Spectral Independence in High-Dimensional Expanders and Applications to the Hardcore Model

Nima Anari
Department of Computer Science
Stanford University
Stanford, CA, USA
anari@cs.stanford.edu

Kuikui Liu
Computer Science & Engineering
University of Washington
Seattle, WA, USA
liukui17@cs.washington.edu

Shayan Oveis Gharan
Computer Science & Engineering
University of Washington
Seattle, WA, USA
shayan@cs.washington.edu

2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS) | 978-1-7281-9621-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/FOCS46700.2020.00125

***Abstract*—We say a discrete probability distribution over subsets of a finite ground set is spectrally independent if an associated pairwise influence matrix has a bounded largest eigenvalue for the distribution and all of its conditional distributions. We prove that if a distribution is spectrally independent, then the corresponding high dimensional simplicial complex is a local spectral expander. Using a line of recent works on mixing time of high dimensional walks on simplicial complexes [KM17; DK17; KO18; AL20], this implies that the corresponding Glauber dynamics mixes rapidly and generates (approximate) samples from the given distribution.**

As an application, we show that natural Glauber dynamics mixes rapidly (in polynomial time) to generate a random independent set from the hardcore model up to the uniqueness threshold. This improves the quasi-polynomial running time of Weitz's deterministic correlation decay algorithm [Wei06] for estimating the hardcore partition function, also answering a long-standing open problem of mixing time of Glauber dynamics [LV97; LV99; DG00; Vig01; Eft+16].

***Keywords*-approximate counting; Markov chain Monte Carlo; Glauber dynamics; spectral independence; high-dimensional expanders; correlation decay**

I. Introduction

Suppose we have a ground set $[n] = \{1, \dots, n\}$ of elements. Let $\mu : 2^{[n]} \to \mathbb{R}_+$ be a probability distribution on subsets of $[n]$. We say μ is d-homogeneous, if for every $S \in \operatorname{supp}\{\mu\}$, we have $|S| = d$. When the choice of μ and $[n]$ are clear from context, we will write $\Pr[i] = \Pr_{S\sim\mu}[i \in S]$ and $\Pr[\bar{j}] = \Pr_{S\sim\mu}[i \notin S]$. The following definitions are crucial in our paper.

Definition I.1 ((Signed) Pairwise Influence Matrix)**.** *Fix a distribution μ on subsets of a ground set $[n]$. We define the pairwise correlation matrix $\Psi_\mu \in \mathbb{R}^{n\times n}$ by*

$$\Psi_\mu(i,j) \stackrel{\text{def}}{=} \Pr[j \mid i] - \Pr[j \mid \bar{i}]$$

for $i \neq j$, and $\Psi_\mu(i,i) = 0$ for all $i = 1, \dots, n$. We refer to the entry $\Psi_\mu(i,j)$ as the pairwise influence of i on j.

Note this differs from existing definitions of "influence"; see Section I-D for further discussion. One may also view Ψ_μ is a matrix of pairwise correlations.

Definition I.2 (Spectral Independence)**.** *We say a probability distribution μ on subsets of $[n]$ is η-*spectrally independent *if $\lambda_{\max}(\Psi_\mu) \le \eta$. Note that since the maximum eigenvalue is always at most the maximum absolute row/column sum, we have μ is η-spectrally independent if either*

$$\sum_{j\neq i} |\Psi_\mu(i,j)| \le \eta \qquad \text{or} \qquad \sum_{j\neq i} |\Psi_\mu(j,i)| \le \eta,$$

We say μ is $(\eta_0, \dots, \eta_{n-2})$-spectrally independent if μ is η_0-independent, for all $0 \le i < n$, $\{\mu | i \text{ in/out}\}$ is η_1-independent, for all i, j, $\{\mu | i \text{ in/out}, j \text{in/out}\}$ is η_2-independent, and so on.

Note that for any i, we always have $\eta_i \le n - i - 1$; the smaller η_i's are, the more independent μ is. Ideally, we are interested in distributions where $\eta_0, \dots, \eta_{n-2} \le O(1)$ independent of n. Observe that if μ is a product distribution, then it is $(0, \dots, 0)$-independent.

Let us explain a more interesting example. Recall that a probability distribution μ is *negatively correlated* if for all $i \neq j$, we have $\Pr[i|j] \le \Pr[i]$. If μ is d-homogeneous and all measures obtainable from μ by conditioning are negatively correlated, then μ is $(1, 1, \dots, 1)$-spectrally independent.

For a bad example, consider the distribution μ which places $1/2$ probability to both $\{1, \dots, \frac{n}{2}\}$ and $\{\frac{n}{2}+1, \dots, n\}$. In this case , $\lambda_{\max}(\Psi_\mu) = n - 1$.

Given a probability distribution μ we can define a Markov chain called the *Glauber dynamics* to generate samples from μ as follows: Given a set $S \in \operatorname{supp}\{\mu\}$, we choose a uniformly random element i and we transition to

$$\begin{cases} S \setminus \{i\} & \text{with prob } \frac{\mu(S\setminus\{i\})}{\mu(S\setminus\{i\})+\mu(S\cup\{i\})} \\ S \cup \{i\} & \text{o.w.,} \end{cases}$$

It turns out that this chain has the right stationary distribution.

The following is our main technical theorem.

Theorem I.3 (Main)**.** *For any $(\eta_0, \dots, \eta_{n-2})$-spectrally independent distribution $\mu : 2^{[n]} \to \mathbb{R}_+$, the natural Glauber*

2575-8454/20/$31.00 ©2020 IEEE
DOI 10.1109/FOCS46700.2020.00125

dynamics (defined above) has spectral gap at least

$$\frac{1}{n}\prod_{i=0}^{n-2}\left(1-\frac{\eta_i}{n-i-1}\right)$$

We note that prior works [FM92; AGR16] show that as long as the distribution μ, and all its conditional distributions, satisfying certain negative correlation properties, then a very similar Markov chain mixes rapidly. In our setting, negative correlation is equivalent to all entries of Ψ_μ being nonpositive. Thus, in a similar spirit to spectral negative dependence [Ana+19], one may view spectral independence and Theorem I.3 as also relaxing these negative correlation requirements to allow for some positive correlation between elements, while still providing mixing time guarantees.

In the following sections we will explain an application of the above theorem in bounding the mixing time of the Glauber dynamics for sampling independent sets from the hardcore distribution. The proof of Theorem I.3 uses recent connections developed by the authors and collaborators between analysis of Markov chains and the field of high dimensional expanders [Ana+19].

A. Connections to High Dimensional Simplicial Complexes

Let us first phrase our main contribution in the language of high-dimensional expanders. For a ground set $U = [n]$ of elements, a *simplicial complex* X is a downward closed family of subsets of U. Sets in X are also called *faces* of X. The dimension of a face in X is its size. For an integer k, we write $X(k)$ to denote all faces of X of size k. We say X is *pure* if all maximal faces have the same size. The dimension of X is the size of the maximum face in X. For a pure d-dimensional simplicial complex X, we say X is d-*partite* if U can be partitioned into sets $U_1, \ldots, U_d$ such that every maximal face σ has exactly one element of each U_i.

