ON THE MIXING TIME OF GLAUBER DYNAMICS FOR THE HARD-CORE AND RELATED MODELS ON $G(n, d/n)$

CHARILAOS EFTHYMIOU* AND WEIMING FENG†

ABSTRACT. We study the single-site Glauber dynamics for the fugacity λ, Hard-core model on the random graph $G(n, d/n)$. We show that for the typical instances of the random graph $G(n, d/n)$ and for fugacity $\lambda < \frac{d^d}{(d-1)^{d+1}}$, the mixing time of Glauber dynamics is $n^{1+O(1/\log\log n)}$.

Our result improves on the recent elegant algorithm in [Bezáková, Galanis, Goldberg Štefankovič; ICALP'22]. The algorithm there is a MCMC based sampling algorithm, but it is not the Glauber dynamics. Our algorithm here is *simpler*, as we use the classic Glauber dynamics. Furthermore, the bounds on mixing time we prove are smaller than those in Bezáková et al. paper, hence our algorithm is also *faster*.

The main challenge in our proof is handling vertices with unbounded degrees. We provide stronger results with regard the spectral independence via branching values and show that the our Gibbs distributions satisfy the approximate tensorisation of the entropy. We conjecture that the bounds we have here are optimal for $G(n, d/n)$.

As corollary of our analysis for the Hard-core model, we also get bounds on the mixing time of the Glauber dynamics for the Monomer-dimer model on $G(n, d/n)$. The bounds we get for this model are slightly better than those we have for the Hard-core model

* University of Warwick, Coventry, CV4 7AL, UK. Email: Charilaos.Efthymiou@warwick.ac.uk.
† University of Edinburgh, Edinburgh, EH8 9AB, UK. Email: wfeng@ed.ac.uk.

1. INTRODUCTION

The Hard-core model and the related problem of the geometry of independent sets on the spare random graph $G(n, d/n)$ is a fundamental area of study in discrete mathematics [Fri90, DM11], in computer science they are studied in the context of the random *Constraint Satisfaction Problems* [CE15, GS14], while in statistical physics they are studied as instances of *disordered systems*. Using the so-called *Cavity method* [KMR$^+$07, BKZZ13], physicists make some impressive predictions about the independent sets of $G(n, d/n)$, such as higher order replica symmetry breaking etc. Physicists' predictions are (typically) mathematically non-rigorous. Most of these predictions about independent sets still remain open as basic natural objects in the study such as the partition function, or the free energy are extremely challenging to analyse.

The Hard-core model with fugacity $\lambda > 0$, is a distribution over the *independent sets* of an underlying graph G such that every independent set σ is assigned probability measure $\mu(\sigma)$ which is proportional to $\lambda^{|\sigma|}$, where $|\sigma|$ is the cardinality of σ. Here, we consider the case where the underlying graph is a typical instance of the sparse random graph $G(n, d/n)$. This is the random graph on n vertices, while each edge appears independently with probability $p = d/n$. The quantity $d > 0$ corresponds to the *expected degree*. For us here the expected degree is a bounded constant, i.e., we have $d = \Theta(1)$, hence the graph is sparse.

Our focus is on approximate sampling from the aforementioned distribution using *Glauber dynamics*. This is a classic, very popular, algorithm for approximate sampling. The popularity of this process, mainly, is due to its simplicity and the strong approximation guarantees that provides. The efficiency of Glauber dynamics for sampling is studied by means of the *mixing time*.

Recently, there has been an "explosion" of results about the mixing time of Glauber dynamics for *worst-case* instances the problem, e.g. [ALO20, CLV21, CFYZ22, Eft22]. Combined with the earlier hardness results in [Sly10, SS14, GŠV16] one could claim that for worst-case instances the behaviour of Glauber dynamics for the Hard-core model, but also the related approximate sampling-counting problem, is well understood. Specifically, for the graphs of maximum degree Δ, Glauber dynamics exhibits $O(n \log n)$ mixing time for any fugacity $\lambda < (\Delta - 1)^{\Delta - 1}/(\Delta - 2)^{\Delta}$, while the hardness results support that this region of λ is best possible.

The aforementioned upper bound on λ coincides with the *critical point* for the uniqueness/non-uniqueness phase transition of the Hard-core model on the infinite Δ-regular tree [Kel85]. At this point in the discussion, perhaps, it is important to note the dependency of the critical point on the *maximum degree*. This is the point where the situation with the random graph $G(n, d/n)$ differentiates from the worst case one.

For $G(n, d/n)$ and for the range of the expected degree d we consider here, typically, almost all of the vertices in the graph, e.g., say 99%, are of degree very close to d. On the other hand, the maximum degree of $G(n, d/n)$ is as large as $\Theta(\frac{\log n}{\log \log n})$, i.e., it is *unbounded*. In light of this observation, it is natural to expect that the Glauber dynamics on the Hard-core model mixes fast for values of the fugacity that depend on the *expected degree*, rather the maximum degree. Note that, this implies to use Glauber dynamics to sample from the Hard-core model with fugacity λ taking *much larger* values than what the worst-case bound implies.

For $d > 1$, let $\lambda_c(d) = \frac{d^d}{(d-1)^{(d+1)}}$. One of the main result in our paper is as follows: we show that for any $d > 1$ and for typical instances of $G(n, d/n)$, the Glauber dynamics on the Hard-core with any fugacity $\lambda < \lambda_c(d)$, exhibits mixing time which is $n^{1+\frac{C}{\log \log n}} = n^{1+o(1)}$, for some absolute constant $C > 0$ which depends only on λ and d.

It is our *conjecture* that the bound on the mixing time is tight. Furthermore, following intuitions from [CE15], as well as from *statistical physics* predictions in [BKZZ13], it is our *conjecture* that the bound $\lambda_c(d)$ on the fugacity λ is also tight, in the following sense: for $\lambda > \lambda_c(d)$ it is not precluded that there is a region where efficient approximate sampling is possible, however, the approximation guarantees are *weaker* than those we have here.

Our result improves on the elegant sampling algorithm that was proposed recently in [BGGŠ22] for the same distribution, i.e., the Hard-core model on $G(n, d/n)$. That algorithm, similarly to the one we consider here, relies on the Markov Chain Monte Carlo method. The authors use Spectral Independence [ALO20, CLV21] to show that the underlying Markov chain exhibits mixing time which is $O\left(n^{1+\theta}\right)$ for any $\lambda < \lambda_c(d)$ and arbitrary small consant $\theta > 0$. The idea that underlies the algorithm in [BGGŠ22] is reminiscent of the *variable marking* technique that was introduced in [Moi19] for approximate counting with the Lovász Local Lemma, and was further exploited in [FGYZ21a, FHY21, JPV21a, GGGY21]. Here, we use a different, more straightforward, approach and analyse directly the Glauber dynamics.

Note that both algorithms, i.e., here and in [BGGŠ22], allow for the same range for the fugacity λ. On the other hand, the algorithm we study here is the (much simpler) Glauber dynamics, while the running time guarantees we obtain here are asymptotically better.

Previous works in the area, i.e., even before [BGGŠ22], in order to prove their results and avoid the use of maximum degree, have been focusing on various parameters of $G(n, d/n)$ such as the expected degree [EHSV18], or the connective constant [SSŠY17]. Which, as it turns out are not that different with each other. Here, we utilise the notion of *branching value*, which is somehow related to the previous ones.

The notion of the branching value as well as its use for establishing Spectral Independence was introduced in [BGGŠ22]. Unfortunately, the result there were not sufficiently strong to imply rapid mixing of Glauber dynamics. Their analytic tools for Spectral independence (and others) seems to not be able to handle all that well vertices with unbounded degree. Here we derive stronger results for Spectral independence than those in [BGGŠ22] in the sense that they are more *general* and more *accurate*. Specifically, in our analysis we are able to accommodate vertices of *all degrees*, while we use a more elaborate matrix norm to establish spectral independence, reminiscent of those introduced in [Eft22]. Furthermore, we utilise results from [CFYZ22] that allow us deal with the unbounded degrees of the graph in order to establish our rapid mixing results.

2. Results

Consider the fixed graph $G = (V, E)$ on n vertices. Given the parameter $\lambda > 0$, which we call *fugacity*, we define the *Hard-core* model $\mu = \mu_{G,\lambda}$ to be a distribution on the independent sets of the graph G, Specifically, every independent set σ is assigned probability measure $\mu(\sigma)$ defined by

$$\mu(\sigma) \propto \lambda^{|\sigma|} \ , \tag{1}$$

where $|\sigma|$ is equal to the size of the independent set σ.

We use $\{\pm 1\}^V$ to encode the configurations of the Hard-core model, i.e., the independent sets of G. Particularly, the assignment $+1$ implies that the vertex is in the independent set, while -1 implies the opposite. We often use physics' terminology where the vertices with assignment $+1$ are called "occupied", whereas the vertices with -1 are "unoccupied".

We use the discrete time, (single site) *Glauber dynamics* to approximately sample from the aforementioned distributions. Glauber dynamics is a Markov chain with state space the support of the distribution μ. Typically, we assume that the chain starts from an arbitrary configuration $X_0 \in \{\pm 1\}^V$. For $t \geq 0$, the transition from the state X_t to X_{t+1} is according to the following steps:

(1) Choose uniformly at random a vertex v.
(2) For every vertex w different than v, set $X_{t+1}(w) = X_t(w)$.
(3) Set $X_{t+1}(v)$ according to the marginal of μ at v, conditional on the neighbours of v having the configuration specified by X_{t+1}.

It is standard that when a Markov chain satisfies a set of technical conditions called *ergodicity*, then it converges to a unique stationary distribution. For the cases we consider here, Glauber dynamics is trivially ergodic, while the stationary distribution is the corresponding Hard-core model μ.

Let P be the transition matrix of an ergodic Markov chain $\{X_t\}$ with a finite state space Ω and equilibrium distribution μ. For $t \geq 0$ and $\sigma \in \Omega$, let $P^t(\sigma, \cdot)$ denote the distribution of X_t when the initial state of the chain satisfies $X_0 = \sigma$. The *mixing time* of the Markov chain $\{X_t\}_{t\geq 0}$ is defined by

$$T_{\rm mix} = \max_{\sigma\in\Omega} \min\left\{ t>0 \mid \|P^t(\sigma,\cdot)-\mu\|_{\rm TV} \leq \frac{1}{2{\rm e}} \right\} .$$

Our focus is on the mixing time of Glauber dynamics for the Hard-core model for the case where the underlying graph is a typical instance of $G(n, d/n)$, where the expected degree $d > 0$ is a assumed to be a fixed number.

2.1. **Mixing Time for Hard-core Model.** For $z > 1$, we let the function $\lambda_c(z) = \frac{z^z}{(z-1)^{(z+1)}}$. It is a well-known result from [Kel85] that the uniqueness region of the Hard-core model on the k-ary tree, where $k \geq 2$, holds for any λ such that

$$\lambda < \lambda_c(k) .$$

The following theorem is the main result of this work.

Theorem 2.1. *For fixed $d > 1$ and any $\lambda < \lambda_c(d)$, there is a constant $C > 0$ such that the following is true: Let $\mu_{\boldsymbol{G}}$ be the Hard-core model with fugacity λ on the graph $\boldsymbol{G} \sim G(n, d/n)$. With probability $1 - o(1)$ over the instances of $\boldsymbol{G}$, Glauber dynamics on $\mu_{\boldsymbol{G}}$ exhibits mixing time*

$$T_{\rm mix} \leq n^{\left(1+\frac{C}{\log\log n}\right)} .$$

2.2. **Extensions to Monomer-dimer Model.** Utilising the techniques we develop in order to prove Theorem 2.1, we get mixing time bounds for the Glauber dynamics on the Monomer-Dimer model on $G(n, d/n)$.

Given a fixed graph $G = (V, E)$ and a parameter $\lambda > 0$, which we call *edge weight*, we define the Monomer-Dimer model $\mu = \mu_{G,\lambda}$ to be a distribution on the *matchings* of the graph G such that every matching σ is assigned probability measure $\mu(\sigma)$ defined by

$$\mu(\sigma) \propto \lambda^{|\sigma|} , \tag{2}$$

where $|\sigma|$ is equal to the number of edges in the matching σ.

Note that the Hard-core model considers configurations on the vertices of G, while the Monomer-Dimer model considers configurations on the edges. Similarly to the independent sets, we use $\{\pm 1\}^E$ to encode the matchings of G. Specifically, the assignment $+1$ on the edge e implies that the edge is in matching, while -1 implies the opposite.

For the Monomer-Dimer model the definition of Glauber dynamics $\{X_t\}_{t\geq 0}$ extends in the natural way. That is, assume that the chain starts from an arbitrary configuration $X_0 \in]\pm 1^E$. For $t \geq 0$, the transition from the state X_t to X_{t+1} is according to the following steps:

(1) Choose uniformly at random an edge e.
(2) For every edge f different than e, set $X_{t+1}(f) = X_t(f)$.
(3) Set $X_{t+1}(e)$ according to the marginal of μ at e, conditional on the neighbours of e having the configuration specified by X_{t+1}.

We consider the case of the Monomer-Dimer distribution where the underlying graph is an instance of $G(n, d/n)$. We prove the following result.

Theorem 2.2. *For fixed $d > 1$ and any $\lambda > 0$, there is a constant $C > 0$ such that the following is true: Let $\mu_{\boldsymbol{G}}$ be the Monomer-dimer model with edge weight λ on the graph $\boldsymbol{G} \sim G(n, d/n)$. With probability $1 - o(1)$ over the instances of $\boldsymbol{G}$, Glauber dynamics on $\mu_{\boldsymbol{G}}$ exhibits mixing time*

$$T_{\rm mix} \leq n^{\left(1+C\sqrt{\frac{\log\log n}{\log n}}\right)} .$$

The proof of Theorem 2.2 can be found in Section 8.

For the Monomer-dimer model on general graphs, the best-known result is the $\tilde{O}(n^2 m)$ mixing time of the Jerrum-Sinclair chain [JS89], where $m = |E|$ is the number of edges. For graphs with bounded maximum degree $\Delta = O(1)$, the spectral independence technique proved the $O(n \log n)$ mixing time of Glauber dynamics [CLV21]. However, this result cannot be applied directly to the random graph $G(n, d/n)$, because the maximum degree of a random graph is typically unbounded. For the Monomer-dimer model on $G(n, d/n)$, [BGGŠ22] gave a sampling algorithm with running time $n^{1+\theta}$, where $\theta > 0$ is an arbitrarily small constant, and [JPV21b] also proved the $n^{2+o(1)}$ mixing time of Glauber dynamics in a special case $\lambda = 1$. Our result in Theorem 2.2 proves the $n^{1+o(1)}$ mixing time of Glauber dynamics, which improves all the previous results for the Monomer-dimer model on the random graph $G(n, d/n)$ with constant λ.

We remark that for the Monomer-dimer model, we actually proved the $n^{1+o(1)}$ mixing time of Glauber dynamics on *all* graphs satisfying $\Delta \log^2 \Delta = o(\log^2 n)$. See Theorem 8.1 for a more general result.

Note that, apart from Section 8, the rest of the paper focuses on the Hard-core model, i.e., proving Theorem 2.1.

Notation. Suppose that we are given a Gibbs distribution μ on the graph $G = (V, E)$. We denote with Ω the support of μ.

Suppose that Ω is a set of configuration at the vertices of G. Then, for any $\Lambda \subseteq V$ and any $\tau \in \{\pm 1\}^{\Lambda}$, we let $\mu^{\Lambda,\tau}$ (or μ^{τ} if Λ is clear from the context) denote the distribution μ conditional on that the configuration at Λ is τ. Alternatively, we use the notation $\mu(\cdot \mid (\Lambda, \tau))$ for the same conditional distribution. We let $\Omega^{\tau} \subseteq \Omega$ be the support of $\mu^{\Lambda,\tau}$. We call τ *feasible* if Ω^{τ} is nonempty.

For any subset $S \subseteq V$, let μ_S denote the marginal of μ at S, while let Ω_S denote the support of μ_S. In a natural way, we define the conditional marginal. That is, for $\Lambda \subseteq V \setminus S$ and $\sigma \in \{\pm 1\}^{\Lambda}$, we let $\mu_S^{\Lambda,\sigma}$ (or μ_S^{σ} if Λ is clear from the context) denote the marginal at S conditional on the configuration at Λ being σ. Alternatively we use $\mu_S(\cdot \mid (\Lambda, \sigma))$ for μ_S^{σ}. We let Ω_S^{σ} denote the support of μ_S^{σ}.

