  $\delta \in (0, 1)$  if

$$(3) \quad |F'_d(\hat{x}_d)| = \frac{d(1 - \beta\gamma)\hat{x}_d}{(\beta\hat{x}_d + 1)(\hat{x}_d + \gamma)} \leq 1 - \delta, \quad \text{where } \hat{x}_d \text{ is the unique fixed point of } F_d(x) = \lambda \left( \frac{\beta x + 1}{x + \gamma} \right)^d.$$

The property of being  $d$ -unique corresponds to the uniqueness of Gibbs measure on  $(d + 1)$ -regular tree. It was well known that sampling in anti-ferromagnetic two-spin systems on  $\Delta$ -regular graphs is intractable if  $(\beta, \gamma, \lambda)$  is not  $(\Delta - 1)$ -unique [SS12, GŠV15]. We consider the following criterion for two-spin systems.

**Condition 1.2 (uniqueness criterion).** Let  $\delta \in (0, 1)$ . The anti-ferromagnetic two-spin system specified by  $(\beta, \gamma, \lambda)$  on graph  $G = (V, E)$  with maximum degree  $\Delta \geq 3$  satisfies:

- when  $\gamma \leq 1$ :  $(\beta, \gamma, \lambda)$  is  $(\Delta - 1)$ -unique with gap  $\delta$ ;
- when  $\gamma > 1$ :  $(\beta, \gamma, \lambda)$  is  $(\Delta - 1)$ -unique with gap  $\delta$  and  $G$  is  $\Delta$ -regular.

For the classes of anti-ferromagnetic two-spin systems satisfying such uniqueness criterion, we show the following optimal bounds on the MLS constant and the mixing time of Glauber dynamics.

**Theorem 1.3 (main theorem: two-spin systems).** Let  $\delta \in (0, 1)$ . There exists a  $C(\delta) = \exp(O(1/\delta))$  such that for every anti-ferromagnetic two-spin system with  $n$  vertices that satisfies Condition 1.2 with gap  $\delta$ , the modified log-Sobolev constant  $\rho^{\text{GD}}$  of the Glauber dynamics satisfies

$$\rho^{\text{GD}} \geq \frac{1}{C(\delta)n}.$$

Consequently, let

$$\alpha = \begin{cases} \gamma + \gamma^{-1} + 2 & \text{if } \beta = 0 \\ \beta^{-1} + 2 & \text{if } \beta > 0 \end{cases},$$

the mixing time of the lazy Glauber dynamics is bounded by

$$T_{\text{mix}}(\varepsilon) \leq C(\delta)n \left( 2 \log n + \log \log \alpha + \log \log \left( \lambda + \frac{1}{\lambda} \right) + 200 \right) \log \frac{1}{\varepsilon}.$$

Due to the hardness results in [SS12, GŠV15], Theorem 1.3 gives sharp computational phase transitions, since sampling in not- $(\Delta - 1)$ -unique  $\Delta$ -regular anti-ferromagnetic two-spin systems is intractable.

**Remark 1.4** (mixing time of the continuous-time Glauber dynamics). Theorem 1.3 gives a lower bound on the modified log-Sobolev constant. By the well-known result in [BT06], Theorem 1.3 also implies the

$$C(\delta)n \left( 2 \log n + \log \log \alpha + \log \log \left( \lambda + \frac{1}{\lambda} \right) + \log \frac{1}{2\varepsilon^2} \right)$$

mixing time upper bound for the continuous-time Glauber dynamics. See Section 3.1 for the definitions of the continuous-time Glauber dynamics and its mixing time.

**Remark 1.5** (comparison to the up-to- $\Delta$ -uniqueness). The uniqueness condition (Condition 1.2) assumed by Theorem 1.3 slightly deviates from the *up-to- $\Delta$ -uniqueness* (i.e.  $d$ -unique for all  $1 \leq d \leq \Delta - 1$ ) assumed in e.g. [LLY13, CLV20, CLV21, CFYZ21b] for spin systems with  $\Delta$ -bounded maximum degree.

It is known that  $|F'_d(\hat{x}_d)|$  in (3) is monotonically increasing in  $d$  if and only if  $\gamma \leq 1$  (Proposition C.1). Therefore, when  $\gamma \leq 1$ , being  $(\Delta - 1)$ -unique immediately implies the up-to- $\Delta$ -uniqueness; and in contrast when  $\gamma > 1$ , the property of being  $d$ -unique may no longer be monotone in  $d$ . And hence:

- Case ( $\gamma \leq 1$ ): the uniqueness condition assumed by Theorem 1.3 is the same as the up-to- $\Delta$ -uniqueness on instances with  $\Delta$ -bounded max-degree, as in [LLY13, CLV20, CLV21, CFYZ21b];
- Case ( $\gamma > 1$ ): the uniqueness condition assumed by Theorem 1.3 is restricted to the regular graphs, but it can give strictly broader regime than the up-to- $\Delta$ -uniqueness.

To the best of our knowledge, this is the first time that a strictly stronger critical algorithm is proved on regular graphs than on general graphs, for anti-ferromagnetic two-spin systems.

Both the hardcore and anti-ferromagnetic Ising models fall into the *strictly* anti-ferromagnetic case where  $\gamma \leq 1$ . Hence the following corollaries hold, whose formal proofs are given in Appendix A.

**Corollary 1.6** (hardcore model). *Let  $\delta \in (0, 1)$ . There exists a  $C(\delta) = \exp(O(1/\delta))$  such that for every hardcore model on  $n$ -vertex graph  $G = (V, E)$  with maximum degree  $\Delta \geq 3$  and fugacity  $\lambda \leq (1 - \delta)\lambda_c(\Delta) = (1 - \delta) \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta}$ , the mixing time of the lazy Glauber dynamics is bounded by*

$$T_{\text{mix}}(\varepsilon) \leq C(\delta)n (2 \log n + 200) \log \varepsilon^{-1}.$$

**Corollary 1.7** (anti-ferromagnetic Ising model). *Let  $\delta \in (0, 1)$ . There exists a  $C(\delta) = \exp(O(1/\delta))$  such that for every anti-ferromagnetic Ising model with temperature  $\beta \in (0, 1)$  and external field  $\lambda > 0$  on  $n$ -vertex graph  $G = (V, E)$  with maximum degree  $\Delta \geq 3$  that satisfies either one of the followings:*

- $\beta \geq \frac{\Delta-2+\delta}{\Delta-\delta}$ ;
- $\beta < \frac{\Delta-2+\delta}{\Delta-\delta}$  and  $\lambda \in (0, \lambda_c] \cup [\bar{\lambda}_c, \infty)$ , where  $\lambda_c = \lambda_c(\delta, \beta)$  and  $\bar{\lambda}_c = \bar{\lambda}_c(\delta, \beta)$  that satisfy  $\lambda_c \leq \bar{\lambda}_c$  and  $\lambda_c \bar{\lambda}_c = 1$ , are the critical thresholds for  $\lambda$  in anti-ferromagnetic Ising model [LLY13, SST14];

the mixing time of lazy Glauber dynamics is bounded by

$$T_{\text{mix}}(\varepsilon) \leq C(\delta)n \left( 2 \log n + \log \log (\beta^{-1} + 3) + \log \log (\lambda + \lambda^{-1}) + 200 \right) \log \varepsilon^{-1}.$$

Note that the Ising uniqueness regime in Corollary 1.7 is much broader than the regime  $\beta \in [\frac{\Delta-2+\delta}{\Delta-\delta}, 1)$  assumed in [CFYZ21a, AJK<sup>+</sup>21b] for the anti-ferromagnetic case, which corresponds to the uniqueness regime for all external fields  $\lambda$ . In fact, before this work, proving optimal mixing times for  $\lambda$ -dependent uniqueness regimes was a major challenge to the current techniques [AJK<sup>+</sup>21b].

The modified log-Sobolev inequalities (MLSI) are very powerful. For example, by the Herbst argument (e.g. [CGM21, Lemma 15]), the MLSI in Theorem 1.3 also implies the following concentration bound.

**Corollary 1.8.** *Let  $\delta \in (0, 1)$ . There exists a  $C = \exp(O(1/\delta))$  such that for every anti-ferromagnetic two-spin system with  $n$  vertices, if Condition 1.2 is satisfied with gap  $\delta$ , then it holds for the Gibbs distribution  $\mu$  and for any observable function  $f : \Omega(\mu) \rightarrow \mathbb{R}$  and any  $\alpha \geq 0$  that*

$$\Pr_{x \sim \mu} \left[ \left| f(x) - \mathbb{E}_\mu[f] \right| \geq \alpha \right] \leq 2 \exp\left(-\frac{\alpha^2}{2Cnv(f)}\right),$$

where  $v(f)$  is the maximum of one-step variances,

$$v(f) \triangleq \max_{x \in \Omega(\mu)} \left\{ \sum_{y \in \Omega(\mu)} P(x, y) (f(x) - f(y))^2 \right\},$$

where  $P$  denotes the transition matrix of the Glauber dynamics over  $\mu$ .

**1.2. Results for general distributions.** Let  $\mu$  be a distribution over  $\{-1, +1\}^n$  and let  $\Omega(\mu)$  be its support. Given  $\Lambda \subseteq [n]$ , we use  $\mu_\Lambda$  to denote the marginal distribution on  $\Lambda$  projected from  $\mu$ , and we write  $\mu_i = \mu_{\{i\}}$  for  $i \in [n]$ . Given any partial configuration  $\sigma \in \Omega(\mu_\Lambda)$  where  $\Lambda \subseteq [n]$ , we use  $\mu^\sigma$  to denote the conditional distribution over  $\{-1, +1\}^n$  induced by  $\mu$  conditional on  $\sigma$ , and for  $x \in \Omega(\mu_i^\sigma)$ , we use  $\mu^{\sigma \wedge i \leftarrow x}$  to denote the conditional distribution obtained from  $\mu^\sigma$  by further conditioning on the spin of  $i \in [n]$  being fixed as  $x$ .

The notion of *spectral independence* was introduced by Anari, Liu and Oveis Gharan in [ALO20]. We use the absolute version of the spectral independence considered in [FGYZ21, CFYZ21b].

**Definition 1.9** (spectral independence (absolute version)). Let  $\mu$  be a distribution over  $\{-1, +1\}^n$ . For any  $\Lambda \subseteq [n]$ ,  $\sigma \in \Omega(\mu_\Lambda)$ , the *absolute influence matrix*  $\Psi_{\mu^\sigma} \in \mathbb{R}_{\geq 0}^{n \times n}$  is defined as

$$\forall i, j \in [n], \quad \Psi_{\mu^\sigma}(i, j) \triangleq \max_{x, y \in \Omega(\mu_i^\sigma)} d_{\text{TV}}\left(\mu_j^{\sigma \wedge i \leftarrow x}, \mu_j^{\sigma \wedge i \leftarrow y}\right),$$

where  $d_{\text{TV}}(\cdot, \cdot)$  denotes the total variation distance. Let  $\eta > 0$ . The distribution  $\mu$  is said to be  $\eta$ -*spectrally independent* (SI) if for any  $\Lambda \subseteq V$ , any  $\sigma \in \Omega(\mu_\Lambda)$ , the spectral radius of the influence matrix  $\Psi_{\mu^\sigma}$  satisfies

$$\rho(\Psi_{\mu^\sigma}) \leq \eta.$$

It was known that assuming constant marginal lower bound, the spectral independence can guarantee the optimal mixing of Glauber dynamics [CLV21, BCC<sup>+</sup>21]. In fact, MLSIs have been proved assuming the same marginal lower bound [Mar19, SS20]. However, such strong condition on marginal bounds does not hold in general for spin systems with unbounded maximum degree, and it is a major open problem to prove MLSI and optimal mixing time for such models.

We introduce the following notion that weakens the marginal lower bound condition.

**Definition 1.10** (marginal stability). Let  $\zeta > 0$ . A distribution  $\mu$  over  $\{-1, +1\}^n$  is said to be  $\zeta$ -*marginally stable* if for any  $i \in [n]$ , any  $S \subseteq \Lambda \subseteq [n] \setminus \{i\}$ , and any  $\sigma \in \Omega(\mu_\Lambda)$ ,

$$R_i^\sigma \leq \zeta \quad \text{and} \quad R_i^\sigma \leq \zeta \cdot R_i^{\sigma_S},$$

where  $R_i^\sigma \triangleq \frac{\mu_i^\sigma(+1)}{\mu_i^\sigma(-1)}$  denotes the marginal ratio, and  $R_i^{\sigma_S}$  is accordingly defined for  $\sigma_S$ .

The marginal lower bound assumption imposes a lower bound  $b$  for the marginal probability  $\mu_i^\sigma(x) > b$  for all possible spins  $x$ . The marginal stability, weakens this to the following properties combined:

- (1) a one-sided marginal lower bound, to ensure that  $\mu_i^\sigma(-1)$  is not too small;
- (2) a one-sided decay of correlation, to ensure that pinning does not greatly increase the marginal ratio.

Furthermore, it is crucial that the one-sided directions in above Item 1 and Item 2 are consistent. Intuitively, the direction towards +1 means more critical and hence needs to be controlled (while being applied to two-spin systems with a different direction of criticality, the roles of +1/−1 are flipped in a preprocessing).

Such condition ingeniously captures the subcritical two-spin systems. On one hand, it is strong enough, together with the spectral independence property to guarantee the optimal mixing of Glauber dynamics. On the other hand, it is also weak enough to be satisfied by the subcritical anti-ferromagnetic two-spin systems.

In order to deduce optimal mixing times from spectral independence and marginal stability, we need these properties to hold for all subcritical external fields. Given a distribution  $\mu$  over  $\{-1, +1\}^n$  and a vector  $\lambda = (\phi_v)_{v \in [n]} \in \mathbb{R}_{>0}^n$  that specifies the local fields, we use  $(\lambda * \mu)$  to denote the distribution obtained from “magnetizing”  $\mu$  with the local fields in  $\lambda$ . Formally:

$$(4) \quad \forall \sigma \in \{-1, +1\}^n, \quad (\lambda * \mu)(\sigma) \propto \mu(\sigma) \prod_{i \in [n]: \sigma_i = +1} \lambda_i.$$

In particular, if  $\lambda_i = \theta \in \mathbb{R}_{>0}$  for all  $i \in [n]$  for some scalar  $\theta \in \mathbb{R}_{>0}$ , we simply write  $(\theta * \mu)$  for  $(\lambda * \mu)$ .

We formalize the following sufficient condition for a MLSI for Glauber dynamics.

**Condition 1.11.** *Let  $\eta > 1, \varepsilon > 0, \zeta > 1$  be parameters. Then  $\mu$  is a distribution over  $\{-1, +1\}^n$  that satisfies:*

- (1)  $(\lambda * \mu)$  is  $\eta$ -spectrally independent for all  $\lambda \in (0, 1 + \varepsilon]^n$ ;
- (2)  $(\lambda * \mu)$  is  $\zeta$ -marginally stable for all  $\lambda \in (0, 1]^n$ .

An ( $\varepsilon = 0$ ) variant of Condition 1.11 (1) was used in [CFYZ21b] and called *complete spectral independence*.

Recall that  $\rho^{\text{GD}}(\mu)$  denotes the modified log-Sobolev constant of Glauber dynamics on  $\mu$  and  $\mu^\sigma$  is the conditional distribution over  $\{-1, +1\}^n$  induced by  $\mu$  conditional on  $\sigma$ . We further denote by  $\rho_{\min}^{\text{GD}}(\mu)$  the minimum modified log-Sobolev constant for  $\mu^\sigma$  over all possible  $\sigma$ :

$$(5) \quad \rho_{\min}^{\text{GD}}(\mu) \triangleq \min_{\Lambda \subseteq [n]} \min_{\sigma \in \Omega(\mu_\Lambda)} \rho^{\text{GD}}(\mu^\sigma).$$

**Theorem 1.12** (main theorem: general). *For any distribution  $\mu$  over  $\{-1, +1\}^n$ , if  $\mu$  satisfies Condition 1.11 with  $\eta > 1, \varepsilon > 0, \zeta > 1$ , then the following holds for the modified log-Sobolev constants for Glauber dynamics:*

$$\forall \theta \in (0, 1), \quad \rho^{\text{GD}}(\mu) \geq \left(\frac{\theta}{\varepsilon}\right)^{30\eta + \frac{\log(4\zeta)}{\log(1+\varepsilon)}} \rho_{\min}^{\text{GD}}(\theta * \mu).$$

Theorem 1.12 is a boosting theorem for modified log-Sobolev inequality (MLSI). By choosing a suitable constant gap  $\theta$ , the MLSI for the original near-critical distribution  $\mu$  is reduced, by losing a constant factor, to the MLSI for the magnetized distribution  $\theta * \mu$  that falls into a subcritical regime, where the minimum MLS constant  $\rho_{\min}^{\text{GD}}(\theta * \mu)$  is easier to analyze. A similar boosting theorem for the Poincaré constant (spectral gap) was established in [CFYZ21b], essentially by assuming the spectral independence part of Condition 1.11. Here we prove a similar boosting for the MLSI by further assuming the marginal stability.

**Remark 1.13** (applications to spin systems). When applying Theorem 1.12 to anti-ferromagnetic two-spin systems, as in [CFYZ21b], we can first preprocess the distribution  $\mu$  by properly flipping the roles of spins in  $\{+1, -1\}$  for each vertex, so that after the preprocessing, the distribution will only get “less critical” by decreasing the local field at every vertex. We then formally verify the spectral independence and the marginal stability properties in Condition 1.11 for the flipped distribution assuming the uniqueness.

Theorem 1.12 can then be applied to boosting the MLSI in the subcritical regime up to the uniqueness threshold, where the MLSI in the subcritical regime can be obtained from, for example, the result on the Ricci curvature in [EHMT17]. This proves Theorem 1.3. The detailed analysis is given in Section 8.

1.3. **Background and related work.** The computational phase transition for sampling and counting in two-spin systems has drawn considerable studies [JS93, LV99, Vig99, DG00, GJP03]. The  $d$ -uniqueness (Definition 1.1) represents the uniqueness of infinite-volume Gibbs measure on  $(d+1)$ -regular tree [Wei05, LLY13, SST14]. Initiated in a seminal work of Weitz [Wei06], correlation-decay based deterministic counting algorithms were given for anti-ferromagnetic two-spin systems with  $\Delta$ -bounded maximum degree that satisfy the up-to- $\Delta$ -uniqueness (being  $d$ -unique for all  $d < \Delta$ ) [Wei06, LLY12, SST14, LLY13]. These algorithms run in time  $n^{O(\log \Delta)}$ . Together with the hardness results in the non-uniqueness regime [Sly10, SS12, GŠV15], this gives a computational phase transition for spin systems with constant maximum degree.

Due to lower bound results in [HS07, DP11], the optimal mixing time of Glauber dynamics is  $O(n \log n)$ . It is also widely believed that such optimal mixing time should hold for the two-spin systems in the uniqueness regime. Proving such conjectures is extremely challenging. A substantial body of research works have dedicated to this. Using coupling based techniques, optimal  $O(n \log n)$  mixing times were proved assuming girth lower bound [HV06, EHŠ<sup>+</sup>19] or for Ising models with constant maximum degree [MS13].

The spectrum based techniques tries to lower bound the spectral gap of Glauber dynamics. In a seminal work [ALO20], Anari, Liu, and Oveis Gharan introduced the concept of spectral independence and applied the tools from high-dimensional expander walks developed in [ALOV19, AL20] to relate the spectral gap to the decay of correlation properties. For anti-ferromagnetic two-spin systems satisfying the uniqueness condition with gap  $\delta$ , the spectral gap was improved from  $n^{-O(1/\delta)}$  [ALO20, CLV20], to  $\Delta^{-O(1/\delta)} n^{-1}$  [CLV21, JPV21], and finally to  $(1/\delta)^{-O(1/\delta)} n^{-1}$  [CFYZ21b] which was optimal in  $n$  for arbitrary maximum degree  $\Delta$ . However, as spectral gaps, they are not sufficient for optimal  $O(n \log n)$  mixing time.

Entropy based techniques that could prove modified log-Sobolev inequalities (MLSI) were considered [CGM21, ALOV20]. Although modified log-Sobolev (MLS) constants can give tight bounds on mixing times, they are notoriously difficult to analyze. In many previous works [CMT15, FM16, EHMT17, Mar19, SS20, Con20], the optimal  $\Omega(n^{-1})$  MLS bounds were proved only in the regimes where more standard techniques such as coupling could also work. Perhaps the first breakthrough to this was the one achieved by Chen, Liu and Vigoda [CLV21]: there and in a follow-up work [BCC<sup>+</sup>21], a  $(b/\Delta)^{O(1/(\delta b))} n^{-1}$  MLS bound was proved for anti-ferromagnetic two-spin systems satisfying the uniqueness condition with gap  $\delta$ , assuming marginal lower bound  $b$ . This MLS bound beats the coupling in regimes and is optimal in  $n$  for constant  $\Delta$ . However, the reliance on margin bound results in a bad dependence on the max-degree  $\Delta$ .

Recently, Anari, Jain, Koehler, Pham, and Vuong [AJK<sup>+</sup>21a] proposed the notion of *entropic independence*, which was crucial for removing the reliance on marginal lower bound in CLV’s argument [CLV21], by assuming spectral independence for all fields. This was followed by [AJK<sup>+</sup>21b] and [CFYZ21a], where both works used the field dynamics invented in [CFYZ21b] to connect the entropic independence to the MLS constant of Glauber dynamics. Both succeeded in proving optimal mixing for Ising models in a uniqueness regime that holds for all external fields. Moreover, [AJK<sup>+</sup>21b] also achieves  $O(n \log n)$  mixing for a modified version of Glauber dynamics called *balanced Glauber dynamics* on the hardcore model. A major obstacle for the current techniques is to prove *optimal* mixing of Glauber dynamics for spin systems with *unbounded* maximum degree in *field-dependent* uniqueness regimes, which is typical for computational phase transitions for anti-ferromagnetic two-spin systems.

**Concurrent work.** When preparing the current paper, we were informed by Yuansi Chen and Ronen Eldan about their concurrent work [CE22], in which they independently prove the same optimal mixing bound for the hardcore model, through a more abstract framework called “localization schemes”.

## 2. OUTLINE OF PROOFS

In this section, we outline our proof of Theorem 1.12.

2.1. **Product domination and block factorization.** The spectral independence and marginal stability in Condition 1.11 together ensure a property of product domination, which plays a key role in the proof.

The probability generating function  $g_\mu$  for a distribution  $\mu$  over  $\{-1, +1\}^n$  is defined by

$$g_\mu(z_1, z_2, \dots, z_n) \triangleq \sum_{\sigma \in \{-1, +1\}^n} \mu(S) \prod_{i \in [n]: \sigma_i = +1} z_i.$$

We introduce the following concept of product distribution.

**Definition 2.1** (product domination). Let  $\alpha \in (0, 1)$  be real. A distribution  $\mu$  over  $\{-1, +1\}^n$  is said to be  $(1/\alpha)$ -product dominated if for all  $(z_1, \dots, z_n) \in \mathbb{R}_{>0}^n$ ,

$$g_\mu(z_1^\alpha, z_2^\alpha, \dots, z_n^\alpha)^{\frac{1}{\alpha}} \leq \prod_{i=1}^n (\mu_i(+1)z_i + \mu_i(-1)).$$

Furthermore,  $\mu$  is  $(1/\alpha)$ -product dominated on  $D \subseteq \mathbb{R}_{>0}^n$  if the above holds for all  $(z_1, \dots, z_n) \in D$ .

The product domination asserts that the generating function of  $\mu$  is “ $\alpha$ -fractionally” dominated by that of a product distribution, in which the  $i$ -th variable takes the value  $+1$  independently with probability  $\mu_i(+1)$ . Such  $\alpha$ -fractional form  $g_\mu(z_1^\alpha, z_2^\alpha, \dots, z_n^\alpha)^{1/\alpha}$  of generating function has appeared in the definition of fractionally log-concave distributions [AASV21, AJK<sup>+</sup>21a]. More significantly, product domination gives an equivalent characterization of the entropic independence introduced in [AJK<sup>+</sup>21a]. More precisely,  $\mu$  is  $(1/\alpha)$ -product dominated if and only if its homogenization  $\mu^{\text{hom}}$  is  $(1/\alpha)$ -entropically independent. Such equivalence and formal definitions of entropic independence and homogenization are given in Section 4. Product domination is also related to a Brascamp-Lieb type inequality. See Remark 4.5 for the detailed relation among product domination, entropic independence, and a specific Brascamp-Lieb type inequality.

We show that this product domination property is guaranteed by Condition 1.11. For technical reasons, we will show that a weakening of Condition 1.11 is sufficient to guarantee the product domination.

**Condition 2.2.** Let  $\eta > 1$ ,  $\varepsilon > 0$ ,  $\zeta > 1$  be parameters. The  $\mu$  is a distribution over  $\{-1, +1\}^n$  that satisfies:

- (1)  $(\lambda * \mu)$  is  $\eta$ -spectrally independent for all  $\lambda \in (0, 1 + \varepsilon]^n$ ;
- (2)  $\mu$  is  $\zeta$ -marginally stable.

Condition 2.2 (complete spectral independence with marginal stability) weakens Condition 1.11 because it only requires the marginal stability to hold for  $\mu$  itself but not for  $(\lambda * \mu)$  with other external fields  $\lambda$ .

**Lemma 2.3.** For any distribution  $\mu$  over  $\{-1, +1\}^n$ , if  $\mu$  satisfies Condition 2.2 with  $\eta > 1$ ,  $\varepsilon > 0$  and  $\zeta > 1$ , then for any  $\Lambda \subseteq [n]$  and  $\sigma \in \Omega(\mu_\Lambda)$ , the distribution  $\mu_{[n] \setminus \Lambda}^\sigma$  is  $(1/\alpha)$ -product dominated, where

$$(6) \quad \alpha = \min \left\{ \frac{1}{2\eta}, \frac{\log(1 + \varepsilon)}{\log(1 + \varepsilon) + \log 2\zeta} \right\}.$$

Lemma 2.3 is proved in Section 5. More precisely, the complete spectral independence (Item 1 in Condition 2.2) guarantees that  $\mu$  is  $(1/\alpha)$ -product dominated on the domain  $D = (0, (1 + \varepsilon)^{1/\alpha}]^n$ , and the marginal stability (Item 2 in Condition 2.2) allows to extend such product domination from  $D$  to  $\mathbb{R}_{>0}^n$ .

The product domination property is closely related to the *uniform block factorization of entropy*, which gives rise to the entropy decay and MLSI for uniform block dynamics.

**Definition 2.4** (uniform block factorization of entropy [CP20]). Let  $n \geq 1$  and  $1 \leq \ell \leq n$  be two integers, and  $C > 0$ . A distribution  $\mu$  over  $\{-1, +1\}^n$  is said to satisfy the  $\ell$ -uniform block factorization of entropy with parameter  $C$  if for all  $f : \Omega(\mu) \rightarrow \mathbb{R}_{\geq 0}$ ,

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

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

We show that the product domination with all pinnings implies the block factorization of entropy.



**Lemma 2.5.** *Let  $\mu$  be a distribution over  $\{-1, +1\}^n$  and  $\alpha \in (0, 1)$ . If for any  $\Lambda \subseteq [n]$  and any  $\sigma \in \Omega(\mu_\Lambda)$ , the conditional marginal distribution  $\mu_{[n] \setminus \Lambda}^\sigma$  is  $(1/\alpha)$ -product dominated, then for every integer  $\ell \geq 1/\alpha$ , the distribution  $\mu$  satisfies  $\ell$ -uniform block factorization of entropy with parameter  $C = (\frac{\ell n}{\ell})^{1/\alpha+1}$ .*

Lemma 2.5 is proved in Section 4. The proof is based on the aforementioned equivalence between product domination and entropic independence, while the latter is known to guarantee the uniform block factorization of entropy [AJK<sup>+</sup>21a].

Lemma 2.3 and Lemma 2.5 together show that Condition 2.2 guarantees that  $\mu$  satisfies the  $\ell$ -uniform block factorization of entropy, which is sufficient to imply the MLSI for the  $\ell$ -uniform block dynamics on  $\mu$  [CP20]. In order to enhance this to the MLSI for single-site Glauber dynamics without resorting to marginal lower bound, we further apply the  $k$ -transformation introduced in [CFYZ21b].

**2.2.  $k$ -transformation and boosting of MLSI.** The  $k$ -transformation operation for a multi-dimensional probability distribution with Boolean domain is formally defined as follows.

**Definition 2.6** ( $k$ -transformation [CFYZ21b]). Let  $\mu$  be a distribution over  $\{-1, +1\}^n$  and  $k \geq 1$  an integer. The  $k$ -transformation of  $\mu$  gives a distribution  $\mu_k = \text{Trans}(\mu, k)$  over  $\{-1, +1\}^{n \times k}$ , constructed as follows.

Let  $X \sim \mu$ . Then  $\mu_k = \text{Trans}(\mu, k)$  is the distribution of  $Y \in \{-1, +1\}^{n \times k}$  constructed as follows:

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

The  $k$ -transformation defines a sort-of lifting operation on  $\mu$ . It effectively replaces every  $i \in [n]$  with a gadget, which is essentially a hardcore model on  $k$ -clique. The limiting object of the uniform block dynamics on  $\mu_k$  when  $k \rightarrow \infty$  gives the *field dynamics* process introduced in [CFYZ21b].

The significance of such lifting operations on  $\mu$  is that the uniform block factorizations of entropy for  $\mu_k$  for all sufficiently large  $k$  give rise to a boosting of MLSI for Glauber dynamics, which holds with no further restriction on marginals.

**Lemma 2.7.** *Let  $\mu$  be a distribution over  $\{-1, +1\}^n$ ,  $\theta \in (0, 1)$  and  $C > 0$ . If there is a finite  $K_0 = K_0(\mu, \theta, C)$  such that for all integers  $k \geq K_0$ , the distribution  $\mu_k = \text{Trans}(\mu, k)$  satisfies  $\lceil \theta kn \rceil$ -uniform block factorization of entropy with parameter  $C$ , then the Glauber dynamics on  $\mu$  has the modified log-Sobolev constant*

$$\rho^{\text{GD}}(\mu) \geq \frac{\rho_{\min}^{\text{GD}}(\theta * \mu)}{C}.$$

Lemma 2.7 states a boosting of MLSI: it transforms MLSIs for a sub-critical  $\theta * \mu$  to that for a near critical  $\mu$ , given the uniform block factorizations of  $k$ -transformed distributions  $\mu_k$  for all sufficiently large  $k$ . Similar boosting results have been established in [AJK<sup>+</sup>21b] for the 1-uniform block factorization of entropy (a.k.a. approximate tensorization of entropy) and in [CFYZ21b] for Poincaré constants (spectral gaps).

Lemma 2.7 is proved by establishing a “magnetized block factorization of entropy” (Definition 6.1) for  $\mu$ . A magnetized block factorization on  $\mu$  can be obtained from a natural projection of the uniform block factorization on  $\mu_k$  back to  $\mu$  and letting  $k \rightarrow \infty$ . The magnetized block factorization of entropy could be seen as a variant of block factorization of entropy for the *field dynamics* introduced in [CFYZ21b], and it intrinsically relates the MLS constant for  $\mu$  to those for the magnetized distribution  $\theta * \mu$ . A formal proof of Lemma 2.7 is given in Section 6.

It only remains to guarantee that the  $k$ -transformed distribution  $\mu_k$  indeed satisfies the desired uniform block factorization of entropy for all sufficiently large  $k$ . From the argument in Section 2.1, this holds as long as Condition 2.2 can be verified for all such  $\mu_k$ , which is guaranteed by the following lemma.

**Lemma 2.8.** *For any distribution  $\mu$  over  $\{-1, +1\}^n$ , if  $\mu$  satisfies Condition 1.11 with  $\eta > 1$ ,  $\varepsilon > 0$  and  $\zeta > 1$ , then there exists a finite  $k_0 = 10(1+\varepsilon)(1+\zeta)$  such that for all integers  $k \geq k_0$ , the distribution  $\mu_k = \text{Trans}(\mu, k)$  satisfies Condition 2.2 with parameters  $(\eta', \varepsilon', \zeta')$ , where  $\eta' = 2\eta + 5$ ,  $\varepsilon' = \varepsilon$  and  $\zeta' = 2\zeta$ .*

Lemma 2.8 basically says that Condition 1.11 is almost invariant under  $k$ -transformation (it is not exactly invariant because Condition 2.2 is still weaker than Condition 1.11). This can be formally verified by using a natural coupling between  $\mu$  and  $\mu_k$ . The formal proof is in Section 7.

**2.3. Proof of main theorem.** We now prove Theorem 1.12. Fix  $\eta > 1$ ,  $\varepsilon > 0$ ,  $\zeta > 1$  and distribution  $\mu$ . Assume that  $\mu$  satisfies Condition 1.11 with parameters  $\eta, \varepsilon$  and  $\zeta$ . Due to Lemma 2.8, there is a finite  $k_0 = 10(1 + \varepsilon)(1 + \zeta)$  such that for all integers  $k \geq k_0$ , Condition 2.2 holds for  $\mu_k = \text{Trans}(\mu, k)$  with parameters  $\eta' = 2\eta + 5$ ,  $\varepsilon' = \varepsilon$  and  $\zeta' = 2\zeta$ , which according to Lemma 2.3 and Lemma 2.5, means that for

$$\alpha = \min \left\{ \frac{1}{4\eta + 10}, \frac{\log(1 + \varepsilon)}{\log(1 + \varepsilon) + \log(4\zeta)} \right\},$$

for all integers  $\ell \geq 4\eta + 11 + \frac{\log(4\zeta)}{\log(1+\varepsilon)} \geq 1/\alpha$ ,  $\mu_k$  satisfies  $\ell$ -uniform block factorization of entropy with parameter

$$C = \left( \frac{enk}{\ell} \right)^{1/\alpha+1} \leq \left( \frac{enk}{\ell} \right)^{4\eta+12+\frac{\log(4\zeta)}{\log(1+\varepsilon)}}.$$

Fix an arbitrary  $\theta \in (0, 1)$ . Consider  $\ell = \lceil \theta kn \rceil$ . For all sufficiently large integers  $k \geq k_0$  satisfying  $\theta kn \geq 4\eta + \frac{\log 4\zeta}{\log(1+\varepsilon)} + 11$ ,  $\mu_k$  satisfies the  $\lceil \theta kn \rceil$ -uniform block factorization of entropy with parameter

$$C = \left( \frac{enk}{\lceil \theta kn \rceil} \right)^{4\eta+12+\frac{\log(4\zeta)}{\log(1+\varepsilon)}} \leq \left( \frac{e}{\theta} \right)^{4\eta+\frac{\log(4\zeta)}{\log(1+\varepsilon)}+12} \leq \left( \frac{e}{\theta} \right)^{30\eta+\frac{\log(4\zeta)}{\log(1+\varepsilon)}}$$

This holds for all  $k \geq K_0$ , where  $K_0 = \max \left\{ 10(1 + \varepsilon)(1 + \zeta), \frac{4\eta}{\theta n} + \frac{\log(4\zeta)}{\theta n \log(1+\varepsilon)} + \frac{11}{\theta n} \right\}$ . Note that  $K' = K_0(\theta)$  is finite because  $\varepsilon > 0$ ,  $\zeta > 1$  and  $\eta > 1$  are fixed parameters. Theorem 1.12 follows from Lemma 2.7.

**2.4. Open problems.** Compared to the classes of two-spin systems resolved in [LLY13, CFYZ21b], the uniqueness criterion (Condition 1.2) used in Theorem 1.3 still leaves open the optimal mixing for the following classes of anti-ferromagnetic two-spin systems  $(\beta, \gamma, \lambda)$  on  $G$ :

- $\gamma > 1$ , the graph  $G$  is a general graph with maximum degree  $\Delta$ , and  $(\beta, \gamma, \lambda)$  is  $d$ -unique for all  $1 \leq d < \Delta$  (i.e. up-to- $\Delta$ -unique),

that is, the “skewed” anti-ferromagnetic case (where  $\gamma > 1$ ) on general irregular graphs.

The main technical difficulty for this case is that the MLS constant for such case was very much understudied, even in much sub-critical regimes, so there lacks a MLSI for the “easier” regime from where we can apply our boosting theorem (Theorem 1.12) for MLSI.

Another minor technical difficulty is that the marginal stability asserted by Condition 1.11 does not hold for this case in general. However, we believe that this can be circumvented because what Condition 1.11 asserts is just a sufficient condition for the product domination property. And we conjecture that for any anti-ferromagnetic two-spin system satisfying the up-to- $\Delta$ -uniqueness, the  $k$ -transformed distribution  $\mu_k$  satisfies the proper product domination property for all sufficiently large  $k$ . Verifying this conjecture would imply the optimal mixing for the above “skewed” anti-ferromagnetic two-spin systems on general irregular graphs, if a MLSI is also proved in an easier regime when assuming the up-to- $\Delta$ -uniqueness.