We will often weight the maximal faces of a pure d-dimensional simplicial complex X by some function $w : X(d) \to \mathbb{R}_{>0}$. This induces weights on all faces of X via

$$w(\tau) = \sum_{\sigma \in X(d):\sigma \supset \tau} w(\sigma) \tag{1}$$

For a face τ of X, the *link* of τ is the simplicial complex $X_\tau = \{\sigma \setminus \tau : \sigma \in X, \sigma \supset \tau\}$. We endow the maximal faces of X_τ with the weight $w_\tau(\sigma) = w(\tau \cup \sigma)$.

The 1-*skeleton* of link X_τ of τ is a weighted graph defined as follows: For every element $i \in U$, such that $\{i\} \in X_\tau$ we have a vertex. We connect two vertices i, j if $\{i, j\} \in X_\tau$ and the weight of the edges is $w_\tau(\{i, j\})$. We will let P_τ denote the simple random walk on the 1-skeleton of X_τ.

We also define a random walk on the maximal faces of X by a two-step process. If the walk is currently at some $\sigma \in X(d)$, we transition by

1) removing a uniformly random element $i \in \sigma$
2) adding a random $j \notin \sigma \setminus \{i\}$ to $\sigma \setminus \{i\}$ with probability proportional to $w(\sigma \cup \{j\} \setminus \{i\})$

Note that there is always a nonzero probability staying at σ in a given step.

The transition probability matrix $P_d^\vee$ of this random walk may be written down as

$$P_d^\vee(\sigma, \sigma') = \begin{cases} \sum_{\tau \subset \sigma : |\tau| = d-1} \frac{w(\sigma)}{d \cdot w(\tau)}, & \text{if } \sigma = \sigma' \\ \frac{w(\sigma')}{d \cdot w(\sigma \cap \sigma')}, & \text{if } |\sigma \cap \sigma'| = d-1 \\ 0, & \text{o.w.} \end{cases}$$

where we recall that $w(\tau) = \sum_{\sigma \in X(d):\sigma \supset \tau} w(\sigma)$.

Given a distribution μ on subsets of $[n]$, define a pure n-dimensional n-partite simplicial complex X^μ as follows: Let the ground set of elements be $\{1, \overline{1}, 2, \overline{2}, \ldots, n, \overline{n}\}$ with n parts $U_1 = \{1, \overline{1}\}, U_2 = \{2, \overline{2}\}, \ldots, U_n = \{n, \overline{n}\}$. For every set $S \in \mathrm{supp}\{\mu\}$ we add a maximal face σ_S which has i for every $i \in S$ and $\bar{i}$ for every $i \notin S$. We assign a weight to σ_S given by $w(\sigma_S) = \mu(S)$. We turn this into a simplicial complex by taking downward closure of all maximal faces. Note that in this case, $P_n^\vee$ describes exactly the Glauber dynamics for sampling $S \subset [n]$ with probability proportional to $\mu(S)$.

We are now ready to define the notion of high-dimensional expansion that we will use, which was first introduced in [DK17; KM17; KO18; Opp18].

Definition I.4 (Local Spectral Expander; [KO18])**.** *Let X be a pure d-dimensional simplex complex. We say a face τ of X is an α-spectral expander if the second largest eigenvalue of the simple (non-lazy) random walk on the 1-skeleton of X_τ is at most α. We say X is $(\alpha_0, \ldots, \alpha_{d-2})$-local spectral expander if for all $0 \le k \le d-2$, every $\tau \in X(k)$ is an α_k-spectral expander.*

We prove the following theorem making connection between spectral independence of probability distributions and local spectral expanders.

Theorem I.5. *For any $(\eta_0, \ldots, \eta_{n-2})$-spectrally independent distribution $\mu : 2^{[n]} \to \mathbb{R}_+$, the pure n-dimensional n-partite simplicial complex X^μ is a $(\frac{\eta_0}{n-1}, \frac{\eta_1}{n-2}, \ldots, \frac{\eta_{n-2}}{1})$-local spectral expander.*

We note that there are strong theorems in the literature of high-dimensional expanders [Opp18] which show that if the $(d-2)$-dimensional faces of a pure d-dimensional complex X are α-spectral expanders for $\alpha \le 1/2d$, then every face of X is a 2α-spectral expander. However, such theorems fail dramatically when the $(d-2)$-dimensional faces have spectral expansion, say, $1/2$.

Here, the main new ingredient is to show that as long as the underlying distribution μ is spectrally independent for $\eta_0, \ldots, \eta_{n-2} \le O(1)$, then we get better and better spectral expansion as we go to lower dimensional faces of the underlying weighted simplicial complex X^μ.

The key usefulness of local spectral expansion lies in the following local-to-global theorem, which may be used to bound $\lambda_2(P_d^\vee)$. A weaker version of this result was already proved in [KO18].

Theorem I.6 ([AL20]). *Consider a pure d-dimensional simplicial complex X with weights w. If (X, w) is a $(\alpha_0, \ldots, \alpha_{d-2})$-local spectral expander, then*

$$\lambda_2(P_d^\vee) \leq 1 - \frac{1}{d}\prod_{k=0}^{d-2}(1-\alpha_k)$$

Remark I.7. For instance, if there is a constant α such that (X, w) is a $\left(\frac{\alpha}{d-1}, \frac{\alpha}{d-2}, \ldots, \frac{\alpha}{2}, \frac{\alpha}{1}\right)$-local spectral expander, then we would obtain $\lambda_2(P_d^\vee) \leq 1 - \frac{1}{d^{1+\alpha}}$. This is precisely what we do for the hardcore model.

Proof of Theorem I.3: By Theorem I.5, spectral independence of μ implies strong local spectral expansion of X^μ. Theorem I.6 then furnishes the spectral gap of the Glauber dynamics, which we recall is described by $P_n^\vee$. ∎

It now remains to prove Theorem I.5.

B. Application to Sampling from Hardcore Distribution

Our main application of the above machinery is to generate random samples from the hardcore distribution. Given a graph $G = (V, E)$, and a parameter $\lambda > 0$, sample an independent set I with probability $\lambda^{|I|}/Z_G(\lambda)$, where

$$Z_G(\lambda) = \sum_{I \subset V \text{ independent}} \lambda^{|I|}$$

is the normalizing constant, a.k.a., the partition function. Exact computation of $Z_G(\lambda)$ is #P-Hard [Val79; Vad95; Gre00] even when the input graphs have special structure [Vad02] and hence, we can only hope for efficient approximation algorithms.

Studying the hardcore model has been pivotal in helping us understand the relationship between phase transitions in statistical physics and phase transitions in efficient approximability. Specifically, has been known since [Kel85] that there is a critical threshold $\lambda_c(\Delta) \stackrel{\text{def}}{=} \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta} \approx \frac{e}{\Delta-2}$ for which the Gibbs distribution is unique on the infinite Δ-regular tree if and only if $\lambda < \lambda_c(\Delta)$. The case $\lambda < \lambda_c(\Delta)$ exactly corresponds to the regime where the "influence" of a vertex u on another vertex v decays exponentially fast in the distance between u, v. This is known to physicists as the uniqueness regime for the hardcore model. On the flip side, $\lambda > \lambda_c(\Delta)$ exactly corresponds to the regime where long-range correlations persist in the model.