All the above notation for configurations on the vertices of G can be extended naturally for configurations on the edges of the graph G. We omit presenting it, because it is very similar to the above.

2.3. **Hard-core Model - Entropy Tensorisation for Rapid Mixing.** We prove Theorem 2.1 by exploiting the notion of *approximate tensorisation of the entropy*.

Let μ be a distribution with support $\Omega \subseteq \{\pm 1\}^V$. For any function $f : \Omega \to \mathbb{R}_{\geq 0}$, we let $\mu(f) = \sum_{x \in \Omega} \mu(x) f(x)$, i.e., $\mu(f)$ is the expected value of f with respect to μ. Define the entropy of f with respect to μ by

$$\mathrm{Ent}_{\mu}(f) = \mu\left(f \log \tfrac{f}{\mu(f)}\right) \quad ,$$

where we use the convention that $0 \log 0 = 0$.

Let $\tau \in \Omega_{V \setminus S}$ for some $S \subset V$. Define the function $f_{\tau} : \Omega_S^{\tau} \to \mathbb{R}_{\geq 0}$ by having $f_{\tau}(\sigma) = f(\tau \cup \sigma)$ for all $\sigma \in \Omega_S^{\tau}$ *. Let $\mathrm{Ent}_S^{\tau}(f_{\tau})$ denote the entropy of f_{τ} with respect to the conditional distribution μ_S^{τ}. Furthermore, we let

$$\mu(\mathrm{Ent}_S(f)) = \sum_{\tau \in \Omega_{V \setminus S}} \mu_{V \setminus S}(\tau) \mathrm{Ent}_S^{\tau}(f_{\tau}) \ ,$$

i.e., $\mu(\mathrm{Ent}_S(f))$ is the average of the entropy $\mathrm{Ent}_S^{\tau}(f_{\tau})$ with respect to the measure $\mu_{V \setminus S}(\cdot)$. When $S = \{v\}$, i.e., the set S is a singleton, we abbreviate $\mu(\mathrm{Ent}_{\{v\}}(f))$ to $\mu(\mathrm{Ent}_v(f))$.

Definition 2.3 (Approximate Tensorisation of Entropy)**.** *A distribution μ with support $\Omega \subseteq \{\pm 1\}^V$ satisfies the approximate tensorisation of entropy with constant $C > 0$ if for all $f : \Omega \to \mathbb{R}_{\geq 0}$ we have that*

$$\mathrm{Ent}_{\mu}(f) \leq C \cdot \sum_{v \in V} \mu\left(\mathrm{Ent}_v(f)\right) \quad .$$

*With a slight abuse of notation we use $\tau \cup \sigma$ to indicate the configuration what agrees with τ at S and with σ at $V \setminus S$.

On can establish bounds on the mixing time of Glauber dynamics by means of the approximate tensorisation of entropy of the equilibrium distribution μ. Specifically, if μ satisfies the approximate tensorisation of entropy with constant C, then after every transition of Glauber dynamics, the Kullback–Leibler divergence† between the current distribution and the stationary distribution decays by a factor which is at least $(1-C/n)$, where $n = |V|$ is the number of variables.

As far as the mixing time of Glauber dynamics is concerned, if a distribution μ satisfies the approximate tensorisation of entropy with parameter C then we have following well known relation (e.g. see [CLV21, Fact 3.5]),

$$T_{\text{mix}} \le \left\lceil Cn \left(\log\log \frac{1}{\mu_{\min}} + \log(2) + 2 \right) \right\rceil , \qquad \text{where } \mu_{\min} = \min_{x \in \Omega} \mu(x) \ . \tag{3}$$

In light of the above, Theorem 2.1 follows as a corollary from the following result.

Theorem 2.4 (Hard-core Model Tensorisation)**.** *For any fixed $d > 1$ and any $\lambda < \lambda_c(d)$, there is a constant $A > 0$ that depends only on d and λ such that the following is true:*

Let $\mu_{\boldsymbol{G}}$ be the Hard-core model with fugacity λ on the graph $\boldsymbol{G} \sim G(n, d/n)$. With probability $1 - o(1)$ over the instances of $\boldsymbol{G}$, $\mu_{\boldsymbol{G}}$ satisfies the approximate tensorisation of entropy with parameter $n^{A/\log\log n}$.

Proof of Theorem 2.1. Theorem 2.1 follows from Theorem 2.4 and (3).

Specifically, plugging the result from Theorem 2.4 into (3) we get the following: with probability $1-o(1)$ over the instances of $\boldsymbol{G}$ we have that

$$\begin{aligned} T_{\text{mix}} &\le n^{1+\frac{A}{\log\log n}} \left(\log\log \frac{1}{\mu_{\min}} + \log(2) + 2 \right) \\ &\le n^{1+\frac{A}{\log\log n}} \left(\log\log \left(1 + \lambda + \lambda^{-1}\right)^n + \log(2) + 2 \right) \\ &= n^{1+\frac{A}{\log\log n}} \left(\log n + \log\log(1 + \lambda + \lambda^{-1}) \right) \le n^{1+\frac{2A}{\log\log n}} \ . \end{aligned}$$

For the second derivation, we note that for the Hard-core distribution $\mu = \mu_{\boldsymbol{G}}$, we have that $\mu_{\min}$ is at least $\min\{1, \lambda^n\}/(1+\lambda)^n$, which implies that $\mu_{\min} \ge (1 + \lambda + \lambda^{-1})^{-n}$.

Note that Theorem 2.1 follows from the above, by setting $C = 2A$. □

3. Our Approach & Contributions

In this section we describe our approach towards establishing our results. Our focus is on the Hard-core model.

3.1. **Tensorisation and Block-Factorisation of Entropy.** We establish the tensorisation of the entropy, described in Theorem 2.4, by exploiting the recently introduced notion of *block factorisation of entropy* in [CP20]. Specifically, we build on the framework introduced in [CLV21] to relate the tensorisation and the block factorisation of the entropy.

The framework in [CLV21] relies on the assumption that the maximum degree of the underlying graph is bounded. Otherwise, the results it implies are not strong. In our setting here, a vanilla application of this approach would not be sufficient to give the desirable bounds on the tensorisation constant due to the fact that the typical instances of $G(n, d/n)$ have unbounded maximum degree. To this end, we employ techniques from [CFYZ22].

Given the graph $G = (V, E)$, and the integer $\ell \ge 0$, we let $\binom{V}{\ell}$ denote all subsets $S \subseteq V$ with $|S| = \ell$.

†For discrete probability distributions P and Q on a discrete space $\mathcal{X}$, the Kullback–Leibler divergence is defined by $D_{\text{KL}}(P||Q) = \sum_{x\in\mathcal{X}} P(x) \log \frac{P(x)}{Q(x)}$.

Definition 3.1 (ℓ-block Factorisation of Entropy). *Let μ be a distribution over $\{\pm 1\}^V$ and $1 \leq \ell \leq |V| = n$ be an integer. The distribution μ satisfies the ℓ block factorisation of entropy with parameter C if for all $f : \Omega \to \mathbb{R}_{\geq 0}$ we have that*

$$\mathrm{Ent}_\mu(f) \leq \frac{C}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \mu\left(\mathrm{Ent}_S(f)\right) \quad . \tag{4}$$

The notion of the ℓ block factorisation of entropy generalises that of the approximate tensorisation of entropy. Specifically, a distribution that satisfies the $\ell = 1$ block factorisation of entropy with parameter C, also satisfies the approximate tensorisation of entropy with parameter C/n.

As far as the Hard-core model on $G(n, d/n)$ is concerned, we show the following theorem, which is one of the main technical results in our paper.

Theorem 3.2. *For fixed $d > 1$ and any $0 < \lambda < \lambda_c(d)$, consider $\boldsymbol{G} \sim G(n, d/n)$ and let $\mu_{\boldsymbol{G}}$ be the Hard-core model on $\boldsymbol{G}$ with fugacity λ. With probability $1 - o(1)$ over the instances of $\boldsymbol{G}$ the following is true: There is a constant $K = K(d, \lambda) > 0$, such that for*

$$\frac{1}{\alpha} = K \tfrac{\log n}{\log \log n} \quad ,$$

for any $1/\alpha \leq \ell < n$, $\mu_{\boldsymbol{G}}$ satisfies the ℓ-block factorisation of entropy with parameter $C = (\frac{en}{\ell})^{1+1/\alpha}$.

Let us have a high level overview of how we use the ℓ-block factorisation and particularly Theorem 3.2 to establish our entropy tensorisation result in Theorem 2.4.

Note that Theorem 3.2 essentially implies the following: Suppose that $G = (V, E)$ is a *typical instance* of $G(n, d/n)$. Then, the Hard-core model μ on G, with fugacity $\lambda < \lambda_c(d)$, is such that for any $f : \Omega \to \mathbb{R}_{>0}$ we have

$$\mathrm{Ent}_\mu(f) \leq \left(\frac{\mathrm{e}}{\theta}\right)^{1+1/\alpha} \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \mu\left(\mathrm{Ent}_S(f)\right) \quad , \tag{5}$$

where $\ell = \lceil \theta n \rceil$ and $\theta \in (0, 1)$ is a constant satisfying $\lceil \theta n \rceil \geq 1/\alpha = \Omega(\log n / \log \log n)$.

Let $G[S]$ be the subgraph of G that is induced by the vertices in the set S. On the RHS of (5), the entropy is evaluated with respect to conditional distributions μ_S^τ, which is the Hard-core model on the subgraph $G[S]$ given the boundary condition τ on $V \setminus S$.

We let $C(S)$ denote the set of connected components in $G[S]$. With a slight abuse of notation, we use $U \in C(S)$ to denote the set of vertices in the component U, as well. It is not hard to see that the Hard-core model μ_S^τ, for $\tau \in \Omega_{V \setminus S}$, factorises as a product distribution over Gibbs marginals at the components $U \in C(S)$, i.e.,

$$\mu_S^\tau = \bigotimes_{U \in C(S)} \mu_U^\tau \quad .$$

We use the following result for the factorisation of entropy on product distributions [Ces01, CMT15, CLV21].

Lemma 3.3 ([CLV21, Lemma 4.1]). *For any $S \subseteq V$, any $\tau \in \Omega_{V \setminus S}$, any $f : \Omega_S^\tau \to \mathbb{R}_{\geq 0}$,*

$$\mathrm{Ent}_S^\tau(f) \leq \sum_{U \in C(S)} \mu_S^\tau[\mathrm{Ent}_U(f)] \quad .$$

Combining Lemma 3.3 and (5) we get that

$$\mathrm{Ent}_\mu(f) \leq \left(\frac{\mathrm{e}}{\theta}\right)^{1+1/\alpha} \mathbb{E}_{\boldsymbol{S} \sim \binom{V}{\ell}} \left[\textstyle\sum_{U \in C(\boldsymbol{S})} \mu\left(\mathrm{Ent}_U(f)\right)\right] \quad , \tag{6}$$

where $\boldsymbol{S} \sim \binom{V}{\ell}$ denotes that $\boldsymbol{S}$ is a uniformly random element from $\binom{V}{\ell}$.

The above step allows us to reduce the proof of approximate tensorisation to that of the components in $C(\boldsymbol{S})$. We choose the parameter $\ell = \lceil \theta n \rceil$ so that the connected components in $C(\boldsymbol{S})$ are typically small.

In light of the above, Theorem 2.4 follows by establishing two results: The first one is to derive a bound on the constant of the approximate tensorisation of entropy for the components of size k in $C(\boldsymbol{S})$, for each $k > 0$. The second result is to derive tail bounds on the size of the components in $C(\boldsymbol{S})$ for $\boldsymbol{S} \sim \binom{V}{\ell}$.

Lemma 3.4. *For any fixed $d > 0$, for any $\lambda < \lambda_c(d)$, consider $\boldsymbol{G} \sim G(n, d/n)$. With probability $1 - o(1)$ over the instances of $\boldsymbol{G}$, the following is true:*

For any $k \geq 1$ and $H \subseteq V$ such that $|H| = k$, the Hard-core model μ_H on $\boldsymbol{G}[H]$ with fugacity λ satisfies the approximate tensorization of entropy with constant

$$\mathrm{AT}(k) \leq \min\left\{2k^2\left(1+\lambda+1/\lambda\right)^{2k+2},\ 3\log\left(1+\lambda+1/\lambda\right)\cdot\left((1+\lambda)k\right)^{2+2\eta}\right\}\quad, \tag{7}$$

where $\eta = B(\log n)^{1/r}$, while $B = B(d, \lambda)$ and $r = r(d) \in (1, 2)$ are constants that depend on d, λ.

As far as size of the components in $C(\boldsymbol{S})$ is concerned, we use the following result from [BGGŠ22].

Lemma 3.5 ([BGGŠ22])**.** *Let $d > 1$ be a constant. There is a constant $L = L(d)$ such that the following holds with probability at least $1 - o(1)$ over the $\boldsymbol{G} \sim G(n, d/n)$. Let $\boldsymbol{S} \sim \binom{V}{\ell}$, while let $C_v \subseteq \boldsymbol{S}$ be the set of vertices that are in the same component as vertex v in $\boldsymbol{G}[\boldsymbol{S}]$. For any integer $k \geq \log n$, it holds that*

$$\Pr[|C_v| = k] \leq (2\mathrm{e})^{\mathrm{e}Lk}\left(\frac{2\ell}{n}\right)^k \leq (2\mathrm{e})^{\mathrm{e}Lk}\left(2\theta\right)^k \quad .$$

Theorem 2.4 follows by combining Theorem 3.2, with Lemmas 3.5 and 3.4. For a full proof of Theorem 2.4, see Section 5.

3.2. **Spectral Independence with Branching Values.** An important component in our analysis is to establish Spectral Independence bounds for the Hard-core model on typical instances of $G(n, d/n)$.

For worst-case graph instances (i.e., non random), typically, we establish Spectral Independence for a region of the parameters of the Gibbs distribution which is expressed in terms of the maximum degree $\varDelta$ of the underlying graph G. As far as $G(n, d/n)$ is concerned, the maximum degree does not seem to be the appropriate graph parameter to consider for this problem.

Here, we utilise the notion of branching value. The notion of the branching value as well as its use for establishing Spectral Independence was introduced in [BGGŠ22]. Unfortunately, the result there were not sufficiently strong to imply rapid mixing of Glauber dynamics. Here we derive stronger results for Spectral independence than those in [BGGŠ22] in the sense that they are more *general* and more *accurate*. Specifically, in our analysis we are able to accommodate vertices of *all degrees*, while we use a more elaborate matrix norm to establish spectral independence, reminiscent of those introduced in [Eft22]. Furthermore, we utilise results from [CFYZ22] that allow us deal with the unbounded degrees of the graph in order to establish our rapid mixing results.

Before getting to further details in our discussion, let us first introduce some basic notions. We start with the *pairwise influence matrix* $\mathcal{I}_G^{\Lambda,\tau}$ and the related notion of Spectral Independence. These notions were first introduced in [ALO20]. In this paper, we use the absolute version introduced in [FGYZ21b].

Consider a *fixed* graph $G = (V, E)$. Assume that we are given a Gibbs distribution μ on the configuration space $\{\pm 1\}^V$. We define the pairwise influence matrix $\mathcal{I}_G^{\Lambda,\tau}$ as follows: for a set of vertices $\Lambda \subset V$ and a configuration τ at Λ, the matrix $\mathcal{I}_G^{\Lambda,\tau}$ is indexed by the vertices in $V \setminus \Lambda$, while for any two vertices, different with each other $v, w \in V \setminus \Lambda$, if w can take both values ± 1 given τ, we have that

$$\mathcal{I}_G^{\Lambda,\tau}(w, u) = ||\mu_u(\cdot \mid (\Lambda, \tau), (\{w\}, +)) - \mu_u(\cdot \mid (\Lambda, \tau), (\{w\}, -))||_{\mathrm{TV}} \quad ; \tag{8}$$

if w can only take one value in ± 1 given τ, we have $\mathcal{I}_G^{\Lambda,\tau}(w, u) = 0$. Also, we have that $\mathcal{I}_G^{\Lambda,\tau}(w, w) = 0$ for all $w \in V \setminus \Lambda$. That is, the diagonal of $\mathcal{I}_G^{\Lambda,\tau}$ is always zero.