### 3. PRELIMINARIES

**3.1. Mixing time and modified log-Sobolev constant.** Let  $\Omega$  be a finite state space and  $(X_t)_{t \in \mathbb{Z}_{\geq 0}}$  a discrete-time Markov chain on  $\Omega$  with transition matrix  $P$ . The Markov chain  $P$  is *irreducible*, if for any  $\sigma, \tau \in \Omega$ , there exists  $t \in \mathbb{Z}_{\geq 0}$ , such that  $P^t(\sigma, \tau) > 0$ . The Markov chain  $P$  is *aperiodic*, if  $\gcd\{t \in \mathbb{Z}_{>0} \mid P^t(\sigma, \sigma) > 0\} = 1$  holds for any  $\sigma \in \Omega$ . The fundamental theorem of Markov chains says that an irreducible and aperiodic Markov chain  $P$  converges to a unique stationary distribution  $\mu$  over  $\Omega$  such that  $\mu P = \mu$ . The Markov chain  $P$  is *reversible* with respect to  $\mu$ , if the detailed balance equation  $\mu(\sigma)P(\sigma, \tau) = \mu(\tau)P(\tau, \sigma)$

holds for all  $\sigma, \tau \in \Omega$ . Such  $\mu$  satisfying the detailed balance equation is a stationary distribution of  $P$ . The mixing time of a discrete-time Markov chain  $P$  with stationary distribution  $\mu$  is defined as

$$T_{\text{mix}}(\varepsilon) = \max_{X_0 \in \Omega} \min \{t \in \mathbb{Z}_{\geq 0} \mid d_{\text{TV}}(P^t(X_0, \cdot), \mu) \leq \varepsilon\},$$

where  $d_{\text{TV}}$  denotes the *total variation distance* and is defined by

$$d_{\text{TV}}(\mu, \nu) = \frac{1}{2} \sum_{\sigma \in \Omega} |\mu(\sigma) - \nu(\sigma)|.$$

The continuous-time Markov chain is defined as follows. Let matrix  $Q : \Omega \times \Omega \rightarrow \mathbb{R}_{\geq 0}$  denote the *transition rate*. We remark that the row sum of  $Q$  may *not* be 1. The continuous-time Markov chain is a stochastic process  $(Y_t)_{t \in \mathbb{R}_{\geq 0}}$ . For any  $t > 0$ ,  $Y_t$  follows the distribution  $H_t(Y_0, \cdot)$  and  $H_t = \exp(Lt) = \sum_{k=0}^{\infty} \frac{t^k L^k}{k!}$  is called the *heat kernel*, where the *generator*  $L$  of the continuous time Markov chain is an operator defined by for any  $\psi : \Omega \rightarrow \mathbb{R}$ ,

$$L\psi(x) = \sum_{y \in \Omega} Q(x, y)(\psi(y) - \psi(x)).$$

Let  $P$  be the transition matrix of a discrete-time Markov chain. The continuous-time version  $(Y_t)_{t \in \mathbb{R}_{\geq 0}}$  of  $P$  is the continuous-time Markov chain with transition rate  $Q = P$ . In this case,  $L = P - I$ . For any  $t > 0$ ,  $Y_t$  follows the distribution  $P^N(Y_0, \cdot)$ , where  $N$  follows the Poisson distribution with mean value  $t$ . In particular, if  $P$  is the non-lazy Glauber dynamics on  $\mu$ , its continuous-time version is called the continuous-time Glauber dynamics on  $\mu$ . The mixing time of a continuous-time Markov chain is defined by

$$T_{\text{mix}}(\varepsilon) = \max_{Y_0 \in \Omega} \inf \{t \in \mathbb{R}_{\geq 0} \mid d_{\text{TV}}(Y_t, \mu) \leq \varepsilon\}.$$

Analysis of the mixing time can be done through establishing certain functional inequalities, such as Poincaré inequalities and modified log-Sobolev inequalities (MLSI). Let  $\mu$  be the stationary distribution of Markov chain  $P$  on state space  $\Omega$ , and  $(\mathbb{R}^{\Omega}, \langle \cdot, \cdot \rangle_{\mu})$  be the corresponding space with the inner-product

$$\langle f, g \rangle_{\mu} \triangleq \sum_{\sigma \in \Omega} \mu(\sigma) f(\sigma) g(\sigma), \quad \forall f, g \in \mathbb{R}^{\Omega}.$$

The *Dirichlet form*  $\mathcal{E}_P(\cdot, \cdot)$  is defined by

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

The *modified log-Sobolev (MLS) constant* is defined by

$$\rho = \rho^P(\mu) \triangleq \inf \left\{ \frac{\mathcal{E}_P(f, \log f)}{\mathbf{Ent}_{\mu}[f]} \mid f \in \mathbb{R}_{>0}^{\Omega} \text{ and } \mathbf{Ent}_{\mu}[f] \neq 0 \right\},$$

where  $\mathbf{Ent}_{\mu}[f] \triangleq \mathbb{E}_{\mu}[f \log f] - \mathbb{E}_{\mu}[f] \log \mathbb{E}_{\mu}[f]$  and we assume  $0 \log 0 = 0$ .

The following relation between modified log-Sobolev constant and mixing time was known.

**Proposition 3.1** ([BT06]). *Let  $P$  be an irreducible, aperiodic and reversible discrete-time Markov chain on finite state space  $\Omega$  with stationary distribution  $\mu$ . The mixing time of the continuous version of  $P$  satisfies*

$$T_{\text{mix}}(\varepsilon) \leq \frac{1}{\rho^P(\mu)} \left( \log \log \frac{1}{\mu_{\min}} + \log \frac{1}{2\varepsilon^2} \right),$$

where  $\mu_{\min} \triangleq \min_{\sigma \in \Omega} \mu(\sigma)$ .

**Proposition 3.2.** *Let  $P$  be an irreducible, aperiodic and reversible discrete-time Markov chain on finite state space  $\Omega$  with stationary distribution  $\mu$ . Let  $P_{zz} = \frac{1}{2}(I + P)$  denote the discrete-time lazy version of  $P$ . The mixing time of  $P_{zz}$  satisfies*

$$T_{\text{mix}}(\varepsilon) \leq \frac{4}{\rho^P(\mu)} \left( \log \log \frac{1}{\mu_{\min}} + 20 \right) \log \frac{1}{\varepsilon},$$

where  $\mu_{\min} \triangleq \min_{\sigma \in \Omega} \mu(\sigma)$ .

The proof of Proposition 3.2 is standard. We conclude the proof in Appendix A.

**3.2. Entropic independence.** Let  $\mu : \binom{[n]}{k} \rightarrow \mathbb{R}_{\geq 0}$  be a distribution over all  $k$ -subsets of  $[n]$ . We call such distribution a *homogeneous distribution*.

Let  $\Omega \subseteq \binom{[n]}{k}$  denote the support of  $\mu$ . Let  $X$  be the downward closure of  $\Omega$ . 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  $\mu$ . For any *face*  $\alpha \in X$ , let  $|\alpha|$  denote the *dimension* of  $\alpha$ . For any integer  $0 \leq j \leq k$ , let  $X(j)$  denote all the faces in  $X$  with dimension  $j$ .

**Definition 3.3** (down/up walk). Let  $X$  be the simplicial complexes generated by a homogeneous distribution  $\mu : \binom{[n]}{k} \rightarrow \mathbb{R}_{\geq 0}$ . Let  $0 \leq j < k$  be an integer.

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

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

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

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

The following definition of *entropic independence* was introduced in [AJK<sup>+</sup>21a].

**Definition 3.4** (entropic independence [AJK<sup>+</sup>21a]). Let  $\alpha \in (0, 1)$ . A distribution  $\mu$  over  $\binom{[n]}{k}$  is said to be  $(1/\alpha)$ -*entropically independent* if for any distribution  $\nu$  over  $\Omega(\mu)$ ,

$$D_{\text{KL}}(\nu D_{k \rightarrow 1} \parallel \mu D_{k \rightarrow 1}) \leq \frac{1}{\alpha k} D_{\text{KL}}(\nu \parallel \mu).$$

Let  $\mu$  be a distribution over  $\binom{[n]}{k}$ . For a set  $R \subseteq [n]$  satisfying  $\Pr_{S \sim \mu}[R \subseteq S] > 0$ , we use  $\mu^R$  to denote the *link* of  $\mu$  produced by  $R$ . This notion was used in, for examples, [Opp18, KO20, ALOV19, ALO20, AJK<sup>+</sup>21a]. Formally,  $\mu^R$  is a distribution over  $X_R = \left\{ S \setminus R \mid S \in \binom{[n]}{k} \wedge S \supseteq R \right\}$  defined by

$$(7) \quad \forall T \in X_R, \quad \mu^R(T) \propto \mu(T \uplus R).$$

The following relative entropy decay result was implied by the entropic independence property [AJK<sup>+</sup>21a]. We say a property holds for all links of  $\mu$ , if it holds for  $\mu^R$  for all  $R \subseteq [n]$  with  $\Pr_{S \sim \mu}[R \subseteq S] > 0$ .

**Theorem 3.5** ([AJK<sup>+</sup>21a, Theorem 5]). *Let  $\mu$  be a distribution over  $\binom{[n]}{k}$  and  $\alpha \in (0, 1)$ . If the  $(1/\alpha)$ -entropic independence holds for all links of  $\mu$ , then for any integer  $0 \leq j \leq k - \lceil 1/\alpha \rceil$ , any distribution  $\nu$  over  $\Omega(\mu)$ ,*

$$D_{\text{KL}}(\nu D_{k \rightarrow j} \parallel \mu D_{k \rightarrow j}) \leq \left( 1 - \kappa \left( j, k, \frac{1}{\alpha} \right) \right) D_{\text{KL}}(\nu \parallel \mu),$$

where

$$(8) \quad \kappa(j, k, c) \triangleq \frac{(k+1-j-c)^{c-\lceil c \rceil} \prod_{i=0}^{\lceil c \rceil - 1} (k-j-i)}{(k+1)^c}.$$

Let  $\mu_{(k)} = \mu$ . For any integer  $0 \leq j < k$ , let  $\mu_{(j)} = \mu_{(k)} D_{k \rightarrow j}$ . Given any function  $f = f^{(k)} : X(k) \rightarrow \mathbb{R}_{\geq 0}$ , for any integer  $0 \leq j < k$ , define  $f^{(j)} : X(j) \rightarrow \mathbb{R}_{\geq 0}$  by that  $f^{(j)} = U_{j \rightarrow k} f^{(k)}$ .

**Lemma 3.6.** Let  $\mu = \mu_{(k)}$  be a distribution over  $\binom{[n]}{k}$ . Let  $0 \leq j < k$  and  $\kappa \in (0, 1)$ . Assume that for any distribution  $\nu$  absolutely continuous with respect to  $\mu$ , it holds that  $D_{\text{KL}}(\nu D_{k \rightarrow j} \parallel \mu D_{k \rightarrow j}) \leq (1 - \kappa) D_{\text{KL}}(\nu \parallel \mu)$ . Then for any function  $f^{(k)} : X(k) \rightarrow \mathbb{R}_{\geq 0}$ ,

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

*Proof.* The proof is standard. We include it here for completeness. First note that if  $f^{(k)} = 0$ , then the lemma holds trivially. Without loss of generality, we assume  $f \neq 0$ . By homogeneity, we may further assume  $\mathbb{E}_{\mu_{(k)}} [f^{(k)}] = 1$ . Let  $\nu$  be a distribution over  $X(k)$  defined by  $\nu(\sigma) = \mu(\sigma) f^{(k)}(\sigma)$  for all  $\sigma \in X(k)$ .

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

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

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

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

**3.3. Multivariate hypergeometric distribution.** Let  $V$  be a set of  $n$  buckets, each of them has  $k$  balls. Suppose we pick  $\ell$  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  $\mathbf{a} = (a_v)_{v \in V}$  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

$$(9) \quad \Omega(\Pi_{V,k,\ell}) \triangleq \left\{ \mathbf{a} = (a_v)_{v \in V} \mid \sum_{v \in V} a_v = \ell \text{ and } \forall v \in V, a_v \in \mathbb{Z}_{\geq 0} \right\},$$

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

$$(10) \quad \Pi_{V,k,\ell}(\mathbf{a}) = \frac{\prod_{v \in V} \binom{k}{a_v}}{\binom{kn}{\ell}}.$$

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

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

*Proof.* For each bucket  $v \in V$ , let  $(v, 1), (v, 2), \dots, (v, k)$  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  $a_v$ .  $\square$

#### 4. FACTORIZATION OF ENTROPY VIA PRODUCT DOMINATION

In this section, we prove Lemma 2.5, the implication from the product domination to the uniform block factorization of entropy. We first prove an equivalence relation (Lemma 4.4) between product domination (Definition 2.1) and entropic independence (Definition 3.4); and then the uniform block factorization of entropy is established through the entropic independence (Proposition 4.6).

**4.1. Product domination and entropic independence.** Recall the following equivalent algebraic definition of entropic independence [AJK<sup>+</sup>21a].

**Definition 4.1** (algebraic definition of entropic independence [AJK<sup>+</sup>21a]). Let  $\alpha \in (0, 1)$ . A homogeneous distribution  $\pi$  over  $\binom{[n]}{k}$  is  $(1/\alpha)$ -entropically independent if

$$(11) \quad \forall (z_1, \dots, z_n) \in \mathbb{R}_{>0}^n, \quad g_\pi(z_1^\alpha, \dots, z_n^\alpha)^{\frac{1}{k\alpha}} \leq \frac{1}{k} \sum_{i=1}^n \Pr_{S \sim \pi} [i \in S] z_i.$$

Furthermore,  $\pi$  is  $(1/\alpha)$ -entropically independent on  $D \subseteq \mathbb{R}_{>0}^n$ , if above holds for all  $(z_1, \dots, z_n) \in D$ .

**Remark 4.2.** The original algebraic definition of entropic independence [AJK<sup>+</sup>21a, Theorem 4] requires (11) to hold for all  $(z_1, \dots, z_n) \in \mathbb{R}_{\geq 0}^n$ . The two definitions are equivalent by continuity.

**Remark 4.3.** Definition 4.1 is equivalent to Definition 3.4 (see [AJK<sup>+</sup>21a, Theorem 4]).

For any distribution  $\mu$  over  $\{-1, +1\}^n$ . The *homogenization* of  $\mu$ , denoted by  $\mu^{\text{hom}}$ , is a distribution over  $\binom{[n] \cup [\bar{n}]}{n}$ , where  $[\bar{n}] = \{\bar{1}, \bar{2}, \dots, \bar{n}\}$ . For any configuration  $\sigma \in \{-1, +1\}^n$ , we define

$$S_\sigma \triangleq \{i \mid \sigma_i = +1\} \cup \{\bar{i} \mid \sigma_i = -1\},$$

then the homogenization  $\mu^{\text{hom}}$  is defined by

$$\forall \sigma \in \{-1, +1\}^n, \quad \mu^{\text{hom}}(S_\sigma) \triangleq \mu(\sigma),$$

and  $\mu^{\text{hom}}(T) = 0$  for any  $T$  that cannot be expressed as  $S_\sigma$  for some  $\sigma \in \{-1, +1\}^n$ .

The following lemma gives the relation between product domination and entropic independence.

**Lemma 4.4.** Let  $\mu$  be a distribution  $\mu$  over  $\{-1, +1\}^n$  and  $\alpha \in (0, 1)$ . Let  $D \subseteq \mathbb{R}_{>0}^n$  and define

$$D^{\text{hom}} = \left\{ (z_1, z_2, \dots, z_n, z_{\bar{1}}, z_{\bar{2}}, \dots, z_{\bar{n}}) \in \mathbb{R}_{>0}^{2n} \mid \left( \frac{z_1}{z_{\bar{1}}}, \frac{z_2}{z_{\bar{2}}}, \dots, \frac{z_n}{z_{\bar{n}}} \right) \in D \right\}.$$

$\mu$  is  $(1/\alpha)$ -product dominated on  $D$  if and only if  $\mu^{\text{hom}}$  is  $(1/\alpha)$ -entropically independent on  $D^{\text{hom}}$ .

In particular,  $\mu$  is  $(1/\alpha)$ -product dominated if and only if  $\mu^{\text{hom}}$  is  $(1/\alpha)$ -entropically independent.

**Remark 4.5.** The product domination is closely related to a specific Brascamp-Lieb type inequality [BCLM11, BCC<sup>+</sup>21], which says that there is a constant  $C > 0$  such that

$$(12) \quad \forall \{f_u\}_{u \in [n]}, \text{ where } f_u : \Omega(\mu_u) \rightarrow \mathbb{R}, \quad \mathbb{E}_{\sigma \sim \mu} \left[ \prod_{u \in V} f_u(\sigma_u) \right] \leq \left( \prod_{u \in V} \mathbb{E}_{\sigma_u \sim \mu_u} [|f_u(\sigma_u)|^C] \right)^{1/C}.$$

By taking  $C = 1/\alpha$  and  $f_u(x) = z_u^{\frac{x+1}{2C}}$ , it is straightforward to see that the product domination of  $\mu$  is a special case of (12). It was known that the entropic independence of  $\mu^{\text{hom}}$  is equivalent to the Brascamp-Lieb type inequality in (12), which can be proved via the subadditivity of entropy [CC09, BCC<sup>+</sup>21]. As a consequence, the entropic independence of  $\mu^{\text{hom}}$  implies the product domination of  $\mu$ . To the best of our knowledge, the inverse part is not known before and is established here for the first time.

*Proof.* Denote  $\pi = \mu^{\text{hom}}$ . We first prove the sufficiency. Note that we have

$$g_\pi(x_1, x_2, \dots, x_n, 1, 1, \dots, 1) = g_\mu(x_1, x_2, \dots, x_n),$$

$\Pr_{S \sim \pi} [i \in S] = \mu_i(+1)$  and  $\Pr_{S \sim \pi} [\bar{i} \in S] = \mu_i(-1)$ . Therefore,  $\mu$  being  $(1/\alpha)$ -product dominated on  $D$  means that the generating function of  $\pi = \mu^{\text{hom}}$  satisfies that for all  $(x_1, x_2, \dots, x_n) \in D$ ,

$$g_\pi(x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha, 1, \dots, 1)^{\frac{1}{\alpha}} \leq \prod_{i=1}^n \left( \Pr_{S \sim \pi} [i \in S] x_i + \Pr_{S \sim \pi} [\bar{i} \in S] \right).$$

Hence, for any  $(z_1, z_2, \dots, z_n, z_{\bar{1}}, z_{\bar{2}}, \dots, z_{\bar{n}}) \in D^{\text{hom}}$ ,

$$g_\pi \left( \left( \frac{z_1}{z_{\bar{1}}} \right)^\alpha, \left( \frac{z_2}{z_{\bar{2}}} \right)^\alpha, \dots, \left( \frac{z_n}{z_{\bar{n}}} \right)^\alpha, 1, \dots, 1 \right)^{\frac{1}{\alpha}} \leq \prod_{i=1}^n \left( \Pr_{S \sim \pi} [i \in S] \frac{z_i}{z_{\bar{i}}} + \Pr_{S \sim \pi} [\bar{i} \in S] \right).$$

Multiplying both sides by  $(z_{\bar{1}} z_{\bar{2}} \dots z_{\bar{n}})$  gives

$$(13) \quad g_\pi(z_1^\alpha, z_2^\alpha, \dots, z_n^\alpha, z_{\bar{1}}^\alpha, z_{\bar{2}}^\alpha, \dots, z_{\bar{n}}^\alpha)^{\frac{1}{\alpha}} \leq \prod_{i=1}^n \left( \Pr_{S \sim \pi} [i \in S] z_i + \Pr_{S \sim \pi} [\bar{i} \in S] z_{\bar{i}} \right).$$

Applying AM-GM inequality, for any  $(z_1, z_2, \dots, z_n, z_{\bar{1}}, z_{\bar{2}}, \dots, z_{\bar{n}}) \in D^{\text{hom}}$ ,

$$\begin{aligned} g_\pi(z_1^\alpha, z_2^\alpha, \dots, z_n^\alpha, z_{\bar{1}}^\alpha, z_{\bar{2}}^\alpha, \dots, z_{\bar{n}}^\alpha)^{\frac{1}{n\alpha}} &\leq \prod_{i=1}^n \left( \Pr_{S \sim \pi} [i \in S] z_i + \Pr_{S \sim \pi} [\bar{i} \in S] z_{\bar{i}} \right)^{\frac{1}{n}} \\ (\text{by AM-GM}) &\leq \frac{1}{n} \sum_{i=1}^n \left( \Pr_{S \sim \pi} [i \in S] z_i + \Pr_{S \sim \pi} [\bar{i} \in S] z_{\bar{i}} \right), \end{aligned}$$

which implies that  $\pi = \mu^{\text{hom}}$  is  $(1/\alpha)$ -entropically independent on  $D^{\text{hom}}$  by Definition 4.1.

Next, we prove the necessity. Fix arbitrary  $(x_1, x_2, \dots, x_n) \in D$ . Define  $z_1, \dots, z_n$  and  $z_{\bar{1}}, \dots, z_{\bar{n}}$  respectively as

$$\forall i \in [n], \quad z_i = \frac{x_i}{x_i \Pr_{S \sim \pi} [i \in S] + \Pr_{S \sim \pi} [\bar{i} \in S]} \quad \text{and} \quad z_{\bar{i}} = \frac{1}{x_i \Pr_{S \sim \pi} [i \in S] + \Pr_{S \sim \pi} [\bar{i} \in S]}.$$

It is straightforward to verify that  $(z_1, \dots, z_n, z_{\bar{1}}, \dots, z_{\bar{n}}) \in D^{\text{hom}}$  and  $\Pr_{S \sim \pi} [i \in S] z_i + \Pr_{S \sim \pi} [\bar{i} \in S] z_{\bar{i}} = 1$ . Therefore,  $\pi = \mu^{\text{hom}}$  being  $(1/\alpha)$ -entropically independent means that the generating function of  $\pi$  satisfies

$$g_\pi(z_1^\alpha, \dots, z_n^\alpha, z_{\bar{1}}^\alpha, \dots, z_{\bar{n}}^\alpha)^{\frac{1}{\alpha}} \leq \left( \frac{1}{n} \sum_{i=1}^n \left( \Pr_{S \sim \pi} [i \in S] z_i + \Pr_{S \sim \pi} [\bar{i} \in S] z_{\bar{i}} \right) \right)^n = 1.$$

Note that  $\prod_{i=1}^n (\Pr_{S \sim \pi} [i \in S] z_i + \Pr_{S \sim \pi} [\bar{i} \in S] z_{\bar{i}}) = 1$ . Therefore,

$$g_\pi(z_1^\alpha, \dots, z_n^\alpha, z_{\bar{1}}^\alpha, \dots, z_{\bar{n}}^\alpha)^{\frac{1}{\alpha}} \leq 1 = \prod_{i=1}^n \left( \Pr_{S \sim \pi} [i \in S] z_i + \Pr_{S \sim \pi} [\bar{i} \in S] z_{\bar{i}} \right).$$

Dividing both sides by  $(z_{\bar{1}} z_{\bar{2}} \dots z_{\bar{n}})$  gives

$$g_\pi \left( \left( \frac{z_1}{z_{\bar{1}}} \right)^\alpha, \left( \frac{z_2}{z_{\bar{2}}} \right)^\alpha, \dots, \left( \frac{z_n}{z_{\bar{n}}} \right)^\alpha, 1, \dots, 1 \right)^{\frac{1}{\alpha}} \leq \prod_{i=1}^n \left( \Pr_{S \sim \pi} [i \in S] \frac{z_i}{z_{\bar{i}}} + \Pr_{S \sim \pi} [\bar{i} \in S] \right).$$

Note that  $\frac{z_i}{z_{\bar{i}}} = x_i$  and recall that  $g_\pi(x_1, x_2, \dots, x_n, 1, \dots, 1) = g_\mu(x_1, x_2, \dots, x_n)$ ,  $\Pr_{S \sim \pi} [i \in S] = \mu_i(+1)$  and  $\Pr_{S \sim \pi} [\bar{i} \in S] = \mu_i(-1)$ . We have the following holds for all  $(x_1, x_2, \dots, x_n) \in D$

$$g_\mu(x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha)^{\frac{1}{\alpha}} \leq \prod_{i=1}^n (\mu_i(+1)x_i + \mu_i(-1)),$$

which implies  $\mu$  is  $(1/\alpha)$ -product dominated on  $D$ . □

**4.2. Entropic independence and block factorization of entropy.** We now use the entropic independence to obtain the uniform block factorization of entropy. Recall that the link of a distribution is defined in (7).

**Proposition 4.6.** *Let  $\mu$  be a distribution over  $\{-1, +1\}^n$  and  $\pi = \mu^{\text{hom}}$  over  $\binom{[n] \cup [\bar{n}]}{n}$  its homogenization. Let  $\alpha \in (0, 1)$ . If  $(1/\alpha)$ -entropic independence holds for all links of  $\pi$ , then for any  $\lceil 1/\alpha \rceil \leq \ell \leq n$ ,  $\mu$  satisfies  $\ell$ -uniform block factorization of entropy with  $C = \kappa(n - \ell, n, 1/\alpha)^{-1}$ , where  $\kappa(\cdot)$  is defined in (8).*

The proof of Proposition 4.6 is standard. We include it here for completeness.

*Proof of Proposition 4.6.* Fix any function  $f : \Omega(\mu) \rightarrow \mathbb{R}_{\geq 0}$ . We construct  $f^{(n)} : \Omega(\pi) \rightarrow \mathbb{R}_{\geq 0}$  as that  $f^{(n)}(S_\sigma) = f(\sigma)$  for all  $\sigma \in \Omega(\mu)$ , where  $S_\sigma = \{i \mid \sigma_i = +1\} \cup \{\bar{i} \mid \sigma_i = -1\}$ . Let  $X$  denote the simplicial complexes generated by  $\pi$ . Let  $U$ . and  $D$ . denote the up walk and down walk on  $X$  (Definition 3.3). Let  $\pi_{(n)} = \pi$  and  $\pi_{(j)} = \pi_{(n)} D_{n \rightarrow j}$  for all  $0 \leq j < n$ . Let  $f^{(j)} = U_{j \rightarrow n} f^{(n)}$  for all  $0 \leq j < n$ .

Recall the notation  $\mu[\mathbf{Ent}_S [f]]$  used in Definition 2.4:

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

The following lemma is proved in [CLV21] (see the proof of Lemma 2.6 in the full version of [CLV21]).

**Lemma 4.7** ([CLV21]). *Let  $\mu$  be a distribution over  $\{-1, +1\}^n$ . For any  $0 \leq j \leq n$ , it holds that*

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

Note that all conditional marginal distributions induced by  $\pi$  are  $(1/\alpha)$ -entropically independent. By Theorem 3.5 and Lemma 3.6, for any  $f : \Omega(\mu) \rightarrow \mathbb{R}_{\geq 0}$  and  $0 \leq j \leq n - \lceil 1/\alpha \rceil$ ,

$$(15) \quad \mathbf{Ent}_{\pi_{(j)}} [f^{(j)}] \leq (1 - \kappa(j, n, 1/\alpha)) \mathbf{Ent}_{\pi_{(n)}} [f^{(n)}],$$

where  $\kappa(\cdot)$  is defined in (8). Hence, for any  $\lceil 1/\alpha \rceil \leq \ell \leq n$ ,

$$\begin{aligned} \mathbf{Ent}_\mu [f] &\stackrel{(*)}{=} \mathbf{Ent}_{\pi_{(n)}} [f^{(n)}] = \mathbf{Ent}_{\pi_{(n)}} [f^{(n)}] - \mathbf{Ent}_{\pi_{(n-\ell)}} [f^{(n-\ell)}] + \mathbf{Ent}_{\pi_{(n-\ell)}} [f^{(n-\ell)}] \\ &\text{(by (15) and Lemma 4.7)} \leq \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{[n]}{\ell}} \mu[\mathbf{Ent}_S [f]] + (1 - \kappa(n - \ell, n, 1/\alpha)) \mathbf{Ent}_{\pi_{(n)}} [f^{(n)}] \\ &\stackrel{(\star)}{=} \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{[n]}{\ell}} \mu[\mathbf{Ent}_S [f]] + (1 - \kappa(n - \ell, n, 1/\alpha)) \mathbf{Ent}_\mu [f] \end{aligned}$$

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

$$\mathbf{Ent}_\mu [f] \leq \frac{\kappa(n - \ell, n, \frac{1}{\alpha})^{-1}}{\binom{n}{\ell}} \sum_{S \in \binom{[n]}{\ell}} \mu[\mathbf{Ent}_S [f]]. \quad \square$$

**4.3. Block factorization of entropy via product domination.** We are now ready to prove Lemma 2.5.

*Proof of Lemma 2.5.* We interpret  $\mu$  as a distribution over the power set  $2^{[n]}$ .

Let  $\pi = \mu^{\text{hom}}$  over  $\binom{[n] \cup [\bar{n}]}{n}$  be its homogenization. There is a one-to-one correspondence between conditional distribution in  $\mu$  and links of  $\pi$ . Recall the link defined in (7). Fix any link  $\pi^R$  of  $\pi$ . It is straightforward to verify that there exists a partial configuration  $\sigma \in \Omega(\mu_\Lambda)$  such that

$$\left( \mu_{V \setminus \Lambda}^\sigma \right)^{\text{hom}} = \pi^R.$$

By Lemma 4.4, assumption of Lemma 2.5 and the monotonicity of entropic independence (see Definition 3.4), the  $\lceil 1/\alpha \rceil$ -entropic independence holds for all links of  $\pi$ . Due to Proposition 4.6, for any  $\lceil 1/\alpha \rceil \leq \ell \leq n$ , the distribution  $\mu$  satisfies  $\ell$ -uniform block factorization of entropy with parameter

$$C = \frac{1}{\kappa(n - \ell, n, \lceil 1/\alpha \rceil)} = \binom{n}{\lceil 1/\alpha \rceil} / \binom{\ell}{\lceil 1/\alpha \rceil} \leq \left( \frac{en}{\lceil 1/\alpha \rceil} \right)^{\lceil 1/\alpha \rceil} / \left( \frac{\ell}{\lceil 1/\alpha \rceil} \right)^{\lceil 1/\alpha \rceil} = \left( \frac{en}{\ell} \right)^{\lceil 1/\alpha \rceil} \leq \left( \frac{en}{\ell} \right)^{\frac{1}{\alpha} + 1}. \quad \square$$



## 5. PRODUCT DOMINATION FROM marginally STABLE SPECTRAL INDEPENDENCE

In this section, we prove Lemma 2.3, establishing of the product domination property (Definition 2.1) through the spectral independence and marginal stability properties guaranteed in Condition 2.2.

We first define the complete spectral independence, which will be used in the following sections.

**Definition 5.1** (complete spectral independence). Let  $\eta > 1$  and  $\varepsilon > 0$ . A distribution  $\mu$  over  $\{-1, +1\}^n$  is said to be  $(\eta, \varepsilon)$ -completely spectrally independent, if  $(\lambda * \mu)$  is  $\eta$ -spectrally independent for any  $\lambda \in (0, 1 + \varepsilon]^n$ .

Define a function  $F_{\mu, \alpha} : \mathbb{R}_{>0}^n \rightarrow \mathbb{R}$  by

$$(16) \quad F_{\mu, \alpha}(z_1, z_2, \dots, z_n) \triangleq \frac{g_{\mu}(z_1^{\alpha}, z_2^{\alpha}, \dots, z_n^{\alpha})^{\frac{1}{\alpha}}}{\prod_{i=1}^n (\mu_i(+1)z_i + \mu_i(-1))}.$$

It is not hard to see that  $F_{\mu, \alpha} \leq 1$  implies that  $\mu$  is  $(1/\alpha)$ -product dominated. Moreover, the following lemmas transform Condition 2.2 to the following conditions regarding function  $F_{\mu, \alpha}$ .

**Lemma 5.2.** *Let  $\eta > 1$  and  $\varepsilon > 0$ . If a distribution  $\mu$  over  $\{-1, +1\}^n$  is  $(\eta, \varepsilon)$ -completely spectrally independent, then for any  $0 < \alpha \leq 1/(2\eta)$ , it holds that  $F_{\mu, \alpha}(x) \leq 1$  for all  $x \in (0, (1 + \varepsilon)^{1/\alpha}]^n$ .*

**Lemma 5.3.** *Let  $\zeta > 1$ . If a distribution  $\mu$  over  $\{-1, +1\}^n$  is  $\zeta$ -marginally stable, then for any  $\alpha \in (0, 1)$ , any  $x \in \mathbb{R}_{>0}^n$ , any  $i \in [n]$ , if  $x_i \geq (2\zeta)^{\frac{1}{1-\alpha}}$ , then*

$$\left. \frac{\partial F_{\mu, \alpha}}{\partial z_i} \right|_{z=x} \leq 0.$$

The complete spectral independence implies the product domination in  $(0, (1 + \varepsilon)^{1/\alpha}]^n$  through Lemma 5.2, which is extended to all  $\mathbb{R}_{>0}^n$  through the monotonicity in Lemma 5.3 implied by the marginal stability.

The proof of Lemma 5.2 and Lemma 5.3 are given in Section 5.1 and Section 5.2 respectively.

It remains to ensure the complete spectral independence and the marginal stability in Condition 2.2 closed under pinning, which is straightforward because their definitions already consider all pinnings.

**Fact 5.4.** *Let  $\eta > 1, \zeta$  and  $\varepsilon > 0$ . If a distribution  $\mu$  over  $\{-1, +1\}^n$  is  $(\eta, \varepsilon)$ -completely spectrally independent and  $\zeta$ -marginally stable, then these properties also hold for  $\mu_{[n] \setminus \Lambda}^{\sigma}$  for arbitrary  $\Lambda \subseteq [n]$  and  $\sigma \in \Omega(\mu_{\Lambda})$ .*

*Proof of Lemma 2.3.* Denote  $\nu = \mu_{[n] \setminus \Lambda}^{\sigma}$ . Without loss of generality, suppose  $[n] \setminus \Lambda = [m] = \{1, 2, \dots, m\}$ . By the definition of the function  $F_{\nu, \alpha}$ , it suffices to show that

$$\forall x \in \mathbb{R}_{>0}^m, \quad F_{\nu, \alpha}(x) \leq 1.$$

Denote  $D = (0, (1 + \varepsilon)^{1/\alpha}]^m$ . By Fact 5.4 and Lemma 5.2, since  $\alpha \leq 1/(2\eta)$ ,  $F_{\nu, \alpha}(x) \leq 1$  for any  $x \in D$ .

Therefore, it remains to take care of those  $x \notin D$ . Fix an arbitrary  $x \in \mathbb{R}_{>0}^m \setminus D$ . Define  $\tilde{x} \in \mathbb{R}_{>0}^m$  as that  $\tilde{x}_i = \min\{x_i, (1 + \varepsilon)^{1/\alpha}\}$  for all  $i \in [m]$ . Obviously  $\tilde{x} \in D$ , and hence  $F_{\nu, \alpha}(\tilde{x}) \leq 1$ . We only need to show that  $F_{\nu, \alpha}(x) \leq F_{\nu, \alpha}(\tilde{x})$ .

Denote  $M = \{i \in [m] \mid x_i > (1 + \varepsilon)^{1/\alpha}\}$ . By mean value theorem, there exists  $\theta \in (0, 1)$  such that

$$F_{\nu, \alpha}(x) - F_{\nu, \alpha}(\tilde{x}) = \left\langle (x - \tilde{x}), \nabla F_{\nu, \alpha}(\theta x + (1 - \theta)\tilde{x}) \right\rangle = \sum_{i \in M} (x_i - \tilde{x}_i) \left. \frac{\partial F}{\partial z_i} \right|_{z=\theta x + (1-\theta)\tilde{x}},$$

where the last equation holds because  $x_i = \tilde{x}_i$  for all  $i \in [m] \setminus M$ . Fix any  $i \in M$ . It holds that  $x_i > \tilde{x}_i = (1 + \varepsilon)^{1/\alpha}$ , thus  $(\theta x + (1 - \theta)\tilde{x})_i > (1 + \varepsilon)^{1/\alpha}$ . Note that by the choice of  $\alpha$  in (6), it holds that

$$(17) \quad (2\zeta)^{\frac{1}{1-\alpha}} \leq (1 + \varepsilon)^{1/\alpha} < (\theta x + (1 - \theta)\tilde{x})_i.$$

Combining (17) with Fact 5.4 and Lemma 5.3, we have

$$F_{\nu, \alpha}(x) - F_{\nu, \alpha}(\tilde{x}) = \sum_{i \in M} (x_i - \tilde{x}_i) \left. \frac{\partial F}{\partial z_i} \right|_{z=\theta x + (1-\theta)\tilde{x}} \leq 0.$$

Hence  $F_{v,\alpha}(x) \leq F_{v,\alpha}(\tilde{x}) \leq 1$ . This concludes the proof.  $\square$

**5.1. Fractional log-concavity from complete spectral independence (proof of Lemma 5.2).** The following lemma was implicit in [AJK<sup>+</sup>21b, AASV21, AJK<sup>+</sup>21a].

**Lemma 5.5.** *Let  $\eta > 1$  and  $\varepsilon > 0$ . If a distribution  $\mu$  over  $\{-1, +1\}^n$  is  $(\eta, \varepsilon)$ -completely spectrally independent, then for any  $0 < \alpha \leq 1/(2\eta)$ , the function  $\log g_{\mu^{\text{hom}}}(z_1^\alpha, \dots, z_n^\alpha, z_{\bar{1}}^\alpha, \dots, z_{\bar{n}}^\alpha)$  is concave on*

$$\Lambda_{\alpha,\varepsilon} \triangleq \left\{ (z_1, \dots, z_n, z_{\bar{1}}, \dots, z_{\bar{n}}) \in \mathbb{R}_{>0}^{2n} \mid \forall i \in [n], 0 < z_i \leq z_{\bar{i}}(1 + \varepsilon)^{1/\alpha} \right\},$$

where  $\mu^{\text{hom}}$  is  $\mu$ 's homogenization over  $\binom{[n] \cup [\bar{n}]}{n}$  and  $g_{\mu^{\text{hom}}}$  is its generating function.

The concavity property in Lemma 5.5 is called the “ $\alpha$ -fractional log-concavity” of  $\mu^{\text{hom}}$  [AASV21, AJK<sup>+</sup>21a].