In the seminal work of Weitz [Wei06], it was shown that for any $\lambda < \lambda_c(\Delta)$ and fixed constant Δ, there exists a deterministic fully polynomial time approximation scheme (FPTAS) for estimating $Z_G(\lambda)$. Immediately following, a sequence of results [SS14; Gal+14; GŠV15; GŠV16] beginning with the seminal work of Sly [Sly10] proved a matching lower bound for the case $\lambda > \lambda_c(\Delta)$. There is no fully polynomial randomized approximation scheme (FPRAS) for estimating $Z_G(\lambda)$ on graphs of maximum degree $\leq \Delta$ when $\lambda > \lambda_c(\Delta)$ unless NP = RP. This rigorously established the first example where the statistical physics phase transition coincides with a computational complexity phase transition.

Weitz's algorithm is based on the correlation decay framework which was later on developed for estimating partition functions of two state spin systems [LLY12; LLY13; SST14]. More recently, a new framework was established based on Barvinok's polynomial interpolation method [Bar16b; Bar16a; PR17; PR19] where Weitz's result was reproved using a different deterministic algorithm which only uses the knowledge of connected subgraphs of G of diameter $O_{\epsilon,\delta}(\log n)$ [PR17]. All of these methods suffer from a quasi-polynomial running time when the input graph has unbounded max-degree. Specifically, if $\lambda = (1-\delta)\lambda_c(\Delta)$, then there is a constant $C(\delta)$ such that Weitz's correlation decay algorithm returns a $(1 \pm \epsilon)$-multiplicative approximation of $Z_G(\lambda)$ in time $O\left((n/\epsilon)^{C(\delta)\log\Delta}\right)$. In particular, due to the exponential dependence in $\log\Delta$, Weitz's algorithm does not run in polynomial time for graphs with unbounded maximum degree. Roughly speaking, the main difficulty is that in order to estimate the partition function, one needs to estimate the marginal probabilities of vertices within $O(1/n)$-error, and to do that one needs to look at $O(\log n)$-depth neighborhood of vertices which leads to a quasi-polynomial number of operations on graphs of max-degree polynomial in n.

On the other hand, it is conjectured that the natural Glauber dynamics mixes in polynomial time up to the uniqueness threshold. But to this date after a long line of works [LV97; LV99; DG00; Vig01] this was only shown up to $\frac{2}{\Delta-2}$ for general graphs and up to the uniqueness threshold for special families of graphs [Wei04; Wei06; Res+13; Eft+16]. We use the result of the previous sections to prove that for *any graph* the Glauber dynamics mix in polynomial time up to the uniqueness threshold.

For sampling from the hardcore model, the Glauber dynamics can be described via the following two-step process. To make a transition from an independent set I to another,

1) Select a uniformly random vertex $v \in V$.
2) If $v \in I$, remove v from I with probability $\frac{1}{1+\lambda}$, and keep it otherwise.
3) If $v \notin I$ and v is not a neighbor of some $u \in I$, add v to I with probability $\frac{\lambda}{1+\lambda}$, and leave it otherwise.

It is clear that this process is reversible. It is also clear that this Markov chain is connected, since there is a path from every independent set to the empty independent set $\emptyset$. Hence, these dynamics have a unique stationary distribution π, and the distribution of the chain converges to stationarity in total variation distance as the number of steps goes to infinity. Finally, by checking the detailed balance condition that the stationary distribution π of the Glauber dynamics is exactly

the Gibbs distribution μ. Our goal is to bound the ϵ-total variation mixing time of the Glauber dynamics starting from any state τ, which is given by

$$t_\tau(\epsilon) = \min\left\{t \in \mathbb{Z}_{\geq 0} : \left\|P^t(\tau,\cdot) - \pi\right\|_1\right\} \leq \epsilon$$

where P denotes the transition probability matrix describing the chain. Here, $P^t(\tau,\cdot)$ gives the distribution at time t of the chain started at τ.

Theorem I.8. *There is a function $C : [0,1] \to \mathbb{R}_{>0}$ such that for every graph $G = (V,E)$ with maximum degree $\leq \Delta$, every $0 < \delta < 1$, and $\lambda = (1-\delta)\lambda_c(\Delta)$, the associated hardcore distribution μ is $(\eta_0,\ldots,\eta_{n-2})$-spectrally independent where $\eta_i \leq \min\left\{C(\delta), \frac{\lambda}{1+\lambda}(n-i-1)\right\}$ for every $0 \leq i \leq n-2$.*

Combined with Theorem I.3, we obtain fast mixing for the Glauber dynamics for sampling independent sets according to the hardcore distribution whenever $\lambda < \lambda_c(\Delta)$ (for the precise mixing time, see Remark I.10).

Corollary I.9. *For every $\delta > 0$, there exists a fully polynomial randomized approximation scheme for estimating $Z_G(\lambda)$ at $\lambda = (1-\delta)\lambda_c(\Delta)$ on any graph G with maximum degree $\leq \Delta$.*

Remark I.10. For $0 < \delta < 1$, $\lambda = (1-\delta)\lambda_c(\Delta)$ and a graph $G = (V,E)$ maximum degree Δ, the Glauber dynamics from any starting state τ has mixing time

$$t_\tau(\epsilon) \leq O\left(((1+\lambda)\cdot n)^{1+C(\delta)} \cdot \log\left(\frac{1}{\epsilon\cdot\mu(\tau)}\right)\right)$$

To be explicit, the constant $C(\delta)$ obeys the bound

$$C(\delta) \leq \exp(O(1/\delta))$$

Note $\mu(\emptyset) = \frac{1}{Z_G(\lambda)} \geq \frac{1}{(1+\lambda)^n}$ so that

$$t_\emptyset(\epsilon) \leq O\left(n^{2+C(\delta)} \cdot \log\left(\frac{1}{\epsilon}\right)\right)$$

The key advantage of our result is that the running time has no dependence on Δ. Furthermore, $\lambda_c(\Delta) \leq 4$ for all $\Delta \geq 3$ so we may treat λ as bounded above by a constant. Hence, only the gap parameter δ matters.

C. Related Works

The question of building deterministic approximation algorithms for estimating $Z_G(\lambda)$ on bounded degree graphs has been settled. The seminal work of Weitz [Wei06] proved that there is an FPTAS on graphs of maximum degree $\leq \Delta$ whenever $\lambda < \lambda_c(\Delta)$ using the correlation decay method. The polynomial interpolation method of [Bar16a] has also since been demonstrated to give an FPTAS [PR17; PR19] as well.