Recall that, above, $\mu_u(\cdot \mid (\Lambda, \tau), (\{w\}, 1))$ is the Gibbs marginal that vertex u, conditional that the configuration at Λ is τ and the configuration at w is 1. We have the analogous for $\mu_u(\cdot \mid (\Lambda, \tau), (\{w\}, -1))$.

Definition 3.6 (Spectral Independence)**.** *For a real number $\eta > 0$, the Gibbs distribution μ_G on $G = (V, E)$ is η-spectrally independent, if for every $0 \le k \le |V| - 2$, $\Lambda \subseteq V$ of size k and $\tau \in \{\pm 1\}^{\Lambda}$ the spectral radius of $\mathcal{I}_G^{\Lambda,\tau}$ satisfies that $\rho(\mathcal{I}_G^{\Lambda,\tau}) \le \eta$.*

We bound the spectral radius of $\mathcal{I}_G^{\Lambda,\tau}$ by means of matrix norms. Specifically, we use the following norm of $\mathcal{I}_G^{\Lambda,\tau}$

$$\left|\left| D^{-1} \cdot \mathcal{I}_G^{\Lambda,\tau} \cdot D \right|\right|_\infty , \tag{9}$$

where D is the diagonal matrix indexed by the vertices in $V \setminus \Lambda$ such that

$$D(u,u) = \begin{cases} \deg_G(v)^{1/\chi} & \text{if } \deg_G(v) \ge 1 \\ 1 & \text{if } \deg_G(v) = 0 \end{cases} , \tag{10}$$

where the parameter χ is being specified later.

Let $G = (V, E)$ be a fixed graph. For any vertex $v \in V$ and integer $\ell \ge 0$, we use $N_{v,\ell}$ to denote the number of simple paths with $\ell+1$ vertices that start from v in graph G. By definition, we have that $N_{v,0} = 1$.

Definition 3.7 (d-branching value)**.** *Let $d \ge 1$ be a real number and $G = (V, E)$ be a graph. For any vertex $v \in V$, the d-branching value S_v is defined by $\sum_{\ell \ge 0} N_{v,\ell}/d^{\ell}$.*

We establish spectral independence results that utilise the notion of d-branching value that are similar to the following one.

Theorem 3.8. *Let $d > 1$ be a real number and $G = (V, E)$ be a graph. Let μ_G be the Hard-core model with fugacity $\lambda < \lambda_c(d)$. For any $\alpha > 0$ such that the d-branching value $S_v \le \alpha$ for all $v \in V$ the following is true: μ_G is η-spectrally independent for*

$$\eta \le C_0 \cdot \alpha^{1/r} ,$$

where $C_0 = C_0(d, \lambda)$, while the quantity $r = r(d) \in (1, 2)$ are constants.

Theorem 3.8 is a special case of a stronger result we obtain, i.e., Theorem 6.2. Also, note that Theorem 3.8 is *not* necessarily about $G(n, d/n)$. As a matter of fact in order to use the above result for $G(n, d/n)$ we need to establish bounds on its branching value. To this end, we use the following result from [BGGŠ22].

Lemma 3.9 ([BGGŠ22, Lemma 9])**.** *Let $d \ge 1$. For any fixed $d' > d$, with probability $1 - o(1)$ over $\boldsymbol{G} \sim G(n, d/n)$, the d'-branching factor of every vertex in G is at most $\log n$.*

It is worth mentioning that Lemma 3.9, here, is a *weaker* version of Lemma 9 in [BGGŠ22], i.e., we do not really need the full strength of the result there.

Concluding this short introductory section about Spectral Independence, let us remark that for our results we work with the so-called *Complete Spectral Independence* for the Hard-core model, introduced in [CFYZ21, CFYZ22]. This is more general a notion compared to the (standard) Spectral Independence. For further discussion see Section 4.2.

4. Entropy Factorisation from Stability and Spectral Independence

In this section we establish the ℓ-block factorisation of entropy for the Hard-core model on $G(n, d/n)$ as it is described in Theorem 3.2. To this end, we employ techniques from [CFYZ22]. This means that we study the Hard-core model on $G(n, d/n)$ in terms of the stability of ratios of the marginals and the so-called Complete Spectral Independence.

4.1. **Ratios of Gibbs Marginals & Stability.** Consider the *fixed* graph $G = (V, E)$ and a Gibbs distribution μ on this graph. For a vertex $w \in V$, the region $K \subseteq V \setminus \{w\}$ and $\tau \in \{\pm 1\}^K$, we consider the *ratio of marginals* at w denoted as $R^{K,\tau}(w)$ such that

$$R_G^{K,\tau}(w) = \frac{\mu_w(+1 \mid K, \tau)}{\mu_w(-1 \mid K, \tau)} \ . \tag{11}$$

Recall that $\mu_w(\cdot \mid K, \tau)$ denotes the marginal of the Gibbs distribution $\mu(\cdot \mid K, \tau)$ at vertex w. Also, note that the above allows for $R^{K,\tau}(w) = \infty$, e.g., when $\mu_w(-1 \mid K, \tau) = 0$ and $\mu_w(+1 \mid K, \tau) \neq 0$.

Definition 4.1 (Marginal stability)**.** *Let $\zeta > 0$ be a real number. The Gibbs distribution μ_G on $G = (V, E)$ is called ζ-marginally stable if for any $w \in V$, for any $\Lambda \subset V$, for any configuration σ at Λ and any $S \subseteq \Lambda$ we have that*

$$R_G^{\Lambda,\tau}(w) \leq \zeta \qquad \textit{and} \qquad R_G^{\Lambda,\tau}(w) \leq \zeta \cdot R_G^{S,\tau_S}(w) \ . \tag{12}$$

As far as the stability of the Hard-core marginals at $G(n, d/n)$ is concerned, we prove the following result.

Theorem 4.2 (Stability Hard-Core Model)**.** *For any fixed $d > 0$, for any $\lambda < \lambda_c(d)$, consider $\boldsymbol{G} \sim G(n, d/n)$ and let $\mu_{\boldsymbol{G}}$ be the Hard-core model on $\boldsymbol{G}$ with fugacity λ. With probability $1 - o(1)$ over the instances $\boldsymbol{G}$, $\mu_{\boldsymbol{G}}$ is $2(1+\lambda)^{\frac{2\log n}{\log\log n}}$-marginally stable.*

Proof. Let $\zeta = 2(1+\lambda)^{2\frac{\log n}{\log\log n}}$. Also, let $N(w)$ be the set of the neighbours of w.

For any $\Lambda \subseteq V$ and any $\tau \in \{\pm\}^{\Lambda}$, we have that $\mu_w(+1 \mid \Lambda, \tau) \leq \frac{\lambda}{1+\lambda}$. One can see that the equality holds if $N(w) \subseteq \Lambda$ and for every $u \in N(w)$ we have that $\tau(w) = -1$. Noting that $R_{\boldsymbol{G}}^{\Lambda,\tau}(w)$ is increasing in the value of the Gibbs marginal $\mu_w(+1 \mid \Lambda, \tau)$, it is immediate that

$$\Pr\left[R_{\boldsymbol{G}}^{\Lambda,\tau}(w) \leq \lambda < \zeta \quad \forall \Lambda \subseteq V\right] = 1 \ . \tag{13}$$

It remains to show that

$$\Pr\left[R_{\boldsymbol{G}}^{\Lambda,\tau}(w) \leq \zeta \cdot R_{\boldsymbol{G}}^{S,\tau_S}(w) \quad \forall \Lambda \subset V, \ \forall S \subset \Lambda\right] = 1 - o(1) \ . \tag{14}$$

In light of (13), (14) follows by showing that

$$\Pr\left[R_{\boldsymbol{G}}^{S,\tau_S} > 2\lambda\,(1+\lambda)^{-2\frac{\log n}{\log\log n}} \quad \forall \Lambda \subset V, \ \forall S \subset \Lambda\right] = 1 - o(1) \ . \tag{15}$$

If there is $u \in N(w)$ such that $\tau(u) = +1$, then $R_{\boldsymbol{G}}^{\Lambda,\tau}(w) = 0$ and (14) holds trivially since $R_{\boldsymbol{G}}^{S,\tau_S}(w) \geq 0$. We focus on the case that all vertices $u \in N(w) \cap \Lambda$ satisfy $\tau(u) = -1$.

Let $\mathcal{E}$ be the event that none of the vertices in $N(w)$ is occupied, while let γ_S be the probability of the event $\mathcal{E}$ under the Gibbs distribution $\mu(\cdot \mid S, \tau_S)$. It is standard to show that

$$R_{\boldsymbol{G}}^{S,\tau_S} = \frac{\frac{\lambda}{1+\lambda}\gamma_S}{1 - \frac{\lambda}{1+\lambda}\gamma_S} \ .$$

Noting that the function $f(x) = \frac{x}{1-x}$ is increasing in $x \in (0,1)$, while $\gamma_S \geq (\frac{1}{1+\lambda})^{\deg_{\boldsymbol{G}}(w)}$, we have that

$$R_{\boldsymbol{G}}^{S,\tau_S} \geq \frac{\frac{\lambda}{1+\lambda}(\frac{1}{1+\lambda})^{\deg_{\boldsymbol{G}}(w)}}{1 - \frac{\lambda}{1+\lambda}(\frac{1}{1+\lambda})^{\deg_{\boldsymbol{G}}(w)}} = \frac{\lambda}{(1+\lambda)^{\deg_{\boldsymbol{G}}(w)+1} - \lambda} \ .$$

From the above it is immediate to get (15). Specifically, it follows from the above inequality and Lemma B.1 which implies that for any fixed number $\epsilon > 0$, the maximum degree in $\boldsymbol{G}$ is less than $(1+\epsilon)\frac{\log n}{\log\log n}$ with probability $1 - o(1)$.

This concludes the proof of Theorem 4.2. □

4.2. **(Complete) Spectral Independence.** The notions of the pairwise influence matrix $\mathcal{I}_G^{\Lambda,\tau}$ and the Spectral Independence, as we introduce them in Section 3.2, are typically used to establish bounds on the spectral gap for Glauber dynamics and hence derive bounds on the mixing time of the chain.

The authors in [CLV21], make a further use of Spectral Independence to obtain the approximate tensorisation of entropy. Unfortunately, a vanilla application of their technique is not sufficient to prove our tensorisation results, mainly, because of the unbounded degrees we typically have in $G(n, d/n)$.

In this work, we exploit ideas from [CLV21] together with the related notion of the *Complete Spectral Independence*, in order to establish our factorisation results for the entropy in Theorem 3.2. Specifically, we utilise the connection between complete spectral independence and the ℓ block factorisation of entropy that was established in [CFYZ22] (see further details in the following section).

Since the notions of the pairwise influence matrix $\mathcal{I}_G^{\Lambda,\tau}$ and the Spectral Independence are so important, let us recall them once more, even though they have already been defined in Section 3.2. Consider a *fixed* graph $G = (V, E)$. Assume that we are given a Gibbs distribution μ on the configuration space $\{\pm 1\}^V$.

We define the pairwise influence matrix $\mathcal{I}_G^{\Lambda,\tau}$ as follows: for a set of vertices $\Lambda \subset V$ and a configuration τ at Λ, the matrix $\mathcal{I}_G^{\Lambda,\tau}$ is indexed by the vertices in $V \setminus \Lambda$, while for any two vertices $v, w \in V \setminus \Lambda$, different with each other, if w can take both values ± 1 given τ, we have that

$$\mathcal{I}_G^{\Lambda,\tau}(w, u) = ||\mu_u(\cdot \mid (\Lambda, \tau), (\{w\}, +)) - \mu_u(\cdot \mid (\Lambda, \tau), (\{w\}, -))||_{\mathrm{TV}} \quad ; \tag{16}$$

if w can only take one value in ± 1 given τ, we have $\mathcal{I}_G^{\Lambda,\tau}(w, u) = 0$. Also, we have that $\mathcal{I}_G^{\Lambda,\tau}(w, w) = 0$ for all $w \in V \setminus \Lambda$. That is, the diagonal of $\mathcal{I}_G^{\Lambda,\tau}$ is always zero.

Recall that, above, $\mu_u(\cdot \mid (\Lambda, \tau), (\{w\}, 1))$ is the Gibbs marginal that vertex u, conditional that the configuration at Λ is τ and the configuration at w is 1. We have the analogous for $\mu_u(\cdot \mid (\Lambda, \tau), (\{w\}, -1))$.

Definition 4.3 (Spectral Independence)**.** *For a real number $\eta > 0$, the Gibbs distribution μ_G on $G = (V, E)$ is η-spectrally independent, if for every $0 \le k \le |V| - 2$, $\Lambda \subseteq V$ of size k and $\tau \in \{\pm 1\}^\Lambda$ the spectral radius of $\mathcal{I}_G^{\Lambda,\tau}$ satisfies that $\rho(\mathcal{I}_G^{\Lambda,\tau}) \le \eta$.*

We proceed to introduce the Complete Spectral Independence. First, consider the notion of the Magnetising operation.

Definition 4.4 (Magnetising operation)**.** *Let μ_G be a Gibbs distribution on the graph $G = (V, E)$. For any local fields $\vec{\phi} \in \mathbb{R}_{>0}^V$, the magnetised distribution $\vec{\phi} * \mu$ satisfies*

$$\forall \sigma \in \{\pm 1\}^V, \quad (\vec{\phi} * \mu)(\sigma) \propto \mu(\sigma) \prod_{v \in V : \sigma_v = +1} \phi_v \ .$$

We denote $\vec{\phi} * \mu$ by $\phi * \mu$ if $\vec{\phi}$ is a constant vector with value ϕ.

Suppose that μ is the Hard-core model on G with fugacity λ. It is immediate that the magnetisied distribution $\vec{\phi} * \mu$ can be viewed as the *non-homogenious* Hard-core model such that each vertex v has its own fugacity $\lambda_v = \lambda \cdot \phi_v$.

Definition 4.5 (Complete Spectral Independence)**.** *For two reals $\eta > 0$ and $\xi > 0$, the Gibbs distribution μ_G on $G = (V, E)$ is (η, ξ)-completely spectrally independent, if the magnetised distribution $\vec{\phi} * \mu$ is η-spectrally independent for all $\vec{\phi} \in (0, 1 + \xi]^V$.*

As far as the Hard-core model on the random graph $G(n, d/n)$ is concerned, we prove the following result.

Theorem 4.6. *For any fixed $d > 1$ and $\lambda < \lambda_c(d)$, there exist bounded constants $r = r(d, \lambda) \in (1, 2)$, $B = B(d, \lambda) > 0$ and $s = s(d, \lambda) > 0$ such that the following holds:*

Consider $\boldsymbol{G} \sim G(n, d/n)$ and let $\mu_{\boldsymbol{G}}$ be the Hard-core model on $\boldsymbol{G}$ with fugacity λ. With probability $1 - o(1)$ over the instances of $\boldsymbol{G}$, $\mu_{\boldsymbol{G}}$ is $(B \cdot (\log n)^{1/r}, s)$-completely spectrally independent.

The proof of Theorem 4.6 appears in Section 6.

4.3. **Entropy Block Factorisation - Proof of Theorem 3.2.** The following theorem, from [CFYZ22], allows us to derive a bound on the ℓ- block factorisation parameter of the entropy by using the result in Theorem 4.2 for the stability of Gibbs marginals and the result in Theorem 4.6 for Complete Spectral Independence.

Theorem 4.7 ([CFYZ22, Lemma 2.3])**.** *Let $\eta > 0, \xi > 0$ and $\zeta > 0$ be parameters. Let μ_G be a Gibbs distribution on $G = (V, E)$. If μ_G is (η, ξ)-completely spectrally independent and ζ-marginally stable, then for any $1/\alpha \le \ell < n$, μ_G satisfies the ℓ block factorisation of entropy with parameter $C = (\frac{en}{\ell})^{1+1/\alpha}$, where*

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

Proof of Theorem 3.2. From Theorem 4.6 we have the following: with probability $1-o(1)$ over the instances of $\boldsymbol{G}$ we have that $\mu_{\boldsymbol{G}}$ is (η, s)-completely spectrally independent where $s = s(d, \lambda)$ is *constant*, while

$$\eta = B \cdot (\log n)^{1/r} = o\left(\frac{\log n}{\log\log n}\right),$$

where $B = B(d, \lambda)$ and $r = r(d, \lambda) \in (1, 2)$ are constants specified in the statement of Theorem 4.6. The second equality above follows by noting that $1/r < 1$, bounded away from 1.