*Proof of Lemma 5.2.* Fix  $0 < \alpha \leq 1/(2\eta)$ . Define the 1-homogeneous function as

$$f(z_1, \dots, z_n, z_{\bar{1}}, \dots, z_{\bar{n}}) = g_{\mu^{\text{hom}}}(z_1^\alpha, \dots, z_n^\alpha, z_{\bar{1}}^\alpha, \dots, z_{\bar{n}}^\alpha)^{\frac{1}{\alpha n}}.$$

Note that  $f$  is concave on  $\Lambda_{\alpha,\varepsilon}$ . This is because  $g_{\mu^{\text{hom}}}(z_1^\alpha, \dots, z_n^\alpha, z_{\bar{1}}^\alpha, \dots, z_{\bar{n}}^\alpha)$  is  $\alpha n$ -homogeneous, and by Lemma 5.5, it is also log-concave as a function of  $(z_1, \dots, z_n, z_{\bar{1}}, \dots, z_{\bar{n}}) \in \Lambda_{\alpha,\varepsilon}$ , which implies the concavity of  $f$  on  $\Lambda_{\alpha,\varepsilon}$  by [AJK<sup>+</sup>21a, Lemma 25]. Therefore, for any  $(z_1, \dots, z_n, z_{\bar{1}}, \dots, z_{\bar{n}}) \in \Lambda_{\alpha,\varepsilon}$ ,

$$\begin{aligned} f(z_1, \dots, z_n, z_{\bar{1}}, \dots, z_{\bar{n}}) &\leq f(1, 1, \dots, 1) + \sum_{i \in [n] \cup [\bar{n}]} \frac{\partial f}{\partial z_i}(1, 1, \dots, 1)(z_i - 1) \\ &= \sum_{i \in [n] \cup [\bar{n}]} \frac{\partial f}{\partial z_i}(1, 1, \dots, 1)z_i, \end{aligned}$$

where the equation holds since  $f$  is 1-homogeneous. Note that  $\mu^{\text{hom}}$  is a distribution over  $\binom{[n] \cup [\bar{n}]}{n}$  and  $\frac{\partial f}{\partial z_i}(1, 1, \dots, 1) = \frac{1}{n} \Pr_{S \sim \mu^{\text{hom}}} [i \in S]$  for all  $i \in [n] \cup [\bar{n}]$ . Therefore, for any  $(z_1, \dots, z_n, z_{\bar{1}}, \dots, z_{\bar{n}}) \in \Lambda_{\alpha,\varepsilon}$ ,

$$g_{\mu^{\text{hom}}}(z_1^\alpha, \dots, z_n^\alpha, z_{\bar{1}}^\alpha, \dots, z_{\bar{n}}^\alpha)^{\frac{1}{\alpha n}} = f(z_1, \dots, z_n, z_{\bar{1}}, \dots, z_{\bar{n}}) \leq \frac{1}{n} \sum_{i=1}^n \left( \Pr_{S \sim \mu^{\text{hom}}} [i \in S] + \Pr_{S \sim \mu^{\text{hom}}} [\bar{i} \in S] \right).$$

This means that  $\mu^{\text{hom}}$  is  $(1/\alpha)$ -entropically independent over  $\Lambda_{\alpha,\varepsilon}$ . Then by Lemma 4.4,  $\mu$  is  $(1/\alpha)$ -product dominated on  $\left(0, (1 + \varepsilon)^{\frac{1}{\alpha}}\right)^n$ . By definition of product domination, for any  $(x_1, \dots, x_n) \in \left(0, (1 + \varepsilon)^{\frac{1}{\alpha}}\right)^n$ ,

$$F_{\mu,\alpha}(x_1, x_2, \dots, x_n) = \frac{g_{\mu}(x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha)^{\frac{1}{\alpha}}}{\prod_{i=1}^n (\mu_i(+1)x_i + \mu_i(-1))} \leq 1. \quad \square$$

It remains to formally verify Lemma 5.5. A variant of the lemma was proved in [AJK<sup>+</sup>21b] assuming the spectral domination property for correlation matrix. Lemma 5.5 can be proved in the same way.

**Definition 5.6** (signed correlation matrix [AJK<sup>+</sup>21b]). Let  $\mu$  be a distribution over  $\{-1, +1\}^n$ . The correlation matrix  $\Psi_\mu^{\text{Cor}} \in \mathbb{R}_{\geq 0}^{n \times n}$  is defined by

$$\forall i, j \in [n], \quad \Psi_\mu^{\text{Cor}}(i, j) = \begin{cases} \mu_i(-1) & \text{if } i = j; \\ \mu_j^{i \leftarrow +1}(+1) - \mu_j(+1) & \text{if } i \neq j \text{ and } +1 \in \Omega(\mu_i); \\ 0 & \text{otherwise.} \end{cases}$$

**Definition 5.7** (signed influence matrix [ALO20]). Let  $\mu$  be a distribution over  $\{-1, +1\}^n$ . The signed influence matrix  $\Psi_\mu^{\text{Inf}} \in \mathbb{R}^{n \times n}$  is defined by

$$\forall i, j \in [n], \quad \Psi_\mu^{\text{Inf}}(i, j) = \begin{cases} \mu_j^{i \leftarrow +1}(+1) - \mu_j^{i \leftarrow -1}(+1) & \text{if } i \neq j \text{ and } \Omega(\mu_i) = \{-1, +1\}; \\ 0 & \text{otherwise.} \end{cases}$$

**Remark 5.8.** The influence matrix  $\Psi_\mu$  in Definition 1.9 is satisfies that  $\Psi_\mu(i, j) = |\Psi_\mu^{\text{Inf}}(i, j)|$ .

**Lemma 5.9** ([HJ12, Corollary 8.1.19]). *Let  $A, B \in \mathbb{R}^{n \times n}$  and suppose  $B$  is non-negative. If  $|A| \leq B$ , then  $\rho(A) \leq \rho(|A|) \leq \rho(B)$ .*

The following relation between influence matrix and correlation matrix was proved in [AASV21].

**Lemma 5.10** ([AASV21]). *The spectrum of  $\Psi_{\mu^{\text{Cor}}}^{\text{Cor}}$  is the union of  $\{\lambda_i + 1\}_{1 \leq i \leq n}$  and  $n$  copies of 0, where  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$  are eigenvalues of  $\Psi_\mu^{\text{Inf}}$ .*

*Proof of Lemma 5.5.* By the proof of Proposition 19 in [AJK<sup>+</sup>21b], we only need to verify that for any  $\vec{v} = (v_1, \dots, v_n, v_{\bar{1}}, \dots, v_{\bar{n}}) \in \Lambda_{\alpha, \varepsilon}$ , denoted  $\vec{v}^\alpha = (v_1^\alpha, \dots, v_n^\alpha, v_{\bar{1}}^\alpha, \dots, v_{\bar{n}}^\alpha)$ , it holds that

$$\lambda_{\max} \left( \Psi_{\vec{v}^\alpha * \mu^{\text{hom}}}^{\text{Cor}} \right) \leq \frac{1}{\alpha}.$$

Note that  $\vec{v}^\alpha * \mu^{\text{hom}}$  is the homogenization of  $u * \mu$ , where  $u = (u_i)_{1 \leq i \leq n}$  has  $u_i = \left(\frac{v_i}{v_{\bar{i}}}\right)^\alpha \leq (1 + \varepsilon)$  for all  $i \in [n]$ .

Therefore, we have  $\lambda_{\max} \left( \Psi_{\vec{v}^\alpha * \mu^{\text{hom}}}^{\text{Cor}} \right) = \lambda_{\max} \left( \Psi_{u * \mu}^{\text{Inf}} \right) + 1$  by Lemma 5.10; and  $\lambda_{\max} \left( \Psi_{u * \mu}^{\text{Inf}} \right) \leq \eta$  by Lemma 5.9 and the  $(\eta, \varepsilon)$ -complete spectral independence of  $\mu$ . Together, we have  $\lambda_{\max} \left( \Psi_{\vec{v}^\alpha * \mu^{\text{hom}}}^{\text{Cor}} \right) \leq \eta + 1 \leq \frac{1}{\alpha}$ .  $\square$

**5.2. Monotonicity from marginal stability (proof of Lemma 5.3).** For any  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}_{>0}^n$ ,

$$\left. \frac{\partial F_{\mu, \alpha}}{\partial z_i} \right|_{z=x} = \frac{F_{\mu, \alpha}(x)}{x_i} \left( \frac{x_i^\alpha}{g_\mu(x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha)} \left. \frac{\partial g_\mu}{\partial z_i} \right|_{z=(x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha)} - \frac{\mu_i(+1)x_i}{\mu_i(+1)x_i + \mu_i(-1)} \right).$$

Observe that

$$\frac{x_i^\alpha}{g_\mu(x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha)} \left. \frac{\partial g_\mu}{\partial z_i} \right|_{z=(x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha)} = \frac{\sum_{\sigma: \sigma_i = +1} \mu(\sigma) \prod_{j: \sigma_j = +1} x_j^\alpha}{\sum_{\sigma} \mu(\sigma) \prod_{j: \sigma_j = +1} x_j^\alpha} = (x^\alpha * \mu)_i(+1),$$

where  $x^\alpha = (x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha)$ . Furthermore, it is safe to focus on the case where  $\mu_i(+1) > 0$  and  $\mu_i(-1) > 0$ , because otherwise it would have already implied that  $\left. \frac{\partial F_{\mu, \alpha}}{\partial z_i} \right|_{z=x} = 0$  for all  $x \in \mathbb{R}_{>0}^n$ . Therefore, we have

$$\left( \frac{x_i}{F_{\mu, \alpha}} \right) \cdot \left. \frac{\partial F_{\mu, \alpha}}{\partial z_i} \right|_{z=x} = (x^\alpha * \mu)_i(+1) - \frac{\mu_i(+1)x_i}{\mu_i(+1)x_i + \mu_i(-1)} = \left( 1 + \frac{(x^\alpha * \mu)_i(-1)}{(x^\alpha * \mu)_i(+1)} \right)^{-1} - \left( 1 + \frac{\mu_i(-1)}{\mu_i(+1)x_i} \right)^{-1}.$$

Note that  $(1 + x)^{-1}$  is decreasing in  $x > 0$ . To prove  $\left. \frac{\partial F_{\mu, \alpha}}{\partial z_i} \right|_{z=x} \leq 0$ , it suffices to verify  $\frac{(x^\alpha * \mu)_i(-1)}{(x^\alpha * \mu)_i(+1)} \geq \frac{\mu_i(-1)}{\mu_i(+1)x_i}$ , or equivalently,  $\frac{(x^\alpha * \mu)_i(+1)}{(x^\alpha * \mu)_i(-1)} \leq x_i \frac{\mu_i(+1)}{\mu_i(-1)}$ . Indeed, it holds that

$$\begin{aligned} \frac{(x^\alpha * \mu)_i(+1)}{(x^\alpha * \mu)_i(-1)} &= \frac{\sum_{\sigma \in \Omega(\mu_{[n] \setminus \{i\}})} (x^\alpha * \mu)_{[n] \setminus \{i\}}(\sigma) \cdot (x^\alpha * \mu)_i^\sigma(+1)}{\sum_{\sigma \in \Omega(\mu_{[n] \setminus \{i\}})} (x^\alpha * \mu)_{[n] \setminus \{i\}}(\sigma) \cdot (x^\alpha * \mu)_i^\sigma(-1)} \\ &= \frac{\sum_{\sigma \in \Omega(\mu_{[n] \setminus \{i\}})} (x^\alpha * \mu)_{[n] \setminus \{i\}}(\sigma) \cdot \mu_i^\sigma(+1) \cdot x_i^\alpha}{\sum_{\sigma \in \Omega(\mu_{[n] \setminus \{i\}})} (x^\alpha * \mu)_{[n] \setminus \{i\}}(\sigma) \cdot \mu_i^\sigma(-1)} \\ &\leq x_i^\alpha \max_{\sigma \in \Omega(\mu_{[n] \setminus \{i\}})} \frac{\mu_i^\sigma(+1)}{\mu_i^\sigma(-1)}. \end{aligned}$$

In above, we enumerate all  $\sigma$  in  $\Omega(\mu_{[n] \setminus \{i\}})$  because  $\mu$  and  $(x^\alpha * \mu)$  have the same support. By Definition 1.10, it holds that for all possible partial configuration  $\sigma$ , we have  $R_i^\sigma + R_i^\sigma / R_i \leq 2\zeta$ , where  $R_i^\sigma = \mu_i^\sigma(+1) / \mu_i^\sigma(-1)$ ,

which implies  $R_i^\sigma \leq 2\zeta(1 + 1/R_i)^{-1} = 2\zeta \cdot \mu_i(+1)$ . Hence, it holds that,

$$\frac{(x^\alpha * \mu)_i(+1)}{(x^\alpha * \mu)_i(-1)} \leq x_i^\alpha \max_{\sigma \in \Omega(\mu|_{[n] \setminus \{i\}})} \frac{\mu_i^\sigma(+1)}{\mu_i^\sigma(-1)} \leq x_i^\alpha \cdot 2\zeta \cdot \mu_i(+1) \leq x_i \frac{\mu_i(+1)}{\mu_i(-1)},$$

where the last inequality comes from the fact that  $x_i^{\alpha-1} \cdot 2\zeta \leq 1$ , which is guaranteed by  $x_i \geq (2\zeta)^{\frac{1}{1-\alpha}}$ .

## 6. BOOSTING MODIFIED LOG-SOBOLEV CONSTANT VIA $k$ -TRANSFORMATION

In this section, we prove Lemma 2.7. The key step is to establish the ‘‘magnetized block factorization of entropy’’. Let  $\theta \in (0, 1)$  and  $\pi = \theta * \mu$ . Formally, for every  $\sigma \in \{-1, +1\}^V$ ,

$$(18) \quad \pi(\sigma) = \frac{\mu(\sigma)\theta^{\|\sigma\|_+}}{Z_\pi}, \quad \text{where } Z_\pi \triangleq \sum_{\sigma \in \{-1, +1\}^V} \mu(\sigma)\theta^{\|\sigma\|_+} \text{ and } \|\sigma\|_+ = |\{v \in V \mid \sigma_v = +1\}|.$$

We use  $\text{Bin}(V, 1 - \theta)$  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  $\Lambda \subseteq V$ ,

$$\Pr_{R \sim \text{Bin}(V, 1-\theta)} [R = \Lambda] = (1 - \theta)^{|\Lambda|} \theta^{|V| - |\Lambda|}.$$

**Definition 6.1** (magnetized block factorization of entropy). Let  $\theta \in (0, 1)$  and  $C > 0$ . A distribution  $\mu$  over  $\{-1, +1\}^V$  is said to satisfy  $\theta$ -magnetized block factorization of entropy with parameter  $C$  if the following holds for  $\pi = \theta * \mu$  and for all  $f : \Omega(\mu) \rightarrow \mathbb{R}_{\geq 0}$ :

$$(19) \quad \text{Ent}_\mu [f] \leq C \cdot \frac{Z_\pi}{\theta^{|V|}} \mathbb{E}_{R \sim \text{Bin}(V, 1-\theta)} [\pi_R(\mathbb{1}_R) \cdot \text{Ent}_{\pi^{\mathbb{1}_R}} [f]],$$

assuming  $\pi_R(\mathbb{1}_R) \cdot \text{Ent}_{\pi^{\mathbb{1}_R}} [f] = 0$  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 (19), the entropy is evaluated with respect to the original distribution  $\mu$ ; and in the right-hand-side (RHS), the entropy is evaluated with respect to the conditional distribution  $\pi^{\mathbb{1}_R}$ , where  $\pi = \theta * \mu$  is obtained from  $\mu$  by changing the local fields by a factor  $\theta$ . This definition plays a key role to relate  $\mu$  to  $\theta * \mu$ . Lemma 2.7 can be proved by combining the following two lemmas.

**Lemma 6.2.** *Let  $\theta \in (0, 1)$  and  $C > 0$ . For any distribution  $\mu$  over  $\{-1, +1\}^V$  satisfying  $\theta$ -magnetized block factorization of entropy with parameter  $C$ , the Glauber dynamics on  $\mu$  has the modified log-Sobolev constant*

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

**Lemma 6.3.** *Let  $\mu$  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  $\mu_k = \text{Trans}(\mu, k)$  satisfies  $\lceil \theta kn \rceil$ -uniform block factorization with parameter  $C$ , then  $\mu$  satisfies  $\theta$ -magnetized block factorization of entropy with parameter  $C$ .*

The rest of this section is dedicated to the proofs of Lemma 6.2 and Lemma 6.3.

**6.1. Implication to Modified Log-Sobolev Constant.** In this section, we prove Lemma 6.2. Let  $P = P_\mu^{\text{GD}}$  denote the Glauber dynamics on  $\mu$ . We only need to prove the following lemma.

**Lemma 6.4.** *Let  $\theta \in (0, 1)$  and  $\mu$  be a distribution over  $\{-1, +1\}^V$ . For any function  $f : \{-1, +1\}^V \rightarrow \mathbb{R}_{\geq 0}$ ,*

$$(20) \quad \rho_{\min}^{\text{GD}}(\pi) \cdot \frac{Z_\pi}{\theta^n} \mathbb{E}_{R \sim \text{Bin}(V, 1-\theta)} [\pi_R(\mathbb{1}_R) \cdot \text{Ent}_{\pi^{\mathbb{1}_R}} [f]] \leq \mathcal{E}_P(f, \log f),$$

where  $n = |V|$ ,  $\pi = \theta * \mu$ ,  $\rho_{\min}^{\text{GD}}(\pi)$  is defined in (5), and  $Z_\pi$  is defined in (18).

Lemma 6.2 is a straightforward consequence of Lemma 6.4. The rest of this section is dedicated to the proof of Lemma 6.4.

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

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

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

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

$$\begin{aligned} \mathcal{E}_P(f, \log f) &= \frac{1}{2} \sum_{\sigma, \tau \in \Omega(\mu)} \mu(\sigma)P(\sigma, \tau) (f(\sigma) - f(\tau))(\log f(\sigma) - \log f(\tau)) \\ &= \frac{1}{2} \sum_{\sigma \in \Omega(\mu)} \mu(\sigma) \frac{1}{n} \sum_{v \in V} \sum_{\tau \in \Omega(\mu^{\sigma_{V \setminus v}})} \mu^{\sigma_{V \setminus v}}(\tau) (f(\sigma) - f(\tau))(\log f(\sigma) - \log f(\tau)) \\ &= \frac{1}{n} \sum_{v \in V} \sum_{\chi \in \Omega(\mu_{V \setminus \{v\}})} \mu_{V \setminus \{v\}}(\chi) \cdot \frac{1}{2} \sum_{\sigma, \tau \in \Omega(\mu^\rho)} \mu^\chi(\sigma)\mu^\chi(\tau) (f(\sigma) - f(\tau))(\log f(\sigma) - \log f(\tau)) \\ (21) \quad &= \frac{1}{n} \sum_{v \in V} \sum_{\chi \in \Omega(\mu_{V \setminus \{v\}})} \mu_{V \setminus \{v\}}(\chi) \cdot \mathbf{Cov}_{\mu^\chi}(f, \log f) = \frac{1}{n} \sum_{v \in V} \mu[\mathbf{Cov}_v(f, \log f)], \end{aligned}$$

where we use the following notation

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

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

$$(22) \quad C \cdot \mathbf{Ent}_\mu[f] \leq \mathcal{E}_P(f, \log f) = \frac{1}{n} \sum_{v \in V} \mu[\mathbf{Cov}_v(f, \log f)].$$

Now, we are ready to prove Lemma 6.4 Consider the modified log-Sobolev inequality for the Glauber dynamics on  $\pi^{\mathbb{1}_R}$ . By (22), 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  $\rho^{\text{GD}}(\pi^{\mathbb{1}_R})$ :

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

where  $\rho_{\min}^{\text{GD}}(\pi) \leq \rho^{\text{GD}}(\pi^\sigma)$  for all feasible partial configuration  $\sigma$ , as defined in (5). Recall that  $\text{Bin}(V, 1 - \theta)$  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 (20) can be upper bounded as follows

$$(23) \quad \begin{aligned} &\rho_{\min}^{\text{GD}}(\pi) \cdot \frac{Z_\pi}{\theta^{|V|}} \mathbb{E}_{R \sim \text{Bin}(V, 1 - \theta)} [\pi_R(\mathbb{1}_R) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}}[f]] \\ &\leq \frac{Z_\pi}{\theta^{|V|}} \cdot \frac{1}{n} \sum_{v \in V} \mathbb{E}_{R \sim \text{Bin}(V, 1 - \theta)} [\pi_R(\mathbb{1}_R) \cdot \pi^{\mathbb{1}_R}[\mathbf{Cov}_v(f, \log f)]] . \end{aligned}$$

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

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

where  $\|\sigma\|_+$  denotes the number of vertices with  $+$  pinning, and 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

$$\pi_R(\mathbb{1}_R) = \mathbb{Pr}_{\sigma \sim \pi} \left[ R \subseteq \sigma^{-1}(+1) \right] = \mathbb{Pr}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[ R \subseteq \sigma^{-1}(+1) \right],$$

and  $\pi^{\mathbb{1}_R}[\text{Cov}_v(f, \log f)]$  is well-defined, such that

$$\pi^{\mathbb{1}_R}[\text{Cov}_v(f, \log f)] = \mathbb{E}_{\sigma \sim \pi_{V \setminus \{v\}}^{\mathbb{1}_R}} \left[ \text{Cov}_{\pi^\sigma}(f, \log f) \right] = \mathbb{E}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[ \text{Cov}_{\pi^\sigma}(f, \log f) \mid R \subseteq \sigma^{-1}(+1) \right].$$

Therefore,

$$\begin{aligned}
\pi_R(\mathbb{1}_R) \cdot \pi^{\mathbb{1}_R}[\text{Cov}_v(f, \log f)] &= \mathbb{Pr}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[ R \subseteq \sigma^{-1}(+1) \right] \cdot \mathbb{E}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[ \text{Cov}_{\pi^\sigma}(f, \log f) \mid R \subseteq \sigma^{-1}(+1) \right] \\
&= \mathbb{E}_{\sigma \sim \pi_{V \setminus \{v\}}} \left[ I[R \subseteq \sigma^{-1}(+1)] \cdot \text{Cov}_{\pi^\sigma}(f, \log f) \right].
\end{aligned}$$

- 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

$$\pi_R(\mathbb{1}_R) \cdot \text{Cov}_{\pi^{\mathbb{1}_R}}(f, \log f) = 0;$$

or else, if  $\pi_R(\mathbb{1}_R) > 0$  but  $v \in R$ , then  $\pi^{\mathbb{1}_R}[\text{Cov}_v(f, \log f)]$  is well-defined, but for  $\sigma \sim \pi_{V \setminus \{v\}}^{\mathbb{1}_R}$ , 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.

$$\pi^{\mathbb{1}_R}[\text{Cov}_v(f, \log f)] = \mathbb{E}_{\sigma \sim \pi_{V \setminus \{v\}}^{\mathbb{1}_R}} \left[ \text{Cov}_{\pi^{\mathbb{1}_R \uplus \sigma}}(f, \log f) \right] = 0.$$

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

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

This gives the equation  $(\star)$  in (24). Meanwhile, the other two equations in (24) follows respectively from linearity of expectation and the fact that  $\mathbb{Pr}_{R \subseteq V} [R \subseteq \Lambda] = \theta^{|V| - |\Lambda|}$  for all  $\Lambda \subseteq V$ .

Furthermore, it can be verified that

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

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  $\sigma_\pm^v(v) = \pm 1$ , the second equation is due to the chain rule and the relation between  $\pi$  and  $\mu$  in (18), and the last inequality is due to  $\Omega(\pi) = \Omega(\mu)$ ,  $(f(\sigma_+^v) - f(\sigma_-^v)) (\log f(\sigma_+^v) - \log f(\sigma_-^v)) \geq 0$  and the relaxation

$$(26) \quad \pi_v^\sigma(+1) \leq \mu_v^\sigma(+1),$$

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

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

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

Combining (23), (24), and (25), we have

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

where  $P$  denotes the Glauber dynamics for  $\mu$ . This proves Lemma 6.4.

**6.2. Magnetized Block Factorization of Entropy.** In this section, we prove Lemma 6.3. Let  $\mu$  be a distribution over  $\{-1, +1\}^V$ . Let  $\theta \in (0, 1)$  be a real number, recall that  $\pi = \theta * \mu$  is a distribution and  $Z_\pi$  is the normalization factor defined in (18). Recall that we use  $\text{Bin}(V, 1-\theta)$  to denote the distribution of subset  $R \subseteq V$  that is randomly generated by including each  $v \in V$  into  $R$  independently with probability  $1-\theta$ . Let  $\eta > 0$ . Suppose  $\mu_k$  satisfies the  $\lceil \theta k n \rceil$ -uniform block factorization with constant  $C$  for all integers  $k \geq k_0$ . To prove Lemma 6.3, we need to show that for any  $f : \{-1, +1\}^V \rightarrow \mathbb{R}_{\geq 0}$ , it holds that

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

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

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

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

Recall that

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

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

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

Note that above inequality holds for all  $k \geq k_0$ . Lemma 2.7 is a straightforward consequence of the following two results

$$(27) \quad \mathbf{Ent}_{\mu_k} [f^k] = \mathbf{Ent}_{\mu} [f];$$

$$(28) \quad \lim_{k \rightarrow \infty} \frac{C}{\binom{nk}{\lceil \theta kn \rceil}} \sum_{S \in \binom{V_k}{\lceil \theta kn \rceil}} \mu_k \left[ \mathbf{Ent}_S [f^k] \right] = C \cdot \frac{Z_{\pi}}{\theta^{|V|}} \mathbb{E}_{R \sim \text{Bin}(V, 1-\theta)} \left[ \pi_R(\mathbb{1}_R) \cdot \mathbf{Ent}_{\pi^{\mathbb{1}_R}} [f] \right].$$

In the rest of section, we verify (27) and (28).

6.2.1. *Verification of (27).* By definition, it holds that

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

Note that

$$\begin{aligned} \sum_{\sigma \in \Omega(\mu_k)} \mu_k(\sigma) f^k(\sigma) \log f^k(\sigma) &= \sum_{\alpha \in \Omega(\mu)} \sum_{\sigma \in \Omega(\mu_k): \sigma^* = \alpha} \mu_k(\sigma) f(\alpha) \log f(\alpha) = \sum_{\alpha \in \Omega(\mu)} \mu(\alpha) f(\alpha) \log f(\alpha); \\ \sum_{\sigma \in \Omega(\mu_k)} \mu_k(\sigma) f^k(\sigma) &= \sum_{\alpha \in \Omega(\mu)} \sum_{\sigma \in \Omega(\mu_k): \sigma^* = \alpha} \mu_k(\sigma) f(\alpha) = \sum_{\alpha \in \Omega(\mu)} \mu(\alpha) f(\alpha). \end{aligned}$$

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

6.2.2. *Verification of (28).* We introduce some notations. Let  $\varphi \in [0, 1]^V$  be a vector. For any subset  $\Lambda \subseteq V$ , we use  $\varphi_{\Lambda} \in [0, 1]^V$  denote the vector induced by restricting  $\varphi$  on  $\Lambda$ , i.e.

$$\varphi_{\Lambda}(v) = \begin{cases} \varphi(v) & v \in \Lambda; \\ 1 & v \notin \Lambda. \end{cases}$$

Let  $\varphi_{\Lambda} * \mu$  denote the distribution obtained by imposing local fields  $\varphi_{\Lambda}$  on  $\mu$ , formally

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

The above definition can be viewed as a generalization of Equation (4). Here, we allow  $\varphi_{\Lambda}(v) = 0$  for some vertices  $v \in V$ . The distribution  $\varphi_{\Lambda} * \mu$  is well-defined if and only if  $Z(\varphi_{\Lambda}) > 0$ .

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

**Lemma 6.5.** *Let  $C > 0$  and  $k \geq 1$  and  $1 \leq \ell \leq kn$  be two integers. If  $\mu_k$  satisfies the  $\ell$ -uniform block factorization with constant  $C$ , then for any  $f : \{-1, +1\}^V \rightarrow \mathbb{R}_{\geq 0}$ , let*

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

it holds that

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



where  $\mathbb{1}_R$  is the all-(+1) configuration on  $R$  and  $\tau^{-1}(+1) = \{v \in V \mid \tau_v = +1\}$ , the distribution  $(\mathbf{b}_{V \setminus R} * \mu)^{\mathbb{1}_R}$  is obtained from  $\mathbf{b}_{V \setminus R} * \mu$  conditional on  $\mathbb{1}_R$ , and  $\Pi_{V,k,\ell}(\cdot)$  is the multivariate hypergeometric distribution in (10).

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

**Lemma 6.6.** *Let  $\theta \in (0, 1)$  and  $\pi = \mu^{(\theta)}$ . Let  $f : \{-1, +1\}^V \rightarrow \mathbb{R}_{\geq 0}$  be a function. It holds that*

$$\lim_{k \rightarrow \infty} H_f(\lceil \theta nk \rceil, k) = \frac{Z_\pi}{\theta^{|V|}} \mathbb{E}_{R \sim \text{Bin}(V, 1-\theta)} [\pi_R(\mathbb{1}_R) \cdot \text{Ent}_{\pi^{\mathbb{1}_R}} [f]].$$

Equation (28) can be verified by combining Lemma 6.5 and Lemma 6.6. We verify these two lemmas in the rest of this section.

**6.3. Proof of Lemma 6.5.** For any subset  $\Lambda \subseteq V_k$ , we use  $\mu_{k,\Lambda}$  to denote the marginal distribution on  $\Lambda$  induced from  $\mu_k$ . By definition, we have

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

Recall that  $V_k = V \times [k]$ . For each  $v \in V$ , define  $C_v = \{v_1, v_2, \dots, v_k\}$ . 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  $\phi_\rho \in [0, 1]^{V \setminus F(\rho)}$  such that

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

**Claim 6.7.** *It holds that  $\text{Ent}_{\mu_k^\rho} [f^k] = \text{Ent}_{(\phi_\rho * \mu)^{\mathbb{1}_{F(\rho)}}} [f]$ .*

By Claim 6.7, we have

$$(29) \quad H_f(\ell, k) = \frac{1}{\binom{kn}{\ell}} \cdot \sum_{\sigma \in \Omega(\mu_k)} \mu_k(\sigma) \sum_{S \in \binom{V_k}{\ell}} \text{Ent}_{\mu_{(\phi_\rho, \mathbb{1}_{F(\rho)}})} [f], \quad \text{where } \rho = \rho(\sigma, S) \triangleq \sigma_{V_k \setminus S}.$$

Note that  $S$  is sampled from  $\binom{V_k}{\ell}$  uniformly at random. Let  $\mathbf{a} = (a_v)_{v \in V}$ , where  $a_v = |S \cap C_v|$ . Vector  $\mathbf{a}$  follows the distribution multivariate hypergeometric distribution  $\Pi_{V,k,\ell}$ . We can use the following procedure to generate the random set  $S$ : sample  $\mathbf{a} \sim \Pi_{V,k,\ell}$ ; 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  $\mathbf{b} = \mathbf{a}/k$ , so that  $\phi_\rho(v) = b_v$  for all  $v \in V \setminus F(\rho)$ . By letting  $R = F(\rho)$ , we rewrite equation (29) as

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

Here, vector  $k\mathbf{b}$  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  $\Pr[R = F(\rho)] = 0$ . 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  $\mu^{(b_{V \setminus R}), \mathbb{1}_R}$  will always be well defined.

To generate a random configuration  $\sigma \sim \mu_k$ , one can sample  $\tau \sim \mu$ , then transform  $\tau$  to  $\sigma$  by Definition 2.6. It turns out that  $F(\sigma)$  can be uniquely determined by  $\tau$ , 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  $\sigma$  such that  $\sigma^* = \tau$ . We have

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

By definition, it is easy to verify that

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

This implies that

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

This proves Lemma 6.5.

*Proof of Claim 6.7.* By [CFYZ21b, Lemma 6.2], it is straightforward to verify

$$(30) \quad \forall \xi \in \Omega(\mu), \quad (\phi_\rho * \mu)^{\mathbb{1}_{F(\rho)}}(\xi) = \sum_{\tau \in \Omega(\mu_k): \tau^* = \xi} \mu_k^\rho(\tau).$$

By definition, it holds that

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

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 (30), it holds that

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

Similarly, it holds that

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

Note that  $\Omega((\phi_\rho * \mu)^{\mathbb{1}_{F(\rho)}}) \subseteq \Omega(\mu)$ . This implies  $\mathbf{Ent}_{\mu_k^\rho} [f^k] = \mathbf{Ent}_{(\phi_\rho * \mu)^{\mathbb{1}_{F(\rho)}}} [f]$ .  $\square$

6.4. **Proof of Lemma 6.6.** Fix the distribution  $\mu$ , parameters  $\theta \in (0, 1)$ , and a function  $f : \{-1, +1\}^V \rightarrow \mathbb{R}_{\geq 0}$ . We define the following function  $H : [0, 1]^V \rightarrow \mathbb{R}$ :

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

Let  $\ell = \lceil \theta nk \rceil$ . We have

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

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

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

**Claim 6.8.** For any  $\varepsilon > 0$ , any  $\delta > 0$ , there exists  $K > 0$  such that for all  $k \geq K$ ,  $\Pr_{k\mathbf{b} \sim \Pi_{V, k, \ell}} [\mathcal{B}_{\delta, k}] \leq \varepsilon$ .

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

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

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

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

where  $M$  depends only on  $\mu$  and  $f$ . By Claim 6.8, there exists  $K > 0$  such that for all  $k \geq K$ , it holds that  $\Pr_{k\mathbf{b} \sim \Pi_{V, k, \ell}} [\mathcal{B}_{\delta, k}] \leq \frac{\varepsilon}{2 \max\{H(\boldsymbol{\theta}), M\}}$ . Hence for all  $k \geq K$ , we have

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

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

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

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

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

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

*Proof of Claim 6.8.* Recall that  $n = |V|$  and  $\ell = \lceil \theta kn \rceil$ . Observe that when  $k \geq \frac{2}{\delta}$ ,

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

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  $k \geq K_0$ ,

$$(32) \quad \forall v \in V, \quad \Pr_{kb \sim \Pi_{V,k,\ell}} \left[ \left| b_v - \frac{\lceil \theta kn \rceil}{kn} \right| \geq \frac{\delta}{2} \right] \leq 2 \exp\left(\frac{-\delta^2 k}{2}\right) \leq \frac{\varepsilon}{n},$$

where the last inequality holds because  $k$  is sufficiently large. Combining (31) and (32) proves the claim.  $\square$

## 7. INVARIANTS OF $k$ -TRANSFORMATION

In this section, we prove Lemma 2.8, that the spectral independence and marginal stability properties stated in Condition 1.11 are roughly invariant under  $k$ -transformation (Definition 2.6).

This is proved by two lemmas. Recall that concept of complete spectral independence (Definition 5.1).

**Lemma 7.1.** *Let  $\eta, \varepsilon > 0$ . If a distribution  $\mu$  over  $\{-1, +1\}^n$  is  $(\eta, \varepsilon)$ -completely spectrally independent and*

$$\mu_{-1}^{\min} \triangleq \min_{i \in [n]} \min_{\sigma \in \Omega(\mu_{[n] \setminus \{i\}})} \mu_i^\sigma(-1) > 0,$$

*then there exists a finite  $k_0 = 10(1 + \varepsilon)/\mu_{-1}^{\min}$  such that for all integers  $k \geq k_0$ , the  $k$ -transformed distribution  $\mu_k = \text{Trans}(\mu, k)$  is  $(2\eta + 5, \varepsilon)$ -completely spectrally independent.*

To state the next lemma, we also define the concept of complete marginal stability.

**Definition 7.2** (complete marginal stability). Let  $\zeta > 1$ . A distribution  $\mu$  over  $\{-1, +1\}^n$  is said to be completely  $\zeta$ -marginally stable if  $(\lambda * \mu)$  is  $\zeta$ -marginally stable for any  $\lambda \in (0, 1]^n$ .

**Lemma 7.3.** *Let  $\zeta > 1$ . If a distribution  $\mu$  over  $\{-1, +1\}^n$  is completely  $\zeta$ -marginally stable, then for any integer  $k \geq 1$ , the  $k$ -transformed distribution  $\mu_k = \text{Trans}(\mu, k)$  is  $2\zeta$ -marginally stable.*

Lemma 2.8 follows immediately from Lemma 7.1 and Lemma 7.3.

*Proof of Lemma 2.8.* By Condition 1.11,  $\mu$  is completely  $\zeta$ -marginally stable. Then for any  $i \in [n]$  and  $\sigma \in \Omega(\mu_{[n] \setminus \{i\}})$ , it holds that

$$\mu_i^\sigma(-1) = \left( \frac{\mu_i^\sigma(+1)}{\mu_i^\sigma(-1)} + 1 \right)^{-1} \geq \frac{1}{1 + \zeta}.$$

Let  $k_0 = 10(1 + \varepsilon)(1 + \zeta)$ . By Lemma 7.1, the distribution  $\mu_k$  is  $(2\eta + 5, \varepsilon)$ -completely spectrally independent for all  $k \geq k_0$ . By Lemma 7.3, the distribution  $\mu_k$  is  $4\zeta^2$ -marginally stable for all  $k \geq 1$ .  $\square$

**7.1. Complete spectral independence of  $\mu_k$  (proof of Lemma 7.1).** The *correlation matrix* (Definition 5.6) was introduced in [AASV21]. We consider the *absolute correlation matrix*.

**Definition 7.4** (absolute correlation matrix [AJK<sup>+</sup>21b]). Let  $\mu$  be a distribution over  $\{-1, +1\}^{[n]}$ . The absolute correlation matrix  $\Psi_\mu^{\text{AbsCor}} \in \mathbb{R}_{\geq 0}^{[n] \times [n]}$  is defined by

$$\forall i, j \in [n], \quad \Psi_\mu^{\text{AbsCor}}(i, j) \triangleq \left| \Psi_\mu^{\text{Cor}}(i, j) \right|,$$

where  $\Psi_\mu^{\text{Cor}}$  is the correlation matrix in Definition 5.6.

The spectral independence (Definition 1.9) and complete spectral independence (Definition 5.1) are defined using absolute influence matrix. Similarly, we can define using absolute correlation matrix.