For studying the mixing time of the Glauber dynamics in the uniqueness regime, there has been a long line of work starting with [LV97; LV99; DG00; Vig01]. For general graphs, the state-of-the-art was given by [Vig01], which showed the Glauber dynamics mixes in $O(n\log n)$ steps when $\lambda < \frac{2}{\Delta-2}$. A more recent result of [Eft+16] shows that for any $0 < \delta < 1$, there is a $\Delta_0(\delta)$ such that for any $\Delta \geq \Delta_0(\delta)$ and $\lambda = (1-\delta)\lambda_c(\Delta)$, the Glauber dynamics mixes in $O(n\log n)$ steps for graphs with maximum degree Δ and girth ≥ 7. Much more is known for line graphs [JS89], amenable graphs [Dye+02; Wei04; Wei06], trees [MSW03; MSW04; Wei04], planar graphs [Hay06], graphs with large girth [HV05], $\mathbb{Z}^2$ [VVY13; Res+13; Bla+13; Bla+19], and $G(n,d/n)$ random graphs (or, more generally, graphs with bounded connective constant) [MS08; MS13; SSY13; Sin+15].

On the hardness side, exact computation of $Z_G(\lambda)$ is known to be #P-Hard [Val79; Vad95; Gre00], even for very restricted families of graphs [Vad02]. For hardness of approximation, [LV97] showed there exists a constant $c > 0$ such that there is no FPRAS for estimating $Z_G(1)$ when $\lambda > c/\Delta$ unless NP = RP. For the case of evaluating $Z_G(1)$, this was improved in [DFJ02], which showed that there is no FPRAS for estimating $Z_G(1)$ on graphs with maximum degree exceeding 25 unless NP = RP. [DFJ02] further showed that the Glauber dynamics has exponential mixing time for $\Delta \geq 6$. [MWW07] provided further evidence the Markov chain techniques are likely to fail for sampling from the Gibbs distribution when $\lambda > \lambda_c(\Delta)$. These results were dramatically improved in the work of [Sly10] (and further refined by follow-up works [SS14; Gal+14; GŠV15; GŠV16]), which showed that unless NP = RP, there is no FPRAS for estimating $Z_G(\lambda)$ on graphs of maximum degree $\leq \Delta$ when $\lambda > \lambda_c(\Delta)$.

D. Relation to Existing Definitions of Influence

For spin systems, our pairwise influence matrix is reminiscent but different from the Dobrushin influence matrix used in [Hay06; DGJ09] and the works [Dob70; DS85a; DS85b; DS87]. Specifically, the (i,j)th entry of the Dobrushin influence matrix considered in these prior works is given by the maximum absolute difference $\left|\Pr[j \mid i,\tau] - \Pr[j \mid \bar{i},\tau]\right|$ over all partial assignments τ of the remaining ground elements excluding i,j. In the case of the hardcore distribution for an input graph $G = (V,E)$ with fugacity $\lambda > 0$, this influence matrix is exactly $\frac{\lambda}{1+\lambda}A$, where A is the adjacency matrix of G (see, for instance, [Hay06]). On the other hand, our pairwise influence matrix Ψ_μ may have nonzero entries for $u,v \in V$ not connected by an edge, since $\Psi_\mu(u,v)$ considers the marginal of v conditioned only on u or $\overline{u}$, with the assignment for other elements left undetermined. Furthermore, our method requires understanding exponentially many pairwise influence matrices, one for each conditional distribution, while all variants of the Dobrushin condition only require analyzing a single Dobrushin influence matrix.

E. Subsequent Works

Finally, we mention several follow-up works applying and extending the notion of spectral independence we introduce this paper. The first is the work by [CLV20], where they obtained rapid mixing of the Glauber dynamics for all two-state spin systems in the correlation decay regime. [Che+20; Fen+20] extended our notion of spectral independence to multi-state spin systems, and obtained new mixing results for the Glauber dynamics for sampling q-colorings on triangle-free graphs.

F. Proof Overview

For a face σ of X^μ, recall P_σ denotes the transition probability matrix of the simple random walk on the 1-skeleton of X^μ_σ. Our first technical contribution is the following.

Theorem I.11. *For every distribution μ over subsets of a ground set $[n]$, the eigenvalues of Ψ_μ are real. Furthermore, we have the identity $\lambda_2(P_\emptyset) = \frac{1}{n-1} \cdot \lambda_{\max}(\Psi_\mu)$.*

Given this, we may now prove Theorem I.5.

Proof of Theorem I.5: Since Theorem I.11 holds for any distribution μ, it in particular holds for all conditional distributions of μ. Now, observe that conditioning on an element i being "in" corresponds exactly to taking the link of X^μ w.r.t. i. Similarly, conditioning on an element i being "out" corresponds exactly to taking the link of X^μ w.r.t. $\bar{i}$. The result then follows by definition of spectral independence and local spectral expansion. ∎

The proof of Theorem I.11 hinges on the observation that for each element $i \in [n]$, no face of X^μ can contain both i and $\bar{i}$. In particular, there is no edge connecting i and $\bar{i}$ in the 1-skeleton of X^μ, for each i. Thus, there are n parts, one corresponding to each element of $[n]$, such that all edges only go between parts. This *n-partite* structure of the 1-skeleton of X^μ induces $n-1$ additional "trivial" eigenvalues, besides the trivial eigenvalue of 1, in the transition matrix $P_\emptyset$. This is, in fact, a generalization of the fact that the transition matrix of a bipartite graph always also has eigenvalue -1. We show that Ψ_μ is essentially equal to $P_\emptyset$ projected away from these n trivial eigenvalues; see Claim III.2 and Claim III.3 in Section III for more details.

We apply these results to the hardcore distribution over independent sets of an input graph $G = (V, E)$. Theorem I.11 tells us that to bound $\lambda_2(P_\emptyset)$, it suffices to bound $\lambda_{\max}(\Psi_\mu)$. We show how to bound $\lambda_{\max}(\Psi_\mu)$ by bounding $\sum_{u \in V : u \neq v} |\Psi_\mu(u, v)|$ for any vertex $v \in V$. In particular, we have the following two bounds.

Lemma I.12. *Consider the hardcore distribution μ on independent sets of a graph $G = (V, E)$ on n vertices. Then for every $v \in V$, and every $\lambda > 0$, we have the bound*

$$\sum_{u \in V : u \neq v} |\Psi_\mu(u, v)| \leq \frac{\lambda}{1+\lambda} \cdot (n-1)$$

Proof: Observe that the maximum probability that a given vertex is placed in a random independent set is at most $\frac{\lambda}{1+\lambda}$. In particular, $\Pr[v \mid u], \Pr[v \mid \overline{u}] \in \left[0, \frac{\lambda}{1+\lambda}\right]$ so that $|\Psi_\mu(u, v)| \leq \frac{\lambda}{1+\lambda}$ for every $u \neq v$. The claim follows. ∎

Theorem I.13. *There exists a function $C : [0, 1] \to \mathbb{R}_{>0}$ such that for every graph $G = (V, E)$ of maximum degree $\leq \Delta$, every vertex $v \in V$, every $0 < \delta < 1$, and $\lambda = (1-\delta)\lambda_c(\Delta)$, we have the following bound,*

$$\sum_{u \in V : u \neq v} |\Psi_\mu(u, v)| \leq C(\delta)$$

To be explicit, $C(\delta)$ satisfies $C(\delta) \leq \exp(O(1/\delta))$.