Furthermore, from Theorem 4.2 we have the following: With probability $1 - o(1)$ over the instances of $\boldsymbol{G}$, the distribution $\mu_{\boldsymbol{G}}$ is ζ-marginally stable, where

$$\zeta \le 2(1+\lambda)^{2\frac{\log n}{\log\log n}} \ .$$

In light of all the above, the theorem follows by plugging the above values into Theorem 4.7. □

5. Approximate Tensorisation of Entropy

In this section we prove our results related to the approximate tensorisation of the entropy. These are Theorem 2.4 and Lemma 3.4.

5.1. **Proof of Theorem 2.4.** In this section we give the full proof of Theorem 2.4. Recall the high level description of the steps we follow towards this endeavour in Section 3.1.

Proof of Theorem 2.4. From Theorem 3.2 we have the following: For $d > 1$ and $\lambda < \lambda_c(d)$, consider $\boldsymbol{G} \sim G(n, d/n)$, while let $\mu = \mu_{\boldsymbol{G}}$ be the Hard-core model on $\boldsymbol{G}$ with fugacity λ. Let the number $\theta = \theta(d, \lambda)$ in the interval $(0, 1)$ be a parameter whose value is going to be specified later. Then, with probability $1 - o(1)$ over the instances of $\boldsymbol{G}$, for $\ell = \lceil \theta n \rceil$ and for any $f : \Omega \to \mathbb{R}_{>0}$ we have that

$$\mathrm{Ent}_\mu(f) \le \left(\frac{\mathrm{e}}{\theta}\right)^{1+1/\alpha} \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \mu\left(\mathrm{Ent}_S(f)\right). \tag{17}$$

Recall that $C(S)$ denotes the set of connected components in $\boldsymbol{G}[S]$, the subgraph that is induced by vertices in S. With a slight abuse of notation, we use $U \in C(S)$ to denote the set of vertices in the component U. By the conditional independence property of the Gibbs distribution and Lemma 3.3, we have

$$\begin{aligned}
\mathrm{Ent}_\mu(f) &\le \left(\frac{\mathrm{e}}{\theta}\right)^{1+1/\alpha} \frac{1}{\binom{n}{\ell}} \sum_{S\in\binom{V}{\ell}} \sum_{U\in C(S)} \mu\left(\mathrm{Ent}_U(f)\right) \\
\text{(by Lemma 3.4)} \quad &\le \left(\frac{\mathrm{e}}{\theta}\right)^{1+1/\alpha} \frac{1}{\binom{n}{\ell}} \sum_{S\in\binom{V}{\ell}} \sum_{U\in C(S)} \mathrm{AT}(|U|) \sum_{v\in U} \mu[\mathrm{Ent}_v(f)] \\
&\le \left(\frac{\mathrm{e}}{\theta}\right)^{1+1/\alpha} \sum_{v\in V} \mu[\mathrm{Ent}_v(f)] \sum_{k\ge 1} \mathrm{AT}(k)\Pr[|C_v|=k] \ , \qquad (18)
\end{aligned}$$

where C_v is the connected component in $\boldsymbol{G}[S]$, where S is sampled from $\binom{V}{\ell}$ uniformly at random. In order to bound the innermost summation on the R.H.S. of (18) we distinguish two cases for k. For $1 \le k \le \log n$, we use the trivial bound $\Pr[|C_v|=k] \le 1$, while Lemma 3.4 implies that

$$\begin{aligned}
\sum_{k=1}^{\log n} \mathrm{AT}(k)\Pr[|C_v|=k] &\le \sum_{k=1}^{\log n} \mathrm{AT}(k) = \sum_{k=1}^{\log n} 3\log\left(1+\lambda+\lambda^{-1}\right)\cdot((1+\lambda)k)^{2+2\eta} \\
&\le 3\log\left(1+\lambda+\lambda^{-1}\right)\cdot \log n \cdot ((1+\lambda)\log n)^{2+2\eta} \\
&\le 3\log\left(1+\lambda+\lambda^{-1}\right)\cdot ((1+\lambda)\log n)^{3+2\eta} \ ,
\end{aligned}$$

where $\eta = B(\log n)^{1/r}$, for constants $B = B(d,\lambda)$ and $r = r(d) \in (1,2)$. Elementary calculations imply that

$$\sum_{k=1}^{\log n} \mathrm{AT}(k)\Pr[|C_v|=k] \le 3\log\left(1+\lambda+\lambda^{-1}\right)\cdot((1+\lambda)\log n)^{3+2\eta} \le n^x \ , \qquad (19)$$

for $x = o\left(\frac{1}{\log\log n}\right)$.

For $k \ge \log n$, we use the bound in Lemma 3.5 for $\Pr[|C_v|=k]$, while from Lemma 3.4 we have

$$\sum_{k\ge \log n} \mathrm{AT}(k)\Pr[|C_v|=k] \le 2k^2\left(1+\lambda+\lambda^{-1}\right)^{2k+2}(2\mathrm{e})^{\mathrm{e}Lk}(2\theta)^k \ ,$$

where $L = L(d)$ is the parameter in Lemma 3.5. We choose sufficiently small $\theta = \theta(d,\lambda)$ such that

$$\forall k \ge 1, \quad 2k^2\left(1+\lambda+\lambda^{-1}\right)^{2k+2}(2\mathrm{e})^{\mathrm{e}Lk}(2\theta)^k \le (1/2)^k \ .$$

This implies that

$$\sum_{k\ge \log n} \mathrm{AT}(k)\Pr[|C_v|=k] \le \sum_{k\ge\log n}\left(\frac{1}{2}\right)^k \le 1 \ . \qquad (20)$$

Plugging (19), (20) into (18), we get the following: With probability $1-o(1)$ over the instances of $\boldsymbol{G}$ we have that

$$\mathrm{Ent}_\mu(f) \le \left(\frac{\mathrm{e}}{\theta}\right)^{1+1/\alpha}\left(n^{\left(\frac{1}{\log\log n}\right)}+1\right)\sum_{v\in V}\mu[\mathrm{Ent}_v(f)] \ .$$

Since, by Theorem 3.2 we have that $\frac{1}{\alpha} = K(\frac{\log n}{\log\log n})$, for a constant $K = K(d,\lambda)$, and $\theta = \theta(d,\lambda)$ is also a constant, the above inequality can be written as follows: there is a constant $A = A(d,\lambda)$ such that

$$\mathrm{Ent}_\mu(f) \le n^{\left(\frac{A}{\log\log n}\right)}\sum_{v\in V}\mu[\mathrm{Ent}_v(f)] \ .$$

The above concludes the proof of Theorem 2.4. □

5.2. **Proof of Lemma 3.4.** Lemma 3.4 follows as a corollary from the following result.

Lemma 5.1. *Let $\lambda > 0$, let the graph $G = (V, E)$, while let μ_G be the Hard-core model on G with fugacity λ. Let $M \subseteq V$ be a subset of vertices and $\sigma \in \{\pm 1\}^{V \setminus M}$ on $V \setminus M$. The conditional distribution $\mu_M(\cdot \mid V \setminus M, \sigma)$ satisfies the approximate tensorisation of entropy with constant*

$$C = 2|M|^2 \left(1 + \lambda + 1/\lambda\right)^{2|M|+2} \quad . \tag{21}$$

Furthermore, if μ is η-spectrally independent, for some number η, then $\mu_M(\cdot \mid V \setminus M, \sigma)$ satisfies the approximate tensorisation of entropy with constant

$$C = 3 \log\left(1 + \lambda + 1/\lambda\right) \cdot \left((1+\lambda)|M|\right)^{2+2\eta} \quad . \tag{22}$$

In light of Lemma 5.1 the proof of Lemma 3.4 is straightforward. In what follows, we provide its proof for the sake of completeness.

Proof of Lemma 3.4. The first bound in (7) follows directly from the first part of Lemma 5.1. We now prove the second bound in (7).

By Theorem 4.6, with probability $1 - o(1)$ we have that $\mu_{\boldsymbol{G}}$ is $(B(\log n)^{1/r}, s)$-completely spectrally independent, where $r = r(d, \lambda) \in (1, 2)$, $B = B(d, \lambda) > 0$ and $s = s(d, \lambda) > 0$. As a consequence, $\mu_{\boldsymbol{G}}$ is $B(\log n)^{1/r}$-spectrally independent with probability $1 - o(1)$. Hence, the second bound in (7) follows from the second part of Lemma 5.1. □

Proof of Lemma 5.1. Note that this result is for a fixed graph G. Also, since we talk about tensorisation of entropy, w.l.o.g. assume that $|M| \geq 2$.

For brevity, we use π to denote the distribution $\mu_M(\cdot \mid V \setminus M, \sigma)$. We use $\Omega \subseteq \{\pm 1\}^M$ to denote the support of the distribution π. Let $P : \Omega \times \Omega \to \mathbb{R}_{\geq 0}$ denote the transition matrix of the Glauber dynamics on π. It is elementary to verify that the Glauber dynamics on Ω is ergodic.

The Poincaré inequality for π is that

$$\forall f : \Omega \to \mathbb{R}_{\geq 0} : \quad \gamma \mathrm{Var}_\pi[f] \leq \frac{1}{|M|} \sum_{v \in \Lambda} \pi[\mathrm{Var}_v(f)] \quad ,$$

where γ is the *Poincaré constant*, a.k.a. spectral gap of P. The log-Sobolev inequality for π is that

$$\forall f : \Omega \to \mathbb{R}_{\geq 0} : \quad \alpha \mathrm{Ent}_\pi[f] \leq \frac{1}{|M|} \sum_{v \in \Lambda} \pi[\mathrm{Var}_v(\sqrt{f})] \quad ,$$

where α is the *log-Sobolev constant*.

The following well-known inequality from [CMT15, Proposition 1.1] relates C, the approximate tensorisation constant we want to bound, and the log-Sobolev constant α. We have that

$$C \leq \frac{1}{\alpha |M|} \quad .$$

From [DSC96, Corollary A.4] we have the following relation between γ and α

$$\alpha \geq \frac{1 - 2\pi_{\min}}{\log(1/\pi_{\min} - 1)} \cdot \gamma, \qquad \text{where } \pi_{\min} = \min_{\sigma \in \Omega} \pi(\sigma) \quad .$$

We may assume $|\Omega| \geq 3$, as $|\Omega| \leq 2$ is the trivial case in which $C \leq 1$. It holds that $\pi_{\min} \leq \frac{1}{3}$. Combining the above two inequalities together, we have

$$C \leq \frac{3 \log(1/\pi_{\min})}{|M| \gamma} \quad . \tag{23}$$

Next, we use Cheeger's inequality to derive a crude lower bound on the spectral gap γ. Recall that Cheeger's inequality implies that

$$\gamma \geq \Phi^2/2 \quad , \tag{24}$$

where

$$\Phi = \min_{\Omega_0 \subseteq \Omega : \pi(\Omega_0) \le \frac{1}{2}} \frac{1}{\pi(\Omega_0)} \sum_{x \in \Omega_0} \sum_{y \in \Omega \setminus \Omega_0} \pi(x) P(x, y) \ ,$$

and (as defined above) P is the transition matrix of the Glauber dynamics.

Since the Glauber dynamics is ergodic, for any $\Omega_0 \subset \Omega$, there exists $x \in \Omega_0$ and $y \in \Omega \setminus \Omega_0$ such that $P(x, y) > 0$. By the definition of the Glauber dynamics, such x and y satisfy $P(x, y) \ge \frac{1}{|\Lambda|} \min\{\frac{1}{1+\lambda}, \frac{\lambda}{1+\lambda}\}$. Hence, we have that

$$\Phi \ge \frac{2\pi_{\min}}{|M|} \min\left\{\frac{1}{1+\lambda}, \frac{\lambda}{1+\lambda}\right\} \ . \tag{25}$$

Plugging (24) and (25) into (23), we get that

$$C \le \frac{3}{2} \frac{|M| \log(1/\pi_{\min})}{\left(\pi_{\min} \cdot \min\left\{\frac{1}{1+\lambda}, \frac{\lambda}{1+\lambda}\right\}\right)^2} \ .$$

Finally, for the Hard-core model, the quantity $\pi_{\min}$ can be lower bounded as

$$\pi_{\min} \ge \frac{\min\{1, \lambda^{|M|}\}}{(1+\lambda)^{|M|}} \ , \tag{26}$$

which implies

$$C \le 2|M|^2 \left(1 + \lambda + 1/\lambda\right)^{2|M|+2} \ .$$

The above proves (21).

We proceed with the proof of (22). Recall that, now, we further assume that μ is η-spectrally independent. From the definition of spectral independence, it is straightforward to see that π is also spectrally independent. Hence, for any $0 \le k \le |M| - 2$, any $S \subseteq M$ with $|S| = k$ and any feasible pinning $\sigma \in \{\pm 1\}^S$, the spectral radius of the influence matrix $\mathcal{I}_\pi^{S,\sigma}$ satisfies

$$\rho(\mathcal{I}_\pi^{S,\sigma}) \le \eta \ .$$

Furthermore, given any condition σ on S, for any $v \notin S$, it holds that $\pi_v^{S,\sigma}(+1) \le \frac{\lambda}{1+\lambda}$. This implies that for any $u, v \notin S$, it holds that

$$\mathcal{I}_\pi^{S,\sigma}(u, v) \le \frac{\lambda}{1+\lambda} \ ,$$

which it turn implies that

$$\frac{\rho(\mathcal{I}_\pi^{S,\sigma})}{|M| - k - 1} \le \frac{\lambda}{1+\lambda} \ .$$

By [FGYZ21b, Theorem 3.2], the spectral gap γ can be lower bounded by

$$\gamma \ge \left(\frac{1}{(1+\lambda)|M|}\right)^{2+2\eta} \ .$$

For the sake of completeness Theorem 3.2 from [FGYZ21b] can be found in the Appendix as Lemma A.1.

Plugging the above into (23), we get that

$$C \le \frac{3\log(1/\pi_{\min})}{|M|} \cdot ((1+\lambda)|M|)^{2+2\eta} \ .$$

Finally, by the lower bound in (26), we have

$$C \le 3\log\left(1 + \lambda + 1/\lambda\right) \cdot ((1+\lambda)|M|)^{2+2\eta} \ .$$

The above concludes the proof of Lemma 3.4. □

6. Proof of Theorem 4.6 - Complete Spectral Independence

In this section we establish that on typical instances of $\boldsymbol{G} \sim G(n, d/n)$ the Hard-core model μ with fugacity $\lambda < \lambda_c(d)$ exhibits complete spectral independence in the way that is specified in Theorem 4.6.

As a first step we establish spectral independence bounds for the Hard-core model on a *fixed* graph of a given d-branching value (see Definition 3.7), for some $d > 1$. Critically, the fugacity λ is upper bounded by $\lambda_c(d)$.

We introduce the non-homogenious Hard-core model, i.e., every vertex u has its own fugacity λ_u. Specifically, for $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_v)_{v\in V}$ such that $\boldsymbol{\lambda}_v \in \mathbb{R}_{>0}$, for all $v \in V$, we define the non-homogenious Hard-core mode $\mu_{G,\boldsymbol{\lambda}}$ such that for any $\sigma \in \{\pm\}^V$ we have that

$$\mu_{G,\boldsymbol{\lambda}}(\sigma) \propto \prod\nolimits_{v\in V:\sigma_v=+} \boldsymbol{\lambda}_v \quad . \tag{27}$$

For the sake brevity, in what follows, we let $||\boldsymbol{\lambda}||_\infty$ denote the maximum value over $\{\boldsymbol{\lambda}_v\}_{v\in V}$.

Note that we employ the *potential method* in order to establish our spectral independence. Specifically, we use the following lemma from [SSŠY17].