**Definition 7.5** (limited correlation). Let  $\eta, \varepsilon > 0$ . A distribution  $\mu$  is said to have  $\eta$ -limited correlation if for any  $\Lambda \subseteq V$ , any  $\sigma \in \Omega(\mu_\Lambda)$ , the spectral radius of the absolute correlation matrix  $\Psi_{\mu^\sigma}^{\text{AbsCor}}$  satisfies

$$\rho\left(\Psi_{\mu^\sigma}^{\text{AbsCor}}\right) \leq \eta.$$

$\mu$  is said to have  $(\eta, \varepsilon)$ -complete limited correlation if  $(\lambda * \mu)$  has  $\eta$ -limited correlation for all  $\lambda \in (0, 1 + \varepsilon]^n$ .

**Lemma 7.6.** Let  $\eta, \varepsilon > 0$ . Let  $\mu$  be a distribution over  $\{-1, +1\}^{[n]}$ . If  $\mu$  is  $(\eta, \varepsilon)$ -completely spectrally independent, then  $\mu$  has  $(\eta + 1, \varepsilon)$ -complete limited correlation.

**Lemma 7.7.** Let  $\eta, \varepsilon > 0$ . Let  $\mu$  be a distribution over  $\{-1, +1\}^{[n]}$  satisfying

$$(33) \quad \mu_{-1}^{\min} \triangleq \min_{v \in [n]} \min_{\sigma \in \Omega(\mu_{[n] \setminus \{v\}})} \mu_v^\sigma(-1) > 0.$$

Let  $k_0 = k_0(\mu, \varepsilon) = 10(1 + \varepsilon)/\mu_{-1}^{\min} > 0$  be a finite real number. If  $\mu$  has  $(\eta, \varepsilon)$ -complete limited correlation, then for each integer  $k \geq k_0$ , it holds that  $\mu_k$  is  $(2\eta + 3, \varepsilon)$ -completely spectrally independent.

Lemma 7.1 is a straightforward corollary of the above two lemmas. We then prove Lemma 7.6 in Section 7.1.1, and prove Lemma 7.7 in Section 7.1.2 respectively.

7.1.1. *Proof of Lemma 7.6.* The following lemma is a well-known fact for non-negative matrix.

**Lemma 7.8** ([HJ12, Lemma 8.3.1]). Let  $A \in \mathbb{R}_{\geq 0}^{n \times n}$  be a non-negative matrix. The spectral radius  $\rho(A)$  equals to the maximum eigenvalue  $\lambda_{\max}(A)$ . Consequently,  $\rho(A + I) = \rho(A) + 1$ .

The next lemma is the relation between the influence matrix and the correlation matrix. Recall that the signed influence matrix is defined in Definition 5.7.

**Lemma 7.9.** Let  $\mu$  be a distribution over  $\{-1, +1\}^{[n]}$  satisfying  $\mu_i(-1) > 0$  for all  $i \in [n]$ , it holds that

$$\Psi_\mu^{\text{Inf}} = \text{diag}^{-1}\left(\{\mu_i(-1)\}_{i \in [n]}\right) \Psi_\mu^{\text{Cor}} - I,$$

where  $\text{diag}^{-1}\left(\{\mu_i(-1)\}_{i \in [n]}\right)$  is a diagonal matrix satisfying  $\text{diag}^{-1}\left(\{\mu_i(-1)\}_{i \in [n]}\right)(i, i) = \frac{1}{\mu_i(-1)}$ , and  $I$  is the  $n$ -by- $n$  identity matrix.

*Proof.* Suppose  $i \neq j$ . If  $\Omega(\mu_i) = \{-1\}$  or  $\Omega(\mu_i) = \{+1\}$ , then it holds that  $\Psi_\mu^{\text{Inf}}(i, j) = \Psi_\mu^{\text{Cor}}(i, j) = 0$ . Suppose  $\Omega(\mu_i) = \{-1, +1\}$ . It holds that

$$\begin{aligned} \Psi_\mu^{\text{Inf}}(i, j) &= \mu_j^{i \leftarrow +1}(+1) - \mu_j^{i \leftarrow -1}(+1) \\ &= \frac{\Pr_{X \sim \mu} [X_i = +1 \wedge X_j = +1]}{\Pr_{X \sim \mu} [X_i = +1]} - \frac{\Pr_{X \sim \mu} [X_j = +1] - \Pr_{X \sim \mu} [X_i = +1 \wedge X_j = +1]}{\Pr_{X \sim \mu} [X_i = -1]} \\ &= \frac{\Pr_{X \sim \mu} [X_i = +1 \wedge X_j = +1] - \Pr_{X \sim \mu} [X_i = +1] \Pr_{X \sim \mu} [X_j = +1]}{\Pr_{X \sim \mu} [X_i = -1] \Pr_{X \sim \mu} [X_i = +1]} \\ &= \frac{1}{\mu_i(-1)} \Psi_\mu^{\text{Cor}}(i, j). \end{aligned}$$

By definition, if  $i = j$ , then  $\Psi_\mu^{\text{Inf}}(i, i) = 0$ , and thus we have  $\Psi_\mu^{\text{Inf}} = \text{diag}^{-1}\left(\{\mu_i(-1)\}_{i \in [n]}\right) \Psi_\mu^{\text{Cor}} - I$ .  $\square$

Now, we are ready to prove Lemma 7.6.

*Proof of Lemma 7.6.* We use  $\Psi^{\text{AbsInf}}$  to denote the absolute influence matrix in Definition 1.9. By Definition 5.1 and Definition 7.4, it suffices to prove that for any distribution  $\mu$  over  $\{-1, +1\}^{[n]}$ , it holds that

$$(34) \quad \rho\left(\Psi_{\mu}^{\text{AbsCor}}\right) \leq \rho\left(\Psi_{\mu}^{\text{AbsInf}}\right) + 1.$$

Lemma 7.6 is a straightforward corollary of the above inequality.

Note that for any  $i \in [n]$  such that  $\mu_i(-1) = 0$ , the  $i$ -th row and  $i$ -th column in  $\Psi_{\mu}^{\text{Cor}}$  and  $\Psi_{\mu}^{\text{Inf}}$  are all 0. Hence, it suffices to consider  $\Psi_{\mu_S}^{\text{Cor}}$  and  $\Psi_{\mu_S}^{\text{Inf}}$ , where  $S = \{i \in [n] \mid \mu_i(-1) > 0\}$ . Without loss of generality, we can assume that the distribution  $\mu$  satisfies  $\mu_i(-1) > 0$  for all  $i \in [n]$ . By Lemma 7.9,

$$\Psi_{\mu}^{\text{Cor}} = \text{diag}\left(\{\mu_i(-1)\}_{i \in [n]}\right) (\Psi_{\mu}^{\text{Inf}} + I)$$

Note that  $\Psi_{\mu}^{\text{Inf}}(i, i) = 0$  for all  $i \in [n]$ . For any  $i, j \in [n]$ , it holds that

$$\Psi_{\mu}^{\text{AbsCor}}(i, j) \leq \Psi_{\mu}^{\text{AbsInf}}(i, j) + I(i, j)$$

because  $0 < \mu_i(-1) \leq 1$  for all  $i \in [n]$ . This implies (34).  $\square$

**7.1.2. Proof of Lemma 7.7.** We use the following definitions and lemmas to prove Lemma 7.7. Let  $\mu$  be a distribution over  $\{-1, +1\}^n$ . For any integer  $k \geq 1$ , let  $\mu_k$  denote the  $k$ -transformation of  $\mu$  (Definition 2.6). We use  $V = [n]$  to denote the variable set of  $\mu$  and  $V_k = [n] \times [k]$  to denote the variable set of  $\mu_k$ . For each  $v \in [n]$  and  $i \in [k]$ , we use  $v_i$  to denote the pair  $(v, i) \in V_k$ . For any  $\Lambda \subseteq V_k$ , we use  $\mu_{k, \Lambda}$  to denote the marginal distribution on  $\Lambda$  projected from  $\mu_k$ . We simply denote  $\mu_{k, \{v_i\}}$  by  $\mu_{k, v_i}$ .

**Lemma 7.10.** *Let  $\eta, \varepsilon > 0$ . If  $\mu$  has  $(\eta, \varepsilon)$ -complete limited correlation, then for any integer  $k \geq 1$ ,  $\mu_k$  has  $(\eta + 2, \varepsilon)$ -complete limited correlation.*

**Lemma 7.11.** *Let  $\varepsilon \geq 0$  and  $k \in \mathbb{Z}_{>0}$ . Let  $\mathbf{z} \in (0, \varepsilon]^{V_k}$ ,  $\Lambda \subseteq V_k$ ,  $v_i \in V_k \setminus \Lambda$ . For any  $\sigma \in \Omega(\mu_{k, \Lambda})$  where  $\mu_{k, v_i}^{\sigma}(+1) > 0$ , there exist  $\mathbf{x} \in (0, \varepsilon]^V$  satisfying  $x_v = 1$ , a subset  $R \subseteq V$  satisfying  $v \notin R$ , and a partial configuration  $\tau \in \Omega(\mu_R)$  such that*

$$\frac{(\mathbf{z} * \mu_k)_{v_i}^{\sigma}(-1)}{(\mathbf{z} * \mu_k)_{v_i}^{\sigma}(+1)} = \frac{k}{z_{v_i}} \left( \frac{(\mathbf{x} * \mu)_{v_i}^{\tau}(-1)}{(\mathbf{x} * \mu)_{v_i}^{\tau}(+1)} + \frac{1}{k} \sum_{v_j \in C_v \setminus (\Lambda \cup \{v_i\})} z_{v_j} \right),$$

where  $C_v \triangleq \{v_i \mid i \in [k]\}$ .

Lemma 7.10 can be proved by going through the proof of [AJK<sup>+</sup>21b, Proposition 26]. Lemma 7.11 is a technical lemma that relates  $\mu$  to  $\mu_k$  with local fields and pinnings. We first use Lemma 7.10 and Lemma 7.11 to prove Lemma 7.7, and then prove Lemma 7.10 and Lemma 7.11.

*Proof of Lemma 7.7 assuming Lemma 7.10 and Lemma 7.11.* By Lemma 7.10, we know that  $\mu_k$  has  $(\eta + 2, \varepsilon)$ -complete limited correlation. Fix  $\Lambda \subseteq V_k$ ,  $\sigma \in \Omega(\mu_{k, \Lambda})$ , and  $\mathbf{z} \in (0, 1 + \varepsilon]^{V_k}$ . It holds that

$$\rho\left(\Psi_{(\mathbf{z} * \mu_k)^{\sigma}}^{\text{AbsCor}}\right) \leq \eta + 2.$$

Let  $\pi = (\mathbf{z} * \mu_k)_{V_k \setminus \Lambda}^{\sigma}$ , which is obtained by projecting  $(\mathbf{z} * \mu_k)^{\sigma}$  on subset  $V_k \setminus \Lambda$ . By the definition of absolute correlation matrix, for any  $v_i \in \Lambda$ , the row and the column in  $\Psi_{(\mathbf{z} * \mu_k)^{\sigma}}^{\text{AbsCor}}$  corresponding to  $v_i$  only contain zeros. We have

$$\rho\left(\Psi_{\pi}^{\text{AbsCor}}\right) \leq \eta + 2.$$

Let  $k_0(\mu, \varepsilon) = 10(1 + \varepsilon)/\mu_{-1}^{\min}$ , where  $\mu_{-1}^{\min}$  is defined in (33). Note that  $k_0$  is finite because  $\mu_{-1}^{\min} > 0$ . For all integer  $k \geq k_0$ , we claim that

$$(35) \quad \min_{v_i \in V_k \setminus \Lambda} \pi_{v_i}(-1) \geq \frac{2}{3}.$$

Recall that we use  $\Psi_{\pi}^{\text{AbsInf}}$  to denote the absolute influence matrix in Definition 1.9. Then, by Lemma 7.9,

$$\rho\left(\Psi_{\pi}^{\text{AbsInf}}\right) = \rho\left(\text{diag}^{-1}(\{\pi_{v_i}(-1)\}_{v_i \in V_k \setminus \Lambda}) \cdot \Psi_{\pi}^{\text{AbsCor}} - I\right)$$

Note that the diagonal of  $\text{diag}^{-1}(\{\pi_{v_i}(-1)\}_{v_i \in V_k \setminus \Lambda}) \cdot \Psi_{\pi}^{\text{AbsCor}} - I$  are a set of zeros. By Lemma 5.9 and (35),

$$\rho\left(\Psi_{\pi}^{\text{AbsInf}}\right) \leq \rho\left(\frac{3}{2}\Psi_{\pi}^{\text{AbsCor}}\right) = \frac{3}{2}\rho\left(\Psi_{\pi}^{\text{AbsCor}}\right) \leq 2\eta + 3.$$

By the definition of absolute influence matrix, for any  $v_i \in \Lambda$ , the row and the column in  $\Psi_{(z * \mu_k)^\sigma}^{\text{AbsInf}}$  corresponding to  $v_i$  only contain zeros. We have

$$\rho\left(\Psi_{(z * \mu_k)^\sigma}^{\text{AbsInf}}\right) \leq 2\eta + 3.$$

Finally, we only need to verify inequality 35. To do this, we only need to show that for each  $v_i \in V_k \setminus \Lambda$ , it holds that

$$\frac{(z * \mu_k)_{v_i}^\sigma(+1)}{(z * \mu_k)_{v_i}^\sigma(-1)} \leq \frac{1}{2}.$$

When  $\mu_{k,v_i}^\sigma(+1) = 0$ , this holds trivially. Otherwise when  $\mu_{k,v_i}^\sigma(+1) > 0$ , by Lemma 7.11, there exists  $\mathbf{x} \in \mathbb{R}_{>0}^{[n]}$  where  $x_v = 1$  and a feasible partial configuration  $\tau \in \Omega(\mu_R)$ , where  $R \subseteq V$  and  $v \notin R$ , such that

$$\begin{aligned} \frac{(z * \mu_k)_{v_i}^\sigma(+1)}{(z * \mu_k)_{v_i}^\sigma(-1)} &= \frac{z_{v_i}}{k} \left( \frac{(\mathbf{x} * \mu)_{v_i}^\tau(-1)}{(\mathbf{x} * \mu)_{v_i}^\tau(+1)} + \frac{1}{k} \sum_{v_j \in C_v \setminus (\Lambda \cup \{v_i\})} z_{v_j} \right)^{-1} \\ &\leq \frac{z_{v_i}}{k} \frac{(\mathbf{x} * \mu)_{v_i}^\tau(+1)}{(\mathbf{x} * \mu)_{v_i}^\tau(-1)} \\ (*) &\leq \frac{1 + \varepsilon}{k} \max_{\sigma \in \Omega(\mu_{V \setminus \{v\}})} \frac{\mu_v^\sigma(+1)}{\mu_v^\sigma(-1)} \\ &\leq \frac{1 + \varepsilon}{k} \max_{\sigma \in \Omega(\mu_{V \setminus \{v\}})} \frac{1}{\mu_v^\sigma(-1)} = \frac{1 + \varepsilon}{k\mu_{-1}^{\min}}, \end{aligned}$$

which is less than  $\frac{1}{2}$  when  $k \geq \frac{10(1+\varepsilon)}{\mu_{-1}^{\min}}$ . Inequality (\*) holds because (1)  $0 < z_{v_i} \leq 1 + \varepsilon$ ; (2) the fact that the value of  $v$  is not fixed by  $\tau$ ; (3)  $x_v = 1$ .  $\square$

A version of Lemma 7.10 with signed correlation matrix was proved in [AJK<sup>+</sup>21b]. We give a proof of Lemma 7.10 by applying the same argument there. We include the proof for completeness.

*Proof of Lemma 7.10.* Let  $\mu$  be a distribution over  $\{-1, +1\}^V$ , where  $V = [n]$ . Define a more general  $\vec{k} = (k_1, k_2, \dots, k_n) \in \mathbb{Z}_{>0}$  transformation, which transforms  $\mu$  to a new distribution  $\mu_{\vec{k}}$ , where  $\mu_{\vec{k}}$  is defined over  $\{-1, +1\}^{V_{\vec{k}}}$  and  $V_{\vec{k}} = \{(i, j) \mid 1 \leq i \leq n, 1 \leq j \leq k_i\}$ . For each  $v \in [n]$ ,  $i \in [k_v]$ , we use  $v_i$  to denote  $(v, i)$ . To sample  $Y \sim \mu_{\vec{k}}$ , we first sample  $X \sim \mu$ , and then for any  $v \in V$

- if  $X_v = -1$ , then let  $Y_{v_i} = -1$  for all  $i \in [k_v]$ ;
- if  $X_v = +1$ , then sample  $j^* \in [k_v]$  u.a.r., set  $Y_{v_{j^*}} = +1$  and  $Y_{v_j} = -1$  for all  $j \in [k_v] \setminus \{j^*\}$ .

It is straightforward to verify the  $k$ -transformation in Definition 2.6 is a special case when  $\vec{k}$  is a constant vector with value  $k$ .

We prove the following results. For any  $\vec{k} \in \mathbb{Z}_{>0}^n$ , any  $\mathbf{y} \in \mathbb{R}_{>0}^{k_1 + \dots + k_n}$ , it holds that

$$(36) \quad \rho\left(\Psi_{\mathbf{y} * \mu_{\vec{k}}}^{\text{AbsCor}}\right) \leq \rho\left(\Psi_{\mathbf{x} * \mu}^{\text{AbsCor}}\right) + 2, \text{ where } \forall i \in [n], \mathbf{x}_i \triangleq \frac{1}{k_i} \sum_{j=1}^{k_i} y_{(i,j)}.$$

We first use (36) to prove the lemma. We need to prove that for any partial configuration  $\sigma \in \{-1 + 1\}^\Lambda$  of  $\mu_k$ , where  $\Lambda \subseteq V_k$ , it holds that for any  $z \in (0, 1 + \varepsilon]^{n^k}$ ,

$$\rho\left(\Psi_{(z * \mu_k)^\sigma}^{\text{AbsCor}}\right) = \rho\left(\Psi_{z * \mu_k^\sigma}^{\text{AbsCor}}\right) \leq \eta + 2.$$

By [AJK<sup>+</sup>21b, Lemma 15], for any feasible condition  $\sigma \in \{-1 + 1\}^\Lambda$  with respect to  $\mu_k$ , there exists a feasible condition  $\tau$  with respect to  $\mu$ , local fields  $\lambda \in (0, 1]^n$  together with a vector  $\vec{k} \in \mathbb{Z}_{>0}$  such that

$$\rho\left(\Psi_{\mu_k^\sigma}^{\text{AbsCor}}\right) = \rho\left(\Psi_{(\lambda * \mu^\tau)_{\vec{k}}}^{\text{AbsCor}}\right),$$

where  $(\lambda * \mu^\tau)_{\vec{k}}$  is obtained by applying  $\vec{k}$ -transformation on  $\lambda * \mu^\tau$ . Using (36) on  $\mu_k^\sigma$  implies that

$$\rho\left(\Psi_{z * \mu_k^\sigma}^{\text{AbsCor}}\right) \leq \rho\left(\Psi_{(\mathbf{x} \odot \lambda) * \mu^\tau}^{\text{AbsCor}}\right) + 2,$$

where for all  $i \in [n]$ ,  $x_i \triangleq \frac{1}{k_i} \sum_{j=1}^{k_i} z_{(i,j)}$  and  $(\mathbf{x} \odot \lambda) \in (0, 1 + \varepsilon]^n$  satisfying  $(\mathbf{x} \odot \lambda)_v = x_v \lambda_v \leq 1 + \varepsilon$ . Since  $\mu$  has  $(\eta, \varepsilon)$ -complete limited correlation, we have

$$\rho\left(\Psi_{\mu_k^\sigma}^{\text{AbsCor}}\right) = \rho\left(\Psi_{(\lambda * \mu^\tau)_{\vec{k}}}^{\text{AbsCor}}\right) \leq \rho\left(\Psi_{(\mathbf{x} \odot \lambda) * \mu^\tau}^{\text{AbsCor}}\right) + 2 = \rho\left(\Psi_{((\mathbf{x} \odot \lambda) * \mu)^\tau}^{\text{AbsCor}}\right) + 2 \leq \eta + 2.$$

Now, we only need to verify Equation (36). For convenience, we denote  $\Psi_{\mathbf{y} * \mu_{\vec{k}}}^{\text{AbsCor}}$  as  $\Psi_{\vec{k}}$  and  $\Psi_{\mathbf{x} * \mu}^{\text{AbsCor}}$  as  $\Psi$  respectively. Without loss of generality, we may assume  $+1 \in \Omega(\mu_i)$  for all  $i \in [n]$ . Suppose  $+1 \notin \Omega(\mu_i)$  for some  $i \in [n]$ . Then the  $i$ -th row and the  $i$ -th column of  $\Psi$  are all zeros, and the rows  $(i, j)$  and columns  $(i, j)$  for  $j \in [k_i]$  in  $\Psi_{\vec{k}}$  are all zeros. Hence, the variable  $i$  and all variables  $(i, j)$  for  $j \in [k_i]$  have fixed value and they do not affect the spectral radiuses of correlation matrices. In this case, we can simply consider the distribution  $\mu_{[n] \setminus \{i\}}$  and its transformations.

Let  $\widehat{\Psi}_{\vec{k}}$  be a matrix with the same size as  $\Psi_{\vec{k}}$  defined as

$$\forall u, v \in [n], i \in [k_u], j \in [k_v], \quad \widehat{\Psi}_{\vec{k}}(u_i, v_j) \triangleq \begin{cases} 1 + (\mathbf{y} * \mu_{\vec{k}})_{u_i(+1)}, & u = v \text{ and } i = j; \\ \Psi_{\vec{k}}(u_i, v_j), & \text{otherwise.} \end{cases}$$

From this definition, we know that  $\Psi_{\vec{k}}(u_i, v_j) \leq \widehat{\Psi}_{\vec{k}}(u_i, v_j)$ , and by Lemma 5.9, it holds that

$$(37) \quad \rho\left(\Psi_{\vec{k}}\right) \leq \rho\left(\widehat{\Psi}_{\vec{k}}\right).$$

Let  $\widehat{\Psi}$  be another matrix with the same size as  $\Psi$  defined as

$$\forall u, v \in [n], \quad \widehat{\Psi}(u, v) \triangleq \sum_{h \in [k_v]} \widehat{\Psi}_{\vec{k}}(u_1, v_h).$$

In the above definition,  $\widehat{\Psi}(u, v)$  is the sum over all  $\widehat{\Psi}_{\vec{k}}(u_1, v_h)$  for  $h \in [k_v]$ . The following claim shows that the  $u_1$  in the definition can be replaced by any  $u_i$  for  $i \in [k_u]$ . The claim will be proved later.

**Claim 7.12.** *For any  $u, v \in [n]$  and  $i \in [k_u]$ , it holds that  $\widehat{\Psi}(u, v) = \sum_{h \in [k_v]} \widehat{\Psi}_{\vec{k}}(u_i, v_h)$ .*

To prove (36), we prove the following two inequalities

$$(38) \quad \rho\left(\widehat{\Psi}\right) \leq \rho(\Psi) + 2$$

$$(39) \quad \rho\left(\widehat{\Psi}_{\vec{k}}\right) \leq \max\left\{\rho\left(\widehat{\Psi}\right), 1\right\}.$$

Combining (37), (38) and (39), we have

$$\rho\left(\Psi_{\vec{k}}\right) \leq \rho\left(\widehat{\Psi}_{\vec{k}}\right) \leq \max\left\{\rho\left(\widehat{\Psi}\right), 1\right\} \leq \rho(\Psi) + 2,$$

which proves (36).



We first prove (38). By the definition of  $\widehat{\Psi}$ , we know that for any  $u, v \in [n]$ , if  $u = v$ , it holds that

$$\widehat{\Psi}(u, u) = 1 + (\mathbf{y} * \mu_{\bar{k}})_{u_1(+1)} + \sum_{j \in [k_u] \setminus \{1\}} \Psi_{\bar{k}}(u_1, u_j) \stackrel{(*)}{\leq} 1 + \sum_{j \in [k_u]} (\mathbf{y} * \mu_{\bar{k}})_{u_j(+1)}.$$

where  $(*)$  holds because it is straightforward to see  $\Psi_{\bar{k}}(u_1, u_j) = (\mathbf{y} * \mu_{\bar{k}})_{u_j(+1)}$  for all  $j \in [k_u] \setminus \{1\}$ . We have the following claim about the distribution  $(\mathbf{y} * \mu_{\bar{k}})$ .

**Claim 7.13.** *For any distinct  $u, v \in [n]$ , any  $i \in [k_u]$  and  $j \in [k_v]$ , it holds that*

$$(\mathbf{y} * \mu_{\bar{k}})_{v_j}^{u_i \leftarrow +1(+1)} = \frac{y_{v_j} (\mathbf{x} * \mu)_v^{u_i \leftarrow +1(+1)}}{\sum_{\ell \in [k_v]} y_{v_\ell}} \quad \text{and} \quad (\mathbf{y} * \mu_{\bar{k}})_{v_j(+1)} = \frac{y_{v_j} (\mathbf{x} * \mu)_v(+1)}{\sum_{\ell \in [k_v]} y_{v_\ell}},$$

where  $\mathbf{x}$  is defined in (36).

We first prove the lemma assuming Claim 7.13, and then prove Claim 7.13. By Claim 7.13, it is straightforward to verify that

$$(40) \quad \widehat{\Psi}(u, u) = 1 + \sum_{j \in [k_u]} (\mathbf{y} * \mu_{\bar{k}})_{u_j(+1)} = 1 + (\mathbf{x} * \mu)_u(+1).$$

For any  $u, v \in [n]$ , if  $u \neq v$ , it holds that

$$(41) \quad \begin{aligned} \widehat{\Psi}(u, v) &= \sum_{h \in [k_v]} \widehat{\Psi}_{\bar{k}}(u_1, v_h) = \sum_{h \in [k_v]} |(\mathbf{y} * \mu_{\bar{k}})_{v_h}^{u_1 \leftarrow +1(+1)} - (\mathbf{y} * \mu_{\bar{k}})_{v_h(+1)}| \\ &= \sum_{h \in [k_v]} \left| \frac{y_{v_h} ((\mathbf{x} * \mu)_v^{u_1 \leftarrow +1(+1)} - (\mathbf{x} * \mu)_v(+1))}{\sum_{\ell \in [k_v]} y_{v_\ell}} \right| = |(\mathbf{x} * \mu)_v^{u_1 \leftarrow +1(+1)} - (\mathbf{x} * \mu)_v(+1)|, \end{aligned}$$

where (41) holds because of Claim 7.13, which implies that

$$(42) \quad \forall u, v \in [n] \text{ with } u \neq v, \quad \widehat{\Psi}(u, v) = \Psi(u, v).$$

Combining (40) and (42), we have

$$\forall u, v \in [n], \quad \widehat{\Psi}(u, v) \leq \Psi(u, v) + 2I(u, v).$$

Since both  $\widehat{\Psi}$  and  $\Psi$  are non-negative matrices, by Lemma 5.9, (38) holds.

Now, we prove (39). By (42) and the definition of correlation matrix, it is straightforward to verify  $\text{diag}(\{(\mathbf{x} * \mu)_i(+1)\}_{i \in [n]}) \widehat{\Psi}$  is a symmetric matrix.

Hence,  $\widehat{\Psi}$  has an orthogonal eigenbasis  $f^1, \dots, f^n$  with respect to the inner product  $\langle \cdot, \cdot \rangle_{(\mathbf{x} * \mu)}$  with corresponding real eigenvalues  $\lambda_1 \geq \dots \geq \lambda_n$ , where the inner product  $\langle \cdot, \cdot \rangle_{(\mathbf{x} * \mu)}$  is defined by

$$\forall a, b \in \mathbb{R}^{[n]}, \quad \langle a, b \rangle_{(\mathbf{x} * \mu)} = \sum_{i=1}^n a_i b_i (\mathbf{x} * \mu)_i(+1).$$

One could verify that for each eigenvector  $f^t = (f_1^t, \dots, f_n^t) \in \mathbb{R}^n$ , the vector

$$F^t = \underbrace{(f_1^t, \dots, f_1^t)}_{k_1 \text{ times}}, \underbrace{(f_2^t, \dots, f_2^t)}_{k_2 \text{ times}}, \dots, \underbrace{(f_n^t, \dots, f_n^t)}_{k_n \text{ times}}$$

is an eigenvector of  $\widehat{\Psi}_{\bar{k}}$  with eigenvalue  $\lambda_t$ . That is, for any  $u \in [n]$ ,  $i \in [k_u]$ , it holds that

$$\left( \widehat{\Psi}_{\bar{k}}(u_i, \cdot) F^t \right)_{u_i} = \sum_{v \in [n]} f_v^t \sum_{j \in [k_v]} \widehat{\Psi}_{\bar{k}}(u_i, v_j) \stackrel{(*)}{=} \sum_{v \in [n]} \widehat{\Psi}(u, v) f_v^t = \widehat{\Psi}(u, \cdot) F^t = \lambda_t f_u^t = \lambda_t F_{u_i}^t.$$

Equation  $(\star)$  holds due to Claim 7.12, i.e.

$$\forall i \in [k_u], \quad \widehat{\Psi}(u, v) = \sum_{j \in [k_v]} \widehat{\Psi}_k^-(u_1, v_j) \stackrel{(*)}{=} \sum_{j \in [k_v]} \widehat{\Psi}_k^-(u_i, v_j).$$

Moreover, for each  $u \in [n]$ , if we pick  $f[u]^1, f[u]^2, \dots, f[u]^{k_u-1}$  as an orthogonal basis with respect to the inner product  $\langle \cdot, \cdot \rangle_{(y_{u_i})_{i \in [k_u]}}$  of the vector space  $\left\{ f \in \mathbb{R}^{k_u} \mid \langle f, 1 \rangle_{(y_{u_i})_{i \in [k_u]}} = 0 \right\}$ , where the inner product is defined by  $\langle a, b \rangle_{(y_{u_i})_{i \in [k_u]}} = \sum_{i=1}^{k_u} a(i)b(i)y_{u_i}$ , and then the vector

$$F[u]^t = (\underbrace{0, \dots, 0}_{k_1 \text{ times}}, \underbrace{0, \dots, 0}_{k_{u-1} \text{ times}}, f[u]_1^t, f[u]_2^t, \dots, f[u]_{k_u}^t, \underbrace{0, \dots, 0}_{k_{u+1} \text{ times}}, \dots, \underbrace{0, \dots, 0}_{k_n \text{ times}})$$

is an eigenvector of  $\widehat{\Psi}_k^-$  with eigenvalue 1. This is because:

(1) for any  $i \in [k_u]$ , it holds that

$$(\widehat{\Psi}_k^- F[u]^t)_{u_i} = \sum_{j \in [k_u]} \widehat{\Psi}_k^-(u_i, u_j) F[u]_{u_j}^t = F[u]_{u_i}^t + \sum_{j \in [k_u]} (\mathbf{y} * \mu_{\bar{k}}^-)_{u_j} (+1) F[u]_{u_j}^t = F[u]_{u_i}^t,$$

where the last equation holds because

$$\sum_{j \in [k_u]} (\mathbf{y} * \mu_{\bar{k}}^-)_{u_j} (+1) F[u]_{u_j}^t = \frac{\mu_u(+1)}{Z \cdot k_u} \sum_{j \in [k_u]} y_{u_j} F[u]_{u_j}^t = \frac{\mu_u(+1)}{Z \cdot k_u} \langle f[u]^t, 1 \rangle_{(y_{u_i})_{i \in [k_u]}} = 0,$$

where  $Z$  is defined by

$$\begin{aligned} (43) \quad Z &\triangleq \sum_{Y \in \Omega(\mu_{\bar{k}}^-)} \mu_{\bar{k}}^-(Y) \prod_{v_i \in Y^{-1}(+1)} y_{v_i} \\ &= \sum_{X \in \Omega(\mu)} \mu(X) \prod_{v \in X^{-1}(+1)} \sum_{i \in [k_v]} \frac{y_{v_i}}{k_v} \\ &= \sum_{X \in \Omega(\mu)} \mu(X) \prod_{v \in X^{-1}(+1)} x_v; \end{aligned}$$

(2) for any  $v \neq u$  and  $i \in [k_v]$ , it hold that

$$\begin{aligned} (\widehat{\Psi}_k^- F[u]^t)_{v_i} &= \sum_{j \in [k_u]} \widehat{\Psi}_k^-(v_i, u_j) F[u]_{u_j}^t \\ &= \sum_{j \in [k_u]} \left| (\mathbf{y} * \mu_{\bar{k}}^-)_{u_j}^{v_i \leftarrow +1} (+1) - (\mathbf{y} * \mu_{\bar{k}}^-)_{u_j} (+1) \right| f[u]_j^t \\ \text{(by Claim 7.13)} &= \sum_{j \in [k_u]} \left| \frac{y_{u_j} ((\mathbf{x} * \mu)_u^{v_i \leftarrow +1} (+1) - (\mathbf{x} * \mu)_u (+1))}{\sum_{\ell \in [k_u]} y_{u_\ell}} \right| f[u]_j^t \\ &= |(\mathbf{x} * \mu)_u^{v_i \leftarrow +1} (+1) - (\mathbf{x} * \mu)_u (+1)| \langle f[u]^t, 1 \rangle_{(y_{u_j})_{j \in [k_u]}} \\ &= 0. \end{aligned}$$

Finally, note that

$$\{F^1, \dots, F^n\} \cup \left\{ F[u]^1, \dots, F[u]^{k_u-1} \right\}_{u=1}^n$$

forms an orthogonal eigenbasis of  $\widehat{\Psi}_k$  with respect to the inner product  $\langle \cdot, \cdot \rangle_{(\mathbf{y} * \mu_k^-)}$ , where the inner product is defined by  $\langle a, b \rangle_{(\mathbf{y} * \mu_k^-)} = \sum_{v_i \in V_k^-} a(v_i) b(v_i) (\mathbf{y} * \mu_k^-)_{v_i} (+1)$ . Formally, for any distinct  $1 \leq u, v \leq n$

$$\begin{aligned} \langle F^u, F^v \rangle_{(\mathbf{y} * \mu_k^-)} &= \sum_{w=1}^n f_w^u f_w^v \sum_{i \in [k_w]} (\mathbf{y} * \mu_k^-)_{w_i} (+1) \\ &= \frac{1}{Z} \sum_{w=1}^n f_w^u f_w^v \sum_{X \in \Omega(\mu): X_w=1} \mu(X) \prod_{w' \in \sigma^{-1}(X)} \frac{y_{w'}}{k_{w'}} \\ &= \sum_{w=1}^n f_w^u f_w^v (\mathbf{x} * \mu)_w (+1) = \langle f^u, f^v \rangle_{(\mathbf{x} * \mu)} \\ &= 0. \end{aligned}$$

For any  $u \in [n]$ , any  $v \in [n]$  and  $i \in [k_v - 1]$ ,

$$\begin{aligned} \langle F^u, F[v]^i \rangle_{(\mathbf{y} * \mu_k^-)} &= f_v^u \sum_{j \in [k_v]} f[v]_j^i (\mathbf{y} * \mu_k^-)_{v_j} (+1) \\ &= \frac{f_v^u \mu_v (+1)}{Z k_v} \sum_{j \in [k_v]} f[v]_j^i y_{v_j} \\ &= \frac{f_v^u \mu_v (+1)}{Z k_v} \langle f[v]^i, 1 \rangle_{(y_{v_j})_{j \in [k_v]}} \\ &= 0. \end{aligned}$$

For any  $u \in [n]$ , any distinct  $i, j \in [k_u - 1]$ ,

$$\begin{aligned} \langle F[u]^i, F[u]^j \rangle_{(\mathbf{y} * \mu_k^-)} &= \sum_{\ell \in [k_u]} f[u]_\ell^i f[u]_\ell^j (\mathbf{y} * \mu_k^-)_{u_\ell} (+1) \\ &= \frac{\mu_u (+1)}{Z k_u} \sum_{\ell \in [k_u]} f[u]_\ell^i f[u]_\ell^j y_{u_\ell} \\ &= \frac{\mu_u (+1)}{Z k_u} \langle f[u]^i, f[u]^j \rangle_{(y_{u_\ell})_{\ell \in [k_u]}} \\ &= 0. \end{aligned}$$

For any distinct  $v, u \in [n]$ ,  $i \in [k_v - 1]$  and  $j \in [k_u - 1]$ , it is straightforward to verify

$$\langle F[u]^i, F[v]^j \rangle_{(\mathbf{y} * \mu_k^-)} = 0.$$

Hence, the spectrum of  $\widehat{\Psi}_k$  is

$$\{\lambda_1, \dots, \lambda_n\} \cup \left\{ \mathbf{1}^{(t)} \right\}_{t=1}^{\sum_{u \in [n]} k_u - n},$$

So, we know that  $\rho(\widehat{\Psi}_k) \leq \max\{\rho(\widehat{\Psi}), 1\}$ . □

*Proofs of Claim 7.12 and Claim 7.13.* We first prove Claim 7.13, then use Claim 7.13 to prove Claim 7.12. By the definition of conditional probability, we have

$$(\mathbf{y} * \mu_{\vec{k}})^{u_i \leftarrow 1}(+1) = \left( \sum_{\substack{\sigma \in \Omega(\mu_{\vec{k}}): \\ \sigma(u_i)=+1 \wedge \sigma(v_j)=+1}} \mu_{\vec{k}}(\sigma) \prod_{w_\ell \in \sigma^{-1}(+1)} y_{w_j} \right) / \left( \sum_{\substack{\sigma \in \Omega(\mu_{\vec{k}}): \\ \sigma(u_i)=+1}} \mu_{\vec{k}}(\sigma) \prod_{w_\ell \in \sigma^{-1}(+1)} y_{w_j} \right)$$