Remark I.14. We believe $C(\delta) \leq O(1/\delta)$ is possible, which we show is tight in the full version of the paper. We leave this as an open problem. We note that in follow-up work, [CLV20] shows that $\sum_{u \in V : u \neq v} |\Psi_\mu(v, u)| \leq O(1/\delta)$. In other words, they analyze the total pairwise influence of a vertex, while we analyze the total pairwise influence on a vertex.

The key here is that we only need to understand the total sum of correlations between pairs of vertices. This is in contrast to strong spatial mixing results, where one has to analyze the correlation of any subset of vertices on another given vertex.

To prove Theorem I.13, first, we take advantage of the self-avoiding walk tree construction introduced in [Wei06] to reduce to a problem on trees. Then, we give an method to decouple the influence of a set of vertices S on a vertex v into the sum of the single-vertex pairwise influences of each $u \in S$ on v. The primary takeaway from these two steps is that it suffices to control the total pairwise influence of vertices on the root in any rooted tree of maximum degree $\leq \Delta$.

To control correlations between vertices and the root, we leverage the well-known tree recursion, which expresses the marginal of the root in terms of the marginals of its children. We amortize the total pairwise influence of all vertices at a fixed distance from the root using the potential method [LLY12; LLY13; Res+13; SSY13; SST14; Sin+15]; we refer to [Sri14] for further discussion of the potential method. This allows us to show a strong kind of correlation decay, where the total pairwise influence of all vertices at a fixed distance decays as the distance grows. After sharing a preliminary draft of this paper, it was pointed out to us by Eric Vigoda and Zongchen Chen that the notion of correlation decay we prove is very similar to the notion of *aggregate strong spatial mixing* (for trees) studied in [MS13; BCV20].

Now, observe that Theorem I.8 simply follows from Lemma I.12 and Theorem I.8. As a consequence, all we are left to do is to prove Theorem I.11 and Theorem I.13, which we do in the remainder of the paper.

G. Structure of the Paper

In Section II, we review necessary background in the theory of Markov chains, and correlation decay. In Section III, we sketch the proof of Theorem I.11. In Section IV and Section V, we sketch the proof of Theorem I.13. We leave detailed proofs to the full version of the paper [ALO20].

II. Preliminaries

First, let us establish some notational conventions. Unless otherwise specified, all logarithms are in base e. All vectors are assumed to be column vectors. For two vectors $\phi, \psi \in \mathbb{R}^n$, we use $\langle \phi, \psi \rangle$ to denote the standard Euclidean inner product between ϕ and ψ. We use $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geqslant 0}$ to denote the set of positive and nonnegative real numbers, respectively, and $[n]$ to denote $\{1, \dots, n\}$.

A. Markov Chains and Random Walks

We consider a Markov chain as a triple (Ω, P, π) where Ω denotes a (finite) state space, $P \in \mathbb{R}_{\geqslant 0}^{\Omega \times \Omega}$ denotes a transition probability matrix and $\pi \in \mathbb{R}_{\geqslant 0}^{\Omega}$ denotes a stationary distribution of the chain (which will be unique for all chains we consider).

A chain (Ω, P, π) is *reversible* if there is a nonzero nonnegative function $f : \Omega \to \mathbb{R}_{\geqslant 0}$ such that for any pair of states $\tau, \sigma \in \Omega$, $f(\tau)P(\tau, \sigma) = f(\sigma)P(\sigma, \tau)$. If this condition is satisfied, then f is proportional to π. In this paper we only work with reversible Markov chains. Note that being reversible means that the transition matrix P is self-adjoint w.r.t. the inner product $\langle \phi, \psi \rangle_f = \phi^\top \operatorname{diag}(f)\psi$. For any reversible Markov chain (Ω, P, π), the largest eigenvalue of P is 1. We let $\lambda^*(P)$ denote the second largest eigenvalue of P in absolute value. That is, if $-1 \leq \lambda_n \leq \cdots \leq \lambda_1 = 1$ are the eigenvalues of P, then $\lambda^*(P) = \max\{|\lambda_2|, |\lambda_n|\}$.

Theorem II.1 ([DS91]). *For any reversible irreducible Markov chain (Ω, P, π), $\epsilon > 0$, and any starting state $\tau \in \Omega$,*

$$t_\tau(\epsilon) \leq \frac{1}{1 - \lambda^*(P)} \cdot \log\left(\frac{1}{\epsilon \cdot \pi(\tau)}\right).$$

B. Tree Recurrences for Hardcore Model

Fix a tree T rooted at some vertex r. For a vertex v in T, let $\ell(v)$ denote its distance from the root r. We will sometimes refer to it as the "level" which contains v. For a level ℓ, let $L_r(\ell) = \{v \in T : \ell(v) = \ell\}$. For a vertex $u \in T$, we will write T_u for the subtree of T rooted at u.

A key tool we will need to analyze the hardcore model on trees is given by the tree recurrence. To describe the tree recurrence, we need to consider a change of variables w.r.t. the marginal probabilities. Fix a tree T arbitrarily rooted at some vertex $r \in T$, and an arbitrary boundary condition $p : A \to [0, 1]$ for a subset of remaining vertices A. We write the ratio of conditional probabilities as

$$R^p_{T,r} = \frac{\Pr[r \mid p]}{\Pr[\overline{r} \mid p]} = \frac{\Pr[r \mid p]}{1 - \Pr[r \mid p]}$$

Here, we think of the function $p : A \to [0, 1]$ as fixing the marginal probability of vertices $v \in A$ to $p(v)$. In the special case where p maps all vertices of A to 0 or 1, then p is really a boundary condition in the traditional sense, as p is pinning the vertices of A to be in/out. However, later on, we will need the additional flexibility of pinning the marginal of $v \in A$ to a specific value $p(v) \in [0, 1]$

With this notation in hand, we may write the tree recurrence for the hardcore model as

$$R^p_{T,r} = F(R^p_{T_u,u} : u \in L_r(1)) \stackrel{\text{def}}{=} \lambda \prod_{u \in L_r(1)} \frac{1}{R^p_{T_u,u} + 1} \tag{2}$$

Here, we make a slight abuse of notation by writing p even when considering a subtree T_u; this should be understood as the restriction of p to this subtree. We drop the superscript when p is empty; we also drop the subscript T when the tree is clear from context. In the case of a depth-ℓ complete d-ary tree rooted at r with no boundary conditions, all of the R^σ_u are the same. In this case, the tree recurrence simplifies to a univariate recurrence given by

$$f_d(R) = \lambda\left(\frac{1}{R+1}\right)^d$$

C. Correlation Decay and Weitz's Self-Avoiding Walk Tree

In this subsection, we introduce the necessary notation for describing the correlation decay property for spin systems. We begin by rigorously defining correlation decay for a general distribution μ over subsets of a ground set $[n]$.