Lemma 6.1 ([SSŠY17]). *Let $d > 1$ and $\lambda > 0$ be parameters satisfying $\lambda < \lambda_c(d)$. Let $\chi = \chi(d) \in (1, 2)$ be a parameter defined as $\chi = (1 - \frac{d-1}{2}\log(1 + \frac{1}{d-1}))^{-1}$. Let $a = \frac{\chi}{\chi-1}$ and $\Phi(x) = \frac{1}{\sqrt{x(x+1)}}$. There exists $0 < \kappa = \kappa(\lambda) < 1/d$ such that the following holds for any integer $k \geq 1$: for any $x_1, x_2, \ldots, x_k \geq 0$ and $x = \lambda\prod_{i=1}^{k}\frac{1}{1+x_i}$, it holds that $\Phi(x)^a\sum_{i=1}^{k}(\frac{x}{(1+x_i)\Phi(x_i)})^a \leq \kappa^{a/\chi}$.*

We derive the following general result for spectral independence for the non-homogenous Hard-core model on a fixed graph G it terms of its d-branching value.

Theorem 6.2. *Let $d > 1$ be a real number. For the graph $G = (V, E)$, let $\mu_{G,\boldsymbol{\lambda}}$ be the non-homogenious Hard-core model such that $||\boldsymbol{\lambda}||_\infty < \lambda_c(d)$.*

For any $\alpha > 0$ such that the d-branching value $S_v \leq \alpha$ for all $v \in V$ the following is true: $\mu_{G,\boldsymbol{\lambda}}$ is η-spectrally independent for

$$\eta \leq W \cdot \alpha^{1/\chi} \quad ,$$

where $W = W(d, ||\boldsymbol{\lambda}||_\infty)$, while the quantity $\chi = \chi(d) \in (1, 2)$ is from Lemma 6.1.

Note that Theorem 6.2 is about spectral independence, of the non-homogenious Hard-core model, while Theorem 4.6 is about complete spectral independence of the (homogenious) Hard-core model.

We prove Theorem 4.6 from Theorem 6.2 and Lemma 3.9.

Proof of Theorem 4.6. Recall that for Theorem 4.6 we assume that $\lambda < \lambda_c(d)$. This implies that we can choose $\tilde{d}$ such that $\tilde{d} > d$, while

$$\lambda < \lambda_c(\tilde{d}) \quad .$$

Note that the inequality is strict. The above follows by noting that the function $\lambda_c(x)$ is continuous and strictly decreasing for $x > 1$. Also, choose $\tilde{\lambda}$ such that

$$\lambda < \tilde{\lambda} < \lambda_c(\tilde{d}) \quad .$$

Note that both quantities $\tilde{\lambda}$ and $\tilde{d}$ depend only on d, λ, i.e., we have that $\tilde{d} = \tilde{d}(d, \lambda)$ and $\tilde{\lambda} = \tilde{\lambda}(d, \lambda)$.

We prove the theorem by showing that with probability $1 - o(1)$ over the instances of $\boldsymbol{G}$ we have the following: the Hard-core model on $\boldsymbol{G}$ with fugacity λ, denoted as $\mu_{\boldsymbol{G}}$, is (η, s)-completely spectrally independent, where

$$\eta = B \cdot (\log n)^{\frac{1}{r}} \qquad \text{and} \qquad s = \tilde{\lambda}/\lambda - 1 \quad , \tag{28}$$

for constants $B = B(d, \lambda) > 0$ and $r = r(d) \in (1, 2)$.

Consider first the fixed graph $G=(V,E)$ and external fields $\vec{\phi} \in (0, 1+s]^V$. It is straightforward that the distribution $\vec{\phi} * \mu_G$ corresponds to the non-homogenious Hard-core model $\mu_{G,\boldsymbol{\lambda}}$ such that $||\boldsymbol{\lambda}||_\infty \leq \tilde{\lambda} < \lambda_c(\tilde{d})$.

Theorem 6.2 implies the following for $\mu_{G,\boldsymbol{\lambda}}$: for any $\alpha > 0$ such that the $\tilde{d}$-branching value $S_v \leq \alpha$ for all $v \in V$, the distribution $\mu_{G,\boldsymbol{\lambda}}$ is ζ-spectral independent where

$$\zeta \leq W \cdot \alpha^{1/\chi} \ , \tag{29}$$

for $W = W(\tilde{d}, ||\boldsymbol{\lambda}||_\infty)$ specified in Theorem 6.2, while the quantity $\chi = \chi(\tilde{d}) \in (1,2)$ is from Lemma 6.1.

The above imply that μ_G, the (homogenous) Hard-core model on G with fugacity λ is $(B \cdot \alpha^{1/\chi}, s)$-completely spectrally independent.

Note that since W depends on $\tilde{d}$ and $\tilde{\lambda} = ||\boldsymbol{\lambda}||_\infty$, which in turn depend on d, λ, we have that W depends only on d and λ. With the same argument, the parameter χ, above, depends only on d and λ.

As far as $\boldsymbol{G} \sim G(n, d/n)$ is concerned, we work as follows: since, $\tilde{d} > d$, Lemma 3.9 implies that with probability $1 - o(1)$ over the instances of $\boldsymbol{G}$ the $\tilde{d}$-branching value S_v, for all $v \in V$, satisfies $S_v \leq \log n$. In light of this observation, the theorem follows from (29) by having $B = W(\tilde{d}, ||\boldsymbol{\lambda}||_\infty)$ and $r = \chi(\tilde{d})$. ☐

6.1. **A Bound on the Eigenvalue via Weighted Total Influence - Proof of Theorem 6.2.** We use the following result to prove Theorem 6.2.

Theorem 6.3. *Let $d > 1$ be a real number and $G = (V, E)$ be a graph. For $\boldsymbol{\lambda} \in \mathbb{R}^V_{>0}$, let $\mu_{G,\boldsymbol{\lambda}}$ be the non-homogenious Hard-core model on G, while assume that $||\boldsymbol{\lambda}||_\infty < \lambda_c(d)$. Also, let $\mathcal{I}_G$ be the influence matrix induced by $\mu_{G,\boldsymbol{\lambda}}$.*

For any $\alpha > 0$ such that the d-branching value $S_v \leq \alpha$ for all $v \in V$ the following is true: There exists a constant $W = W(d, ||\boldsymbol{\lambda}||_\infty)$ such that

$$\forall r \in V, \quad \sum_{u \in V} \mathcal{I}_G(r,u) \cdot \deg_G(u)^{1/\chi} \leq W \cdot (\alpha \cdot \deg_G(r))^{1/\chi} \quad ,$$

while $\chi = \chi(d) \in (1,2)$ is from Lemma 6.1.

Theorem 6.3 is proved in Section 7. Now, we are ready to prove Theorem 6.2.

Proof of Theorem 6.2. To prove Theorem 6.2 we focus on the spectral radius of $\mathcal{I}^{\Lambda,\sigma}_G$, i.e., $\rho(\mathcal{I}^{\Lambda,\sigma}_G)$, and show that for any choice of $\Lambda \subset V$ and $\sigma \in \{\pm 1\}^\Lambda$ we have

$$\rho(\mathcal{I}^{\Lambda,\sigma}_G) \leq W \cdot \alpha^{1/\chi} \ , \tag{30}$$

Theorem 6.2 follows immediately once we show the above.

Before proving (30), let us make some useful observations. Suppose that we have the non-homogenous Hard-core model with fugacities $\boldsymbol{\lambda} \in \mathbb{R}^V_{>0}$ on the graph $G = (V, E)$, while at the set of vertices Λ we have the configuration τ. Then, it is elementary to verify that this distribution is identical to the non-homogenous Hard-core model on the graph $G' = (V', E')$ with fugacities $(\boldsymbol{\lambda}_v)_{v \in V'}$, such that G' is obtained from G by working as follows: we remove from G every vertex w which either belongs to Λ, or has a neighbour $u \in \Lambda$ such that $\tau(u) = +1$, i.e., u is "occupied" under τ.

Additionally, consider the matrices $\mathcal{I}^{\Lambda,\sigma}_G$ and $\mathcal{I}_{G'}$ induced by the aforementioned Gibbs distributions, respectively. It is not hard to see that $\mathcal{I}_{G'}$ is a principal submatrix of $\mathcal{I}^{\Lambda,\sigma}_G$ obtained by removing columns and rows that correspond to vertices $w \in V \setminus \Lambda$ that have a neighbour $u \in \Lambda$ such that $\tau(u) = +1$. Note that the rows and columns we remove from $\mathcal{I}^{\Lambda,\sigma}_G$ in order to obtain $\mathcal{I}_{G'}$ consist of entries which are zero.

Using the above observation, it is an easy exercise in linear algebra to verify that any additional eigenvalues that $\mathcal{I}^{\Lambda,\tau}_G$ might have, compared to $\mathcal{I}_{G'}$, these can only be equal to zero. Hence, we derive the following relation for the spectral radii of the two matrices:

$$\rho(\mathcal{I}^{\Lambda,\tau}_G) = \rho(\mathcal{I}_{G'}) \ .$$

Furthermore, note that the branching value is non-increasing when removing vertices. Hence, if we have $G, \boldsymbol{\lambda}, d$ and α that satisfy the conditions in Theorem 6.2, then $G', (\boldsymbol{\lambda}_v)_{v\in V'}, d$ and α satisfy the same conditions, as well.

In light of all the above and without loss of generality, we can ignore the pinning and consider the influence matrix $\mathcal{I}_G$.

Hence, instead of proving (30) we consider the following equivalent problem. Consider the graph $G = (V, E)$, while we have $\alpha > 0$ and $d > 1$ such that the d-branching value $S_v \le \alpha$ for all $v \in V$. Also, for $\boldsymbol{\lambda} \in \mathbb{R}^V_{>0}$ such that $||\boldsymbol{\lambda}||_\infty < \lambda_c(d)$ consider $\mu = \mu_G$ the non-homogenous Hard-core model on G with fugacities $\boldsymbol{\lambda}$, while let $\mathcal{I}_G$ be the corresponding pairwise influence matrix (without boundary conditions).

It suffices to show that

$$\rho(\mathcal{I}_G) \le W\alpha^{1/\chi} \ , \tag{31}$$

for W, α, χ specified in the statement of Theorem 6.2.

To this end, we use the matrix norm introduced in Section 3.2. Specifically, it is standard that

$$\rho(\mathcal{I}_G) \le \left|\left|D^{-1}\cdot \mathcal{I}_G \cdot D\right|\right|_\infty \ , \tag{32}$$

where D is the diagonal matrix indexed by the vertices in $V \setminus \Lambda$ such that

$$D(u,u) = \begin{cases} \mathrm{deg}_G(v)^{1/\chi} & \text{if } \mathrm{deg}_G(v) \ge 1 \\ 1 & \text{if } \mathrm{deg}_G(v) = 0 \ , \end{cases}$$

while $\chi = \chi(d) \in (1, 2)$ is from Lemma 6.1.

Noting that $\mathcal{I}_G, D$ are non-negative matrices, from the definition of the matrix norm $||\cdot||_\infty$, we have that

$$\begin{aligned} \left|\left|D^{-1}\cdot \mathcal{I}_G \cdot D\right|\right|_\infty &= \max_{v\in V} \sum_{u\in V} \frac{\mathcal{I}_G(v,u)\cdot D(u,u)}{D(v,v)} \\ &= \max_{\substack{v\in V: \\ \mathrm{deg}_G(v)>0}} \sum_{u\in V} \frac{\mathcal{I}_G(v,u)\cdot \mathrm{deg}_G(u)^{1/\chi}}{\mathrm{deg}_G(v)^{1/\chi}} \\ &\le W\alpha^{1/\chi} \ . \end{aligned}$$

The second equality follows from the observation that for the isolated vertices v, i.e., $\mathrm{deg}_G(v) = 0$, we have that $\mathcal{I}_G(v, u) = 0$ for all $u \in V$. The last inequality is due to Theorem 6.3. Hence, (31) follows by plugging the above into (32).

The theorem follows. □

7. Bound the total influence via self-avoiding walk tree - Proof of Theorem 6.3

Let $G = (V, E)$ be a graph, while assume there is a total order for the vertices in V.

A self-avoiding walk (SAW) in G is a path $v_1, v_2, \ldots, v_\ell$ in G such that $v_i \ne v_j$ for all $i \ne j$. Fix a vertex $r \in V$. We define the SAW-tree $T_{\mathrm{SAW}}(r)$, the *tree of self-avoiding walks*, starting from r, as follows: Consider the set consisting of every walk $v_0, \ldots, v_\ell$ in the graph G that emanates from vertex r, i.e., $v_0 = r$, while one of the following two holds

K.1: $v_0, \ldots, v_\ell$ is a self-avoiding walk,
K.2: $v_0, \ldots, v_{\ell-1}$ is a self-avoiding walk, while there is $j \le \ell - 3$ such that $v_\ell = v_j$.

Each one of the walks in the set corresponds to a vertex in $T_{\mathrm{SAW}}(r)$. Two vertices in $T_{\mathrm{SAW}}(r)$ are adjacent if the corresponding walks are adjacent. Note that two walks in the graph G are considered to be adjacent if one extends the other by one vertex [‡].

[‡]E.g. the walks $P' = w_0, w_1, \ldots, w_\ell$ and $P = w_0, w_1, \ldots, w_\ell, w_{\ell+1}$ are adjacent with each other.

We also use the following terminology: for vertex u in $T_{\rm SAW}(r)$ that corresponds to the walk $v_0, \dots, v_\ell$ in G we say that "u is a *copy* of vertex v_ℓ in $T_{\rm SAW}(r)$". For every vertex v in the graph G, we use C_v to denote the set of copies of v in the SAW-tree $T_{\rm SAW}(r)$.

For $\boldsymbol{\lambda} \in \mathbb{R}^V_{>0}$, consider the non-homogenious Hard-core model $\mu_{G,\boldsymbol{\lambda}}$ on the graph G. We specify the non-homogenious Hard-core model $\mu_{T,\boldsymbol{\lambda}}$ on $T_{\rm SAW}(r)$ such that all the vertices in C_v have fugacity $\boldsymbol{\lambda}_v$, for all $v \in V$.

We use Λ to denote the set of cycle-closing vertices in SAW-tree $T_{\rm SAW}(r)$, i.e., those that correspond to the paths of the kind **K.2**. Let $\sigma \in \{\pm 1\}^\Lambda$ denote the pinning induced by the SAW-tree obtained by working as follows: for $z \in \Lambda$ that corresponds to the path $w_0, \dots w_\ell$ we set $\sigma(z)$ such that

(a) -1 if $w_\ell > w_{\ell-1}$,

(b) $+1$ otherwise.

Λ is a subset of the leaves of $T_{\rm SAW}(r)$, and hence, σ is a pinning of a subset of the leaves. Note that, potentially, there are leaves in T which do not belong to Λ. These are copies of vertices in G which are of degree 1.

The above construction gives rise to the conditional Hard-core distribution $\mu^{\Lambda,\sigma}_{T,\boldsymbol{\lambda}}$ on the SAW-tree $T = T_{\rm SAW}(r)$. Let $\mathcal{I}^{\Lambda,\sigma}_T$ denote the influence matrix induced by $\mu^{\Lambda,\sigma}_{T,\boldsymbol{\lambda}}$.

Recall that $\mathcal{I}_G$ corresponds to the the influence matrix induced by $\mu_{G,\boldsymbol{\lambda}}$. We have the following result that relates the influence matrices $\mathcal{I}_G$ and $\mathcal{I}^{\Lambda,\sigma}_T$.

Lemma 7.1 ([CLV20, Lemma 8]). *For every vertex $v \neq r$ in G, it holds that*

$$\mathcal{I}_G(r, v) = \sum_{u \in C_v} \mathcal{I}^{\Lambda,\sigma}_T(r, u) \ .$$

The above result is very useful in that it allows us to study the influence matrix $\mathcal{I}_G$ by means of the matrix $\mathcal{I}^{\Lambda,\sigma}_T$ which is much simpler to analyse due to the tree underlying structure. In light of the above, we also use the following result from [ALO20].