The numerator equals to

$$\begin{aligned} & \sum_{\substack{\tau \in \Omega(\mu): \\ \tau(u)=+1 \wedge \tau(v)=+1}} \frac{\mu(\tau) y_{u_i} y_{v_j}}{k_u k_v} \prod_{w \in \tau^{-1}(+1) \setminus \{u, v\}} \sum_{\ell \in [k_w]} \frac{y_{w_\ell}}{k_w} \\ &= \sum_{\substack{\tau \in \Omega(\mu): \\ \tau(u)=+1 \wedge \tau(v)=+1}} \frac{\mu(\tau) y_{u_i} y_{v_j}}{k_u k_v} \prod_{w \in \tau^{-1}(+1) \setminus \{u, v\}} x_w \\ &= \frac{y_{v_j}}{\sum_{\ell \in [k_v]} y_{v_\ell}} \cdot \frac{y_{u_i}}{\sum_{\ell \in [k_u]} y_{u_\ell}} \cdot \sum_{\substack{\tau \in \Omega(\mu): \\ \tau(u)=+1 \wedge \tau(v)=+1}} \mu(\tau) \prod_{w \in \tau^{-1}(+1)} x_w. \end{aligned}$$

The denominator equals to

$$\sum_{\substack{\tau \in \Omega(\mu): \\ \tau(u)=+1}} \frac{\mu(\tau) y_{u_i}}{k_v} \prod_{w \in \tau^{-1}(+1) \setminus \{u\}} \sum_{\ell \in [k_w]} \frac{y_{w_\ell}}{k_w} = \frac{y_{u_i}}{\sum_{\ell \in [k_u]} y_{u_\ell}} \sum_{\substack{\tau \in \Omega(\mu): \\ \tau(u)=+1}} \mu(\tau) \prod_{w \in \tau^{-1}(+1)} x_w$$

Hence, we have

$$\begin{aligned} (\mathbf{y} * \mu_{\vec{k}})^{u_i \leftarrow 1}(+1) &= \frac{y_{v_j}}{\sum_{\ell \in [k_v]} y_{v_\ell}} \cdot \left( \sum_{\substack{\tau \in \Omega(\mu): \\ \tau(u)=+1 \wedge \tau(v)=+1}} \mu(\tau) \prod_{w \in \tau^{-1}(+1)} x_w \right) / \left( \sum_{\substack{\tau \in \Omega(\mu): \\ \tau(u)=+1}} \mu(\tau) \prod_{w \in \tau^{-1}(+1)} x_w \right) \\ &= \frac{y_{v_j}}{\sum_{\ell \in [k_v]} y_{v_\ell}} \cdot (\mathbf{x} * \mu_{\vec{v}})^{u \leftarrow 1}(+1). \end{aligned}$$

Recall  $Z$  defined in (43). We have

$$\begin{aligned} (\mathbf{y} * \mu_{\vec{k}})_{u_i}(+1) &= \frac{1}{Z} \sum_{\substack{\sigma \in \Omega(\mu_{\vec{k}}): \\ \sigma(u_i)=+1}} \mu_{\vec{k}}(\sigma) \prod_{w_\ell \in \sigma^{-1}(+1)} y_{w_j} \\ &= \frac{1}{Z} \sum_{\substack{\tau \in \Omega(\mu): \\ \tau(u)=+1}} \frac{\mu(\tau) y_{u_i}}{k_u} \prod_{w \in \tau^{-1}(+1) \setminus \{u\}} \sum_{\ell \in [k_w]} \frac{y_{w_\ell}}{k_w} \\ &= \frac{y_{u_i}}{\sum_{\ell \in [k_u]} y_{u_\ell}} \cdot \frac{1}{Z} \cdot \sum_{\substack{\tau \in \Omega(\mu): \\ \tau(u)=+1}} \mu(\tau) \prod_{w \in \tau^{-1}(+1)} x_w \\ &= \frac{y_{u_i}}{\sum_{\ell \in [k_u]} y_{u_\ell}} \cdot \mu_u(+1). \end{aligned}$$

Next, we prove Claim 7.12. By definition, we have for any  $u, v \in [n]$  with  $u \neq v$ , any  $i \in [k_u]$ , we have

$$\begin{aligned}
\widehat{\Psi}(u, v) &= \sum_{h \in [k_u]} \widehat{\Psi}_k^-(u_1, v_h) = \sum_{h \in [k_v]} |(\mathbf{y} * \mu_k^-)_{v_h}^{u_1 \leftarrow +1} (+1) - (\mathbf{y} * \mu_k^-)_{v_h} (+1)| \\
&\stackrel{\text{(by Claim 7.13)}}{=} \sum_{h \in [k_v]} \left| \frac{y_{v_h} ((\mathbf{x} * \mu)_{v_h}^{u \leftarrow +1} (+1) - (\mathbf{x} * \mu)_{v_h} (+1))}{\sum_{\ell \in [k_v]} y_{v_\ell}} \right| \\
&\stackrel{\text{(by Claim 7.13)}}{=} \sum_{h \in [k_v]} |(\mathbf{y} * \mu_k^-)_{v_h}^{u_i \leftarrow +1} (+1) - (\mathbf{y} * \mu_k^-)_{v_h} (+1)| \\
&= \sum_{h \in [k_v]} \widehat{\Psi}_k^-(u_i, v_h).
\end{aligned}$$

For any  $u \in [n]$  and  $i \in [k_u]$ , we have

$$\begin{aligned}
\widehat{\Psi}(u, u) &= \sum_{h \in [k_u]} \widehat{\Psi}_k^-(u_1, u_h) = 1 + (\mathbf{y} * \mu_k^-)_{u_1} (+1) + \sum_{j \in [k_u] \setminus \{1\}} |(\mathbf{y} * \mu_k^-)_{u_j}^{u_1 \leftarrow +1} (+1) - (\mathbf{y} * \mu_k^-)_{u_j} (+1)| \\
&\stackrel{\text{(by } (\mathbf{y} * \mu_k^-)_{u_j}^{u_1 \leftarrow +1} (+1) = 0 \text{)}}{=} 1 + \sum_{j \in [k_u]} (\mathbf{y} * \mu_k^-)_{u_j} (+1) \\
&= 1 + (\mathbf{y} * \mu_k^-)_{u_i} (+1) + \sum_{j \in [k_u] \setminus \{i\}} |(\mathbf{y} * \mu_k^-)_{u_j}^{u_i \leftarrow +1} (+1) - (\mathbf{y} * \mu_k^-)_{u_j} (+1)| \\
&= \sum_{h \in [k_u]} \widehat{\Psi}_k^-(u_i, u_h). \quad \square
\end{aligned}$$

*Proof of Lemma 7.11.* Recall  $V = [n]$ . First, define

$$\begin{aligned}
R_- &\triangleq \{u \in V \mid \forall i \in [k], u_i \in \Lambda \wedge \sigma_{u_i} = -1\}, \\
R_+ &\triangleq \{u \in V \mid \exists i \in [k], u_i \in \Lambda \wedge \sigma_{u_i} = +1\}, \\
(44) \quad R &\triangleq R_- \uplus R_+.
\end{aligned}$$

Let  $\tau \in \{-1, +1\}^R$  be

$$(45) \quad \forall u \in R, \quad \tau_u \triangleq \begin{cases} -1, & u \in R_- \\ +1, & u \in R_+. \end{cases}$$

Now, note that

$$(z * \mu_k)_{v_i}^\sigma (+1) = \frac{\Pr_{Y \sim z * \mu_k} [Y_{v_i} = +1 \wedge Y_\Lambda = \sigma]}{\Pr_{Y \sim z * \mu_k} [Y_\Lambda = \sigma]} \quad \text{and} \quad (z * \mu_k)_{v_i}^\sigma (-1) = \frac{\Pr_{Y \sim z * \mu_k} [Y_{v_i} = -1 \wedge Y_\Lambda = \sigma]}{\Pr_{Y \sim z * \mu_k} [Y_\Lambda = \sigma]}.$$

Note that  $v_i \in V_k \setminus \Lambda$  and  $\mu_{v_i}^\sigma (+1) > 0$ . We first show that  $v \notin R$ . Suppose  $v \in R_+$ . Since  $v_i \in V_k \setminus \Lambda$ , there exists  $v_j \in C_v$  such that  $\sigma_{v_j} = +1$ , and thus  $\mu_{v_i}^\sigma (+1) = 0$ , but  $\mu_{v_i}^\sigma (+1) > 0$ . Suppose  $v \in R_-$ . It must hold that  $v_j \in \Lambda$  for all  $j \in [k]$ , but  $v_i \notin \Lambda$ . Hence, it holds that  $v \notin R$ .

Define the partition function

$$Z \triangleq \sum_{Y \in \Omega(\mu_k)} \mu(Y^*) \prod_{u_j \in V_k: Y_{u_j} = 1} \frac{z_{u_j}}{k}, \quad \text{where } \forall u \in V, Y^*(u) = \begin{cases} +1 & \text{if } \exists j \in [k], Y(u_j) = +1 \\ -1 & \text{if } \forall j \in [k], Y(u_j) = -1. \end{cases}$$

For any  $u \in [n]$ , let  $S_u \triangleq C_u \setminus \Lambda$ , where  $C_u = \{u_i \mid i \in [k]\}$ . We have

$$\begin{aligned}
\Pr_{Y \sim \mathbf{z} * \mu_k} [Y_{v_i} = +1 \wedge Y_\Lambda = \sigma] &= \frac{1}{Z} \sum_{Y \in \Omega(\mu_k)} \mu(Y^*) \cdot \left( \prod_{u_j \in V_k: Y_{u_j}=1} \frac{z_{u_j}}{k} \right) \cdot \mathbb{1}[Y_{v_i} = 1 \wedge Y_\Lambda = \sigma] \\
&= \frac{1}{Z} \sum_{\substack{Y \in \Omega(\mu_k): \\ Y_{v_i}=+1 \wedge Y_\Lambda=\sigma}} \mu(Y^*) \cdot \prod_{u_j \in V_k \setminus \sigma^{-1}(+1): Y_{u_j}=+1} \frac{z_{u_j}}{k} \cdot \prod_{u_j \in \sigma^{-1}(+1)} \frac{z_{u_j}}{k} \\
(*) &= \frac{1}{Z} \sum_{\substack{X \in \Omega(\mu): \\ X_v=+1 \wedge X_R=\tau}} \mu(X) \cdot \frac{z_{v_i}}{k} \cdot \prod_{\substack{u \in V \setminus R: \\ u \neq v \wedge X_u=+1}} \left( \sum_{u_j \in S_u} \frac{z_{u_j}}{k} \right) \cdot \prod_{u_j \in \sigma^{-1}(+1)} \frac{z_{u_j}}{k}.
\end{aligned}$$

In equation (\*), we enumerate all  $X = Y^*$ . Since  $Y_{v_i} = +1$ , it holds that  $X_v = +1$ . For any  $u \in V \setminus R$  and  $u \neq v$ , if  $X_u = Y_u^* = +1$ , we must select one  $u_j \in C_u \setminus \Lambda = S_u$  to set  $Y_{u_j} = +1$ , which gives the factor

$\prod_{\substack{u \in V \setminus R: \\ u \neq v \wedge X_u=+1}} \left( \frac{1}{k} \sum_{u_j \in S_u} z_{u_j} \right)$  in (\*). Similarly, it holds that

$$\begin{aligned}
&\Pr_{Y \sim \mathbf{z} * \mu_k} [Y_{v_i} = -1 \wedge Y_\Lambda = \sigma] \\
&= \frac{1}{Z} \sum_{Y \in \Omega(\mu_k)} \mu(Y^*) \cdot \left( \prod_{\substack{u_j \in V_k \setminus \sigma^{-1}(1): \\ Y_{u_j}=1}} \frac{z_{u_j}}{k} \cdot \prod_{u_j \in \sigma^{-1}(1)} \frac{z_{u_j}}{k} \right) \cdot \mathbb{1}[Y_{v_i} = -1 \wedge Y_\Lambda = \sigma] \\
&= \frac{1}{Z} \sum_{\substack{X \in \Omega(\mu) \\ X_R=\tau}} \mu(X) \cdot \left( \mathbb{1}[X_v = -1] + \mathbb{1}[X_v = +1] \left( \frac{1}{k} \sum_{v_j \in S_v \setminus \{v_i\}} z_{v_j} \right) \right) \cdot \prod_{\substack{u \in V \setminus R: \\ u \neq v \wedge X_u=+1}} \left( \sum_{u_j \in S_u} \frac{z_{u_j}}{k} \right) \cdot \prod_{u_j \in \sigma^{-1}(+1)} \frac{z_{u_j}}{k}.
\end{aligned}$$

Since  $Y_{v_i} = -1$  and  $X = Y^*$ . If  $X_v = -1$ , then  $Y_{v_j} = 0$  for all  $j \in [k]$ ; if  $X_v = +1$ , since  $Y_{v_i} = -1$ , there exists  $v_j \in S_v \setminus \{v_i\}$  such that  $Y_{v_j} = +1$ . This gives the factor  $\left( \mathbb{1}[X_v = -1] + \mathbb{1}[X_v = +1] \left( \frac{1}{k} \sum_{v_j \in S_v \setminus \{v_i\}} z_{v_j} \right) \right)$  in above formula. Hence, it holds that

$$\frac{(z * \mu_k)_{v_i}^\sigma(-1)}{(z * \mu_k)_{v_i}^\sigma(+1)} = \frac{\Pr_{Y \sim \mathbf{z} * \mu_k} [Y_{v_i} = -1 \wedge Y_\Lambda = \sigma]}{\Pr_{Y \sim \mathbf{z} * \mu_k} [Y_{v_i} = +1 \wedge Y_\Lambda = \sigma]} = \frac{k}{z_{v_i}} \left( \frac{(\mathbf{x} * \mu)_v^\tau(-1)}{(\mathbf{x} * \mu)_v^\tau(+1)} + \frac{1}{k} \sum_{v_j \in S_v \setminus \{v_i\}} z_{v_j} \right),$$

where

$$(46) \quad \forall u \in V, \quad x_u \triangleq \begin{cases} \frac{1}{k} \sum_{u_j \in S_u} z_{u_j}, & u \in V \setminus R \text{ and } u \neq v \\ 1, & u \in R \text{ or } u = v. \end{cases}$$

Note that  $\mathbf{z} \in (0, 1 + \varepsilon]^{V_k}$  implies  $\mathbf{x} \in (0, 1 + \varepsilon]^V$ . □

**7.2. Marginal stability of  $\mu_k$  (Proof of Lemma 7.3).** Fix a subset  $\Lambda \subseteq V_k$  and a feasible configuration  $\sigma \in \Omega(\mu_k, V_k \setminus \Lambda)$  on  $V_k \setminus \Lambda$ . Fix a variable  $v_i \in \Lambda$  and a feasible configuration  $\tau \in \Omega(\mu_{k, \Lambda}^\sigma, \Lambda \setminus \{v_i\})$  on  $\Lambda \setminus \{v_i\}$ . Our goal is to verify the following inequalities:

$$(47) \quad \frac{\mu_{k, v_i}^{\sigma \cup \tau}(+1)}{\mu_{k, v_i}^{\sigma \cup \tau}(-1)} \leq \zeta,$$

$$(48) \quad \frac{\mu_{k,v_i}^{\sigma \cup \tau}(+1)}{\mu_{k,v_i}^{\sigma \cup \tau}(-1)} \leq 2\zeta \cdot \frac{\mu_{k,v_i}^\sigma(+1)}{\mu_{k,v_i}^\sigma(-1)}.$$

We first show that (47) and (48) together indeed guarantee the marginal stability of  $\mu_k$  that we want. By (47), we know that for any  $\gamma \in \Omega(\mu_{k,V_k \setminus \{v_i\}})$ , it holds that  $\mu_{k,v_i}^\gamma(+1)/\mu_{k,v_i}^\gamma(-1) \leq \zeta$ , which implies for any partial pinning  $\rho \in \Omega(\mu_{k,S})$ , where  $S \subseteq V_k \setminus \{v_i\}$ ,  $\mu_{k,v_i}^\rho(+1)/\mu_{k,v_i}^\rho(-1) \leq \zeta$ . Next, consider  $H \subseteq S$ . It holds that

$$\frac{\mu_{k,v_i}^\rho(+1)}{\mu_{k,v_i}^\rho(-1)} \leq \max_{\phi \in \Omega(\mu_{k,V_k \setminus (S \cup \{v_i\})}^\rho)} \frac{\mu_{k,v_i}^{\rho \cup \phi}(+1)}{\mu_{k,v_i}^{\rho \cup \phi}(-1)} \leq 2\zeta \frac{\mu_{k,v_i}^{\rho_H}(+1)}{\mu_{k,v_i}^{\rho_H}(-1)},$$

where in the last inequality we use (48) with  $\sigma = \rho_H$  and  $\sigma \cup \tau = \rho \cup \phi$ .

Our proof is reduced to verifying (47) and (48). In the rest part of this section, without loss of generality, we may assume that  $\mu_{k,v_i}^{\sigma \cup \tau}(+1) > 0$  and  $\mu_{k,v_i}^\sigma(+1) > 0$ , since when  $\mu_{k,v_i}^{\sigma \cup \tau}(+1) = 0$ , (47), (48) hold trivially; and note that  $\mu_{k,v_i}^\sigma(+1) = 0$  implies  $\mu_{k,v_i}^{\sigma \cup \tau}(+1) = 0$ .

We first proof (47). Note that  $\sigma \cup \tau$  is a configuration on  $V_k \setminus \{v_i\}$ . We use Lemma 7.11 with  $\varepsilon = 1$  and  $z = 1$ . By Lemma 7.11, there exist  $\mathbf{x} \in (0, 1]^V$  satisfying  $x_v = 1$ , a subset  $R = V \setminus \{v\}$ , and a configuration  $\rho \in \Omega(\mu_R)$  such that

$$(49) \quad \frac{\mu_{k,v_i}^{\sigma \cup \tau}(-1)}{\mu_{k,v_i}^{\sigma \cup \tau}(+1)} = k \cdot \frac{(\mathbf{x} * \mu)_v^\rho(-1)}{(\mathbf{x} * \mu)_v^\rho(+1)}.$$

Specifically, by (44) and (45), we have

$$\begin{aligned} R_- &= \{u \in V \mid \forall j \in [k], u_j \neq v_i \wedge (\sigma \cup \tau)_{u_j} = -1\}, \\ R_+ &= \{u \in V \mid \exists j \in [k], u_j \neq v_i \wedge (\sigma \cup \tau)_{u_j} = +1\}, \\ R &= R_- \cup R_+ = V \setminus \{v\}, \end{aligned}$$

and  $\rho$  defined by

$$\forall u \in R, \quad \rho_u \triangleq \begin{cases} -1, & u \in R_- \\ +1, & u \in R_+ \end{cases}.$$

Since  $x_v = 1$  and  $v \notin R$  have already hold by Lemma 7.11, we have

$$(50) \quad \frac{(\mathbf{x} * \mu)_v^\rho(-1)}{(\mathbf{x} * \mu)_v^\rho(+1)} \stackrel{(*)}{\geq} \frac{1}{\zeta} > 0,$$

where inequality (\*) holds because  $(\mathbf{x} * \mu)$  is  $\zeta$ -marginally stable. This proves (47).

Now, we bound (48). Recall that we assume  $\mu_{k,v_i}^\sigma(+1) > 0$ . By Lemma 7.11, there exist  $\mathbf{x}' \in (0, 1]^V$  satisfying  $x'_v = 1$ , a subset  $R' \subseteq V$  with  $v \notin R'$ , and a configuration  $\rho' \in \Omega(\mu_{R'})$  such that

$$(51) \quad \frac{\mu_{k,v_i}^\sigma(-1)}{\mu_{k,v_i}^\sigma(+1)} = k \left( \frac{(\mathbf{x}' * \mu)_v^{\rho'}(-1)}{(\mathbf{x}' * \mu)_v^{\rho'}(+1)} + \frac{1}{k} \sum_{v_j \in (C_v \cap \Lambda) \setminus \{v_i\}} 1 \right) \leq k \cdot \frac{(\mathbf{x}' * \mu)_v^{\rho'}(-1)}{(\mathbf{x}' * \mu)_v^{\rho'}(+1)} + k = \frac{k}{(\mathbf{x}' * \mu)_v^{\rho'}(+1)}.$$

By (44) and (45), we have

$$\begin{aligned} R'_- &= \{u \in V \mid \forall j \in [k], u_j \in V_k \setminus \Lambda \wedge \sigma_{u_j} = -1\} \\ R'_+ &= \{u \in V \mid \exists j \in [k], u_j \in V_k \setminus \Lambda \wedge \sigma_{u_j} = +1\} \\ R' &= R'_- \cup R'_+ \\ \forall u \in R', \quad \rho'_u &\triangleq \begin{cases} -1, & u \in R'_- \\ +1, & u \in R'_+ \end{cases} \end{aligned}$$

Before we progress, recall that we have assumed  $\mu_{k,v_i}^{\sigma \cup \tau}(+1) > 0$  and  $\mu_{k,v_i}^\sigma(+1) > 0$ . Combining this fact with (49), (51), it holds that  $(\mathbf{x} * \mu)_v^\rho(+1) > 0$  and  $(\mathbf{x}' * \mu)_v^{\rho'}(+1) > 0$ .

Now, in order to prove (48), we claim that

$$(52) \quad \frac{1}{(\mathbf{x}' * \mu)_v^{\rho'}(+1)} \leq 2\zeta \frac{(\mathbf{x} * \mu)_v^\rho(-1)}{(\mathbf{x} * \mu)_v^\rho(+1)}.$$

Combining (49), (51), and (52), it holds that

$$\frac{\mu_{k,v_i}^\sigma(-1)}{\mu_{k,v_i}^\sigma(+1)} \leq 2\zeta \frac{\mu_{k,v_i}^{\sigma \cup \tau}(-1)}{\mu_{k,v_i}^{\sigma \cup \tau}(+1)},$$

and this proves (48).

Now, we only left to prove (52), which, by some calculation, is equivalent to

$$(53) \quad \frac{(\mathbf{x} * \mu)_v^\rho(+1)}{(\mathbf{x} * \mu)_v^\rho(-1)} + \frac{(\mathbf{x} * \mu)_v^\rho(+1)}{(\mathbf{x} * \mu)_v^\rho(-1)} \Big/ \frac{(\mathbf{x}' * \mu)_v^{\rho'}(+1)}{(\mathbf{x}' * \mu)_v^{\rho'}(-1)} \leq 2\zeta.$$

Both the first and the second term could be bounded by the complete marginal stability of  $\zeta$ . In particular,  $(\mathbf{x} * \mu)_v^\rho = (\mathbf{x}' * \mu)_v^\rho$  holds by the fact  $\rho \in \{-1, +1\}^{V \setminus \{v\}}$  and  $x_v = x'_v$ . Therefore, the second term of (53) can be bounded by

$$\frac{(\mathbf{x} * \mu)_v^\rho(+1)}{(\mathbf{x} * \mu)_v^\rho(-1)} \Big/ \frac{(\mathbf{x}' * \mu)_v^{\rho'}(+1)}{(\mathbf{x}' * \mu)_v^{\rho'}(-1)} = \frac{(\mathbf{x}' * \mu)_v^\rho(+1)}{(\mathbf{x}' * \mu)_v^\rho(-1)} \Big/ \frac{(\mathbf{x}' * \mu)_v^{\rho'}(+1)}{(\mathbf{x}' * \mu)_v^{\rho'}(-1)} \leq \zeta,$$

where the inequality holds by  $\rho_{R'} = \rho'$  the  $\zeta$ -marginal stability of  $\mathbf{x}' * \mu$ .

## 8. APPLICATIONS TO ANTI-FERROMAGNETIC TWO-SPIN SYSTEMS

In this section, we apply Theorem 1.12 to anti-ferromagnetic 2-spin systems and prove the lower bound on the modified log-Sobolev (MLS) constant for anti-ferromagnetic two-spin systems in Theorem 1.3. Given the modified log-Sobolev bound, the mixing time bound in Theorem 1.3 is standard, whose calculation is postponed to Appendix A.

Let  $\mathcal{I} = (G = (V, E), \beta, \gamma, \lambda)$  be an anti-ferromagnetic two-spin system with Gibbs distribution  $\mu$ , where

$$(54) \quad 0 \leq \beta \leq \gamma, \lambda, \gamma > 0 \quad \text{and} \quad \beta\gamma < 1.$$

Let  $n = |V|$  and  $\Delta \geq 3$  denote the maximum degree of  $G$ . Suppose that  $\mathcal{I}$  satisfies Condition 1.2, that is:

- $(\beta, \gamma, \lambda)$  is  $(\Delta - 1)$ -unique with gap  $\delta \in (0, 1)$ ;
- $G$  is regular or  $\gamma \leq 1$ .

The following fact is folklore. A formal proof is provided in Appendix C.

**Proposition 8.1.** *Let  $(\beta, \gamma, \lambda)$  satisfy (54). Let  $\Delta \geq 3$  be an integer and  $\delta \in (0, 1)$ . If  $\gamma \leq 1$ , then  $(\beta, \gamma, \lambda)$  is up-to- $\Delta$  unique with gap  $\delta$  if and only if  $(\beta, \gamma, \lambda)$  is  $(\Delta - 1)$ -unique with gap  $\delta$ .*

With this, we can assume that  $\mathcal{I}$  satisfies the following condition that is equivalent to Condition 1.2.

**Condition 8.2.** *Let  $\delta \in (0, 1)$ . The anti-ferromagnetic two-spin system  $\mathcal{I} = (G, \beta, \gamma, \lambda)$  with maximum degree  $\Delta = \Delta_G \geq 3$  satisfies one of the following two conditions*

- $\gamma \leq 1$  and  $(\lambda, \beta, \gamma)$  is up-to- $\Delta$  unique with gap  $\delta$ ;
- $\gamma > 1$ ,  $(\lambda, \beta, \gamma)$  is  $(\Delta - 1)$ -unique with gap  $\delta$ , and  $G$  is  $\Delta$ -regular.

We will show that the modified log-Sobolev constant  $\rho^{\text{GD}}(\mu)$  of Glauber dynamics on  $\mu$  is at least  $\frac{1}{C(\delta)n}$  for some  $C(\delta) = \exp(O(1/\delta))$ .

As a preprocessing of  $\mu$ , we apply the flipping operation used in [CFYZ21b].



**Definition 8.3** (flipping operation). Let  $\mu$  be a distribution over  $\{-1, +1\}^n$ , and  $\chi \in \{-1, +1\}^n$  be a direction vector. The flipped distribution  $\pi = \text{flip}(\mu, \chi)$  over  $\{-1, +1\}^n$  is defined as

$$\forall \sigma \in \{-1, +1\}^n, \quad \pi(\sigma) \triangleq \mu(\chi \odot \sigma),$$

where  $(\chi \odot \sigma)_i \triangleq \chi_i \sigma_i$  for all  $i \in [n]$ .

In particular, if  $\chi_i = \chi \in \{-1, +1\}$  for all  $i \in [n]$ , we denote  $\text{flip}(\mu, \chi)$  by  $\text{flip}(\mu, \chi)$ .

Let  $\chi = \chi(\mathcal{I}) \in \{-1, +1\}$  be a direction indicator defined by

$$(55) \quad \chi \triangleq \begin{cases} +1, & \lambda \leq \left(\frac{\gamma}{\beta}\right)^{\Delta/2}, \\ -1, & \text{otherwise.} \end{cases}$$

Let  $\pi = \text{flip}(\mu, \chi)$ . By definition,  $\pi$  is the Gibbs distribution of  $\mathcal{I}_{\text{flip}} = (G, \bar{\beta}, \bar{\gamma}, \bar{\lambda})$ , where

$$(56) \quad (\bar{\beta}, \bar{\gamma}, \bar{\lambda}) = \begin{cases} (\beta, \gamma, \lambda) & \text{if } \lambda \leq \left(\frac{\gamma}{\beta}\right)^{\Delta/2}, \\ (\gamma, \beta, \frac{1}{\lambda}) & \text{if } \lambda > \left(\frac{\gamma}{\beta}\right)^{\Delta/2}. \end{cases}$$

Note that either  $\mathcal{I}_{\text{flip}} = \mathcal{I}$  or  $\mathcal{I}_{\text{flip}}$  is obtained by flipping the roles between  $-1$  and  $+1$  in  $\mathcal{I}$ . The following two observation about  $\mathcal{I}_{\text{flip}}$  are straightforward to verify.

**Observation 8.4.**  $\bar{\beta} \geq 0, \bar{\gamma} > 0, \bar{\beta}\bar{\gamma} < 1$  and  $0 < \bar{\lambda} \leq \left(\frac{\bar{\gamma}}{\bar{\beta}}\right)^{\Delta/2}$ .

**Observation 8.5.**  $\rho^{\text{GD}}(\mu) = \rho^{\text{GD}}(\pi)$ .

The next lemma analyzes the modified log-Sobolev constant  $\rho^{\text{GD}}(\pi)$  for flipped distribution  $\pi = \text{flip}(\mu, \chi)$ .

**Lemma 8.6.** *Let  $0 < \delta < 1$ . If  $\mathcal{I}$  with Gibbs distribution  $\mu$  satisfies Condition 8.2 with parameter  $\delta$ , then*

$$\rho^{\text{GD}}(\pi) \geq \frac{1}{C(\delta)n},$$

where  $\rho^{\text{GD}}(\pi)$  is the modified log-Sobolev constant for the Glauber dynamics on  $\pi = \text{flip}(\mu, \chi)$  with  $\chi$  defined in (55), and  $C(\delta) = \exp(O(1/\delta))$  is a constant depending only on  $\delta$ .

The MLS bound in Theorem 1.3 is a direct consequence of Proposition 8.1, Observation 8.5 and Lemma 8.6.

Lemma 8.6 can be proved by Theorem 1.12 together with the following three lemmas.

**Lemma 8.7** (complete spectral independence).  $\pi$  is  $(\frac{288}{\delta}, \frac{\delta}{2})$ -completely spectrally independent.

**Lemma 8.8** (complete marginal stability).  $\pi$  is completely  $\exp(12^5)$ -marginally stable.

**Lemma 8.9** (MLSI in subcritical regime). For any  $0 < \theta \leq 12^{-6}$ , it holds that  $\rho_{\min}^{\text{GD}}(\theta * \pi) \geq \frac{1}{4n}$ .

*Proof of Lemma 8.6.* By Lemmas 8.7-8.9, Theorem 1.12 and setting  $\theta = 12^{-6}$ , we have

$$\rho(\pi) \geq 10^{-7 \left( \frac{9000}{\delta} + \frac{10^{10}}{\log(1 + \frac{\delta}{2})} \right)} \frac{1}{4n} \stackrel{(\star)}{\geq} \frac{10^{-10^{12}/\delta}}{n} = \frac{1}{\exp(O(1/\delta))n}$$

where  $(\star)$  is due to that  $\log(1+x) \geq \frac{x}{2}$  for all  $x \in [0, 1]$ . □

**8.1. Verifying complete spectral independence.** In this section, we prove Lemma 8.7. Let  $\mathcal{I} = (G, \beta, \gamma, \lambda)$  be an anti-ferromagnetic two-spin system instance with Gibbs distribution  $\mu$  satisfying Condition 8.2 with parameter  $\delta \in (0, 1)$ . Let  $\pi = \text{flip}(\mu, \chi)$  be the flipped distribution, where  $\chi$  is defined in (55). We have the following lemma. The proof is given in Section 8.1.1.

**Lemma 8.10.** *Let  $\delta \in (0, 1)$  and  $\mathcal{I} = (G, \beta, \gamma, \lambda)$  be an instance of anti-ferromagnetic two-spin systems, then*

- For all  $1 \leq d < \Delta$ ,  $(\beta, \gamma, \lambda)$  is  $d$ -unique with gap  $\delta$  implies  $(\beta, \gamma, (1 + \frac{\delta}{2})^\chi \lambda)$  is  $d$ -unique with gap  $\frac{\delta}{2}$ .
- If  $\Delta$  further satisfies  $\Delta - 1 > (1 - \frac{\delta}{2}) \bar{\Delta}$  where  $\bar{\Delta} = \frac{1 + \sqrt{\beta\gamma}}{1 - \sqrt{\beta\gamma}}$ , then it holds that

$$\begin{aligned} \lambda \leq \left(\frac{\gamma}{\beta}\right)^{\Delta/2} &\implies \left(1 + \frac{\delta}{2}\right)^\chi \lambda < \left(\frac{\gamma}{\beta}\right)^{\Delta/2} \\ \lambda > \left(\frac{\gamma}{\beta}\right)^{\Delta/2} &\implies \left(1 + \frac{\delta}{2}\right)^\chi \lambda > \left(\frac{\gamma}{\beta}\right)^{\Delta/2}. \end{aligned}$$

Next, we need to use the following definition introduced in [CFYZ21b].

**Definition 8.11** (complete spectral independence in a direction). Let  $\eta, \varepsilon \geq 0$  and  $\chi \in \{-1, +1\}^n$ . A distribution  $\mu$  over  $\{-1, +1\}^n$  is said to be  $(\eta, \varepsilon)$ -completely spectrally independent in direction  $\chi$  if  $\theta^\chi * \mu$  is  $(\eta, \varepsilon)$ -spectrally independent for all  $\theta \in (0, 1 + \varepsilon]^V$ , where  $(\theta^\chi)_v = \theta_v^{\chi_v}$  for all  $v \in V$ .

In particular, if  $\chi$  is a constant vector such that  $\chi_v = \chi$  for all  $v \in [n]$ , we say  $\mu$  is  $(\eta, \varepsilon)$ -completely spectrally independent in direction  $\chi$  for simplicity.

We need the following lemma, whose proof is given in Section 8.1.2.

**Lemma 8.12.** *For any anti-ferromagnetic two spin system instance  $\mathcal{I} = (G, \beta, \gamma, \lambda)$  satisfying Condition 8.2 with parameter  $\delta \in (0, 1)$ , let  $\mu$  denote the Gibbs distribution of  $\mathcal{I}$ ,  $\mu$  is  $(\frac{144}{\delta}, 0)$ -completely spectrally independent in direction  $\chi(\mathcal{I})$  defined in (55), formally,*

$$\chi(\mathcal{I}) = \begin{cases} +1, & \lambda \leq \left(\frac{\gamma}{\beta}\right)^{\Delta/2} \\ -1, & \text{otherwise.} \end{cases}$$

Furthermore, if  $\Delta$  satisfies  $\Delta - 1 \leq (1 - \delta) \bar{\Delta}$  where  $\bar{\Delta} = \frac{1 + \sqrt{\beta\gamma}}{1 - \sqrt{\beta\gamma}}$ , above result holds for any  $\chi(\mathcal{I}) \in \{-1, +1\}$ .

We are now ready to prove Lemma 8.7.

*Proof of Lemma 8.7.* Let  $\nu \triangleq (1 + \frac{\delta}{2})^\chi * \mu$ , which is the Gibbs distribution of the anti-ferromagnetic two-spin system  $\mathcal{J} = (G, \beta, \gamma, (1 + \frac{\delta}{2})^\chi \lambda)$ . We prove that  $\nu$  is  $(\frac{288}{\delta}, 0)$ -completely spectrally independent in direction  $\chi = \chi(\mathcal{I})$  defined in (55). By Lemma 8.10 and the fact that  $\mathcal{J}$  shares the same parameters  $\beta, \gamma$  and graph  $G$  with  $\mathcal{I}$ , we know that  $\mathcal{J}$  satisfies Condition 8.2 with parameter  $\frac{\delta}{2}$ . Recall that  $\bar{\Delta} = \frac{1 + \sqrt{\beta\gamma}}{1 - \sqrt{\beta\gamma}}$ . We consider the following two cases.

Case  $\Delta - 1 \leq (1 - \frac{\delta}{2}) \bar{\Delta}$ . By the further more part Lemma 8.12 (remark that we use Lemma 8.12 with parameter  $\delta/2$ ),  $\nu$  is  $(\frac{288}{\delta}, 0)$ -completely spectrally independent in direction  $\chi$ .

Case  $\Delta - 1 > (1 - \frac{\delta}{2}) \bar{\Delta}$ . By Lemma 8.12, the Gibbs distribution  $\nu$  is  $(\frac{288}{\delta}, 0)$ -completely spectrally independent in direction  $\chi(\mathcal{J})$ . The second part of Lemma 8.10 shows that (1) if  $\chi = +1$ , then  $\chi(\mathcal{J}) = +1$  (2) if  $\chi = -1$ , then  $\chi(\mathcal{J}) = -1$ , which implies  $\chi = \chi(\mathcal{J})$ . Hence,  $\nu$  is  $(\frac{288}{\delta}, 0)$ -completely spectrally independent in direction  $\chi$ .

Lastly, we verify that  $\pi$  is  $(\frac{288}{\delta}, \frac{\delta}{2})$ -completely spectrally independent. Recall that  $\Psi^{\text{AbsInf}}$  is the absolute influence matrix defined in Definition 1.9. Let  $\Lambda \subseteq V$  and  $\sigma \in \Omega(\pi_{V \setminus \Lambda})$ , it is straightforward to check that

for any  $\phi \in (0, 1 + \frac{\delta}{2}]^V$ ,

$$\Psi_{(\phi * \pi)_\Lambda}^{\text{AbsInf}} = \Psi_{(\phi^\chi * \mu)_\Lambda^{\chi \circ \sigma}}^{\text{AbsInf}}$$

where  $(\chi \circ \sigma)_v = \chi \cdot \sigma_v$  for  $v \in \Lambda$  and  $(\phi^\chi)_v = \phi_v^{\chi_v}$  for  $v \in V$ . Let  $\theta \in (0, 1]^V$  such that  $\theta_v = \phi_v / (1 + \frac{\delta}{2})$  for all  $v \in V$ , it holds that  $\phi^\chi * \mu = \theta^\chi * ((1 + \frac{\delta}{2})^\chi * \mu) = \theta^\chi * \nu$ , and

$$\Psi_{(\phi * \pi)_\Lambda}^{\text{AbsInf}} = \Psi_{(\theta^\chi * \nu)_\Lambda^{\chi \circ \sigma}}^{\text{AbsInf}}.$$

Since  $\nu$  is  $(\frac{288}{\delta}, 0)$ -completely spectrally independent in direction  $\chi$ ,  $\pi$  is  $(\frac{288}{\delta}, \frac{\delta}{2})$ -completely spectrally independent.  $\square$

8.1.1. *Gap manipulation.* In this section, we prove Lemma 8.10. We need the following result.