Definition II.2 (Spatial Mixing). *Fix a metric $d : [n] \times [n] \to \mathbb{R}_{\geq 0}$. We say a distribution μ on subsets of $[n]$ exhibits weak spatial mixing w.r.t. d with rate $0 < \alpha < 1$ and constant $C > 0$ if for every $i \in [n]$, every $S \subset [n]$ with $i \notin S$, and every pair of assignments $\tau, \sigma : S \to \{in, out\}$, we have*

$$|R^\tau_i - R^\sigma_i| \leq C \cdot \alpha^{d(i,S)}$$

We say μ exhibits strong spatial mixing if instead we can replace $d(i, S)$ by $d(i, S(\tau, \sigma))$, where $S(\tau, \sigma) \subset S$ is the set of elements on which the assignments τ, σ differ.

In the case of a distribution μ on configurations $\sigma : V \to \{0, 1\}$ on a graph $G = (V, E)$ coming from the hardcore model with activity λ, our ground set will consist of the vertices V. Hence, we will take the d to be the shortest path metric in G throughout the paper.

It has been known since the work of Kelly [Kel85] that for the hardcore model, weak spatial mixing on the infinite Δ-regular tree holds exactly when $\lambda < \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta} \stackrel{\text{def}}{=} \lambda_c(\Delta)$. Here, $\lambda_c(\Delta)$ is known as the critical threshold for the hardcore model on graphs of maximum degree $\leq \Delta$. These results have been subsequently extended to all antiferromagnetic two-state spin systems [LLY12; LLY13; SST14].

The way the threshold $\lambda_c(\Delta)$ is derived is by analyzing when $\left|f'_{\Delta-1}(\hat{R}_{\Delta-1})\right|$ is less than 1. It turns out the gap between $\left|f'_{\Delta-1}(\hat{R}_{\Delta-1})\right|$ and 1 governs the rate α in the definition of spatial mixing. [LLY13] quantified this in the following definition.

Definition II.3 (Up-to-Δ Uniqueness [LLY13])**.** *We say the hardcore model with parameter λ is up-to-Δ unique with gap $0 < \delta < 1$ if for every $1 \le d < \Delta$, we have $\left|f'_d(\hat{R}_d)\right| \le 1-\delta$, where $\hat{R}_d$ denotes the unique fixed point of f_d.*

It is not hard to show that up-to-Δ uniqueness with gap $0 < \delta < 1$ is equivalent to $\lambda \le (1-\Theta(\delta)) \cdot \lambda_c(\Delta)$ (see, for instance, the full version of the paper). Hence, throughout the paper whenever one encounters the phrase "up-to-Δ unique with gap $0 < \delta < 1$", one may safely assume $\lambda \le (1-\Theta(\delta))\lambda_c(\Delta)$.

Surprisingly, Weitz [Wei06] managed to show that for the hardcore model, weak spatial mixing actually implies strong spatial mixing with the same rate α, albeit with a worse constant C. This was extended in [LLY13; SST14] to all antiferromagnetic two-state spin systems. The way this was done was to first reduce spatial mixing on a general graph to spatial mixing on an associated tree known as the self-avoiding walk tree [SS05; Wei06], where one can then use the tree recurrence Eq. (2) to understand spatial mixing.

Theorem II.4 (Theorem 3.1 from [Wei06])**.** *Fix a graph $G = (V, E)$ and a vertex $r \in V$. Then there exists a tree $T = T_{\mathrm{SAW}}(G, r)$ whose vertices may be partitioned into parts $\{C(v)\}_{v \in V}$, one for each vertex of G, and a boundary condition τ_{SAW} on vertices of T such that for any partial assignment $\sigma : S \to \{0,1\}$ of vertices in $S \subset V$, we have $\Pr_G[r \mid \sigma] = \Pr_T[r \mid \tau_{\mathrm{SAW}}, \sigma_{\mathrm{SAW}}]$ and $R^{\sigma}_{G,r} = R^{\tau_{\mathrm{SAW}}, \sigma_{\mathrm{SAW}}}_{T,r}$. Here, σ_{SAW} is the partial assignment on vertices of T with $\sigma_{\mathrm{SAW}}(u) = \sigma(v)$ for every $u \in C(v)$ and every $v \in S$.*

Remark II.5. The tree $T = T_{\mathrm{SAW}}(G, r)$ turns out to have a couple convenient properties. For instance, we have $d_G(r, v) = \min\{\ell(u) : u \in C(v)\}$, and that the maximum degree of T equals the maximum degree of G. We also note that in general, the number of vertices of T may be exponentially large in the size of G. We refer the reader to [Wei06] or the full version of the paper [ALO20] for a complete description of T.

For the second step, in the case of the hardcore model, [Wei06] showed that weak spatial mixing on the infinite Δ-regular tree implies strong spatial mixing on all trees of maximum degree $\le \Delta$, and hence, on all graphs of maximum degree $\le \Delta$. To conveniently state the strong spatial mixing result proved in [Wei06], we make the following definition.

Definition II.6. *If T is a tree rooted at $r \in T$, we define $R^{\min}_{T,r}(\ell) = \min_p R^p_{T,r}$, where p is an assignment of marginals of vertices at depth ℓ in T_u. Similarly, define $R^{\max}_{T,r}(\ell)$ to be the maximum such conditional probability ratio. Finally, define $R^{\min}(\ell) = \min_{T,r} R^{\min}_{T,r}(\ell)$ and $R^{\max}(\ell) = \max_{T,r} R^{\max}_{T,r}(\ell)$, where the minimum and maximum are over all trees T rooted at r of maximum degree $\le \Delta$.*

Remark II.7. Essentially, due to the antiferromagnetic nature of the hardcore model, the level-ℓ boundary condition p minimizing $R^p_{T,r}$ is the all-1 configuration if ℓ is odd, and the all-0 configuration if ℓ is even. Determining the configuration achieving $R^{\max}_r(\ell)$ can be done a similar way.

Fact II.8. *We have the inequalities*

1) $0 = R^{\min}(1) \le R^{\max}(1) = \lambda$,
2) $\frac{\lambda}{(1+\lambda)^{\Delta}} = R^{\min}(2) \le R^{\max}(2) = \lambda$,
3) $R^{\min}(\ell) \le R^{\min}(\ell+1)$ *and* $R^{\max}(\ell) \ge R^{\max}(\ell+1)$ *for any* $\ell \ge 1$.

Theorem II.9 (Weak Spatial Mixing Implies Strong Spatial Mixing; [Wei06])**.** *Assume $\lambda = (1-\delta)\lambda_c(\Delta)$ for some $0 < \delta < 1$. Then there exist constants $C > 0$ and $0 < \alpha < 1$ such that for every tree T of maximum degree $\le \Delta$ rooted at some $r \in T$, and every level ℓ, we have the bound*

$$|R^{\min}_{T,r}(\ell) - R^{\max}_{T,r}(\ell)| \le C \cdot \alpha^{\ell}$$

Later on in the paper, we will need more precise control over C, α. However, the above result is sufficient for the present discussion

III. The Eigenvalues of the Pairwise Influence Matrix

Our goal in this section is to prove Theorem I.11. In fact, we completely characterize the spectrum of $P_\emptyset$ in terms of the spectrum of Ψ_μ, which immediately implies Theorem I.11.