Lemma 7.2 ([ALO20, Lemma B.2]). *Consider the tree $T = (V_T, E_T)$ and and let μ be a Gibbs distribution on $\{\pm 1\}^V$. For any three vertices $u, v, w \in V_T$ such that u is on the path from v to w, for any $M \subseteq V \setminus \{u, v, w\}$ and any $\tau \in \{\pm 1\}^M$ we have that*

$$\mathcal{I}^{M,\tau}(v, w) = \mathcal{I}^{M,\tau}(v, u) \cdot \mathcal{I}^{M,\tau}(u, w) \ .$$

Note that in [ALO20] the influence matrix is defined in a slightly different way than what we have here. Specifically, the matrix $\mathcal{I}_G$ we define here can be obtained from the influence matrix in [ALO20] by taking the absolute value of its entries. Even though Lemma 7.2 was proved for the influence matrix in [ALO20], it is straightforward that it also holds for the influence matrix we define here.

An observation that we use is that for all vertices $z \notin \Lambda$ in $T_{\rm SAW}(r)$ have the following property: suppose that $z \in C_w$ for $w \in V$, then we have that

$$\deg_T(z) = \deg_G(w) \ . \tag{33}$$

In light of Lemmas 7.1 and 7.2, Theorem 6.3 follows by bounding the total influence on the tree $T_{\rm SAW}(r)$ from the root. The following is the main technical result in this section.

Proposition 7.3. *Let $d > 1$ be a real number and $T = (V_T, E_T)$ be a tree rooted at $r \in V$. Let $\boldsymbol{\lambda} \in \mathbb{R}^V_{>0}$ such that $||\boldsymbol{\lambda}||_\infty < \lambda_c(d)$. Let $\mu_{T,\boldsymbol{\lambda}}$ be the non-homogenius Hard-core model on T with fugacity $\boldsymbol{\lambda}$. For any $\alpha > 0$ such that the d-branching value $S_r \leq \alpha$, the following is true:*

There exists a constant $D = D(d, ||\boldsymbol{\lambda}||_\infty)$ such that for any Λ subset of the leaves of T, for any $\sigma \in \{\pm 1\}^\Lambda$ we have that

$$\sum_{u \in V \setminus \{r\}} \mathcal{I}^{\Lambda,\sigma}_T(r, u) \cdot \deg_T(u)^{1/\chi} \leq D \cdot (\alpha \cdot \deg_T(r))^{1/\chi} \ \ ,$$

where $\chi = \chi(d) \in (1,2)$ is defined in Lemma 6.1.

Proof of Theorem 6.3. Given graph G, fix a vertex $r \in V$. Our focus is on bounding the weighted sum $\sum_{u \in V} |\mathcal{I}_G(r,u)| \deg_G(u)^{1/\chi}$.

We construct the SAW-tree $T = T_{\mathrm{SAW}}(r)$ together with the pinning σ on a subset of leaf vertices Λ, as we describe at the beginning of Section 7. We have the non-homogenous Hard-core model on T such that every vertex w, a copy of $v \in V$ in T, has fugacity $\boldsymbol{\lambda}_v$. Let $\mathcal{I}_T^{\Lambda,\sigma}$ denote the influence matrix that corresponds to this distribution.

Lemma 7.1 implies that for any $u \in V$ we have

$$\begin{aligned}
\mathcal{I}_G(r,u) \cdot \deg_G(u)^{1/\chi} &= \deg_G(u)^{1/\chi} \cdot \sum_{v \in C_u} \mathcal{I}_T^{\Lambda,\sigma}(r,v) \\
&= \deg_G(u)^{1/\chi} \cdot \sum_{v \in C_u \setminus \Lambda} \mathcal{I}_T^{\Lambda,\sigma}(r,v) \qquad (34)\\
\text{(by (33))} &= \sum_{v \in C_u \setminus \Lambda} \mathcal{I}_T^{\Lambda,\sigma}(r,v) \deg_T(v)^{1/\chi} \\
&= \sum_{v \in C_u} \mathcal{I}_T^{\Lambda,\sigma}(r,v) \deg_T(v)^{1/\chi} \ . \qquad (35)
\end{aligned}$$

In both (34) and (35) we use the observation that for all $w \in C_v \cap \Lambda$, we have $\mathcal{I}_T^{\Lambda,\sigma}(r,w) = 0$, i.e., since the assignment of w is fixed to $\sigma(w)$, the root has zero influence on w. Hence, we conclude that

$$\sum_{u \in V} \mathcal{I}_G(r,u) \cdot \deg_G(u)^{1/\chi} = \sum_{u \in V} \sum_{v \in C_u} \mathcal{I}_T^{\Lambda,\sigma}(r,v) \deg_T(v)^{1/\chi} \ . \qquad (36)$$

We let $\widetilde{d}$ be the solution to the equation $\lambda_c(\widetilde{d}) = \frac{1}{2}(\lambda + \lambda_c(d))$. Note that $\widetilde{d}$ depends only on d and λ. Also, we have that

$$(a) \quad \lambda < \lambda_c(\widetilde{d}) < \lambda_c(d) \qquad \text{and} \qquad (b) \quad \widetilde{d} > d \ .$$

The inequality in (a) follows from the fact that $\lambda_c(\widetilde{d})$ is the average of $\lambda, \lambda_c(d)$ and $\lambda < \lambda_c(d)$. Also, (b) follows from the observation that $\lambda_c(d) > \lambda_c(\widetilde{d})$ and that $\lambda_c(x)$ is monotonically decreasing in $x > 1$.

Let ϕ_r be the $\widetilde{d}$-branching value of r in SAW-tree $T = T_{\mathrm{SAW}}(G,r)$. Let ψ_r be the d-branching value of vertex r in graph G. Note that there is a unique vertex in G that corresponds to the root of the SAW-tree.

We claim that there exists a constant $K = K(d,\lambda) > 1$ such that

$$\phi_r \le K \cdot \psi_r \le K \cdot \alpha, \qquad (37)$$

where α satisfying $\alpha \ge \psi_r$ is specified in the statement of Theorem 6.3.

Before showing that (37) is true, let us show how we can use it to prove Theorem 6.3. Using Proposition 7.3 with $\widetilde{d}$-branching values (note that $\lambda < \lambda_c(\widetilde{d})$), we have the following: there exists constants $D = D(\widetilde{d},\lambda)$ and $\chi = \chi(\widetilde{d})$ such that

$$\begin{aligned}
\sum_{u \in V} \mathcal{I}_T^{(\Lambda,\sigma)}(r,u) \cdot \deg_T(u)^{1/\chi} &\le D\left(\phi_r \cdot \deg_T(r)\right)^{1/\chi} \\
\text{(by (37))} \quad &\le D \cdot \left(K \cdot \alpha \cdot \deg_T(r)\right)^{1/\chi} \ .
\end{aligned}$$

In light of the above, Theorem 6.3 follows by plugging the above inequality into (36) and setting $W = D \cdot K^{1/\chi}$. Note that W is a constant that depends only on d and λ.

We conclude the proof of Theorem 6.3 by showing that (37) is true. Let $N^T_{r,\ell}$ be the number of simple paths in T of length ℓ, starting from the root r. Fix such a path $P = (v_0 = r, v_1, v_2, \ldots, v_\ell)$ in the tree T. It follows from the definition of the SAW-tree T that P corresponds to one of the following two types of paths in G.

Type 1: P is a simple path of length ℓ starting from r in graph G;

Type 2: the prefix $v_0, v_1, \dots, v_{\ell-1}$ is a simple path length $\ell - 1$ starting from r in graph G and v_ℓ is a cycle-closing vertex such that $v_\ell = v_i$ for some $i \le \ell - 3$.

Let $N^G_{r,\ell}$ be the number of length ℓ paths in G that start from r and are of Type 1. Let $\bar{N}^G_{r,\ell}$ be the number of length ℓ paths in G that start from r and are of Type 2. We have the following relation

$$\forall \ell \ge 1, \quad N^T_{r,\ell} = N^G_{r,\ell} + \bar{N}^G_{r,\ell} \;\le\; N^G_{r,\ell} + \ell N^G_{r,\ell-1} \ . \tag{38}$$

For the second inequality, we use the observation that $\bar{N}^G_{r,\ell} \le \ell N^G_{r,\ell-1}$.

Recall that ϕ_r is the $\widetilde{d}$-branching value of r in SAW-tree $T = T_{\text{SAW}}(G, r)$. We have that

$$\begin{aligned} \phi_r &= \sum_{\ell \ge 0} \frac{N^T_{r,\ell}}{\widetilde{d}^\ell} = 1 + \sum_{\ell \ge 1} \frac{N^T_{r,\ell}}{\widetilde{d}^\ell} \\ \text{(by (38))} \quad &\le 1 + \sum_{\ell \ge 1} \frac{N^G_{r,\ell}}{\widetilde{d}^\ell} + \sum_{\ell \ge 0} \frac{(\ell+1) N^G_{r,\ell}}{\widetilde{d}^{\ell+1}} \\ &= \sum_{\ell \ge 0} \frac{N^G_{r,\ell}}{\widetilde{d}^\ell} \left(1 + \frac{\ell+1}{\widetilde{d}}\right) \ . \end{aligned}$$

Since $\widetilde{d} > d > 1$ and $\widetilde{d}$ is determined by d and λ, there exists a constant $K = K(d, \lambda) \ge 1$ such that

$$\forall \ell \ge 0, \quad \frac{1}{\widetilde{d}^\ell} \left(1 + \frac{\ell+1}{\widetilde{d}}\right) \le \frac{K}{d^\ell} \ .$$

In turn, the above implies that

$$\phi_r \le K \cdot \sum_{\ell \ge 0} \frac{N^G_{r,\ell}}{d^\ell} = K \cdot \psi_r \le K \cdot \alpha.$$

This proves (37).

All the above conclude the proof of Theorem 6.3. □

7.1. **Proof of Proposition 7.3.** For $\Lambda \subseteq V_T \setminus \{r\}$ and $\sigma \in \{\pm 1\}^\Lambda$, recall from Section 4.1 the ratio of marginals at the root $R^{K,\tau}_T(r)$ such that

$$R^{\Lambda,\sigma}_T(r) = \frac{\mu_r(+1 \mid \Lambda, \sigma)}{\mu_r(-1 \mid \Lambda, \sigma)} \ . \tag{39}$$

Recall, also, that $\mu_r(\cdot \mid \Lambda, \sigma)$ denotes the marginal of the Gibbs distribution $\mu_{T,\boldsymbol{\lambda}}(\cdot \mid \Lambda, \sigma)$ at the root r.

For a vertex $u \in V_T$, we let T_u be the subtree of T that includes u and all its descendents. We always assume that the root of T_u is the vertex u. With a slight abuse of notation, we let $R^{\Lambda,\sigma}_u$ denote the ratio of marginals at the root for the subtree T_u, where the Gibbs distribution is, now, with respect to T_u, while we impose the boundary condition $\sigma(K \cap T_u)$.

Letting $v_1, v_2, \dots, v_k$ denote the children of v, it is standard to get the following recursion

$$R^{\Lambda,\sigma}_v = \boldsymbol{\lambda}_v \prod_{i=1}^{k} \frac{1}{1 + R^{\Lambda,\sigma}_{v_i}} \ . \tag{40}$$

Let $A \subseteq V_T$ includes all $w \in V_T$ such that the path from the root r to w there is a vertex $v \in \Lambda$. Also, let B includes all $w \in V_T$ such that w has a child $u \in \Lambda$ and $\sigma(u) = +1$. Define the set

$$F = \Lambda \cup A \cup B \ .$$

Given the condition (Λ, σ), it is standard to show that $\mathcal{I}_T(r, w) = 0$, for any $w \in F$. We call the vertices in $V_T \setminus F$ the *free* vertices.

We define a set of parameters $(\alpha_v)_{v\in V}$ as follows: for each leaf vertex v we have $\alpha_v = 1$. For every non-leaf vertex v, let $v_1, v_2, \ldots, v_k$ denote the children of v and define $\alpha_v = 1+\frac{1}{d}\sum_{i=1}^{k}\alpha_{v_i}$. It is straightforward to verify that α_v is the d-branching value of $v \in V$ in the subtree T_v rooted at v. Hence, it holds that $\alpha_r \leq \alpha$.

For the sake of brevity, in what follows, we abbreviate $\mathcal{I}_T^{\Lambda,\sigma}$ and $R_v^{\Lambda,\sigma}$ to $\mathcal{I}_T$ and R_v, respectively. Let $L(h)$ denote the set of all vertices at distance h from the root r. Let $\chi \in (1,2)$ be the parameter and $\Phi(\cdot)$ be the function in Lemma 6.1. We claim that

$$\forall h \geq 1, \qquad \sum_{v\in L(h)\setminus F}\left(\frac{\alpha_v}{\alpha_r}\right)^{1/\chi}\frac{\mathcal{I}_T(r,v)}{R_v\Phi(R_v)} \leq d^{1/\chi}\sqrt{\deg_T(r)}(d\kappa)^{(h-1)/\chi}\ , \tag{41}$$

where

$$\kappa = \sup_{0<x\leq||\boldsymbol{\lambda}||_\infty}\kappa(x) < 1/d, \tag{42}$$

is the parameter, where the function $\kappa(x)$ is specified in Lemma 6.1. Hence, the parameter κ in (41) depends only on $||\boldsymbol{\lambda}||_\infty$. We remark that all the ratios in the above inequality are well-defined because $\alpha_r \geq 1$ and for any free vertex $v \notin F$, it holds that $0 < R_v < \infty$.

Before proving that (41) is true, we show how we can use it to prove the proposition. Note that $\alpha_v \geq 1$ for all $v \in V_T$. Since $d > 1$, we have

$$\alpha_v \geq 1 + \frac{1}{d}(\deg_T(v)-1) \geq \frac{\deg_T(v)}{d}\ .$$

Note that $\mathcal{I}_T(r,v) = 0$ for all fixed $v \in F$. Hence, the weighted total influence can be bounded as follows

$$\begin{aligned}
\sum_{v\in V_T:v\neq r}\mathcal{I}_T(r,v)\cdot\deg_T(v)^{1/\chi} &= \sum_{v\in V_T\setminus F:v\neq r}\mathcal{I}_T(r,v)\cdot\deg_T(v)^{1/\chi}\\
&\leq \sum_{v\in V_T\setminus F:v\neq r}|\mathcal{I}_T(r,v)|(d\alpha_v)^{1/\chi}\\
&= (d\alpha_r)^{1/\chi}\sum_{v\in V_T\setminus F:v\neq r}\left(\frac{\alpha_v}{\alpha_r}\right)^{1/\chi}|\mathcal{I}_T(r,v)|\\
\text{(since } R_v\Phi(R_v)\leq 1\text{)} \qquad &\leq (d\alpha_r)^{1/\chi}\sum_{v\in V_T\setminus F:v\neq r}\left(\frac{\alpha_v}{\alpha_r}\right)^{1/\chi}\frac{|\mathcal{I}_T(r,v)|}{R_v\Phi(R_v)} \qquad (43)\\
&= (d\alpha_r)^{1/\chi}\sum_{h=1}^{\infty}\sum_{v\in L(h)\setminus F}\left(\frac{\alpha_v}{\alpha_r}\right)^{1/\chi}\frac{|\mathcal{I}_T(r,v)|}{R_v\Phi(R_v)}\ .
\end{aligned}$$

Note that in (43), the inequality $R_v\Phi(R_v) \leq 1$ follows from the definition of $\Phi(\cdot)$ in Lemma 6.1. Plugging (41) into the above inequality, we get that

$$\begin{aligned}
\sum_{v\in V:v\neq r}\mathcal{I}_T(r,v)\cdot\deg_T(v)^{1/\chi} &\leq \left(d^2\alpha_r\right)^{1/\chi}\sqrt{\deg_T(r)}\sum_{h=1}^{\infty}(d\kappa)^{(h-1)/\chi}\\
\text{(since } d\kappa<1\text{)} \qquad &\leq (d^2\alpha_r)^{1/\chi}\sqrt{\deg_T(r)}\cdot\frac{1}{1-(d\kappa)^{1/\chi}}\\
\left(\text{by } \frac{1}{2}<\frac{1}{\chi}<1\right) \qquad &\leq \left(d^2\alpha_r\deg_T(r)\right)^{1/\chi}\cdot\frac{1}{1-(d\kappa)^{1/\chi}}\\
\text{(by } \alpha_r\leq\alpha\text{)} \qquad &\leq \frac{d^{2/\chi}\alpha^{1/\chi}}{1-(d\kappa)^{1/\chi}}\cdot\deg_T(r)^{1/\chi}\ .
\end{aligned}$$

The proposition follows by setting the parameter $D = D(d, ||\boldsymbol{\lambda}||_\infty) = \frac{d^{2/\chi}}{1-(d\kappa)^{1/\chi}}$.