**Lemma 8.13** ([CFYZ21b, Proposition 8.6]). *Let  $\beta, \gamma, \lambda$  be real numbers satisfying  $0 \leq \beta \leq \gamma, \gamma > 0, \lambda > 0$  and  $\beta\gamma < 1$ .*

*If  $\beta = 0$ , then the following holds for all integer  $d \geq 1$ :*

- $(0, \gamma, \lambda)$  is  $d$ -unique with gap  $\delta$  iff  $\lambda \leq \lambda_{c, \delta}(d) = \frac{(1-\delta)d^d \gamma^{d+1}}{(d-1+\delta)^{d+1}}$ .

*Assume  $\beta > 0$ . Let  $\bar{\Delta} \triangleq \frac{1+\sqrt{\beta\gamma}}{1-\sqrt{\beta\gamma}}$ . The following hold for all integers  $d \geq 1$ :*

- *If  $d \leq (1-\delta)\bar{\Delta}$ , then  $(\beta, \gamma, \lambda)$  is  $d$ -unique with gap  $\delta$  for all  $\lambda > 0$ .*
- *If  $d > (1-\delta)\bar{\Delta}$ , let  $\zeta_\delta(d) \triangleq d(1-\beta\gamma) - (1-\delta)(1+\beta\gamma)$ ,*

$$x_{1, \delta}(d) = \frac{\zeta_\delta(d) - \sqrt{\zeta_\delta(d)^2 - 4(1-\delta)^2\beta\gamma}}{2(1-\delta)\beta} \quad \text{and} \quad x_{2, \delta}(d) = \frac{\zeta_\delta(d) + \sqrt{\zeta_\delta(d)^2 - 4(1-\delta)^2\beta\gamma}}{2(1-\delta)\beta},$$

*and for  $i \in \{1, 2\}$ , let*

$$\lambda_{i, \delta}(d) = x_{i, \delta}(d) \left( \frac{x_{i, \delta}(d) + \gamma}{\beta x_{i, \delta}(d) + 1} \right)^d.$$

*It holds that  $\lambda_{1, \delta}(d)\lambda_{2, \delta}(d) = \left(\frac{\gamma}{\beta}\right)^{d+1}$  and  $\lambda_{1, \delta}(d) < \left(\frac{\gamma}{\beta}\right)^{(d+1)/2} < \lambda_{2, \delta}(d)$ . And  $(\beta, \gamma, \lambda)$  is  $d$ -unique if and only if  $\lambda \in (0, \lambda_{1, \delta}(d)] \cup [\lambda_{2, \delta}(d), \infty)$ .*

To prove Lemma 8.10, for all  $1 \leq d < \Delta$ , we will show that if  $(\beta, \gamma, \lambda)$  is  $d$ -unique with gap  $\delta$ , then  $(\beta, \gamma, (1 + \frac{\delta}{2})^\chi \lambda)$  is  $d$ -unique with gap  $\delta/2$ . We consider 3 cases: (1)  $\beta = 0$ ; (2)  $\beta > 0$  and  $1 \leq d \leq (1-\delta)\bar{\Delta}$ ; (3)  $\beta > 0$  and  $(1-\delta)\bar{\Delta} < d < \Delta$ .

**Case (1):**  $\beta = 0$ . Since  $\beta = 0$ , it holds that  $\chi = +1$ . By Lemma 8.13, our goal is to show that

$$\left(1 + \frac{\delta}{2}\right) \frac{(1-\delta)d^d \gamma^{d+1}}{(d-1+\delta)^{d+1}} = \left(1 + \frac{\delta}{2}\right) \lambda_{c, \delta}(d) \leq \lambda_{c, \delta/2}(d) = \frac{(1-\frac{\delta}{2})d^d \gamma^{d+1}}{(d-1+\frac{\delta}{2})^{d+1}},$$

which holds because  $(1-\delta)(1+\delta/2) \leq (1-\delta/2)$ .

**Case (2):**  $\beta > 0$  and  $1 \leq d \leq (1-\delta)\bar{\Delta}$ . In this case, Lemma 8.13 tells us that  $(\beta, \gamma, (1 + \frac{\delta}{2})^\chi \lambda)$  is  $d$ -unique with gap  $\delta$ , and hence it is  $d$ -unique with gap  $\frac{\delta}{2}$ .

**Case (3):**  $\beta > 0$  and  $(1-\delta)\bar{\Delta} < d < \Delta$ . Without loss of generality, we assume that  $(\Delta - 1) > (1-\delta)\bar{\Delta}$ .

Fix an integer  $d$  such that  $(1-\delta)\bar{\Delta} < d < \Delta$ . We consider 4 sub-cases: (i)  $\lambda \leq \lambda_{1, \delta}(d)$  and  $\chi = -1$ ; (ii)  $\lambda \geq \lambda_{2, \delta}(d)$  and  $\chi = +1$ . (iii)  $\lambda \leq \lambda_{1, \delta}(d)$  and  $\chi = +1$ ; (iv)  $\lambda \geq \lambda_{2, \delta}(d)$  and  $\chi = -1$ .

Note that without loss of generality, we always assume that  $\lambda_{1, \delta/2}(d)$  and  $\lambda_{2, \delta/2}(d)$  are well defined. Otherwise,  $d < (1-\frac{\delta}{2})\bar{\Delta}$ , and by Lemma 8.13, it holds that  $(\beta, \gamma, (1 + \frac{\delta}{2})^\chi \lambda)$  is  $d$ -unique with gap  $\frac{\delta}{2}$ .

For case (i), it holds that  $(1 + \frac{\delta}{2})^{-1} \lambda \leq \lambda \leq \lambda_{1, \delta}(d)$ , and the proof is done by leveraging Lemma 8.13.

The case (ii) could be proved in the same manner as the case (i).

To prove case (iii), by Lemma 8.13, it suffices for us to show that  $(1 + \frac{\delta}{2})\lambda_{1,\delta}(d) \leq \lambda_{1,\delta/2}(d)$ , which is already done by the previous work [AJK<sup>+</sup>21b, Proof of Proposition 66]. We remark that their proof works for all  $\beta, \gamma > 0$  satisfying  $\beta\gamma < 1$ .

We left to prove case (iv). Note that if we fix the parameter  $d, \delta$ , then  $\lambda_{1,\delta}(d)$  and  $\lambda_{2,\delta}(d)$  are actually functions of  $\beta, \gamma$ . For convenience, we denote them as  $\lambda_{1,\delta}(d; \beta, \gamma)$  and  $\lambda_{2,\delta}(d; \beta, \gamma)$ , respectively. Let  $\beta' = \gamma, \gamma' = \beta$ , it holds that

$$\lambda_{2,\delta}(d; \beta, \gamma) = 1/\lambda_{1,\delta}(d; \beta', \gamma').$$

It suffices to show that  $(1+\delta/2)^{-1}\lambda_{2,\delta}(d; \beta, \gamma) \geq \lambda_{2,\delta/2}(d; \beta, \gamma)$ , which is equivalent to  $(1+\delta/2)\lambda_{1,\delta}(d; \beta', \gamma') \leq \lambda_{1,\delta/2}(d; \beta', \gamma')$ , which is proved in case (iii).

Finally, we prove the second part in Lemma 8.10. Let  $d = \Delta - 1$ , by our assumption in Lemma 8.10, it holds that  $d > (1 - \delta/2)\bar{\Delta} > (1 - \delta)\bar{\Delta}$ . By Lemma 8.13, we have the following two results (1)  $\lambda_{1,\delta}(d)$  and  $\lambda_{2,\delta}(d)$  exist; (2)  $\lambda_{1,\delta/2}(d)$  and  $\lambda_{2,\delta/2}(d)$  exist.

- If  $\lambda \leq (\frac{\gamma}{\beta})^{\Delta/2}$ , then  $\chi = +1$ . By case (3.iii),  $(1 + \frac{\delta}{2})\lambda \leq (1 + \frac{\delta}{2})\lambda_{1,\delta}(d) \leq \lambda_{1,\delta/2}(d) < (\frac{\gamma}{\beta})^{\Delta/2}$ .
- If  $\lambda > (\frac{\gamma}{\beta})^{\Delta/2}$ , then  $\chi = -1$ . Let  $\beta' = \gamma, \gamma' = \beta, \lambda' = 1/\lambda$ , then by case (3.iv), it holds that  $(1 + \frac{\delta}{2})\lambda' \leq (1 + \frac{\delta}{2})\lambda_{1,\delta}(d; \beta', \gamma') \leq \lambda_{1,\delta/2}(d; \beta', \gamma') < (\frac{\gamma'}{\beta'})^{\Delta/2}$ , which implies  $(1 + \frac{\delta}{2})^{-1}\lambda > (\frac{\gamma}{\beta})^{\Delta/2}$ .

**8.1.2. Complete spectral independence of  $\mu$  in direction  $\chi$ .** We prove Lemma 8.12. Fix an anti-ferromagnetic two-spin system instance  $\mathcal{I} = (G, \beta, \gamma, \lambda)$  satisfying Condition 8.2 with parameter  $\delta \in (0, 1)$ . Let  $\mu$  denote the Gibbs distribution of  $\mathcal{I}$ . We prove that  $\mu$  is  $(\frac{14\delta}{\delta}, 0)$ -completely spectrally independent in direction  $\chi$  defined in (55). Fix an arbitrary  $\theta \in (0, 1]^V$ . We show that  $\nu \triangleq \theta^\chi * \mu$  is  $(\frac{7\delta}{\delta}, 0)$ -spectrally independent, which implies the lemma. Note that  $\nu$  is the Gibbs distribution of the two-spin system defined by the tuple  $(G = (V, E), \beta, \gamma, \lambda)$ , where  $\lambda = (\lambda_v)_{v \in V} \in \mathbb{R}_{>0}^V$  satisfies  $\lambda_v = \theta_v^\chi \lambda$ .

First, we introduce some notations and results. For  $\lambda > 0$ , integer  $d \geq 0$ , consider tree recursion for *log-marginal-ratios*  $H_{\lambda,d} : [-\infty, +\infty]^d \rightarrow [-\infty, +\infty]$ ,

$$H_{\lambda,d}(y_1, \dots, y_d) \triangleq \log \lambda + \sum_{i=1}^d \log \left( \frac{\beta e^{y_i} + 1}{e^{y_i} + \gamma} \right)$$

For  $y \in [-\infty, +\infty]$ , let

$$h(y) \triangleq -\frac{(1 - \beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)}.$$

For real number  $\lambda > 0$ , integer  $d > 0$ , we define the intervals  $J_{\lambda,d}$  as follow

$$J_{\lambda,d} = \begin{cases} \left[ -\infty, \log \left( \frac{\lambda}{\gamma^d} \right) \right] & \text{if } \beta = 0; \\ \left[ \log(\lambda\beta^d), \log \left( \frac{\lambda}{\gamma^d} \right) \right] & \text{if } 0 < \beta\gamma \leq 1. \end{cases}$$

Specially, when  $\lambda > 0$  and  $d = 0$ , let  $J_{\lambda,0} = \{\log \lambda\}$ .

We use the following known results about two-spin systems.

**Lemma 8.14** ([CFYZ21b, Theorem 8.8], [CLV20]). *Let  $\nu$  be the Gibbs distribution of a two-spin system defined by graph  $G = (V, E)$ , and parameters  $\beta, \gamma \in \mathbb{R}, \lambda \in \mathbb{R}^V$  such that  $0 \leq \beta \leq \gamma, \gamma > 0, \beta\gamma < 1$ , and  $\lambda_v > 0$  for all  $v \in V$ . For every  $v \in V$ , let  $d_v \triangleq \Delta_v - 1$  where  $\Delta_v$  is the degree of  $v$  in  $G$ . If there exists  $\alpha, c > 0$  such that*

- (1) *for every  $v \in V$  with  $d_v \geq 1$  and every  $(y_1, \dots, y_{d_v}) \in [-\infty, +\infty]^{d_v}$ , it holds that*

$$\sum_{i=1}^{d_v} \sqrt{|h(y)| |h(y_i)|} \leq 1 - \alpha,$$

where  $y = H_{\lambda,d_v}(y_1, \dots, y_{d_v})$ ;

(2) for every  $v \in V$ , every  $y_v \in J_{\lambda_v, d_v}$ , it holds that

$$|h(y_v)| \leq \frac{c}{\Delta},$$

then  $v$  is  $(\frac{2c}{\alpha}, 0)$ -spectrally independent.

**Lemma 8.15** ([CFYZ21b, Theorem 8.11], [LLY13]). *Let  $d \geq 1$  be an integer, and let  $\beta, \gamma, \lambda$  be real numbers satisfying that  $0 \leq \beta \leq \gamma$ ,  $\gamma > 0$ ,  $\lambda > 0$ , and  $\beta\gamma < 1$ . For any  $\delta \in (0, 1)$ , if  $(\beta, \gamma, \lambda)$  is  $d$ -unique with gap  $\delta$ , then for every  $(y_1, \dots, y_d) \in [-\infty, +\infty]^d$  and  $y = H_{\lambda, d}(y_1, y_2, \dots, y_d)$ , it holds that*

$$\sum_{i=1}^d \sqrt{|h(y)| |h(y_i)|} < 1 - \frac{\delta}{2}.$$

**Lemma 8.16** ([CLV20, Lemma 36]). *Let  $\Delta \geq 3$  be an integer, and let  $\beta, \gamma, \lambda$  be real numbers satisfying that  $0 \leq \beta \leq \gamma$ ,  $\gamma > 0$ ,  $\lambda > 0$ , and  $\beta\gamma < 1$ . Suppose  $(\beta, \gamma, \lambda)$  is  $(\Delta - 1)$ -unique. It holds that*

- if  $\gamma \leq 1$ , then for  $0 \leq d < \Delta$ , and every  $y \in J_{\lambda, d}$ , it holds that  $|h(y)| \leq \frac{18}{\Delta}$ ;
- if  $G$  is  $\Delta$ -regular, then for  $d = \Delta - 1$  and every  $y \in J_{\lambda, d}$ , it holds that  $|h(y)| \leq \frac{18}{\Delta}$ .

**Remark 8.17.** The exact statement of Lemma 8.16 is slightly different from [CLV20, Lemma 36], but it can be verified by going through the same proof for [CLV20, Lemma 36]. For completeness, a proof of Lemma 8.16 is provided in Appendix D.

By Lemma 8.14, Lemma 8.15, and Lemma 8.16, to prove that  $v$  is  $(\frac{72}{\delta}, 0)$ -spectrally independent, we only need to prove one of the following two results

- $(\beta, \gamma, \theta^x \lambda)$  is up-to- $\Delta$  unique;
- $G$  is  $\Delta$ -regular and  $(\beta, \gamma, \theta^x \lambda)$  is  $(\Delta - 1)$ -unique.

Note that the spin system  $\mathcal{I} = (G, \beta, \gamma, \lambda)$  in Lemma 8.12 satisfies Condition 8.2 with parameter  $\delta$ . The above two results can be proved by the following lemma.

**Lemma 8.18.** *Let  $0 < \delta < 1$ . Let  $G = (V, E)$  be a graph with maximum degree  $\Delta \geq 3$ . Let  $\beta, \gamma, \lambda$  be real numbers satisfying that  $0 \leq \beta \leq \gamma$ ,  $\gamma > 0$ ,  $\lambda > 0$ , and  $\beta\gamma < 1$ . Let  $\chi$  be the parameter defined in (55),  $\delta \in (0, 1)$ , and  $\theta \in (0, 1]$ , it holds that*

- if  $\gamma \leq 1$ , then  $(\beta, \gamma, \lambda)$  is up-to- $\Delta$  unique with gap  $\delta$  implies  $(\beta, \gamma, \theta^x \lambda)$  is up-to- $\Delta$  unique with gap  $\delta$ ;
- $(\beta, \gamma, \lambda)$  is  $(\Delta - 1)$ -unique with gap  $\delta$  implies  $(\beta, \gamma, \theta^x \lambda)$  is  $(\Delta - 1)$ -unique with gap  $\delta$ .

*Proof.* We prove the first part of the lemma. Assume  $\gamma \leq 1$ . By definition, we need to prove that for every  $1 \leq d < \Delta$ ,  $(\beta, \gamma, \theta^x \lambda)$  is  $d$ -unique with gap  $\delta$ . We consider 3 cases: (1)  $\beta = 0$ ; (2)  $\beta > 0$  and  $d \leq (1 - \delta)\bar{\Delta}$ ; (3)  $\beta > 0$  and  $d > (1 - \delta)\bar{\Delta}$ , where  $\bar{\Delta} \triangleq \frac{1 + \sqrt{\beta\gamma}}{1 - \sqrt{\beta\gamma}}$ .

**Case (1):**  $\beta = 0$ . Fix  $1 \leq d < \Delta$ . In this case, it holds that  $\chi = +1$ . Hence, it holds that  $\theta^x \lambda \leq \lambda \leq \lambda_{c, \delta}$ , where  $\lambda_{c, \delta}$  is defined in Lemma 8.13. By Lemma 8.13, we have  $(\beta, \gamma, \lambda)$  is  $d$ -unique with gap  $\delta$ .

**Case (2):**  $\beta > 0$  and  $d \leq (1 - \delta)\bar{\Delta}$ . In this case,  $(\beta, \gamma, \theta^x \lambda)$  is  $d$ -unique with gap  $\delta$  due to Lemma 8.13.

**Case (3):**  $\beta > 0$  and  $d > (1 - \delta)\bar{\Delta}$ . To handle this case, we need the following result.

**Lemma 8.19** ([LLY13, Lemma 21 (7)]). *Let  $\Delta \geq 3$  be an integer, and let  $\beta, \gamma, \lambda$  be real numbers such that  $0 \leq \beta \leq \gamma \leq 1$ ,  $\gamma > 0$ ,  $\lambda > 0$ . Let  $\delta \in (0, 1)$  be a real number. Then  $(\beta, \gamma, \lambda)$  is up-to- $\Delta$  unique with gap  $\delta$  if and only if  $\lambda \in (0, \lambda_{1, \delta}] \cup [\lambda_{2, \delta}, \infty)$  where*

$$\lambda_{1, \delta} \triangleq \min_{(1-\delta)\bar{\Delta} < d < \Delta} \lambda_{1, \delta}(d)$$

$$\lambda_{2, \delta} \triangleq \max_{(1-\delta)\bar{\Delta} < d < \Delta} \lambda_{2, \delta}(d),$$

where  $\lambda_{1, \delta}(d)$  and  $\lambda_{2, \delta}(d)$  are defined in Lemma 8.13.

Lemma 8.19 can be verified by routinely going through the proof in [LLY13] and taking the gap  $\delta$  into consideration.

We assume that  $(\Delta - 1) > (1 - \delta)\bar{\Delta}$ . Otherwise, the integer  $(1 - \delta)\bar{\Delta} < d < \Delta$  does not exist. If  $\chi = +1$ , then it holds that  $\lambda \leq \left(\frac{\gamma}{\beta}\right)^{\Delta/2}$ . By Lemma 8.13, it holds that  $\lambda \leq \lambda_{1,\delta}(\Delta - 1) < \lambda_{2,\delta}(\Delta - 1) \leq \lambda_{2,\delta}$ . Hence, by Lemma 8.19, we could conclude that  $\lambda \leq \lambda_{1,\delta}$ . Hence for all  $(1 - \delta)\bar{\Delta} < d < \Delta$ , it holds that  $\theta^x \lambda \leq \lambda \leq \lambda_{1,\delta}(d)$ . By Lemma 8.13, it holds that  $(\beta, \gamma, \theta^x \lambda)$  is  $d$ -unique with gap  $\delta$ . The case  $\chi = -1$  can be proved in a similar way.

We prove the second part of the lemma. Again, we consider three cases: (1)  $\beta = 0$ ; (2)  $\beta > 0$  and  $\Delta - 1 \leq (1 - \delta)\bar{\Delta}$ ; (3)  $\beta > 0$  and  $\Delta - 1 > (1 - \delta)\bar{\Delta}$ . Case (1) and (2) follow from the same proof. For case (3), we cannot use Lemma 8.19 because we no longer have  $\gamma \leq 1$ . However, for the second part, we only need to prove  $(\beta, \gamma, \theta^x \lambda)$  is  $(\Delta - 1)$ -unique. If  $\chi = +1$ , then it holds that  $\lambda \leq \left(\frac{\gamma}{\beta}\right)^{\Delta/2}$ . By Lemma 8.13, it holds that  $\lambda \leq \lambda_{1,\delta}(\Delta - 1)$ . By Lemma 8.13, it holds that  $(\beta, \gamma, \theta^x \lambda)$  is  $(\Delta - 1)$ -unique with gap  $\delta$ . The case  $\chi = -1$  can be proved in a similar way.  $\square$

Finally, we prove the *furthermore* part of Lemma 8.12, which states that if  $(\Delta - 1) \leq (1 - \delta)\bar{\Delta}$ , then for  $\theta \in (0, 1]^V$ ,  $\nu \triangleq \theta^x * \mu$  is  $(\frac{144}{\delta}, 0)$ -spectrally independent for all  $\chi \in \{-1, +1\}$ .

**Lemma 8.20** ([CLV20, Lemma 36]). *Let  $\Delta \geq 3$  be an integer, and let  $\beta, \gamma, \lambda$  be real numbers satisfying that  $0 \leq \beta \leq \gamma, \gamma > 0, \lambda > 0, \beta\gamma < 1$ , and  $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$ . For every  $y \in [-\infty, +\infty]$ , it holds that*

$$|h(y)| \leq \frac{1.5}{\Delta}.$$

**Remark 8.21.** Lemma 8.20 is the case S.1 in [CLV20, Lemma 36]. In [CLV20], the result is stated for  $y \in J$  for some interval  $J$ . The proof works for all  $y \in [-\infty, +\infty]$  (see proof of Lemma 36 in [CLV20]).

Note that  $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$  is equivalent to  $(\Delta - 1) < \bar{\Delta}$ , which can be deduced from  $(\Delta - 1) \leq (1 - \delta)\bar{\Delta}$ . Note that the boundedness condition is guaranteed by Lemma 8.20. By Lemma 8.14, Lemma 8.15, and Lemma 8.20, it suffices for us to show that for any  $\theta \in (0, 1]$ , every  $\chi \in \{-1, +1\}$ , and every  $1 \leq d < \Delta$ ,  $(\beta, \gamma, \lambda)$  is  $d$ -unique with gap  $\delta$  implies that  $(\beta, \gamma, \theta^x \lambda)$  is  $d$ -unique, which holds trivially by leveraging Lemma 8.13.

**8.2. Verifying complete marginal stability.** In this section, we prove Lemma 8.8. Recall that  $\mathcal{I} = (G, \beta, \gamma, \lambda)$  is an anti-ferromagnetic two-spin system instance satisfying Condition 8.2 with parameter  $\delta \in (0, 1)$ . Let  $\Delta \geq 3$  denote the maximum degree of  $G$ . Let  $\mu$  denote the Gibbs distribution of  $\mathcal{I}$ . Let  $\pi = \text{flip}(\mu, \chi)$  be the flipped distribution, where  $\chi$  is defined in (55). We show that  $\pi$  is completely  $\exp(12^5)$ -marginally stable.

Recall that  $\pi$  is the Gibbs distribution of  $\mathcal{I}_{\text{flip}} = (G, \bar{\beta}, \bar{\gamma}, \bar{\lambda})$  defined in (56). By Observation 8.4,

$$(57) \quad \bar{\beta} \geq 0, \bar{\gamma} > 0, \bar{\beta}\bar{\gamma} < 1, \text{ and } 0 < \bar{\lambda} \leq \left(\frac{\bar{\gamma}}{\bar{\beta}}\right)^{\Delta/2},$$

To establish the complete marginal stability, we need to show that  $(\phi * \pi)$  is marginally stable for all  $\phi \in (0, 1]^V$ . Equivalently, we consider the more general two-spin system instance  $\mathcal{J} = (G = (V, E), \bar{\beta}, \bar{\gamma}, (\bar{\lambda}_v)_{v \in V})$  with local fields such that

$$(58) \quad \bar{\beta} \geq 0, \bar{\gamma} > 0, \bar{\beta}\bar{\gamma} < 1, \text{ and } 0 < \bar{\lambda}_v \leq \left(\frac{\bar{\gamma}}{\bar{\beta}}\right)^{\Delta/2} \quad \forall v \in V.$$

Let  $\nu$  be the Gibbs distribution of  $\mathcal{J}$ , we will show that  $\nu$  is  $\exp(12^5)$ -marginally stable. Note that  $\mathcal{I} = (G, \beta, \gamma, \lambda)$  satisfies Condition 8.2, which implies

$$(59) \quad G \text{ is regular or } \max\{\bar{\beta}, \bar{\gamma}\} \leq 1.$$

To prove Lemma 8.8, we need the following technical lemmas.

**Lemma 8.22.** For any  $0 < \bar{\lambda}_v \leq \left(\frac{\bar{\gamma}}{\bar{\beta}}\right)^{\Delta/2}$ ,  $(\bar{\beta}, \bar{\gamma}, \bar{\lambda}_v)$  is  $(\Delta - 1)$ -unique (with gap 0) and it holds that

- $\bar{\lambda}_v \bar{\gamma}^{-\Delta} \leq 12^4$ ;
- $\bar{\lambda}_v \bar{\gamma}^{-\Delta} (1 - \bar{\beta} \bar{\gamma}) \leq \frac{12^5}{\Delta}$ .

We remark that compared to the assumption in (2),  $(\bar{\beta}, \bar{\gamma}, \bar{\lambda}_v)$  may not always satisfy  $\bar{\beta} \leq \bar{\gamma}$ , but the definition of the uniqueness condition literally follows Definition 1.1. The uniqueness condition is well-defined because  $F_d(x) = x$  has a unique solution if  $\bar{\beta} \bar{\gamma} < 1$ .

**Lemma 8.23.** Let  $d = \Delta - 1$  and  $\bar{\lambda}_{\max} = \max_{u \in V} \bar{\lambda}_u$ . Let  $v \in V$ ,  $S \subseteq \Lambda \subseteq V \setminus \{v\}$ , and  $\sigma \in \Omega(\mu_\Lambda)$  be a partial pinning. It holds that

$$R_v^\sigma \leq \frac{F_\Delta^{\bar{\lambda}_{\max}}(0)}{F_\Delta^{\bar{\lambda}_{\max}} \circ F_d^{\bar{\lambda}_{\max}}(0)} R_v^{\sigma_S},$$

where  $R_v^\sigma = v_v^\sigma(+1)/v_v^\sigma(-1)$  is the marginal ratio of  $v^\sigma$ , and

$$(60) \quad \forall \lambda > 0, d \in \mathbb{Z}_{>0}, \quad F_d^\lambda(x) = \lambda \left( \frac{\beta x + 1}{x + \gamma} \right)^d,$$

is the uniform tree-recursion function.

**Remark 8.24.** Intuitively, Lemma 8.23 says that the worst case of  $R_v^\sigma/R_v^{\sigma_S}$  is achieved by a  $\Delta$ -regular tree rooted at  $v$ , where  $\sigma_S$  fixes the values of all the vertices in  $\{u \in V \mid \text{dist}_G(u, v) = 2\}$  to  $-1$  and  $\sigma$  further fixes the values of all the vertices in  $\{u \in V \mid \text{dist}_G(u, v) = 1\}$  to  $-1$ .

The proofs of Lemma 8.22 and Lemma 8.23 are deferred to Section 8.2.1 and Section 8.2.2 respectively. We are ready to prove Lemma 8.8.

*Proof of Lemma 8.8.* Let  $\nu$  be the Gibbs distribution of  $\mathcal{J}$ . To prove that  $\pi$  is complete  $\exp(12^5)$ -marginally stable, it suffices for us to show that  $\nu$  is  $\exp(12^5)$ -marginally stable.

Let  $S \subseteq \Lambda \subseteq V$ ,  $v \in V \setminus \Lambda$  and  $\sigma \in \Omega(\nu_{V \setminus \Lambda})$  be a partial configuration on  $V \setminus \Lambda$ . We will show that

$$R_v^\sigma \leq C \text{ and } R_v^\sigma \leq C R_v^{\sigma_S},$$

where  $R_v^\sigma \triangleq \frac{v_v^\sigma(+1)}{v_v^\sigma(-1)}$  denotes the marginal ratio of  $v^\sigma$ , and  $C = \exp(12^5)$  be a universal constant.

For the first part, by considering the worst pinning of all neighbors of  $v$ , we have

$$R_v^\sigma \leq \bar{\lambda}_v \bar{\gamma}^{-\Delta} \leq 12^4 \leq C,$$

where inequalities follow from anti-ferromagnetism and Lemma 8.22 respectively. For the second part, we may assume that  $R_v^{\sigma_S} > 0$ , otherwise  $R_v^\sigma = R_v^{\sigma_S} = 0$ . By Lemma 8.23, it holds that

$$\begin{aligned} \frac{R_v^\sigma}{R_v^{\sigma_S}} &\leq \frac{F_\Delta^{\bar{\lambda}_{\max}}(0)}{F_\Delta^{\bar{\lambda}_{\max}} \circ F_d^{\bar{\lambda}_{\max}}(0)} = \frac{\bar{\lambda}_{\max} \bar{\gamma}^{-\Delta}}{\bar{\lambda}_{\max} \left( \frac{\bar{\beta} \bar{\lambda}_{\max} \bar{\gamma}^{-d+1}}{\bar{\lambda}_{\max} \bar{\gamma}^{-d+\bar{\gamma}}} \right)^\Delta} = \left( \frac{\bar{\lambda}_{\max} + \bar{\gamma}^{d+1}}{\bar{\gamma}(\bar{\lambda}_{\max} \bar{\beta} + \bar{\gamma}^d)} \right)^\Delta \\ &= \left( 1 + \frac{\bar{\lambda}_{\max}(1 - \bar{\beta} \bar{\gamma})}{\bar{\lambda}_{\max} \bar{\beta} \bar{\gamma} + \bar{\gamma}^{d+1}} \right)^\Delta \leq \left( 1 + \frac{\bar{\lambda}_{\max}(1 - \bar{\beta} \bar{\gamma})}{\bar{\gamma}^{d+1}} \right)^\Delta \stackrel{(\star)}{\leq} \left( 1 + \frac{12^5}{\Delta} \right)^\Delta \leq \exp(12^5) = C, \end{aligned}$$

where  $(\star)$  holds by Lemma 8.22. This concludes the proof.  $\square$

8.2.1. *Proof of Lemma 8.22.* In this section, we prove Lemma 8.22. We first show that  $(\bar{\beta}, \bar{\gamma}, \bar{\lambda}_v)$  is  $d$ -unique (with gap 0) for  $d = \Delta - 1$ . Note that by Condition 8.2,  $(\beta, \gamma, \lambda)$  is  $d$ -unique with gap  $\delta$ .

Suppose  $\lambda \leq (\frac{\gamma}{\beta})^{\Delta/2}$ , then we have  $\bar{\lambda} = \lambda, \bar{\beta} = \beta, \bar{\gamma} = \gamma$ . By Lemma 8.13, we know that when  $d \leq (1 - \delta)\bar{\Delta}$ , it holds that  $(\bar{\beta}, \bar{\gamma}, \bar{\lambda}_v)$  is  $d$ -unique with gap  $\delta$ ; and when  $d > (1 - \delta)\bar{\Delta}$ , it holds that  $\bar{\lambda}_v \leq \lambda \leq \lambda_{1,\delta}(d) < (\frac{\gamma}{\beta})^{\Delta/2}$ , which implies  $(\bar{\beta}, \bar{\gamma}, \bar{\lambda}_v)$  is  $d$ -unique with gap  $\delta$ . The  $\lambda > (\frac{\gamma}{\beta})^{\Delta/2}$  case is almost the same by noticing that when we fix  $d$  and  $\delta$ , then  $\lambda_{1,\delta}(d)$  and  $\lambda_{2,\delta}(d)$  are actually functions of  $\beta, \gamma$  that could be written as  $\lambda_{1,\delta}(d; \beta, \gamma), \lambda_{2,\delta}(d; \beta, \gamma)$ , and

$$\lambda \geq \lambda_{2,\delta}(d; \beta, \gamma) \iff \bar{\lambda} \leq \lambda_{1,\delta}(d; \bar{\beta}, \bar{\gamma}),$$

where  $\bar{\lambda} = 1/\lambda, \bar{\beta} = \gamma$ , and  $\bar{\gamma} = \beta$ . Finally, note that since  $(\bar{\beta}, \bar{\gamma}, \bar{\lambda}_v)$  is  $d$ -unique with gap  $\delta$ , it is also  $d$ -unique (with gap 0).

Combining with Observation 8.4, it suffices to prove the following result: for any  $\beta, \gamma, \lambda$  with  $\beta \geq 0, \gamma > 0, \beta\gamma < 1$  and  $0 < \lambda \leq (\gamma/\beta)^{\Delta/2}$  that is  $(\Delta - 1)$ -unique, it holds that  $\lambda\gamma^{-\Delta} \leq 12^4$  and  $\lambda\gamma^{-\Delta}(1 - \beta\gamma) \leq 12^5/\Delta$ . We need the following lemma.

**Lemma 8.25** ([CLV20, Lemma 35]). *Let  $\Delta \geq 3$  be an integer and  $d \triangleq \Delta - 1$ . Let  $\beta, \gamma, \lambda$  be real numbers such that  $\beta \geq 0, \gamma > 0, \beta\gamma < 1, \lambda > 0$  and  $(\beta, \gamma, \lambda)$  is  $d$ -unique (with gap 0).*

- (1) *If  $\beta = 0$ , then we have  $\lambda \leq \frac{4\gamma^{d+1}}{d-1}$ .*
- (2) *If  $\beta > 0$  and  $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$ , it holds that*

$$\text{either } \lambda \leq \frac{18\gamma^{d+1}}{\theta(d)} \quad \text{or } \lambda \geq \frac{\theta(d)}{18\beta^{d+1}},$$

where  $\theta(d) \triangleq d(1 - \beta\gamma) - (1 + \beta\gamma)$ .

**Remark 8.26.** Lemma 35 in [CLV20] further assumes  $\beta \leq \gamma$ . We remark that Lemma 8.25 can be verified by routinely going through the proof in [CLV20].

We first show that  $\lambda\gamma^{-\Delta} \leq 12^4$ . Let  $\bar{\Delta} \triangleq \frac{1+\sqrt{\beta\gamma}}{1-\sqrt{\beta\gamma}}$ . We consider 3 cases: (1)  $\beta > 0$  and  $\Delta < 2\bar{\Delta}$ ; (2)  $\beta > 0$  and  $\Delta \geq 2\bar{\Delta}$ ; (3)  $\beta = 0$ .

**Case (1):  $\beta > 0$  and  $\Delta < 2\bar{\Delta}$ .** Note that we have  $\lambda \leq \left(\frac{\gamma}{\beta}\right)^{\Delta/2}$ , so it holds that

$$\lambda\gamma^{-\Delta} \leq \left(\frac{1}{\beta\gamma}\right)^{\Delta/2} \leq (\beta\gamma)^{-\Delta} \leq (\beta\gamma)^{-2\bar{\Delta}}.$$

Note that from  $\bar{\Delta} = \frac{1+\sqrt{\beta\gamma}}{1-\sqrt{\beta\gamma}}$ , we have  $(\beta\gamma)^{-1} = \left(\frac{\bar{\Delta}+1}{\bar{\Delta}-1}\right)^2$ . Moreover, we have

$$(\beta\gamma)^{-2\bar{\Delta}} = \left(\frac{\bar{\Delta}+1}{\bar{\Delta}-1}\right)^{4\bar{\Delta}} \leq 12^4,$$

where in the last inequality, we use the fact that  $3 \leq \Delta < 2\bar{\Delta}$  which means  $\bar{\Delta} > \frac{3}{2}$ .

**Case (2):  $\beta > 0$  and  $\Delta \geq 2\bar{\Delta}$ .** In this case,  $d \geq \frac{2}{3}\Delta \geq \frac{4}{3}\bar{\Delta} \geq \bar{\Delta}$  is achieved, which means  $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$ . In this case, by Lemma 8.25 with  $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$  and the fact that  $\lambda \leq \left(\frac{\gamma}{\beta}\right)^{\Delta/2}$ , it holds that

$$\lambda\gamma^{-\Delta} \leq \frac{18}{\theta(d)},$$



where  $d \triangleq \Delta - 1$  and  $\theta(d) = d(1 - \beta\gamma) - (1 + \beta\gamma)$ . Note that we have

$$\begin{aligned}\theta(d) &\stackrel{(\star)}{=} d \left( 1 - \left( \frac{\bar{\Delta} - 1}{\bar{\Delta} + 1} \right)^2 \right) - \left( 1 + \left( \frac{\bar{\Delta} - 1}{\bar{\Delta} + 1} \right)^2 \right) \\ &= \frac{4\bar{\Delta}d - 2(\bar{\Delta}^2 + 1)}{(\bar{\Delta} + 1)^2} \stackrel{(*)}{\geq} \frac{\frac{4}{3} \cdot 4\bar{\Delta}^2 - 2(\bar{\Delta}^2 + 1)}{(\bar{\Delta} + 1)^2} \stackrel{(+)}{\geq} \frac{\frac{4}{3}\bar{\Delta}^2}{(\bar{\Delta} + 1)^2} \geq \frac{1}{3},\end{aligned}$$

where  $(\star)$  holds by the fact that  $\beta\gamma = \left( \frac{\bar{\Delta} - 1}{\bar{\Delta} + 1} \right)^2$ ,  $(*)$  holds by the fact that  $d \geq \frac{2}{3}\bar{\Delta} \geq \frac{4}{3}\bar{\Delta}$ ,  $(+)$  holds by the fact that  $\bar{\Delta} \geq 1$ , and the last inequality holds by the fact that the function  $f(x) \triangleq \frac{4x^2}{(x+1)^2}$  is monotone increasing when  $x > 0$  and  $f(1) = 1$ . Hence, in this case, we have  $\lambda\gamma^{-\Delta} \leq 54$ .