Theorem III.1. *The spectrum of $P_\emptyset$ (as a multiset) is precisely the union of the spectrum of $\frac{1}{n-1}\Psi_\mu$ (as a multiset), $n-1$ copies of $-\frac{1}{n-1}$, and an eigenvalue of 1 (corresponding to the top eigenvalue of $P_\emptyset$).*

Note that this also immediately implies Ψ_μ has real eigenvalues, since $P_\emptyset$ has real eigenvalues. The rest of the section is devoted to proving Theorem III.1.

The main idea behind the proof is to relate the spectra of $P_\emptyset$ and of $\frac{1}{n-1}\Psi_\mu$ to an intermediate matrix $M_\emptyset$. This matrix $M_\emptyset$ will be built from $P_\emptyset$ by leveraging knowledge of the "trivial" eigenvalues and eigenvectors induced purely by the n-partite structure of X^μ. Towards this, let us first express the entries of $P_\emptyset$ in a nice form. Observe that using Eq. (1), we have that for $i, j \in [n]$,

$$P_\emptyset(i,j) = \frac{w(\{i,j\})}{w(\{i\})} \cdot \mathbf{1}[i \ne j] = \frac{1}{n-1} \Pr[j \mid i] \cdot \mathbf{1}[i \ne j]$$

Similarly, we have the following for all $i, j \in [n]$.

$$P_\emptyset(i, \bar{j}) = \frac{1}{n-1} \Pr[\bar{j} \mid i] \cdot \mathbf{1}[i \neq j]$$
$$P_\emptyset(\bar{i}, j) = \frac{1}{n-1} \Pr[j \mid \bar{i}] \cdot \mathbf{1}[i \neq j]$$
$$P_\emptyset(\bar{i}, \bar{j}) = \frac{1}{n-1} \Pr[\bar{j} \mid \bar{i}] \cdot \mathbf{1}[i \neq j]$$

Now, we compute the stationary distribution of $P_\emptyset$. Define $\pi \in \mathbb{R}^{2n}$ entrywise by $\pi(i) = \frac{1}{n} \Pr[i]$ and $\pi(\bar{i}) = \frac{1}{n} \Pr[\bar{i}]$. It is easy to see that $P_\emptyset$ is reversible w.r.t. π. Hence, π is indeed stationary w.r.t. $P_\emptyset$. For each element $i \in [n]$, define the vectors $\mathbf{1}^i, \pi^i \in \mathbb{R}^{2n}$ by $\mathbf{1}^i = e_i + e_{\bar{i}}$ and $\pi^i = \pi(i) \cdot e_i + \pi(\bar{i}) \cdot e_{\bar{i}}$. In particular, for each $i \in [n]$, $\mathbf{1}^i, \pi^i$ are vectors which are supported on the two entries corresponding to the two different possible assignments of i. We now define our intermediate matrix as

$$M_\emptyset = P_\emptyset - \frac{n}{n-1} \mathbf{1}\pi^\top + \frac{n}{n-1} \sum_{i=1}^{n} \mathbf{1}^i (\pi^i)^\top$$

We prove the following two claims.

Claim III.2 (Relating $P_\emptyset$ and $M_\emptyset$)**.** *The matrix $P_\emptyset$ has eigenvalue 1 with multiplicity (at least) 1, and eigenvalue $-\frac{1}{n-1}$ with multiplicity (at least) $n-1$. These are the "trivial" eigenvalues of $P_\emptyset$. Furthermore, the spectrum of $M_\emptyset$ (as a multiset) is precisely the spectrum of $P_\emptyset$ with all trivial eigenvalues replaced by n copies of 0.*

Claim III.3 (Relating $M_\emptyset$ and $\frac{1}{n-1}\Psi_\mu$)**.** *The spectrum of $M_\emptyset$ (as a multiset) is precisely the union of the spectrum of $\frac{1}{n-1}\Psi_\mu$ (as a multiset) with n additional copies of 0.*

Theorem III.1 then follows as an immediate consequence of these two claims. We now intuitively sketch why Claim III.2 and Claim III.3, leaving the technical details for the full version of the paper [ALO20].

The main idea behind Claim III.2 is that the vectors $\mathbf{1}^i$ form an orthogonal basis for the span of the right eigenvectors of $P_\emptyset$ corresponding to the eigenvalues 1 and $-\frac{1}{n-1}$ (while the π^i form an orthogonal basis of the corresponding left eigenvectors). The essence of the proof lies in the fact that these "trivial" eigenvectors derive purely from the fact that in the weighted graph with vertex set $\{i, \bar{i} : i \in [n]\}$ corresponding to $P_\emptyset$, there are is no edge between the vertices i and $\bar{i}$, for all $i \in [n]$. This is a generalization of the fact that the random walk matrix of any weighted bipartite graph always has eigenvalue -1, purely due to bipartiteness. This was actually already observed in [Opp18]. One should view $M_\emptyset$ is being defined in a way to "zero out" those eigenvalues. For Claim III.3, the intuition is that Ψ_μ may be obtained from $M_\emptyset$ via orthogonal projection.

Remark III.4. These observations generalize in a straightforward fashion to all partite complexes in the sense that for any d-dimensional d-partite weighted complex (X, w) with parts $U_1, \ldots, U_d$, the indicator vectors $\mathbf{1}^{U_1}, \ldots, \mathbf{1}^{U_d}$ are eigenvectors of $P_\emptyset - \frac{d}{d-1}\mathbf{1}\pi^\top$ with eigenvalue $-\frac{1}{d-1}$. This was also observed in [Opp18].

IV. Influence Decoupling in Weitz's Self-Avoiding Walk Tree

In this section, we take a step towards proving Theorem I.13. Specifically, we focus on bounding

$$\sum_{u \in V : u \neq v} |\Psi_\mu(u, v)|$$

where from now on, we take μ to be the distribution corresponding to the hardcore distribution on input graph $G = (V, E)$ with parameter $\lambda > 0$. Here, the relevant uniqueness threshold is given by $\lambda_c(\Delta) = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta}$.

Before we proceed to bound this quantity for general graphs, we note that one can easily deduce an $O(1)$ upper bound for amenable graphs (i.e. graphs such that the balls around any vertex grows subexponentially fast in the radius) in a black-box fashion directly using strong spatial mixing Definition II.2, thus recovering some of the previously known connections between spatial mixing properties of the hardcore distribution, and temporal mixing of the Glauber dynamics [Dye+02; Wei04]. This class of graphs notably includes lattices such as $\mathbb{Z}^d$, but exclude most graphs such as expanders. Thus, nstead of applying strong spatial mixing as a black-box, we revisit its proof, modifying it as needed.

The high-level strategy is to convert this problem on general graphs to bounding a similar quantity for trees. We do this by leveraging the self-avoiding walk tree construction of [Wei06]. However, since a vertex $u \in G$ may have many copies in the corresponding self-avoiding walk tree $T = T_{\text{SAW}}(G, r)$, we need to "decouple" these copies so as to obtain single-vertex influences again.