We conclude the proof of Proposition 7.3 by showing that (41) is true. We prove (41) by induction on h.

The base case corresponds to $h = 1$. Let $v_1, v_2, \ldots, v_k$ denote all free children of the root r, i.e., $v_i \notin F$. Also, let Γ denote the set of all children of r. Consider the influence from r to a free child $v_i \notin F$. It is standard to show that for the Hard-core model $\mu_{T,\boldsymbol{\lambda}}$ we have that

$$\mathcal{I}_T(r, v_i) = \frac{R_{v_i}}{1 + R_{v_i}} \quad . \tag{44}$$

Hence, we get that

$$\frac{\mathcal{I}_T(r, v_i)}{R_{v_i}\Phi(R_{v_i})} = \frac{\sqrt{R_{v_i}(R_{v_i}+1)}}{R_{v_i}+1} = \sqrt{\frac{R_{v_i}}{R_{v_i}+1}} \leq 1 \quad .$$

The above implies that

$$\begin{aligned}\sum_{v \in L(1)\setminus F} \left(\frac{\alpha_v}{\alpha_r}\right)^{1/\chi} \frac{\mathcal{I}_T(r, v)}{R_v \Phi(R_v)} &\leq \frac{\sum_{i=1}^k \alpha_{v_i}^{1/\chi}}{\alpha_r^{1/\chi}} = \frac{\sum_{i=1}^k \alpha_{v_i}^{1/\chi}}{(1 + \frac{1}{d}\sum_{v_j \in \Gamma} \alpha_{v_j})^{1/\chi}} \\ \text{(by } \{v_1, v_2, \ldots, v_k\} \subseteq \Gamma) \qquad &\leq \frac{\sum_{i=1}^k \alpha_{v_i}^{1/\chi}}{(1 + \frac{1}{d}\sum_{j=1}^k \alpha_{v_j})^{1/\chi}} \\ &\leq d^{1/\chi} \cdot \frac{\sum_{i=1}^k \alpha_{v_i}^{1/\chi}}{(\sum_{j=1}^k \alpha_{v_j})^{1/\chi}} \quad .\end{aligned}$$

Let $q \geq 1$ be such that $\frac{1}{q} + \frac{1}{\chi} = 1$. Note that $q = \frac{\chi}{\chi - 1}$, while, since $\chi \in (1, 2)$, we have that $\frac{1}{q} \leq \frac{1}{2}$. By Hölder's inequality, we get that

$$\sum_{i=1}^k \alpha_{v_i}^{1/\chi} \leq k^{1/q} \left(\sum_{i=1}^k \alpha_v\right)^{1/\chi} \leq \sqrt{\deg_T(r)} \left(\sum_{i=1}^k \alpha_{v_i}\right)^{1/\chi} \quad .$$

This implies that

$$\sum_{v \in L(1)\setminus F} \left(\frac{\alpha_v}{\alpha_r}\right)^{1/\chi} \frac{\mathcal{I}_T(r, v)}{R_v \Phi(R_v)} \leq d^{1/\chi}\sqrt{\deg_T(r)} \cdot (d\kappa)^0 \quad .$$

The above proves the base of the induction.

We now focus on proving the induction step. For $\ell \geq 1$, suppose (41) is true for $h = \ell$. We prove that (41) is also true for $h = \ell + 1$.

For any free vertex $v \in L(\ell + 1) \setminus F$, v's father $u \in L(\ell)$ must be free ($u \notin F$) due to the definition of F. For any $u \in L(\ell) \setminus F$, let $u_1, u_2, \ldots, u_{k(u)}$ denote the free children of the vertex u, where $0 \leq k(u) \leq \deg_T(u)$.

From Lemma 7.2 we have that $\mathcal{I}_T(r, u_i) = \mathcal{I}_T(r, u) \cdot \mathcal{I}_T(u, u_i)$ for all $[i] \in k(u)$. Hence, we get that

$$\sum_{v \in L(\ell+1)\setminus F} \left(\frac{\alpha_v}{\alpha_r}\right)^{1/\chi} \frac{\mathcal{I}_T(r, v)}{R_v \Phi(R_v)} = \sum_{u \in L(\ell)\setminus F} \left(\frac{\alpha_u}{\alpha_r}\right)^{1/\chi} \frac{\mathcal{I}_T(r, u)}{R_u \Phi(R_u)} \sum_{i=1}^{k(u)} \left(\frac{\alpha_{u_i}}{\alpha_u}\right)^{1/\chi} \frac{\mathcal{I}_T(u, u_i)}{R_{u_i}\Phi(R_{u_i})} R_u \Phi(R_u) \quad . \tag{45}$$

From (44), we have

$$\sum_{i=1}^{k(u)} \left(\frac{\alpha_{u_i}}{\alpha_u}\right)^{1/\chi} \frac{\mathcal{I}_T(u,u_i)}{R_{u_i}\Phi(R_{u_i})} R_u\Phi(R_u) \le \sum_{i=1}^{k(u)} \left(\frac{\alpha_{u_i}}{\alpha_u}\right)^{1/\chi} \frac{R_u\Phi(R_u)}{(R_{u_i}+1)\Phi(R_{u_i})}$$

$$\text{(by Hölder's inequality)} \quad \le \left(\frac{\sum_{i=1}^{k(u)} \alpha_{u_i}}{\alpha_u}\right)^{1/\chi} \left(\Phi(R_u)^q \sum_{i=1}^{k(u)} \left(\frac{R_u}{(R_{u_i}+1)\Phi(R_{u_i})}\right)^q\right)^{1/q}, \tag{46}$$

where $q = (1-1/\chi)^{-1}$. Using the recursion in (40), Lemma 6.1 and the definition of κ in (42), we have

$$\left(\Phi(R_u)^q \sum_{i=1}^{k(u)} \left(\frac{R_u}{(R_{u_i}+1)\Phi(R_{u_i})}\right)^q\right)^{1/q} \le \kappa^{1/\chi} .$$

Plugging the above into (46) we get that

$$\sum_{i=1}^{k(u)} \left(\frac{\alpha_{u_i}}{\alpha_u}\right)^{1/\chi} \frac{\mathcal{I}_T(u,u_i)}{R_{u_i}\Phi(R_{u_i})} R_u\Phi(R_u) \le \left(\frac{\sum_{i=1}^{k(u)} \alpha_{u_i}}{\alpha_u}\right)^{1/\chi} \kappa^{1/\chi} \le (d\kappa)^{1/\chi} , \tag{47}$$

where the last inequality holds since

$$\alpha_u = 1 + \frac{1}{d} \sum_{w:w \text{ is a child of } u} \alpha_w \ge \frac{1}{d} \sum_{i=1}^{k(u)} \alpha_{u_i} .$$

Plugging (47) into (45) we get that

$$\sum_{v \in L(\ell+1)\setminus F} \left(\frac{\alpha_v}{\alpha_r}\right)^{1/\chi} \frac{\mathcal{I}_T(r,v)}{R_v\Phi(R_v)} \le (d\kappa)^{1/\chi} \sum_{u\in L(\ell)\setminus F} \left(\frac{\alpha_u}{\alpha_r}\right)^{1/\chi} \frac{\mathcal{I}_T(r,u)}{R_u\Phi(R_u)}$$

$$\le d^{1/\chi}\sqrt{\deg_T(r)}(d\kappa)^{\ell/\chi} .$$

This proves the induction step and concludes the proof (41).

The proposition follows. □

8. Monomer-Dimer Model - Proof of Theorem 2.2

Let us first introduce the notion of line graph. Given a graph $G = (V, E)$, we have the *line graph* L of G such that each vertex in L is an edge in G and $e, f \in E$ are adjancent in L if and only if $e \cap f \neq \emptyset$. Also note that, if Δ is the maximum degree in G, then the maximum degree in L is at most 2Δ.

We use, here the standard observation that the Monomer-Dimer model with edge weight λ, on the graph $G = (V, E)$, corresponds to the Hard-core model on the line graph L with fugacity λ.

We often view the Monomer-Dimer model as a distribution over $\{\pm 1\}^E$, where for any $X \in \{\pm 1\}^E$, any $e \in E$, $X_e = +1$ represents that e is in the matching and $X_e = -1$ represents that e is not in the matching.

Theorem 2.2 is a corollary of the following more general result.

Theorem 8.1. *For any constants $\lambda > 0$, there exist two constants $M_1 = M_1(\lambda), M_2 = M_2(\lambda)$ such that for any graph $G = (V, E)$ with n vertices, maximum degree $\Delta \ge 2$ and $m \ge 10^6(1+\lambda+1/\lambda)^3\Delta^{3/2}$ edges the following is true:*

Let μ_G be the Monomer-Dimer model on G with edge weight λ. Then the Glauber dynamics on μ_G exhibits mixing time such that

$$T_{\text{mix}} \le (M_1\Delta)^{M_2\sqrt{\Delta}} n \log n.$$

Specifically, Theorem 8.1 implies Theorem 2.2 because with probability $1-o(1)$ over the instances of $G(n,d/n)$, the maximum degree is $\Delta=\Theta(\frac{\log n}{\log\log n})$ (e.g. see Lemma B.1), while the number of edges is $\Theta(n)=\omega(\Delta^{3/2})$. Note that the bound on the number of edges is a simple application of Chernoff's bound. Using Theorem 8.1, we know that with probability $1-o(1)$ over the instances of $G(n,d/n)$,

$$T_{\text{mix}} \leq n^{1+C\sqrt{\frac{\log\log n}{\log n}}},$$

for some constant C depending only on λ and d.

From now on, our focus shifts to proving Theorem 8.1. As in the case of the Hard-core model, we consider the more general non-homogenous version of the Monomer-Dimer model. That is, consider the graph $G=(V,E)$ and let $\boldsymbol{\lambda}=(\lambda_e)_{e\in E}\in\mathbb{R}^E_{\geq 0}$ be an assignment of weight to each edge of the graph G. Let $\mu_{\boldsymbol{\lambda}}$ be the distribution over all matchings σ such that $\mu_{\boldsymbol{\lambda}}(\sigma)\propto\prod_{e\in E:\sigma(e)=+1}\lambda_e$.

We have the following result, which can be derived directly from [BGK+07, CLV21].

Lemma 8.2. *For any graph $G=(V,E)$ with the maximum degree Δ, any edge weights $\boldsymbol{\lambda}=(\lambda_e)_{e\in E}\in\mathbb{R}^E_{\geq 0}$, the Gibbs distribution of Monomer-Dimer model specified by G and $\boldsymbol{\lambda}$ is $\left(2\sqrt{1+||\boldsymbol{\lambda}||_\infty\,\Delta}\right)$-spectrally independent.*

Theorem 6.1 in [CLV21] is identical to Lemma 8.2 with the only difference that it is for *homogenous* edge weights, i.e. $\lambda_e=\lambda$ for all $e\in E$. One can extend this theorem and obtain Lemma 8.2 by combining the spectral independence analysis in [CLV21] and the correlation decay analysis in [BGK+07].

Furthermore, we have the following corollary about Complete Spectral Independence.

Corollary 8.3. *For any constant $\lambda>0$ and any graph $G=(V,E)$ with n vertices and maximum degree $\Delta\geq 2$, the following is true:*

Let μ be the Monomer-Dimer model on G with edge weight λ. Then, there exists a constant $K=4\sqrt{1+\lambda}$ such that μ_G is (η,ξ)-completely spectrally independent for

$$\eta\leq K\sqrt{\Delta} \qquad \textit{and} \qquad \xi=1.$$

Proof. Consider $\mu_{\boldsymbol{\lambda}}$, the non-homogenous Monomer-Dimer model on G with edge weights $\boldsymbol{\lambda}\in(0,\lambda(1+\xi)]^E$. It suffices to show that $\mu_{\boldsymbol{\lambda}}$ is spectrally independent with parameter η. Using Lemma 8.2 we have that $\mu_{\boldsymbol{\lambda}}$ is $\left(2\sqrt{1+||\boldsymbol{\lambda}||_\infty\,\Delta}\right)$-spectrally independent. Hence, we have that

$$2\sqrt{1+||\boldsymbol{\lambda}||_\infty\,\Delta}=2\sqrt{1+\lambda(1+\xi)\Delta}=2\sqrt{1+2\lambda\Delta}\leq K\sqrt{\Delta}\ .$$

□

We also derive marginal stability results.

Theorem 8.4 (Stability Monomer-dimer Model)**.** *For any constant $\lambda>0$, for the graph $G=(V,E)$ with n vertices and maximum degree $\Delta\geq 2$, let μ be of the Monomer-Dimer model on G with edge weight λ. Then, we have that μ is $(\lambda+2)^3\Delta^2$-marginally stable.*

Proof. Let $\zeta=(\lambda+2)^3\Delta^2$. For $\Lambda\subseteq E$, a feasible configuration $\tau\in\{\pm\}^\Lambda$ and $e\in E\setminus\Lambda$, let the ratio of Gibbs marginals at e

$$R^{\Lambda,\tau}_G(e)=\frac{\mu^{\Lambda,\tau}_e(+1)}{\mu^{\Lambda,\tau}_e(-1)}\ . \tag{48}$$

Recall that for marginal stability, we need to have that for any $S\subseteq\Lambda$,

$$R^{\Lambda,\tau}_G(e)\leq\zeta \qquad \text{and} \qquad R^{\Lambda,\tau}_G(e)\leq\zeta\cdot R^{S,\tau_S}_G(e)\ . \tag{49}$$

The first bound is easy because $R^{\Lambda,\tau}_G(e)\leq\mu^{\Lambda,\tau}_e(+1)\leq\frac{\lambda}{1+\lambda}\leq\zeta$. We focus on the second one.

Suppose that $e = \{u, w\}$. Let N_u (resp. N_w) be the set of edges incident to u (resp. w) except for the edge e. We may assume that none of the edges in $N_u \cup N_v$ is set to be $+1$ by τ, as otherwise $R_G^{\Lambda,\tau}(e) = 0$ and (49) holds trivially.

Next, we proceed to derive a lower bound for $\mu_e^{S,\tau_S}(+1)$. Let the set $F_u = N_u \setminus S$ and $F_w = N_w \setminus S$. Also, let $F_e = F_u \cup F_w$. We call F_e the set of free edges, since it corresponds the set of edges that are not fixed under τ_S.

Letting $\boldsymbol{\sigma}$ be distributed as in $\mu_{E\setminus S}^{S,\tau_S}$, we have

$$\mu_e^{S,\tau_S}(+1) \geq \left(\frac{\lambda}{1+\lambda}\right) \cdot \Pr[\forall f \in F_e,\ \boldsymbol{\sigma}(f) = -1] \ . \tag{50}$$

We use Ω to denote the support of $\mu_{E\setminus S}^{S,\tau_S}$. We partition Ω into two parts

$$\Omega_- = \{\sigma \in \Omega \mid \forall f \in F_e,\ \sigma(f) = -1\}$$
$$\Omega_+ = \{\sigma \in \Omega \mid \exists f \in F_e,\ \sigma(f) = +1\} \ .$$

For the moment, assume that $\Omega_+ \neq \emptyset$, i.e., the set is non-empty.

For a configuration $\sigma \in \Omega_+$, note there can be at most two edges $f, f' \in F_e$ such $\sigma(f) = \sigma(f') = 1$. Hence, the number of edges in F_e that are set to $+1$, under σ, is at least 1 and at most 2. Furthermore, since σ is a matching, if there are two edges $f, f' \in F_e$ such $\sigma(f) = \sigma(f') = 1$, these must be in different sets, e.g., $f \in F_u$ and $f' \in F_w$.

For $\sigma \in \Omega_+$, let $\eta \in \Omega_-$ be a configuration that agrees with σ on the assignment of the edges outside F_e. Note that, since $\eta \in \Omega_-$, we have $\eta(f) = -1$ for all $f \in F_e$.