**Case (3):  $\beta = 0$ .** In this case, by Lemma 8.25, it holds that

$$\lambda\gamma^{-\Delta} \leq \frac{4}{d-1} \leq 4,$$

where  $d \triangleq \Delta - 1 \geq 2$ .

We next show that  $\lambda\gamma^{-\Delta}(1 - \beta\gamma) \leq \frac{12^5}{\Delta}$ . Let  $\bar{\Delta} \triangleq \frac{1 + \sqrt{\beta\gamma}}{1 - \sqrt{\beta\gamma}}$ , we consider 3 cases: (1)  $\beta > 0$  and  $\Delta < 2\bar{\Delta}$ ; (2)  $\beta > 0$  and  $\Delta \geq 2\bar{\Delta}$ ; (3)  $\beta = 0$ .

**Case (1):  $\beta > 0$  and  $\Delta < 2\bar{\Delta}$ .** First, by the previous result, it holds that  $\lambda\gamma^{-\Delta} \leq 12^4$ . Note that from  $\bar{\Delta} = \frac{1 + \sqrt{\beta\gamma}}{1 - \sqrt{\beta\gamma}}$ , we have  $(\beta\gamma)^{-1} = \left( \frac{\bar{\Delta} + 1}{\bar{\Delta} - 1} \right)^2$  which implies

$$1 - \beta\gamma = 1 - \left( \frac{\bar{\Delta} - 1}{\bar{\Delta} + 1} \right)^2 = \frac{4\bar{\Delta}}{(\bar{\Delta} + 1)^2} \leq \frac{4}{\bar{\Delta}} \leq \frac{8}{\Delta},$$

where in the last inequality, we use the fact that  $\Delta < 2\bar{\Delta}$ . Hence, it holds that

$$\lambda\gamma^{-\Delta}(1 - \beta\gamma) \leq \frac{8 \cdot 12^4}{\Delta}.$$

**Case (2):  $\beta > 0$  and  $\Delta \geq 2\bar{\Delta}$ .** Note that  $\frac{3}{2}d \geq \Delta \geq 2\bar{\Delta}$ , it holds that  $d \geq \frac{4}{3}\bar{\Delta} \geq \bar{\Delta}$ , which means  $\sqrt{\beta\gamma} \leq \frac{\Delta - 2}{\Delta}$ . By Lemma 8.22 with  $\sqrt{\beta\gamma} \leq \frac{\Delta - 2}{\Delta}$  and the fact that  $\lambda \leq \left( \frac{\gamma}{\beta} \right)^{\Delta/2}$ , it holds that:

$$\lambda \leq \frac{18\gamma^{d+1}}{\theta(d)},$$

where  $d \triangleq \Delta - 1$  and  $\theta(d) \triangleq d(1 - \beta\gamma) - (1 + \beta\gamma)$ . This lead us to

$$\begin{aligned}\frac{\lambda(1 - \beta\gamma)}{\gamma^{d+1}} &\leq \frac{18(1 - \beta\gamma)}{\theta(d)} = \frac{18(1 - \beta\gamma)}{d(1 - \beta\gamma) - (1 + \beta\gamma)} = \frac{18}{d - \frac{1 + \beta\gamma}{1 - \beta\gamma}} \\ (\text{by } \Delta = d + 1 \geq 2\bar{\Delta}) &\leq \frac{18}{d - \frac{(1 + \sqrt{\beta\gamma})^2}{(1 + \sqrt{\beta\gamma})(1 - \sqrt{\beta\gamma})}} = \frac{18}{d - \frac{1 + \sqrt{\beta\gamma}}{1 - \sqrt{\beta\gamma}}} \\ &= \frac{18}{d - \bar{\Delta}} \stackrel{(\star)}{\leq} \frac{72}{d} \stackrel{(*)}{\leq} \frac{108}{\Delta},\end{aligned}$$

where  $(\star)$  is deduced by  $d \geq \frac{4}{3}\bar{\Delta}$ , and  $(*)$  is because  $\Delta \leq \frac{3}{2}d$ .

Case (3):  $\beta = 0$ . By Lemma 8.22, it holds that

$$\lambda \leq \frac{4\gamma^{d+1}}{d-1},$$

which will lead us to

$$\frac{\lambda(1-\beta\gamma)}{\gamma^{d+1}} \leq \frac{4}{d-1} \leq \frac{12}{\Delta},$$

where the last inequality comes from the fact that  $3(d-1) \geq \Delta$ .

8.2.2. *Tree recursion analysis.* In order to prove Lemma 8.23, we first introduce the self-avoiding walk tree (SAW) in [Wei06]. Given a graph  $G = (V, E)$  with pinning  $\sigma \in \{-1, +1\}^\Lambda$  on  $\Lambda \subseteq V$ , fields  $\lambda \in \mathbb{R}^V$  and vertex  $v \in V$ , the self-avoiding walk tree  $T = T_{\text{SAW}}(G, v) = (V_T, E_T)$  with fields  $\lambda \in \mathbb{R}^{V_T}$  is recursively constructed as follows.

- (1) If vertex  $v$  is pinned, return the single vertex  $v$  (with field  $\lambda_v$ ).
- (2) Otherwise, let  $u_1, u_2, \dots, u_d$  be the neighbors of  $v$ . For each  $1 \leq i \leq d$ , denote  $G_i$  be the graph obtained by deleting  $v$ , attaching new vertices  $v_j$  with pinning  $-1$  to vertices  $u_j$  for all  $1 \leq j < i$ , and attaching new vertices  $v_j$  with pinning  $+1$  to vertices  $u_j$  for all  $i < j \leq d$ .
- (3) Let  $T$  be a rooted tree at vertex  $v$  (with field  $\lambda_v$ ) with subtrees  $T_1, T_2, \dots, T_d$  (with fields  $\lambda_{T_1}, \lambda_{T_2}, \dots, \lambda_{T_d}$ ), where  $T_i = T_{\text{SAW}}(G_i, u_i)$ .

Furthermore, given fields  $\lambda \in \mathbb{R}^V$  in  $G = (V, E)$ , Observed in [Wei06], the self-avoiding walk tree preserves marginal ratio.

**Proposition 8.27** ([Wei06, LLY13]). *Let  $G = (V, E)$  be a graph,  $\beta \in \mathbb{R}_{\geq 0}, \gamma \in \mathbb{R}_{> 0}, \lambda \in \mathbb{R}_{> 0}^V$  be parameters,  $\sigma \in \{-1, +1\}^\Lambda$  be a valid pinning on  $\Lambda \subseteq V$ , and  $v \in V$  be a vertice. Denote the Gibbs distribution of two-spin model  $(G, \beta, \gamma, \lambda)$  and  $(T = T_{\text{SAW}}(G, v), \beta, \gamma, \lambda_T)$  by  $\mu_G$  and  $\mu_T$  respectively. Then*

$$\frac{\mu_{G,v}(+1)}{\mu_{G,v}(-1)} = \frac{\mu_{T,v}(+1)}{\mu_{T,v}(-1)}.$$

Furthermore, denote the marginal ratio  $\frac{\mu_{T_u,u}(+1)}{\mu_{T_u,u}(-1)}$  by  $R_u$ , where  $T_u$  is the subtree rooted at  $u$  and  $\mu_{T_u,u}$  be the Gibbs distribution of two-spin model  $(T_u, \beta, \gamma, \lambda_{T_u})$ . For all  $u \in T$ , the marginal ratio  $R_u$  satisfies

$$R_u = \lambda \prod_{i=1}^d \left( \frac{\beta R_{u_i} + 1}{R_{u_i} + \gamma} \right),$$

where  $u_1, u_2, \dots, u_d$  denotes the children of  $u$  in  $T_u$ .

*Proof of Lemma 8.23.* Without loss of generality, we may assume  $R_v^{\sigma_S} > 0$ . Denote the neighbors of  $v$  in  $G$  by  $N_G(v)$ . Let

$$S_0 = N_G(v) \setminus S, S_{-1} = N_G(v) \cap \sigma_S^{-1}(-1), \text{ and } S_{+1} = N_G(v) \cap \sigma_S^{-1}(+1).$$

By monotonicity of anti-ferromagnetic two-spin system,

$$(61) \quad R_v^\sigma = \frac{\nu_v^\sigma(+1)}{\nu_v^\sigma(-1)} \leq \bar{\lambda}_v \bar{\gamma}^{-|S_0| - |S_{-1}|} \bar{\beta}^{|S_{+1}|}$$

Let  $T$  be the self-avoiding walk tree of  $G$  with pinning  $\sigma_S$  and  $\mu$  be the Gibbs distribution of two-spin model  $(T, \bar{\beta}, \bar{\gamma}, \bar{\lambda}_T)$ . Let  $N_T(v)$  denote all children of vertex  $v$  in  $T$ . By Proposition 8.27,

$$(62) \quad R_v^{\sigma_S} = \frac{\nu_v^{\sigma_S}(+1)}{\nu_v^{\sigma_S}(-1)} = \frac{\mu_{T,v}(+1)}{\mu_{T,v}(-1)} = \bar{\lambda}_v \prod_{u \in N_T(v)} \left( \frac{\bar{\beta} \frac{\mu_{T_u,u}(+1)}{\mu_{T_u,u}(-1)} + 1}{\frac{\mu_{T_u,u}(+1)}{\mu_{T_u,u}(-1)} + \bar{\gamma}} \right),$$

where  $T_u$  is the subtree rooted at  $u$  and  $\mu_{T_u,u}$  be the Gibbs distribution of two-spin model  $(T_u, \bar{\beta}, \bar{\gamma}, \bar{\lambda}_{T_u})$ . From the construction of  $T$ , we have the following properties.

- (1) There are  $|S_{-1}|$  children of vertex  $v$  with pinning  $-1$ ,  $|S_{+1}|$  children with pinning  $+1$ , and  $|S_0|$  children without pinning;
- (2) For each  $u \in N_T(v)$ ,  $\bar{\lambda}_{T,u} \leq \bar{\lambda}_{\max}$ .

By monotonicity of anti-ferromagnetic two-spin system, for each  $u \in N_T(v)$  without pinning,

$$\frac{\mu_{T_u,u}(+1)}{\mu_{T_u,u}(-1)} \leq \bar{\lambda}_{\max} \bar{\gamma}^{-d_u},$$

where  $d_u \leq d = \Delta - 1$  is the number of children of  $u$  in  $T_u$ . Note that if  $G$  is regular, then  $d_u = d$ . By (59),

$$\frac{\mu_{T_u,u}(+1)}{\mu_{T_u,u}(-1)} \leq \bar{\lambda}_{\max} \bar{\gamma}^{-d} = F_d^{\bar{\lambda}_{\max}}(0).$$

Together with (62) and the monotonicity of anti-ferromagnetic two-spin system,

$$R_v^{\sigma_S} \geq \bar{\lambda}_v \bar{\beta}^{|S_{+1}|} \bar{\gamma}^{-|S_{-1}|} \left( \frac{\bar{\beta} F_d^{\bar{\lambda}_{\max}}(0) + 1}{F_d^{\bar{\lambda}_{\max}}(0) + \bar{\gamma}} \right)^{|S_0|}$$

Combining with (61),

$$\frac{R_v^{\sigma}}{R_v^{\sigma_S}} \leq \left( \frac{\bar{\gamma}(\bar{\beta} F_d^{\bar{\lambda}_{\max}}(0) + 1)}{F_d^{\bar{\lambda}_{\max}}(0) + \bar{\gamma}} \right)^{-|S_0|} \leq \left( \frac{\bar{\gamma}(\bar{\beta} F_d^{\bar{\lambda}_{\max}}(0) + 1)}{F_d^{\bar{\lambda}_{\max}}(0) + \bar{\gamma}} \right)^{-\Delta} = \frac{F_{\Delta}^{\bar{\lambda}_{\max}}(0)}{F_{\Delta}^{\bar{\lambda}_{\max}} \circ F_d^{\bar{\lambda}_{\max}}(0)}.$$

This concludes the proof.  $\square$

**8.3. Modified log-Sobolev constant in subcritical regime.** In this section, we prove Lemma 8.9. In this proof, we consider the continuous-time Markov chain. Let  $\Omega$  be a discrete and finite state space. Let matrix  $Q : \Omega \times \Omega \rightarrow \mathbb{R}_{\geq 0}$  denote the *transition rate*. We remark that the row sum of  $Q$  may *not* be 1. The continuous-time Markov chain is a stochastic process  $(Y_t)_{t \in \mathbb{R}_{\geq 0}}$ . For any  $t > 0$ ,  $Y_t$  follows the distribution  $P_t(Y_0, \cdot)$  and  $P_t = \exp(Lt) = \sum_{k=0}^{\infty} \frac{t^k L^k}{k!}$ , where the *generator*  $L$  of the continuous time Markov chain is an operator defined by for any  $\psi : \Omega \rightarrow \mathbb{R}$ ,

$$L\psi(x) = \sum_{y \in \Omega} Q(x, y)(\psi(y) - \psi(x)).$$

Suppose  $Q$  satisfies the detailed balance equation with respect to the distribution  $b : \Omega \rightarrow \mathbb{R}_{> 0}$ , i.e.

$$\forall x, y \in \Omega, \quad b(x)Q(x, y) = b(y)Q(y, x).$$

The modified log-Sobolev constant for continuous-time Markov chain with rate  $Q$  is defined by

$$\rho(Q) \triangleq \min \left\{ \frac{\mathcal{E}_Q(f, \log f)}{\mathbf{Ent}_b[f]} \mid \forall f : \Omega \rightarrow \mathbb{R}_{\geq 0} : \mathbf{Ent}_b[f] \neq 0 \right\},$$

where the Dirichlet form  $\mathcal{E}_Q(f, \log f)$  is defined by

$$\mathcal{E}_Q(f, \log f) \triangleq \frac{1}{2} \sum_{x, y \in \Omega} b(x)Q(x, y)(f(x) - f(y))(\log f(x) - \log f(y)).$$

Back to our proof, let  $\theta \leq 12^{-6}$  be a constant. Note that  $\theta * \pi$  could be seen as a two-spin system with parameters  $\bar{\beta}, \bar{\gamma}$  and  $\theta \bar{\lambda}$ . Fix  $\Lambda \subseteq [n]$  and  $\sigma \in \Omega(\pi_{\Lambda})$ , let  $\nu \triangleq (\theta * \pi)^{\sigma}$ , we will show that

$$\rho^{\text{GD}}(\nu) \geq \frac{1}{4n}.$$

Let  $\Omega \triangleq \Omega(\nu)$ . For  $i \in [n]$ ,  $\eta_i : \Omega \rightarrow \{-1, +1\}^V$  is defined as

$$\forall x \in \Omega, \forall j \in [n], \quad (\eta_i x)_j \triangleq \begin{cases} x_j & j \neq i \\ -x_j & j = i \end{cases}$$

where for convenience, we denote  $\eta_i(x)$  as  $\eta_i x$ .

The continuous-time Glauber dynamics over  $\nu$  has the transition rate  $Q : \Omega \times \Omega \rightarrow \mathbb{R}_{\geq 0}$  as

$$\forall x \in \Omega, i \in [n] \setminus \Lambda, \quad Q(x, \eta_i x) = \frac{1}{n} \frac{\nu(\eta_i x)}{\nu(x) + \nu(\eta_i x)} = \begin{cases} \frac{1}{n} \frac{\theta \bar{\lambda} \bar{\beta}^{s_i}}{\bar{\lambda} \bar{\beta}^{s_i} + \bar{\gamma}^{\Delta_i - s_i}}, & (\eta_i x)_i = +1 \\ \frac{1}{n} \frac{\bar{\gamma}^{\Delta_i - s_i}}{\theta \bar{\lambda} \bar{\beta}^{s_i} + \bar{\gamma}^{\Delta_i - s_i}}, & (\eta_i x)_i = -1, \end{cases}$$

for any other  $x, y \in \Omega \times \Omega$  not covered by the above case,  $Q(x, y) = 0$ .

where  $\Delta_i$  is the degree of  $i$  in the graph and  $s_i$  denotes the number of +1-neighbors of  $i$  with respect to configuration  $x$ . Now, we consider a tuned version of Glauber dynamics over  $\nu$  whose transition rate  $\widehat{Q} : \Omega \times \Omega \rightarrow \mathbb{R}_{\geq 0}$  is defined as

$$\forall x \in \Omega, i \in [n] \setminus \Lambda, \quad \widehat{Q}(x, \eta_i x) = \begin{cases} \frac{\nu(\eta_i x)}{\nu(x)} = \theta \bar{\lambda} \bar{\gamma}^{-\Delta_i} (\bar{\beta} \bar{\gamma})^{s_i}, & (\eta_i x)_i = +1 \\ 1, & (\eta_i x)_i = -1. \end{cases}$$

for any other  $x, y \in \Omega \times \Omega$  not covered by the above case,  $\widehat{Q}(x, y) = 0$ .

It is straightforward to check that both  $Q$  and  $\widehat{Q}$  are reversible with respect to  $\nu$ .

Note that for  $x \in \Omega$  and  $i \in [n] \setminus \Lambda$ , it holds that

$$Q(x, \eta_i x) = \frac{1}{n} \cdot \frac{\nu(x_-)}{\nu(x_-) + \nu(x_+)} \cdot \widehat{Q}(x, \eta_i x),$$

where  $x_-$  and  $x_+$  are obtained from configuration  $x$  with the  $i$ -th position being modified to  $-1$  and  $+1$ , respectively. If  $\nu(x_-) = 0$ , then it holds that  $x_i = +$  and

$$Q(x, \eta_i x) = \widehat{Q}(x, \eta_i x) = 0.$$

If  $\nu(x_-) > 0$ , we have

$$\frac{\nu(x_-)}{\nu(x_-) + \nu(x_+)} = \frac{1}{1 + \frac{\nu(x_+)}{\nu(x_-)}} \stackrel{(\star)}{\geq} \frac{1}{1 + \theta \bar{\lambda} \bar{\gamma}^{-\Delta}} \geq \frac{1}{1 + \theta \cdot 12^4},$$

where  $(\star)$  could be deduced from  $G$  is  $\Delta$ -regular or  $\bar{\gamma} \leq 1$  (see (59)) and the last inequality holds by Lemma 8.22. Since  $\theta \leq 12^{-4}$ , it holds that for any  $x, y \in \Omega \times \Omega$ ,

$$Q(x, \eta_i x) \geq \frac{1}{2n} \widehat{Q}(x, \eta_i x),$$

which implies

$$\forall f : \Omega \rightarrow \mathbb{R}_{\geq 0}, \quad 2n \cdot \mathcal{E}_Q(f, \log f) \geq \mathcal{E}_{\widehat{Q}}(f, \log f).$$

Hence, it holds that

$$(63) \quad \rho^{\text{GD}}(\nu) \geq \frac{1}{2n} \cdot \rho^{\widehat{\text{GD}}}(\nu),$$

where we use  $\rho^{\text{GD}}(\nu)$  and  $\rho^{\widehat{\text{GD}}}(\nu)$  to denote the modified log-Sobolev constant of the continuous-time Glauber dynamics and continuous-time tuned Glauber dynamics respectively. Remark that by our definition, the discrete-time Glauber dynamics and continuous-time Glauber dynamics have the same modified log-Sobolev constant. Hence, to prove Lemma 8.9, it suffices for us to bound  $\rho^{\widehat{\text{GD}}}(\nu)$ .

We will use the following general result. Let  $T$  be the transition rate of a continuous-time Markov chain  $M$  on  $\mathcal{X}$ . Let  $\mathcal{G}$  be a set of bijective maps from  $\mathcal{X}$  to  $\mathcal{X}$ . We say  $\mathcal{G}$  is a mapping representation of  $T$  if

- for any  $x, y \in \mathcal{X}$  such that  $T(x, y) > 0$ , there exists a unique  $\delta \in \mathcal{G}$  such that  $y = \delta x$ ;
- for any  $\delta \in \mathcal{G}$ , there exists a unique  $\delta^{-1} \in \mathcal{G}$  such that for any  $x \in \mathcal{X}$ ,  $\delta^{-1}(\delta(x)) = x$ .

**Theorem 8.28** ([EHMT17, Theorem 1.1, Theorem 3.9]). *Let  $\mu$  be a distribution over a finite set  $\mathcal{X}$ . Let  $T$  be the transition rate of a continuous time Markov chain  $M$  satisfying the detailed-balance equation with respect to  $\mu$ . Let  $\mathcal{G}$  a mapping representation of  $T$  satisfying  $\alpha\beta x = \beta\alpha x$  for all  $x \in \mathcal{X}, \alpha, \beta \in G$ . If there exist  $H_1 \subseteq \mathcal{X} \times \mathcal{G}$  and  $H_2 \triangleq \{(\alpha x, \alpha^{-1}) \mid (x, \alpha) \in H_1\}$  such that  $H_1 \cap H_2 = \emptyset, H_1 \cup H_2 = \mathcal{X} \times \mathcal{G}$  and*

$$\forall i \in \{1, 2\}, \quad \kappa_i \triangleq \min_{\substack{(x, \alpha) \in H_i \\ T(x, \alpha x) > 0}} \left[ T(x, \alpha x) - \mathbb{1}_{\alpha \neq \alpha^{-1}} T(\alpha x, \alpha(\alpha x)) - \sum_{\eta: \eta \neq \alpha, \alpha^{-1}} \frac{(q - q_*)(\alpha x, \alpha^{-1}, \eta)}{T(x, \alpha x)\mu(x)} \right] \geq 0,$$

where we set  $q(x, \alpha, \eta) \triangleq T(x, \alpha x)T(x, \eta x)\mu(x)$  (we assume  $q(x, \alpha, \eta) = 0$  if  $\mu(x) = 0$ ) and

$$q_*(x, \alpha, \eta) \triangleq \min\{q(x, \alpha, \eta), q(\alpha x, \alpha^{-1}, \eta), q(\eta x, \alpha, \eta^{-1}), q(\alpha \eta x, \alpha^{-1}, \eta^{-1})\}.$$

Then, we have  $\rho^M(\mu) \geq \kappa_1 + \kappa_2$ , where  $\rho^M(\mu)$  denotes the modified log-Sobolev constant of  $M$ .

The above theorem is slightly different from the original theorem in [EHMT17], but it can be proved by going through the proof in [EHMT17]. We give the proof in Appendix B for completeness.

In our proof, we define  $\mathcal{G} = \{\eta_i \mid i \in [n] \setminus \Lambda\}$ . Note that for any  $x, y \in \Omega$  such that  $\widehat{Q}(x, y) > 0$ , it must hold that  $x$  and  $y$  disagree only at one vertex, say  $i \in V$ . We have  $y = \eta_i x$  and  $y \neq \eta_j x$  for all  $j \neq i$ . For any  $\eta_i \in \mathcal{G}$ , it holds that  $\eta_i^{-1} = \eta_i$ . Hence,  $\mathcal{G}$  is a mapping representation of  $\widehat{Q}$ . For any  $\eta_i, \eta_j \in \mathcal{G}$ , any  $x \in \Omega$ , it is straightforward to verify that  $\eta_i \eta_j x = \eta_j \eta_i x$ . We define  $H_1$  and  $H_2$  as

$$\begin{aligned} H_1 &= \{(x, \eta_i) \mid x \in \Omega, i \in [n] \setminus \Lambda, x_i = -1\} \\ H_2 &= \{(x, \eta_i) \mid x \in \Omega, i \in [n] \setminus \Lambda, x_i = +1\}. \end{aligned}$$

It is straightforward to verify  $H_1 \cup H_2 = \Omega \times \mathcal{G}, H_1 \cap H_2 = \emptyset$ , and  $H_2 = \{(\eta_i x, \eta_i^{-1}) \mid (x, \eta_i) \in H_1\}$ . In our application,  $\eta_i = \eta_i^{-1}$  for all  $\eta_i \in \mathcal{G}$ . The  $\kappa_i$  for  $i \in \{1, 2\}$  could be rewritten as

$$\kappa_i = \min_{\substack{(x, \alpha) \in H_i \\ \widehat{Q}(x, \alpha x) > 0}} \left[ \widehat{Q}(x, \alpha x) - \sum_{\eta: \eta \neq \alpha, \alpha^{-1}} \frac{(q - q_*)(\alpha x, \alpha^{-1}, \eta)}{\widehat{Q}(x, \alpha x)v(x)} \right].$$

Besides, definitions of  $H_1$  and  $H_2$  and the reversibility, it is straightforward to verify that

$$\begin{aligned} (64) \quad (A) &= \{(x, \alpha) \mid (x, \alpha) \in H_1 \wedge \widehat{Q}(x, \alpha x) > 0\} = \{(x, \alpha) \mid (\alpha x, \alpha) \in H_2 \wedge \widehat{Q}(\alpha x, x) > 0\} = (B) \\ &\quad (\text{replace } x \text{ with } \alpha x, \alpha \text{ is a bijection}) = \{(\alpha x, \alpha) \mid (x, \alpha) \in H_2 \wedge \widehat{Q}(x, \alpha x) > 0\} \end{aligned}$$

To verify (64), by reversibility,  $v(x)\widehat{Q}(x, \alpha x) = v(\alpha x)\widehat{Q}(\alpha x, x)$ . For  $(x, \alpha) \in H_1$ , it holds that  $v(x) > 0$ , if  $\widehat{Q}(x, \alpha x) > 0$ , then  $v(\alpha x) > 0$  and  $\widehat{Q}(\alpha x, x) > 0$ , which implies  $(x, \alpha) \in B$ , thus  $A \subseteq B$ . Similarly, for any  $(x, \alpha) \in B$ , we can verify that  $(x, \alpha) \in A$ , thus  $B \subseteq A$ . Note that  $\alpha = \alpha^{-1}$  for all  $\alpha \in \mathcal{G}$ . Hence,

$$(65) \quad \forall i \in \{1, 2\}, \quad \kappa_i = \min_{\substack{(x, \alpha) \in H_{3-i} \\ \widehat{Q}(x, \alpha x) > 0}} \left[ \widehat{Q}(\alpha x, x) - \sum_{\eta: \eta \neq \alpha, \alpha^{-1}} \frac{(q - q_*)(x, \alpha, \eta)}{\widehat{Q}(\alpha x, x)v(\alpha x)} \right].$$

To leveraging Theorem 8.28 for the tuned Glauber dynamics, we have the following result. For any two vertices  $i, j$ , we use  $i \sim j$  to denote that  $i$  and  $j$  are adjacent in  $G$ .

**Lemma 8.29.** *Let  $T, v$  in Theorem 8.28 be  $\widehat{Q}, v = (\theta * \pi)^\sigma$ , respectively. We have the following results.*

*Let  $i, j \in [n] \setminus \Lambda$  such that  $i \sim j$  in the graph  $G$  and  $x \in \Omega$  where  $x_i = x_j = -1$ , then we have*

$$q(\eta_i x, \eta_i^{-1}, \eta_j) = q(\eta_j x, \eta_i, \eta_j^{-1}) = q(\eta_i \eta_j x, \eta_i^{-1}, \eta_j^{-1}) = \bar{\beta} \bar{y} \cdot q(x, \eta_i, \eta_j).$$

Moreover when  $i \neq j$  in the graph  $G$ , then for any  $x \in \Omega$ , it holds that

$$q(\eta_i x, \eta_i^{-1}, \eta_j) = q(\eta_j x, \eta_i, \eta_j^{-1}) = q(\eta_i \eta_j x, \eta_i^{-1}, \eta_j^{-1}) = q(x, \eta_i, \eta_j).$$

The proof of Lemma 8.29 is deferred to the end of this section.

We claim that for any  $i \sim j$ ,  $y \in \Omega$ , if  $y_i = +1$  or  $y_j = +1$ , then it holds that

$$(66) \quad (q - q_*)(y, \eta_i, \eta_j) = 0.$$

To verify the claim, we need to consider three cases (1)  $y_i = +1$  and  $y_j = -1$ ; (2)  $y_i = +1$  and  $y_j = +1$ ; (3)  $y_i = -1$  and  $y_j = +1$ . We verify the first case, the other two cases can be verified by a similar argument. Consider the configuration  $x = \eta_i y$ . It holds that  $x \in \Omega$  because  $\theta \bar{\lambda} > 0$  and  $\bar{y} > 0$ . By Lemma 8.29,

$$\begin{aligned} q_*(y, \eta_i, \eta_j) &= \min\{q(y, \eta_i, \eta_j), q(\eta_i y, \eta_i^{-1}, \eta_j), q(\eta_j y, \eta_i, \eta_j^{-1}), q(\eta_i \eta_j y, \eta_i^{-1}, \eta_j^{-1})\} \\ &= \min\{q(y, \eta_i, \eta_j), q(\eta_i y, \eta_i, \eta_j), q(\eta_j y, \eta_i, \eta_j), q(\eta_i \eta_j y, \eta_i, \eta_j)\} \\ &= \min\{q(\eta_i x, \eta_i, \eta_j), q(x, \eta_i, \eta_j), q(\eta_j \eta_i x, \eta_i, \eta_j), q(\eta_j x, \eta_i, \eta_j)\} \\ (*) &= q(\eta_i x, \eta_i, \eta_j) = q(y, \eta_i, \eta_j), \end{aligned}$$

where (\*) holds because Lemma 8.29,  $\eta_i = \eta_i^{-1}$  and  $\eta_j = \eta_j^{-1}$ . Besides, for any  $i \neq j$ , any  $x \in \Omega$ ,

$$(67) \quad (q - q_*)(x, \eta_i, \eta_j) = 0.$$

By (65), (66), (67) and the definitions of  $H_1$  and  $H_2$ , it holds that

$$(68) \quad \kappa_1 = \min_{\substack{(x, \eta_i) \in H_2 \\ \widehat{Q}(x, \eta_i x) > 0}} \left[ \widehat{Q}(\eta_i x, x) - \sum_{j \neq i} \frac{(q - q_*)(x, \eta_i, \eta_j)}{\widehat{Q}(\eta_i x, x) v(\eta_i x)} \right] \stackrel{(*)}{=} \min_{\substack{(x, \eta_i) \in H_2 \\ \widehat{Q}(x, \eta_i x) > 0}} \widehat{Q}(\eta_i x, x) \geq 0,$$

where (\*) holds because  $x_i = +1$  and we can use (66), (67). On the other hand, we have

$$\kappa_2 = \min_{\substack{(x, \eta_i) \in H_1 \\ \widehat{Q}(x, \eta_i x) > 0}} \left[ \widehat{Q}(\eta_i x, x) - \sum_{\substack{j \in [n] \setminus \Lambda \\ j \sim i \\ x_j = -1}} \frac{(1 - \bar{\beta} \bar{y}) q(x, \eta_i, \eta_j)}{\widehat{Q}(\eta_i x, x) v(\eta_i x)} \right] \stackrel{(*)}{=} \min_{\substack{(x, \eta_i) \in H_1 \\ \widehat{Q}(x, \eta_i x) > 0}} \left[ \widehat{Q}(\eta_i x, x) - \sum_{\substack{j \in [n] \setminus \Lambda \\ j \sim i \\ x_j = -1}} \frac{(1 - \bar{\beta} \bar{y}) q(x, \eta_i, \eta_j)}{\widehat{Q}(x, \eta_i x) v(x)} \right],$$

where (\*) holds by reversibility. Note that by definition, it holds that  $q(x, \eta_i, \eta_j) = v(x) \widehat{Q}(x, \eta_i x) \widehat{Q}(x, \eta_j x)$ , which implies

$$(69) \quad \kappa_2 = \min_{(x, \eta_i) \in H_1} \left[ \widehat{Q}(\eta_i x, x) - \sum_{\substack{j \in [n] \setminus \Lambda \\ j \sim i \\ x_j = -1}} (1 - \bar{\beta} \bar{y}) \widehat{Q}(x, \eta_j x) \right] \stackrel{(*)}{\geq} 1 - \Delta \cdot (1 - \bar{\beta} \bar{y}) \cdot \theta \bar{\lambda} \bar{y}^{-\Delta} \geq 1/2,$$

where (\*) is deduced from the fact that  $G$  is  $\Delta$ -regular or  $\bar{y} \leq 1$  and the last inequality holds by Lemma 8.22 and the fact that  $\theta \leq \frac{1}{2} \cdot 12^{-5}$ .

Combining (68) and (69) with Theorem 8.28, it holds that

$$\rho^{\text{GD}}(v) \geq 1/2,$$

which, together with (63), implies that

$$\rho^{\text{GD}}(v) \geq \frac{1}{2n} \rho^{\text{GD}}(v) \geq \frac{1}{4n}.$$

Finally, we finish the proof by proving Lemma 8.29.

*Proof of Lemma 8.29.* We first consider the case where  $i \sim j$  and  $x_i = x_j = -1$ . By definition, we have

$$q(\eta_i x, \eta_i^{-1}, \eta_j) = v(\eta_i x) \widehat{Q}(\eta_i x, x) \widehat{Q}(\eta_i x, \eta_i \eta_j x).$$

Note that if  $v(\eta_i x) = 0$ , then by the definition of two-spin system, it must hold that  $\bar{\beta} = 0$ , we have

$$q(\eta_i x, \eta_i^{-1}, \eta_j) = \bar{\beta} \bar{\gamma} \cdot q(x, \eta_i, \eta_j) = 0.$$

If  $v(\eta_i x) \neq 0$ , we have

$$q(\eta_i x, \eta_i^{-1}, \eta_j) = v(\eta_i x) \cdot 1 \cdot \widehat{Q}(\eta_i x, \eta_i \eta_j x) = v(x) \frac{v(\eta_i x)}{v(x)} \cdot \widehat{Q}(x, \eta_j x) \cdot \bar{\beta} \bar{\gamma} = v(x) \widehat{Q}(x, \eta_i x) \widehat{Q}(x, \eta_j x) \cdot \bar{\beta} \bar{\gamma}.$$

Similarly, it holds that

$$q(\eta_j x, \eta_i, \eta_j^{-1}) = v(x) \widehat{Q}(x, \eta_i x) \widehat{Q}(x, \eta_j x) \cdot \bar{\beta} \bar{\gamma}.$$

Lastly, we analyze  $q(\eta_i \eta_j x, \eta_i^{-1}, \eta_j^{-1})$ . Similarly, we assume  $v(\eta_i \eta_j x) > 0$ , otherwise the result holds trivially.

Note that  $\bar{\lambda} > 0$  and  $\bar{\gamma} > 0$ , thus we have  $v(\eta_i x) > 0$  and  $v(\eta_j x) > 0$ . We have

$$\begin{aligned} q(\eta_i \eta_j x, \eta_i^{-1}, \eta_j^{-1}) &= v(\eta_i \eta_j x) \widehat{Q}(\eta_i \eta_j x, \eta_i x) \widehat{Q}(\eta_i \eta_j x, \eta_j x) \\ &= v(\eta_i \eta_j x) = v(x) \frac{v(\eta_i x)}{v(x)} \frac{v(\eta_j x)}{v(\eta_i x)} \\ &= v(x) \widehat{Q}(x, \eta_i x) \widehat{Q}(\eta_i x, \eta_j x) \\ &= v(x) \widehat{Q}(x, \eta_i x) \widehat{Q}(x, \eta_j x) \cdot \bar{\beta} \bar{\gamma}. \end{aligned} \quad \square$$

We then consider the case where  $i \not\sim j$ . We prove that  $q(\eta_i x, \eta_i^{-1}, \eta_j) = q(x, \eta_i, \eta_j)$ . If  $v(\eta_i x) = 0$ , then it holds that  $q(\eta_i x, \eta_i^{-1}, \eta_j) = q(x, \eta_i, \eta_j) = 0$ . Suppose  $\eta_i x$  is a feasible configuration. We have

$$\begin{aligned} q(\eta_i x, \eta_i^{-1}, \eta_j) &= v(\eta_i x) \widehat{Q}(\eta_i x, x) \widehat{Q}(\eta_i x, \eta_j \eta_i x) \\ \text{(by reversibility)} &= v(x) \widehat{Q}(x, \eta_i x) \widehat{Q}(\eta_i x, \eta_j \eta_i x) \\ (*) &= v(x) \widehat{Q}(x, \eta_i x) \widehat{Q}(x, \eta_j x) \\ &= q(x, \eta_i, \eta_j). \end{aligned}$$

where (\*) holds since  $\widehat{Q}(\eta_i x, \eta_j \eta_i x) = \widehat{Q}(x, \eta_j x)$ . This is because  $i \not\sim j$ , both transitions  $\eta_i x \rightarrow \eta_j \eta_i x$  and  $x \rightarrow \eta_j x$  are to flip the value of  $j$ , and such transition probabilities depend only on the configuration of  $j$  and  $j$ 's neighbors.

The equation  $q(\eta_j x, \eta_i, \eta_j^{-1}) = q(x, \eta_i, \eta_j)$  can be proved in a similar way.

Finally, we prove  $q(\eta_i \eta_j x, \eta_i^{-1}, \eta_j^{-1}) = q(x, \eta_i, \eta_j)$ . Suppose  $v(\eta_i \eta_j x) = 0$ . We have  $q(\eta_i \eta_j x, \eta_i^{-1}, \eta_j^{-1}) = 0$ . There are three cases for  $\eta_i \eta_j x$ : (1) if  $i$  violates the local constraints, then  $v(\eta_i x) = 0$ ; (2) if  $j$  violates the local constraints, then  $v(\eta_j x) = 0$  (3) if some  $k \notin \{i, j\}$  violates the local constraints, then  $v(x) = 0$ . Hence, we have  $q(x, \eta_i, \eta_j) = 0$ . Similarly, if  $v(\eta_i x) = 0$  or  $v(\eta_j x) = 0$ , it holds that  $q(\eta_i \eta_j x, \eta_i^{-1}, \eta_j^{-1}) = q(x, \eta_i, \eta_j) = 0$ . Suppose  $v(\eta_i \eta_j x) \neq 0$ ,  $v(\eta_i x) \neq 0$  and  $v(\eta_j x) \neq 0$ . We have

$$\begin{aligned} q(\eta_i \eta_j x, \eta_i^{-1}, \eta_j^{-1}) &= v(\eta_i \eta_j x) \widehat{Q}(\eta_i \eta_j x, \eta_j x) \widehat{Q}(\eta_i \eta_j x, \eta_i x) \\ \text{(by reversibility)} &= v(\eta_j x) \widehat{Q}(\eta_j x, \eta_i \eta_j x) \widehat{Q}(\eta_i \eta_j x, \eta_i x) \\ \text{(by } i \not\sim j) &= v(\eta_j x) \widehat{Q}(x, \eta_i x) \widehat{Q}(\eta_j x, x) \\ \text{(by reversibility)} &= v(x) \widehat{Q}(x, \eta_i x) \widehat{Q}(x, \eta_j x) \\ &= q(x, \eta_i, \eta_j). \end{aligned}$$

## 9. ACKNOWLEDGEMENTS

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

We would like to thank Heng Guo and Jingcheng Liu 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<sup>+</sup>21a] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence I: Modified log-sobolev inequalities for fractionally log-concave distributions and high-temperature ising models. *arXiv preprint arXiv:2106.04105*, 2021.
- [AJK<sup>+</sup>21b] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence II: Optimal sampling and concentration via restricted modified log-sobolev inequalities. *arXiv preprint arXiv:2111.03247*, 2021.
- [AL20] Vedat Levi Alev and Lap Chi Lau. Improved analysis of higher order random walks and applications. In *STOC*, pages 1198–1211, 2020.
- [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.
- [ALOV19] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials II: high-dimensional walks and an FPRAS for counting bases of a matroid. In *STOC*, pages 1–12, 2019.
- [ALOV20] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials IV: Exchange properties, tight mixing times, and faster sampling of spanning trees. *CoRR*, abs/2004.07220, 2020.
- [BCC<sup>+</sup>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.
- [BCLM11] Franck Barthe, Dario Cordero-Erausquin, Michel Ledoux, and Bernard Maurey. Correlation and Brascamp-Lieb inequalities for Markov semigroups. *Int. Math. Res. Not. IMRN*, (10):2177–2216, 2011.
- [BT06] Sergey G. Bobkov and Prasad Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.*, 19(2):289–336, 2006.
- [CC09] Eric A. Carlen and Dario Cordero-Erausquin. Subadditivity of the entropy and its relation to Brascamp-Lieb type inequalities. *Geom. Funct. Anal.*, 19(2):373–405, 2009.
- [CE22] Yuansi Chen and Ronen Eldan. Localization schemes: A framework for proving mixing bounds for markov chains. *arXiv preprint arXiv:2203.04163*, 2022.
- [CFYZ21a] Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Optimal mixing time for the ising model in the uniqueness regime. *arXiv preprint arXiv:2111.03034*, 2021.
- [CFYZ21b] Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Rapid mixing of Glauber dynamics via spectral independence for all degrees. In *FOCS*, pages 137–148, 2021.
- [CGM21] Mary Cryan, Heng Guo, and Giorgos Mousa. Modified log-Sobolev inequalities for strongly log-concave distributions. *Ann. Probab.*, 49(1):506–525, 2021. (conference version in *FOCS'19*).
- [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.
- [CLV21] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of Glauber dynamics: Entropy factorization via high-dimensional expansion. In *STOC*, 2021. arXiv:2011.02075.



- [CMT15] Pietro Caputo, Georg Menz, and Prasad Tetali. Approximate tensorization of entropy at high temperature. *Ann. Fac. Sci. Toulouse Math.* (6), 24(4):691–716, 2015.
- [Con20] Giovanni Conforti. A probabilistic approach to convex ( $\phi$ )-entropy decay for markov chains. *arXiv preprint arXiv:2004.10850*, 2020.
- [CP20] Pietro Caputo and Daniel Parisi. Block factorization of the relative entropy via spatial mixing. *arXiv preprint arXiv:2004.10574*, 2020.
- [DG00] Martin Dyer and Catherine Greenhill. On markov chains for independent sets. *Journal of Algorithms*, 35(1):17–49, 2000.
- [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.
- [EHMT17] Matthias Erbar, Christopher Henderson, Georg Menz, and Prasad Tetali. Ricci curvature bounds for weakly interacting Markov chains. *Electron. J. Probab.*, 22:Paper No. 40, 23, 2017.
- [EHŠ<sup>+</sup>19] Charilaos Efthymiou, Thomas P. Hayes, Daniel Štefankovič, Eric Vigoda, and Yitong Yin. Convergence of MCMC and loopy BP in the tree uniqueness region for the hard-core model. *SIAM J. Comput.*, 48(2):581–643, 2019. (conference version in FOCS’16).
- [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.
- [FM16] Max Fathi and Jan Maas. Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.*, 26(3):1774–1806, 2016.
- [GJP03] Leslie Ann Goldberg, Mark Jerrum, and Mike Paterson. The computational complexity of two-state spin systems. *Random Structures Algorithms*, 23(2):133–154, 2003.
- [GŠV15] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. *J. ACM*, 62(6):Art. 50, 60, 2015.
- [HJ12] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge university press, 2012.
- [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.
- [HV06] Thomas P. Hayes and Eric Vigoda. Coupling with the stationary distribution and improved sampling for colorings and independent sets. *Ann. Appl. Probab.*, 16(3):1297–1318, 2006. (conference version in SODA’05).
- [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.
- [KO20] Tali Kaufman and Izhar Oppenheim. High order random walks: Beyond spectral gap. *Combinatorica*, 40(1):245–281, 2020.
- [LLY12] Liang Li, Pinyan Lu, and Yitong Yin. Approximate counting via correlation decay in spin systems. In *SODA*, pages 922–940. SIAM, 2012.
- [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.
- [LPW17] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017.
- [LV99] Michael Luby and Eric Vigoda. Fast convergence of the glauber dynamics for sampling independent sets. *Random Structures & Algorithms*, 15(3-4):229–241, 1999.

- [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.
- [Opp18] Izhar Oppenheim. Local spectral expansion approach to high dimensional expanders Part I: Descent of spectral gaps. *Discrete Comput. Geom.*, 59(2):293–330, 2018.
- [Röl07] Adrian Röllin. Translated Poisson approximation using exchangeable pair couplings. *Ann. Appl. Probab.*, 17(5-6):1596–1614, 2007.
- [Sly10] Allan Sly. Computational transition at the uniqueness threshold. In *FOCS*, pages 287–296, 2010.
- [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*).
- [Vig99] Eric Vigoda. Fast convergence of the glauber dynamics for sampling independent sets: Part II. Technical Report TR-99-003, International Computer Science Institute, 1999.
- [Wei05] Dror Weitz. Combinatorial criteria for uniqueness of Gibbs measures. *Random Structures Algorithms*, 27(4):445–475, 2005.
- [Wei06] Dror Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)*, pages 140–149, 2006.

APPENDIX A. MIXING TIME FROM MODIFIED LOG-SOBOLEV CONSTANT

In this section, we prove the mixing time results in Theorem 1.3 and Corollary 1.6. We remark that Corollary 1.7 directly follows from Theorem 1.3.

We first prove Proposition 3.2.

**Lemma A.1** ([LPW17, Theorem 20.3 (ii)]). *Let  $P$  be the transition matrix of an irreducible, aperiodic and reversible Markov chain and  $P_{zz} = (1/2)(I+P)$  be the transition matrix of the lazy version of  $P$ , and  $H_t = e^{t(P-1)}$  be the heat kernel of the continuous version of  $P$ . Let  $\pi$  be their stationary distribution. Let  $m > 0$  be an integer,  $Y \sim \text{Binomial}(4m, \frac{1}{2})$ , and  $\Psi \sim \text{Poisson}(m)$  be random variables. Then for every  $x \in \Omega$ , it holds that*

$$d_{\text{TV}}(P_{zz}^{4m}(x, \cdot), \pi) \leq d_{\text{TV}}(H_m(x, \cdot), \pi) + d_{\text{TV}}(\Pr[Y \in \cdot], \Pr[\Psi + m \in \cdot]).$$

**Lemma A.2** ([Röl07, Example 3.3]). *Let  $m \geq 20$  be an integer,  $Y \sim \text{Binomial}(4m, \frac{1}{2})$ , and  $\Psi \sim \text{Poisson}(m)$  be random variables. Then, it holds that*

$$d_{\text{TV}}(\Pr[Y \in \cdot], \Pr[\Psi + m \in \cdot]) \leq m^{-1/2}.$$

*Proof of Proposition 3.2.* By the mixing result of the continuous version of  $P$  (Proposition 3.1), Lemma A.1, Lemma A.2, and taking  $m = \lceil \frac{1}{\rho^P(\mu)}(\log \log \frac{1}{\mu_{\min}} + 160) \rceil \geq 20$ , we have following mixing time bound for the lazy version of  $P$ :

$$T_{\text{mix}}(1/(2e)) \leq 4m = 4 \left\lceil \frac{1}{\rho^P(\mu)} \left( \log \log \frac{1}{\mu_{\min}} + 160 \right) \right\rceil.$$

Hence, for any  $\varepsilon > 0$ , by the exponential decay of the mixing time, we have

$$T_{\text{mix}}(\varepsilon) \leq \left\lceil 4m \log \frac{1}{\varepsilon} \right\rceil \leq \frac{4}{\rho^P(\mu)} \left( \log \log \frac{1}{\mu_{\min}} + 200 \right) \log \frac{1}{\varepsilon}. \quad \square$$

*Proof of Theorem 1.3.* To prove these corollaries, it only remains to give a lower bound for  $\mu_{\min}$ . Similar to [CFYZ21b], the marginal bound  $b \triangleq \min_{v \in V} \min_{\sigma \in \Omega(\mu_{V \setminus \{v\}})} \min_{c \in \Omega(\mu_v^\sigma)} \mu_v^\sigma(c)$  can be bounded by

$$\frac{1}{b} \leq \begin{cases} (\lambda + \frac{1}{\lambda}) \left( \frac{1}{\gamma} + \gamma + 2 \right)^\Delta, & \beta = 0; \\ (\lambda + \frac{1}{\lambda}) \left( \frac{1}{\beta} + 2 \right)^\Delta, & \beta > 0. \end{cases}$$

Therefore,

$$\begin{aligned} \log \log \frac{1}{\mu_{\min}} &\leq \log n + \log \log \frac{1}{b} \leq \log n + \log \left( n \log \alpha + \log \left( \lambda + \frac{1}{\lambda} \right) \right) \\ &\leq 2 \log n + \log \log \alpha + \log \log \left( \lambda + \frac{1}{\lambda} \right), \end{aligned}$$

where  $\alpha = \begin{cases} \frac{1}{\gamma} + \gamma + 2, & \beta = 0; \\ \frac{1}{\beta} + 2, & \beta > 0. \end{cases}$ . Together with Proposition 3.2, we prove Theorem 1.3. Besides, combining the lower bound with Proposition 3.1, we prove the result claimed in Remark 1.4.  $\square$

*Proof of Corollary 1.6.* Recall that the  $O(n \log n)$  mixing time can be achieved via standard path coupling technique when  $\lambda \leq \frac{1}{2\Delta}$ . Therefore, we may assume that  $\lambda \geq \frac{1}{2\Delta}$ . To prove this corollary, it only remains to give a lower bound for  $\mu_{\min}$ . Similar to [CFYZ21b], the marginal bound

$$b \triangleq \min_{v \in V} \min_{\sigma \in \Omega(\mu_{V \setminus \{v\}})} \min_{c \in \Omega(\mu_v^\sigma)} \mu_v^\sigma(c)$$

can be bounded by

$$b \geq \min \left( \frac{1}{1 + \lambda}, \frac{\lambda}{1 + \lambda} \right) \geq \frac{1}{2\Delta + 1}.$$

Therefore,

$$\log \log \frac{1}{\mu_{\min}} \leq \log n + \log \log \frac{1}{b} \leq \log n + \log \log(2\Delta + 1).$$

Together with Theorem 1.3, we prove this corollary.  $\square$

## APPENDIX B. MODIFIED LOG-SOBOLEV INEQUALITY IN SUB-CRITICAL REGIME

In this section, we prove Theorem 8.28. We need several notations and definitions in [EHMT17].

At first, we will use the following fact about the mapping representation  $\mathcal{G}$ .

**Fact B.1.** *For any function  $F : \mathcal{X} \times \mathcal{G} \rightarrow \mathbb{R}$ , it holds that*

$$\sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} F(x, \delta) T(x, \delta x) \mu(x) = \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} F(\delta x, \delta^{-1}) T(x, \delta x) \mu(x)$$

*Proof.* We have

$$\sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} F(x, \delta) T(x, \delta x) \mu(x) = \sum_{x \in \mathcal{X}} \mu(x) \sum_{\delta \in \mathcal{G}} F(x, \delta) T(x, \delta x) = \sum_{x \in \mathcal{X}} \mu(x) \sum_{\delta \in \mathcal{G}} F(x, \delta^{-1}) T(x, \delta^{-1} x),$$

where the last equation holds because  $\{\delta \mid \delta \in \mathcal{G}\} = \{\delta^{-1} \mid \delta \in \mathcal{G}\}$ . We then have

$$\sum_{x \in \mathcal{X}} \mu(x) \sum_{\delta \in \mathcal{G}} F(x, \delta^{-1}) T(x, \delta^{-1} x) = \sum_{\delta \in \mathcal{G}, x \in \mathcal{X}} \mu(x) F(x, \delta^{-1}) T(x, \delta^{-1} x) = \sum_{\delta \in \mathcal{G}, x \in \mathcal{X}} \mu(\delta x) F(\delta x, \delta^{-1}) T(\delta x, x),$$

where the equation holds because every  $\delta$  is a bijection, thus  $\{x \mid x \in \mathcal{X}\} = \{\delta x \mid x \in \mathcal{X}\}$ . Finally, by reversibility, we have

$$\sum_{\delta \in \mathcal{G}, x \in \mathcal{X}} \mu(\delta x) F(\delta x, \delta^{-1}) T(\delta x, x) = \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} F(\delta x, \delta^{-1}) T(x, \delta x) \mu(x). \quad \square$$

Let  $\mu$  be a distribution over a finite set  $\mathcal{X}$ , and a continuous Markov chain with transition rate  $T$  satisfying the detailed-balance equation with respect to  $\mu$ . Given  $f \in \mathbb{R}^{\mathcal{X}}$ , denote  $\nabla f(x, y) = f(y) - f(x)$ . For each  $\psi \in \mathbb{R}^{\mathcal{X}}$  and  $\rho \in \mathbb{R}_{>0}^{\mathcal{X}}$  satisfying  $\mathbb{E}_{\mu}[\rho] = 1$ , we define  $\mathcal{A}(\rho, \psi)$  and  $\mathcal{B}(\rho, \psi)$  as follows.

$$(70) \quad \mathcal{A}(\rho, \psi) = \frac{1}{2} \sum_{x, y \in \mathcal{X}} \mu(x) T(x, y) (\nabla \psi(x, y))^2 \hat{\rho}(x, y)$$

$$(71) \quad \mathcal{B}(\rho, \psi) = \frac{1}{2} \sum_{x, y \in \mathcal{X}} \mu(x) T(x, y) \left( \frac{1}{2} \hat{L}\rho(x, y) (\nabla \psi(x, y))^2 - \hat{\rho}(x, y) \nabla \psi(x, y) \nabla L\psi(x, y) \right)$$

where

$$\begin{aligned} \hat{\rho}(x, y) &= \theta(\rho(x), \rho(y)), \\ \hat{L}\rho(x, y) &= \partial_1 \theta(\rho(x), \rho(y)) L\rho(x) + \partial_2 \theta(\rho(x), \rho(y)) L\rho(y), \\ \text{and } \theta(x, y) &= \begin{cases} \frac{x-y}{\log x - \log y}, & x \neq y \\ x, & \text{otherwise.} \end{cases} \end{aligned}$$

The relation between  $\mathcal{A}(\rho, \psi)$ ,  $\mathcal{B}(\rho, \psi)$  and modified log-Sobolev inequality was established.

**Proposition B.2** ([EHMT17, Lemma 2.3]). *If for any  $\rho \in \mathbb{R}_{>0}^{\mathcal{X}}$  satisfying  $\mathbb{E}_{\mu}[\rho] = 1$  and  $\psi \in \mathbb{R}^{\mathcal{X}}$ ,*

$$\mathcal{B}(\rho, \psi) \geq \kappa \mathcal{A}(\rho, \psi),$$

*for some  $\kappa \in (0, 1)$ , then the modified log-Sobolev constant is at least  $2\kappa$ .*

Let  $\mathcal{G}$  be a group acting on  $\mathcal{X}$  such that for each  $x, y \in \mathcal{X}$  with transition rate  $T(x, y) > 0$ , there exists a unique  $\delta \in \mathcal{G}$  satisfying  $y = \delta x$ . For each  $\rho \in \mathbb{R}_{>0}^{\mathcal{X}}$  satisfying  $\mathbb{E}_\mu[\rho] = 1$  and  $\psi \in \mathbb{R}^{\mathcal{X}}$ , we may rephrase  $\mathcal{A}(\rho, \psi)$  and  $\mathcal{B}(\rho, \psi)$  as

$$(72) \quad \mathcal{A}(\rho, \psi) = \frac{1}{2} \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) \hat{\rho}(x, \delta x) (\nabla_\delta \psi(x))^2,$$

$$(73) \quad \mathcal{B}(\rho, \psi) = \sum_{x \in \mathcal{X}, \delta, \eta \in \mathcal{G}} \mu(x) T(x, \delta x) T(x, \eta x) B(\rho, \psi)(x, \delta, \eta),$$

$$B(\rho, \psi)(x, \delta, \eta) = \frac{1}{2} (\nabla_\delta \psi(x))^2 \hat{\rho}_1(x, \delta x) \nabla_\eta \rho(x) + \nabla_\delta \psi(x) \nabla_\eta \psi(x) \hat{\rho}(x, \delta x),$$

where  $\nabla_\delta \psi(x) = \psi(\delta x) - \psi(x)$ ,  $\nabla_\rho \psi(x) = \rho(\delta x) - \rho(x)$  and  $\hat{\rho}_i(x, y) = \left. \frac{\partial \rho}{\partial z_i} \right|_{(z_1, z_2) = (\rho(x), \rho(y))}$  for  $i = 1, 2$ .

To verify (73), by the definition of  $\mathcal{B}$ , we have

$$(74) \quad \mathcal{B}(\rho, \psi) = \sum_{x, y \in \mathcal{X}} \mu(x) T(x, y) \left( \frac{1}{4} \hat{L}\rho(x, y) (\nabla \psi(x, y))^2 - \frac{1}{2} \hat{\rho}(x, y) \nabla \psi(x, y) \nabla L \psi(x, y) \right).$$

We have

$$\begin{aligned} & \sum_{x, y \in \mathcal{X}} \mu(x) T(x, y) \hat{L}\rho(x, y) (\nabla \psi(x, y))^2 = \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) (\nabla_\delta \psi(x))^2 (\hat{\rho}_1(x, \delta x) L\rho(x) + \hat{\rho}_2(x, \delta x) L\rho(\delta x)) \\ &= \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) (\nabla_\delta \psi(x))^2 \hat{\rho}_1(x, \delta x) L\rho(x) + \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) (\nabla_\delta \psi(x))^2 \hat{\rho}_2(x, \delta x) L\rho(\delta x) \\ &\stackrel{(*)}{=} \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) (\nabla_\delta \psi(x))^2 \hat{\rho}_1(x, \delta x) L\rho(x) + \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) (\nabla_{\delta^{-1}} \psi(\delta x))^2 \hat{\rho}_2(\delta x, x) L\rho(x) \\ &\stackrel{(\star)}{=} 2 \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) (\nabla_\delta \psi(x))^2 \hat{\rho}_1(x, \delta x) L\rho(x) \\ &= 2 \sum_{x \in \mathcal{X}, \delta, \eta \in \mathcal{G}} \mu(x) T(x, \delta x) T(x, \eta x) (\nabla_\delta \psi(x))^2 \hat{\rho}_1(x, \delta x) \nabla_\eta \rho(x). \quad (\text{by definition of } L\rho), \end{aligned}$$

where  $(*)$  holds because of Fact B.1,  $(\star)$  holds because  $\nabla_{\delta^{-1}} \psi(\delta x) = -\nabla_\delta \psi(x)$  and  $\hat{\rho}_2(\delta x, x) = \hat{\rho}_1(x, \delta x)$ . Besides, we have

$$\begin{aligned} & \sum_{x, y \in \mathcal{X}} \mu(x) T(x, y) \hat{\rho}(x, y) \nabla \psi(x, y) \nabla L \psi(x, y) \\ &= \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) \hat{\rho}(x, \delta x) \nabla \psi_\delta(x) \left( \sum_{\eta \in \mathcal{G}} T(\delta x, \eta \delta x) \nabla_\eta \psi(\delta x) - \sum_{\eta \in \mathcal{G}} T(x, \eta x) \nabla_\eta \psi(x) \right), \end{aligned}$$

Again, by Fact B.1, we have

$$\begin{aligned} & \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) \hat{\rho}(x, \delta x) \nabla \psi_\delta(x) \sum_{\eta \in \mathcal{G}} T(\delta x, \eta \delta x) \nabla_\eta \psi(\delta x) \\ &= \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) \hat{\rho}(\delta x, x) \nabla \psi_{\delta^{-1}}(\delta x) \sum_{\eta \in \mathcal{G}} T(x, \eta x) \nabla_\eta \psi(x) \\ &= - \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) \hat{\rho}(\delta x, x) \nabla \psi_\delta(x) \sum_{\eta \in \mathcal{G}} T(x, \eta x) \nabla_\eta \psi(x) \quad (\text{by } \nabla_{\delta^{-1}} \psi(\delta x) = -\nabla_\delta \psi(x)) \\ &= - \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x) T(x, \delta x) \hat{\rho}(x, \delta x) \nabla \psi_\delta(x) \sum_{\eta \in \mathcal{G}} T(x, \eta x) \nabla_\eta \psi(x), \quad (\text{by } \hat{\rho}(\delta x, x) = \hat{\rho}(x, \delta x)) \end{aligned}$$

which implies

$$\sum_{x,y \in \mathcal{X}} \mu(x)T(x,y)\hat{\rho}(x,y)\nabla\psi(x,y)\nabla L\psi(x,y) = -2 \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x)T(x,\delta x)\hat{\rho}(x,\delta x)\nabla\psi_\delta(x) \sum_{\eta \in \mathcal{G}} T(x,\eta x)\nabla_\eta\psi(x).$$

By (74), we have

$$\mathcal{B}(\rho, \psi) = \sum_{x \in \mathcal{X}, \delta, \eta \in \mathcal{G}} \mu(x)T(x,\delta x)T(x,\eta x)B(\rho, \psi)(x, \delta, \eta).$$

**Lemma B.3** (Lemma 3.6, [EHMT17]). *Let  $\mu$  be a distribution over a finite set  $\mathcal{X}$ . Let  $T$  be the transition rate of a continuous-time Markov chain  $M$  satisfying the detailed-balance equation with respect to  $\mu$ . Let  $\mathcal{G}$  a mapping representation of  $T$ . If  $H$  be a subset of  $\mathcal{X} \times \mathcal{G}$  that satisfies  $H \cup H^{-1} = \mathcal{X} \times \mathcal{G}$ , where  $H^{-1} = \{(\delta x, \delta^{-1}) \mid (x, \delta) \in H\}$ , then for any  $\rho \in \mathbb{R}_{>0}^{\mathcal{X}}$  satisfying  $\mathbb{E}_\mu[\rho] = 1$ , and  $\psi \in \mathbb{R}^{\mathcal{X}}$ ,*

$$\sum_{(x,\delta) \in H} \mu(x)T(x,\delta x)B(\rho, \psi)(x, \delta, \delta) \geq \frac{1}{2}\mathcal{A}(\rho, \psi).$$

*Proof of Lemma B.3.* Note that for any  $x, y \in \mathbb{R}_{>0}$ ,

$$(75) \quad x\partial_1\theta(x, y) + y\partial_2\theta(x, y) = \theta(x, y).$$

The above equation is in [EHMT17, Lemma 3.5]. Therefore, for any  $x \in \mathcal{X}$  and  $\delta \in \mathcal{G}$ , we have

$$\begin{aligned} B(\rho, \psi)(x, \delta, \delta) &= \frac{1}{2} (\nabla_\delta\psi(x))^2 (\hat{\rho}_1(x, \delta x) (\rho(\delta x) - \rho(x)) + 2\hat{\rho}(x, \delta x)) \\ &\stackrel{\text{(by (75))}}{=} \frac{1}{2} (\nabla_\delta\psi(x))^2 (\hat{\rho}_1(x, \delta x)\rho(\delta x) + \hat{\rho}_2(x, \delta x)\rho(\delta x) + \hat{\rho}(x, \delta x)) \\ &\geq \frac{1}{2} (\nabla_\delta\psi(x))^2 \hat{\rho}(x, \delta x), \end{aligned}$$

where the last inequality follows from the fact that  $\partial_1\theta(x, y) + \partial_2\theta(x, y) = \begin{cases} \frac{(x-y)^2}{xy(\log x - \log y)^2}, & x \neq y; \\ 1, & x = y. \end{cases}$  which is non-negative. Furthermore, by reversibility,  $(\nabla_\delta\psi(x))^2 = (\nabla_{\delta^{-1}}\psi(\delta x))^2$  and  $\hat{\rho}(x, \delta x) = \hat{\rho}(\delta x, x)$ ,

$$\begin{aligned} \sum_{(x,\delta) \in H} \mu(x)T(x,\delta x) (\nabla_\delta\psi(x))^2 \hat{\rho}(x, \delta x) &= \sum_{(x,\delta) \in H} \mu(\delta x)T(\delta x, \delta^{-1}(\delta x)) (\nabla_{\delta^{-1}}\psi(\delta x))^2 \hat{\rho}(\delta x, \delta^{-1}(\delta x)) \\ &\stackrel{\text{(by the definition of } H^{-1})}{=} \sum_{(x,\delta) \in H^{-1}} \mu(x)T(x,\delta x) (\nabla_\delta\psi(x))^2 \hat{\rho}(x, \delta x). \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{(x,\delta) \in H} \mu(x)T(x,\delta x)B(\rho, \psi)(x, \delta, \delta) &\geq \frac{1}{2} \sum_{(x,\delta) \in H} \mu(x)T(x,\delta x) (\nabla_\delta\psi(x))^2 \hat{\rho}(x, \delta x) \\ &\geq \frac{1}{4} \sum_{(x,\delta) \in H \cup H^{-1}} \mu(x)T(x,\delta x) (\nabla_\delta\psi(x))^2 \hat{\rho}(x, \delta x) \\ &= \frac{1}{4} \sum_{x \in \mathcal{X}, \delta \in \mathcal{G}} \mu(x)T(x,\delta x) (\nabla_\delta\psi(x))^2 \hat{\rho}(x, \delta x) \\ &\stackrel{\text{(by (72))}}{=} \frac{1}{2}\mathcal{A}(\rho, \psi). \quad \square \end{aligned}$$

We now prove Theorem 8.28

*Proof of Theorem 8.28.* For the mapping  $G$  satisfying the condition in the theorem, the following inequality was proved in [EHMT17] (see proof of Theorem 3.9 in [EHMT17]):

$$(76) \quad \mathcal{B}(\rho, \psi) \geq \sum_{(x, \delta) \in \mathcal{X} \times \mathcal{G}} B(x, \delta, \delta) \left( q(x, \delta, \delta) - \mathbf{1}_{\delta \neq \delta^{-1}} q(\delta x, \delta^{-1}, \delta) - \sum_{\eta: \eta \neq \delta, \delta^{-1}} (q - q_*)(\delta x, \delta^{-1}, \eta) \right).$$

By the definition of  $q$ , reversibility and non-negativity, we have

$$\begin{aligned} \mathcal{B}(\rho, \psi) &\geq \sum_{(x, \delta) \in \mathcal{X} \times \mathcal{G}} \mu(x) T(x, \delta x) B(x, \delta, \delta) \left( T(x, \delta x) - \mathbf{1}_{\delta \neq \delta^{-1}} T(\delta x, \delta(\delta x)) - \sum_{\eta: \eta \neq \delta, \delta^{-1}} \frac{(q - q_*)(\delta x, \delta^{-1}, \eta)}{\mu(x) T(x, \delta x)} \right) \\ &\stackrel{(*)}{=} \sum_{(x, \delta) \in H_1 \cup H_2} \mu(x) T(x, \delta x) B(x, \delta, \delta) \left( T(x, \delta x) - \mathbf{1}_{\delta \neq \delta^{-1}} T(\delta x, \delta(\delta x)) - \sum_{\eta: \eta \neq \delta, \delta^{-1}} \frac{(q - q_*)(\delta x, \delta^{-1}, \eta)}{\mu(x) T(x, \delta x)} \right) \\ &\stackrel{(\star)}{\geq} \kappa_1 \sum_{(x, \delta) \in H_1} \mu(x) T(x, \delta x) B(x, \delta, \delta) + \kappa_2 \sum_{(x, \delta) \in H_2} \mu(x) T(x, \delta x) B(x, \delta, \delta), \end{aligned}$$

where  $(*)$  holds because  $H_1 \cup H_2 = \mathcal{X} \times \mathcal{G}$ , and  $(\star)$  holds because  $H_1 \cap H_2 = \emptyset$ . We now use Lemma B.3. Note that  $H_2 = H_1^{-1}$  and  $H_1 = H_2^{-1}$  and  $H_1 \uplus H_2 = \mathcal{X} \times \mathcal{G}$ . We have

$$\mathcal{B}(\rho, \psi) \geq \frac{\kappa_1 + \kappa_2}{2} \mathcal{A}(\rho, \psi).$$

By Proposition B.2, we conclude the proof.  $\square$

#### APPENDIX C. MONOTONICITY OF UNIQUENESS CONDITION

In this section, we prove a stronger version of Proposition 8.1.

**Proposition C.1** ([LLY13]). *Let  $\beta, \gamma, \lambda$  be parameters of an anti-ferromagnetic system, and  $\hat{x}_d \in \mathbb{R}_{>0}$  be the unique fixed point of recursion  $F_d(x) = \lambda \left( \frac{\beta x + 1}{x + \gamma} \right)^d$  for any  $d \in \mathbb{Z}_{>0}$ . The following statements are equivalent.*

- (1)  $\gamma \leq 1$ ;
- (2)  $|F'_d(\hat{x}_d)|$  is monotone increasing in  $d$ .

*Proof.* When  $\gamma > 1$ , the fixed point  $\hat{x}_d$  satisfies

$$\hat{x}_d = \lambda \left( \frac{\beta \hat{x}_d + 1}{\hat{x}_d + \gamma} \right)^d \leq \frac{\lambda}{\gamma^d}.$$

Hence,

$$0 \leq \lim_{d \rightarrow +\infty} |F'_d(\hat{x}_d)| = \lim_{d \rightarrow +\infty} \frac{d(1 - \beta\gamma)\hat{x}_d}{(\beta\hat{x}_d + 1)(\hat{x}_d + \gamma)} \leq \lim_{d \rightarrow +\infty} \frac{d\lambda}{\gamma^d} = 0.$$

Note that  $F'_1(\hat{x}_1) > 0$ . Therefore,  $F'_d(\hat{x}_d)$  must not monotone increase for all  $d$ .

When  $\gamma \leq 1$ , define  $c(d) : [1, \infty) \rightarrow \mathbb{R}$  as

$$c(d) = \frac{d(1 - \beta\gamma)\hat{x}_d}{(\beta\hat{x}_d + 1)(\hat{x}_d + \gamma)} = \frac{d(1 - \beta\gamma)\hat{x}_d}{p(\hat{x}_d)}, \quad \text{where } p(x) = (\beta x + 1)(x + \gamma).$$

Note that  $c(d) = |F'_d(\hat{x}_d)|$ . Hence, it suffices to show that

$$\gamma \geq 1 \implies \forall d > 0, c'(d) > 0.$$

Let  $q(x) = \frac{\beta x + 1}{x + \gamma}$ . Take the derivative of  $c(d)$  with respect to  $d$ , we have

$$c'(d) = \frac{(1 - \beta\gamma)\hat{x}_d}{p(\hat{x}_d)} \left( 1 - d \ln q(\hat{x}_d) \cdot \frac{\beta\hat{x}_d^2 - \gamma}{p(\hat{x}_d) + d(1 - \beta\gamma)\hat{x}_d} \right).$$

Since  $\frac{(1 - \beta\gamma)\hat{x}_d}{p(\hat{x}_d)} > 0$ , we only need to verify that

$$d \ln q(\hat{x}_d) \cdot \frac{\beta\hat{x}_d^2 - \gamma}{p(\hat{x}_d) + d(1 - \beta\gamma)\hat{x}_d} \leq 1.$$

Let  $x = \hat{x}_d$ . Note that  $d \geq 1$ . It suffices to show that

$$\forall x > 0, \quad (\beta x^2 - \gamma) \ln \left( \frac{\beta x + 1}{x + \gamma} \right) \leq (1 - \beta\gamma)x.$$

Note that  $\beta x^2 - \gamma > 0$  if and only if  $x \geq \sqrt{\gamma/\beta}$ ;  $\ln \left( \frac{\beta x + 1}{x + \gamma} \right) \geq 0$  if and only if  $x \leq \frac{1 - \gamma}{1 - \beta}$ . Since  $0 \leq \beta \leq \gamma \leq 1$  and  $\beta\gamma < 1$ , we have

$$\sqrt{\frac{\gamma}{\beta}} \geq \frac{1 - \gamma}{1 - \beta}.$$

Note that  $(1 - \beta\gamma)x > 0$ , we only need to prove that

$$\forall \frac{1 - \gamma}{1 - \beta} < x < \sqrt{\frac{\gamma}{\beta}}, \quad \ln \left( \frac{x + \gamma}{\beta x + 1} \right) \leq \frac{(1 - \beta\gamma)x}{\gamma - \beta x^2}$$

Note that  $\frac{x + \gamma}{\beta x + 1} = 1 + \frac{(1 - \beta)x + (\gamma - 1)}{\beta x + 1} \leq \exp \left( \frac{(1 - \beta)x + (\gamma - 1)}{\beta x + 1} \right)$ , it suffices to show that

$$\forall \frac{1 - \gamma}{1 - \beta} < x < \sqrt{\frac{\gamma}{\beta}}, \quad \frac{(1 - \beta)x + (\gamma - 1)}{\beta x + 1} \leq \frac{(1 - \beta\gamma)x}{\gamma - \beta x^2}.$$

Note that  $\gamma - \beta x^2 > 0$  if  $x < \sqrt{\frac{\gamma}{\beta}}$ . The above inequality is equivalent to for all  $\frac{1 - \gamma}{1 - \beta} < x < \sqrt{\frac{\gamma}{\beta}}$ ,

$$\beta(\beta - 1)x^3 + \beta\gamma(\beta - 1)x^2 + (\gamma - 1)x + \gamma(\gamma - 1) \leq 0.$$

The above inequality holds because if  $\gamma \leq 1$ , then  $\beta(\beta - 1) < 0$ ,  $\beta\gamma(\beta - 1) < 0$ ,  $\gamma - 1 \leq 0$  and  $\gamma(\gamma - 1) \leq 0$ .  $\square$

#### APPENDIX D. BOUNDEDNESS OF ANTI-FERROMAGNETIC TWO-SPIN SYSTEM

*Proof of Lemma 8.16.* Note that the first part directly follows from Proposition 8.1 and Lemma 36 in [CLV20].

Therefore, we will only focus on the case where  $G$  is  $(\Delta - 1)$ -regular.

First, we prove this for the case where  $\sqrt{\beta\gamma} \geq \frac{\Delta - 2}{\Delta}$ , for all  $y \in [-\infty, +\infty]$ , it holds that

$$|h(y)| = \frac{(1 - \beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)} = \frac{1 - \beta\gamma}{\beta e^y + \gamma e^{-y} + 1 + \beta\gamma} \leq \frac{1 - \beta\gamma}{1 + \beta\gamma + 2\sqrt{\beta\gamma}} = \frac{1 - \sqrt{\beta\gamma}}{1 + \sqrt{\beta\gamma}} \leq \frac{1}{\Delta - 1} \leq \frac{1.5}{\Delta},$$

where in the last inequality, we use the fact that  $\frac{2}{3}\Delta \leq (\Delta - 1)$ .

Now, we only left consider the case where  $\sqrt{\beta\gamma} < \frac{\Delta - 2}{\Delta}$ . When  $\lambda < \lambda_1(\Delta - 1)$ , for any  $y \in J_{\lambda, \Delta - 1}$ ,  $e^y \leq \frac{\lambda}{\gamma^{\Delta - 1}} \leq \frac{18\gamma}{\theta(\Delta - 1)}$ , where the last inequality follows from Lemma 8.25 and  $\theta(d) \triangleq d(1 - \beta\gamma) - (1 + \beta\gamma)$ .

Therefore,  $|h(y)|$  can be bounded as follows

$$|h(y)| = \frac{(1 - \beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)} \leq \frac{1 - \beta\gamma}{\gamma e^{-y} + 1 + \beta\gamma} \leq \frac{1 - \beta\gamma}{\frac{\theta(\Delta - 1)}{18} + 1 + \beta\gamma} = \frac{18(1 - \beta\gamma)}{\Delta(1 - \beta\gamma) + 18\beta\gamma + 16} \leq \frac{18}{\Delta}.$$



Similarly, when  $\lambda > \lambda_2(\Delta - 1)$ , for any  $y \in J_{\lambda, \Delta - 1}$ ,  $e^y \geq \lambda\beta^{\Delta - 1} \geq \frac{\theta(\Delta - 1)}{18\beta}$ , where the last inequality follows from Lemma 8.25. Therefore,

$$|h(y)| = \frac{(1 - \beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)} \leq \frac{1 - \beta\gamma}{\beta e^y + 1 + \beta\gamma} \leq \frac{1 - \beta\gamma}{\frac{\theta(\Delta - 1)}{18} + 1 + \beta\gamma} = \frac{18(1 - \beta\gamma)}{\Delta(1 - \beta\gamma) + 18\beta\gamma + 16} \leq \frac{18}{\Delta}.$$

This concludes the proof of Lemma 8.16. □