Furthermore, noting that η and σ differ only on the configuration of at most two edges, we have that

$$\mu_{E\setminus S}^{S,\tau_S}(\eta) \geq \frac{\mu_{E\setminus S}^{S,\tau_S}(\sigma)}{\max\{\lambda^2, 1\}} \geq \frac{\mu_{E\setminus S}^{S,\tau_S}(\sigma)}{1+\lambda^2} \ . \tag{51}$$

Finally, we note that σ can be uniquely specified by η and the edges in F_e at which the two configurations disagree, i.e., recall that we assumed that σ, τ disagree only at F_e. Hence, using (51), we have that

$$\frac{\sum_{\eta\in\Omega_-} \mu_{E\setminus S}^{S,\tau_S}(\eta)}{\sum_{\tau\in\Omega_+} \mu_{E\setminus S}^{S,\tau_S}(\tau)} \geq \frac{1}{(1+\lambda^2)(|F_u|+1)\times(|F_w|+1)} \geq \frac{1}{(1+\lambda^2)\Delta^2} \ .$$

The above implies the following: for $\Omega_+ \neq \emptyset$, we have that

$$\Pr[\forall f \in F_e,\ \boldsymbol{\sigma}(f) = -1] \geq \left(\frac{1}{1+(1+\lambda^2)\Delta^2}\right) \ , \tag{52}$$

where recall that $\boldsymbol{\sigma}$ be distributed as in $\mu_{E\setminus S}^{S,\tau_S}$. Furthermore, for the case where $\Omega_+ = \emptyset$ it is immediate that $\Pr[\forall f \in F_e,\ \boldsymbol{\sigma}(f) = -1] = 1$.

Hence, we have that

$$\mu_e^{S,\tau_S}(+1) \geq \left(\frac{\lambda}{1+\lambda}\right) \cdot \Pr[\forall f \in F_e,\ \boldsymbol{\sigma}(f) = -1] \geq \left(\frac{\lambda}{1+\lambda}\right) \cdot \left(\frac{1}{1+(1+\lambda^2)\Delta^2}\right).$$

Since $R_G^{\Lambda,\tau}(e)$ is increasing in the value of $\mu_e^{S,\tau_S}(+1)$, we use the above to get that

$$R_G^{S,\tau_S}(e) \geq \frac{\left(\frac{\lambda}{1+\lambda}\right) \cdot \left(\frac{1}{1+(1+\lambda^2)\Delta^2}\right)}{1-\left(\frac{\lambda}{1+\lambda}\right) \cdot \left(\frac{1}{1+(1+\lambda^2)\Delta^2}\right)} \geq \frac{\left(\frac{\lambda}{1+\lambda}\right) \cdot \left(\frac{1}{1+(1+\lambda^2)\Delta^2}\right)}{1-\left(\frac{\lambda}{1+\lambda}\right) \cdot \left(\frac{1}{1+(1+\lambda^2)\Delta^2}\right)} \cdot \frac{R_G^{\Lambda,\tau}(e)}{\lambda}.$$

In the second inequality we use the fact that $R_G^{\Lambda,\tau}(e) \leq \lambda$.

The second inequality in (49) can be proved from the above and noting that

$$\frac{1-\left(\frac{\lambda}{1+\lambda}\right)\cdot\left(\frac{1}{1+(1+\lambda^2)\Delta^2}\right)}{\left(\frac{\lambda}{1+\lambda}\right)\cdot\left(\frac{1}{1+(1+\lambda^2)\Delta^2}\right)}\cdot\lambda\le\zeta=(\lambda+2)^3\Delta^2. \qquad \square$$

This concludes the proof of Theorem 8.4.

Finally, we have the following bound on the approximate tensorisation of entropy.

Lemma 8.5. *Consider the Gibbs distribution μ of the Monomer-dimer model specified by $G=(V,E)$ and edge weight $\lambda>0$. For any $k\ge 1$ and $H\subseteq E$ such that $|H|=k$, any feasible pinning $\tau\in\{\pm\}^{E\setminus H}$, the conditional distribution $\mu_H^{E\setminus H,\tau}$ satisfies the approximate tensorization of entropy with constant*

$$\mathrm{AT}(k)\le 2k^2\left(1+\lambda+1/\lambda\right)^{2k+2} \quad .$$

Similarly to what we have for Theorem 8.4, Lemma 8.5 follows directly from Lemma 3.4 (i.e., the first bound) by utilising the connection between the Monomer-Dimer model on G and the Hard-core model on its line graph L. For this reason, we omit the proof of Lemma 8.5.

We are now ready to prove Theorem 8.1.

Proof of Theorem 8.1 . We set the parameters

$$\eta=4\sqrt{1+\lambda}\cdot\sqrt{\Delta},\quad \xi=1,\quad \zeta=(\lambda+2)^3\Delta^2 \quad .$$

By Corollary 8.3 and Theorem 8.4, the Gibbs distribution μ is (η,ξ)-completely spectrally independent and ζ-marginally stable. Set the parameter

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

It is elementary calculations to verify that

$$\frac{1}{\alpha}\le 100(1+\lambda)\sqrt{\Delta} \quad .$$

Let $m=|E|$ and set $\ell=\lceil\theta m\rceil$, where

$$\theta=\frac{1}{400\mathrm{e}\Delta(1+\lambda+1/\lambda)^2} \quad . \tag{53}$$

Since we assumed that $m\ge 10^6(1+\lambda+1/\lambda)^3\Delta^{3/2}$, it holds that $1/\alpha\le\ell<m$. By Theorem 4.7, the Gibbs distribution μ satisfies the ℓ block factorisation of entropy with parameter

$$C=\left(\frac{em}{\ell}\right)^{1+1/\alpha}\le\left(\frac{e}{\theta}\right)^{1+1/\alpha}\le\frac{1}{2}(A\Delta)^{B\sqrt{\Delta}} \quad , \tag{54}$$

for some constants $A=A(\lambda)$ and $B=B(\lambda)$.

Let Ω denote the support of μ. For any $f:\Omega\to\mathbb{R}_{>0}$ we have that

$$\mathrm{Ent}_\mu(f)\le\frac{C}{\binom{m}{\ell}}\sum_{S\in\binom{E}{\ell}}\mu\left(\mathrm{Ent}_S(f)\right) \quad ,$$

where C is the parameter in (54), while recall that $m=|E|$.

We need to consider the subgraph induced by subset of edges S. For any $S\subseteq E$, let $C(S)$ denotes the set of connected components in $G(V,S)$ which contains at least one edge. With a slight abuse of notation, we use $U\in C(S)$ to denote the set of edges in the component U.

By the conditional independence property of the Gibbs distribution and Lemma 3.3, we have

$$\begin{aligned}
\mathrm{Ent}_\mu(f) &\le \frac{C}{\binom{m}{\ell}} \sum_{S\in\binom{E}{\ell}} \sum_{U\in C(S)} \mu\left(\mathrm{Ent}_U(f)\right) \\
\text{(by Lemma 8.5)} \quad &\le \frac{C}{\binom{m}{\ell}} \sum_{S\in\binom{E}{\ell}} \sum_{U\in C(S)} \mathrm{AT}(|U|) \sum_{e\in U} \mu[\mathrm{Ent}_e(f)] \\
&\le C \sum_{e\in E} \mu[\mathrm{Ent}_e(f)] \sum_{k\ge 1} \mathrm{AT}(k) \Pr[|C_e| = k] \\
&\le C \sum_{e\in E} \mu[\mathrm{Ent}_e(f)] \sum_{k\ge 1} \left(2k^2 \left(1+\lambda+1/\lambda\right)^{2k+2}\right) \Pr[|C_e| = k] \ ,
\end{aligned}$$

where C_e is the connected component in $G(V,S)$ containing e, while S is sampled from $\binom{E}{\ell}$ uniformly at random. At this point we need to bound the probability term $\Pr[|C_e| = k]$.

Consider the line graph L of G. Note that the maximum degree Δ_L of L is at most 2Δ. Furthermore, let v_e denote the vertex in L that corresponds to the edge e in G.

Suppose we sample ℓ vertices $\hat{S}$ uniformly at random from graph L. Let $C(v_e)$ the component in $L[\hat{S}]$ that includes vertex v_e. Then, it is straightforward that the probability $\Pr[|C_e| = k]$ is equal to the probability $\Pr[|C(v_e)| = k]$. Recall that $\Pr[|C_e| = k]$ refers to choosing uniformly random edges from G and $\Pr[|C(v_e)| = k]$ refers to choosing uniformly at random vertices from L.

From [CLV21, Lemma 4.3], we have

$$\begin{aligned}
\Pr[|C_e| = k] = \Pr[|C(v_e)| = k] &\le \frac{\ell}{m}(2\mathrm{e}\Delta_L\theta)^{k-1} \le \frac{\ell}{m}(4\mathrm{e}\Delta\theta)^{k-1} \\
&\le \left(\frac{1}{100(1+\lambda+1/\lambda)^2}\right)^{k-1} \quad ,
\end{aligned}$$

the last inequality follows from that $\ell = \lceil \theta m \rceil$ for θ defined in (53). The above implies that

$$\begin{aligned}
\mathrm{Ent}_\mu(f) &\le C \sum_{e\in E} \mu[\mathrm{Ent}_e(f)] \sum_{k\ge 1} \left(2k^2 \left(1+\lambda+1/\lambda\right)^{2k+2}\right) \left(\frac{1}{100(1+\lambda+1/\lambda)^2}\right)^{k-1} \\
&\le (A\Delta)^{B\sqrt{\Delta}} \sum_{e\in E} \mu[\mathrm{Ent}_e(f)],
\end{aligned}$$

where the last inequality holds by (54). Note that $m \le n\Delta$. We have

$$T_{\mathrm{mix}} \le \left\lceil (A\Delta)^{B\sqrt{\Delta}} m \left(\log\log \frac{1}{\mu_{\min}} + \log(2) + 2\right)\right\rceil \le (M_1\Delta)^{M_2\sqrt{\Delta}} n \log n,$$

where $M_1 = M_1(\lambda)$ and $M_2 = M_2(\lambda)$ are two constants depending only on λ. □

ACKNOWLEDGEMENT

Charilaos Efthymiou is supported by EPSRC New Investigator Award (grant no. EP/V050842/1) and Centre of Discrete Mathematics and Applications (DIMAP), The University of Warwick.

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

Weiming Feng would like to thank Heng Guo for the helpful discussions.

References

[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.

[BGGŠ22] Ivona Bezáková, Andreas Galanis, Leslie Ann Goldberg, and Daniel Štefankovič. Fast sampling via spectral independence beyond bounded-degree graphs. In *ICALP*, volume 229 of *LIPIcs*, pages 21:1–21:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[BGK+07] Mohsen Bayati, David Gamarnik, Dimitriy Katz, Chandra Nair, and Prasad Tetali. Simple deterministic approximation algorithms for counting matchings. In *STOC*, pages 122–127. ACM, 2007.

[BKZZ13] Jean Barbier, Florent Krzakala, Lenka Zdeborová, and Pan Zhang. The hard-core model on random graphs revisited. *Journal of Physics: Conference Series*, 473(1):012021, dec 2013.

[CE15] Amin Coja-Oghlan and Charilaos Efthymiou. On independent sets in random graphs. *Random Struct. Algorithms*, 47(3):436–486, 2015.

[Ces01] Filippo Cesi. Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields. *Probab. Theory Related Fields*, 120(4):569–584, 2001.

[CFYZ21] 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. IEEE, 2021.

[CFYZ22] Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Optimal mixing for two-state anti-ferromagnetic spin systems. In *FOCS*, pages 588–599. IEEE, 2022.

[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.

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

[DM11] Varsha Dani and Cristopher Moore. Independent sets in random graphs from the weighted second moment method. In *RANDOM*, volume 6845 of *Lecture Notes in Computer Science*, pages 472–482. Springer, 2011.

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

[Eft22] Charilaos Efthymiou. Spectral independence beyond uniqueness using the topological method. *CoRR*, abs/2211.03753, 2022.

[EHSV18] Charilaos Efthymiou, Thomas P. Hayes, Daniel Stefankovic, and Eric Vigoda. Sampling random colorings of sparse random graphs. In *SODA*, pages 1759–1771, 2018.

[FGYZ21a] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Fast sampling and counting *k*-sat solutions in the local lemma regime. *J. ACM*, 68(6):40:1–40:42, 2021.

[FGYZ21b] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Rapid mixing from spectral independence beyond the boolean domain. In *SODA*, pages 1558–1577, 2021.

[FHY21] Weiming Feng, Kun He, and Yitong Yin. Sampling constraint satisfaction solutions in the local lemma regime. In *STOC*, pages 1565–1578. ACM, 2021.

[Fri90] Alan M. Frieze. On the independence number of random graphs. *Discret. Math.*, 81(2):171–175, 1990.

[GGGY21] Andreas Galanis, Leslie Ann Goldberg, Heng Guo, and Kuan Yang. Counting solutions to random CNF formulas. *SIAM J. Comput.*, 50(6):1701–1738, 2021.

[GS14] David Gamarnik and Madhu Sudan. Limits of local algorithms over sparse random graphs. In *ITCS*, pages 369–376. ACM, 2014.

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

[JPV21a] Vishesh Jain, Huy Tuan Pham, and Thuy-Duong Vuong. On the sampling Lovász local lemma for atomic constraint satisfaction problems. *arXiv preprint arXiv:2102.08342*, 2021.

[JPV21b] 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.

[JS89] Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM J. Comput.*, 18(6):1149–1178, 1989.

[Kel85] F. P. Kelly. Stochastic models of computer communication systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 47:379–395, 1985.

[KMR+07] Florent Krzakala, Andrea Montanari, Federico Ricci-Tersenghi, Guilhem Semerjian, and Lenka Zdeborová. Gibbs states and the set of solutions of random constraint satisfaction problems. *Proc. Natl. Acad. Sci. USA*, 104(25):10318–10323, 2007.

[Moi19] Ankur Moitra. Approximate counting, the lovász local lemma, and inference in graphical models. *J. ACM*, 66(2):10:1–10:25, 2019.
[Sly10] Allan Sly. Computational transition at the uniqueness threshold. In *FOCS*, pages 287–296, 2010.
[SS14] Allan Sly and Nike Sun. Counting in two-spin models on d-regular graphs. *Ann. Probab.*, 42(6):2383–2416, 2014.
[SSŠY17] Alistair Sinclair, Piyush Srivastava, Daniel Štefankovič, and Yitong Yin. Spatial mixing and the connective constant: optimal bounds. *Probab. Theory Related Fields*, 168(1-2):153–197, 2017.

Appendix A. Bounds on Spectral Gap from Spectral Independence

Let $\mu = \mu_G$ be a Gibbs distribution on graph $G = (V, E)$ with support $\Omega \subseteq \{\pm 1\}^V$. Let P denote the transition matrix of the Glauber dynamics on μ_G. It is well-known that P has non-negative real eigenvalues $1 = \lambda_1 \geq \lambda_2 \geq \ldots \lambda_{|\Omega|} \geq 0$. The spectral gap of P is defined by $1 - \lambda_2$. We have the following relation between spectral independence and spectral gap.

Lemma A.1 ([FGYZ21b, Theorem 3.2]). *Let $\eta \geq 0$ and $0 \leq \varphi < 1$ be two parameters. Let $G = (V, E)$ be a graph with $n = |V|$ vertices. Support the Gibbs distribution μ_G on $G = (V, E)$ satisfies that for every $0 \leq k \leq n - 2$, $\Lambda \subseteq V$ of size k and feasible configuration $\tau \in \{\pm 1\}^\Lambda$,*

$$\rho(|\mathcal{I}_G^{\Lambda,\tau}|) \leq \eta \qquad \textit{and} \qquad \rho(|\mathcal{I}_G^{\Lambda,\tau}|) \leq \frac{\varphi}{n - k - 1}.$$

The spectral gap of Glauber dynamics on μ_G is at least

$$1 - \lambda_2 \geq \frac{(1 - \varphi)^{2+2\eta}}{n^{1+2\eta}}.$$

Appendix B. Structural Properties of $G(n, p)$

Lemma B.1. *For real numbers $d > 0$ and $\epsilon > 0$, let Δ be the maximum degree of the graph $\boldsymbol{G} \sim G(n, d/n)$. Then the following is true: With probability $1 - o(1)$ the graph $\boldsymbol{G}$ satisfies that*

$$(1 - \epsilon)\frac{\log n}{\log \log n} \leq \Delta \leq (1 + \epsilon)\frac{\log n}{\log \log n}.$$

The above result is standard to derive, using the first and second moment method.