Definition IV.1 (R-Pseudoinfluence)**.** *Recall that for a fixed tree T rooted at r with boundary condition $p : A \to [0, 1]$ (where A is a subset of vertices not containing r), we write $R^p_{T,r} = \frac{\Pr[r|p]}{1-\Pr[r|p]}$. For a vertex $v \in T$ with $v \neq r$, we define the R-pseudoinfluence of v on the root r by the quantity*

$$\mathcal{R}^v_{T,r} = \max_p \mathcal{R}^{v,p}_{T,r} \qquad \textit{where} \qquad \mathcal{R}^{v,p}_{T,r} = \left| R^{v^0,p}_{T,r} - R^{v^1,p}_{T,r} \right|$$

and the maximum is taken over all partial assignments $p : L_r(\ell(v)) \setminus \{v\} \to [0, 1]$ of marginal values. Again, we drop the subscript T when the tree is clear from context.

Remark IV.2. It was pointed out to us by Zongchen Chen and Eric Vigoda that our notion of R-pseudoinfluence is very related to the notion of "aggregate strong spatial mixing" used in [MS13] to analyze the Glauber dynamics, and in [BCV20] to analyze the Swendsen-Wang dynamics, both for the ferromagnetic Ising model. In fact, it turns out our result also directly implies aggregate strong spatial mixing for arbitrary trees in the uniqueness regime $\lambda < \lambda_c(\Delta)$.

Our first step is to do the decoupling using the R-pseudoinfluence. The second step is to bound the total R-pseudoinfluence of vertices in a tree on the root. These steps are captured in the following two results. We emphasize Lemma IV.3 is generic, and holds for any two-spin system.

Lemma IV.3 (Decoupling). *Consider the hardcore distribution μ on a graph $G = (V, E)$ with parameter $\lambda > 0$. Fix a vertex $r \in G$ and let $T = T_{\mathrm{SAW}}(G, r)$. Then the following inequality holds:*

$$\sum_{v \in G : v \neq r} |\Psi_\mu(v, r)| \leq 2 \sum_{v \in T : v \neq r} \mathcal{R}_r^v$$

In particular, to bound $\sum_{v \in G : v \neq r} |\Psi_\mu(v, r)|$, it suffices to bound $\sum_{v \in T : v \neq r} \mathcal{R}_r^v$ for every tree T of maximum degree $\leq \Delta$ rooted at r. This motivates the next result.

Proposition IV.4 (R-Pseudoinfluence Bound). *Assume λ is up-to-Δ unique with gap $0 < \delta < 1$. Then for every tree T of maximum degree $\leq \Delta$ rooted at r, we have the bound*

$$\sum_{v \in T : v \neq r} \mathcal{R}_r^v \leq \exp(O(1/\delta))$$

Theorem I.13 follows immediately as a consequence. We leave the proof of Lemma IV.3 to the full version of the paper [ALO20]. The proof of Proposition IV.4 is sketched in the following section.

A. R-Pseudoinfluence Decay

Our goal is now to prove Proposition IV.4. To do this, we write

$$\sum_{v \in T : v \neq r} \mathcal{R}_r^v = \sum_{\ell=1}^{\infty} \sum_{v \in L_r(\ell)} \mathcal{R}_r^v$$

Thus, it suffices to bound $\sum_{v \in L_r(\ell)} \mathcal{R}_r^v$ for each level ℓ. We show that this quantity in fact decays exponentially fast as ℓ increases when $\lambda < \lambda_c(\Delta)$. Specifically, to prove Proposition IV.4, we use the following two lemmas, which precisely quantify the decay rate.

Proposition IV.5 (Decay Rate Bound). *Assume λ is up-to-Δ unique with gap $0 < \delta < 1$. Then there exists $\ell_0 = \Theta(1/\delta)$ such that for every tree T of maximum degree $\leq \Delta$ rooted at r and any $\ell > \ell_0$, we have the bound*

$$\frac{\sum_{v \in L_r(\ell)} \mathcal{R}_r^v}{\max_{u \in L_r(\ell - \ell_0)} \left\{ \sum_{v \in L_u(\ell_0)} \mathcal{R}_u^v \right\}} \leq O(1) \cdot \sqrt{1 - \delta}^{\ell - \ell_0}$$

We prove Proposition IV.5 in the next section. Roughly speaking, the reason for the assumption $\ell > \ell_0$ above is that we can exploit spatial mixing to argue that the marginals of the root is independent of the boundary condition at level ℓ, for a large enough ℓ_0; see Section V for more details. For $\ell < \ell_0$ we use the following lemma, whose proof we leave to the full version of the paper [ALO20].

Lemma IV.6 (Trivial "Decay" Rate). *Assume λ up-to-Δ unique with gap $0 < \delta < 1$. Then for any tree T of maximum degree $\leq \Delta$ rooted at r and any $\ell > 0$, we have*

$$\frac{\sum_{v \in L_r(\ell)} \mathcal{R}_r^v}{\max_{u \in L_r(1)} \left\{ \sum_{v \in L_u(\ell - 1)} \mathcal{R}_u^v \right\}} \leq O(1)$$

Furthermore, for the first level, we have the inequality

$$\sum_{v \in L_r(1)} \mathcal{R}_r^v \leq O(1)$$

These two results together immediately imply Proposition IV.4.

V. Bounding the R-Pseudoinfluence Decay: The Potential Method

Our goal in this section is to prove Proposition IV.5. We use the potential method (otherwise known as the message decay argument), which has been successfully used in [LLY12; LLY13; Res+13; SST14; Sin+15] to establish strong spatial mixing all the way up to the uniqueness threshold. We use the following potential function φ, first introduced in [LLY13].

$$\varphi(R) \stackrel{\text{def}}{=} 2 \log(\sqrt{R} + \sqrt{R + 1})$$
$$\Phi(R) \stackrel{\text{def}}{=} \varphi'(R) = \frac{1}{\sqrt{R(R + 1)}}$$

We note that since Φ is continuous, positive, and decreasing, we have φ is continuously differentiable, strictly monotone increasing and concave. One additional feature of this potential function is that it has no dependence on λ or Δ. While it may be comforting to have an explicit expression for φ, all of our proofs rely at most on the explicit expression for Φ, rather than φ. For the derivation and further discussion of this potential function, we refer the reader to [LLY13].

Definition V.1 (φ-Pseudoinfluence). *For a boundary condition $p : A \to [0, 1]$, where A is a subset of vertices not containing r, let $K_r^p = \varphi(R_r^p)$. Again, we define*

$$\mathcal{K}_r^{v,p} \stackrel{\text{def}}{=} \left| K_r^{v^0,p} - K_r^{v^1,p} \right| = \left| \varphi(R_r^{v^0,p}) - \varphi(R_r^{v^1,p}) \right|$$

Define the φ-pseudoinfluence of a vertex v on r as

$$\mathcal{K}_r^v \stackrel{\text{def}}{=} \max_{p : L_r(\ell(v)) \setminus \{v\} \to [0,1]} K_r^{v,p}$$

Finally, we define

$$K_r^{\min}(\ell) \stackrel{\text{def}}{=} \min_{p : L_r(\ell) \to [0,1]} K_r^p = \varphi(R_r^{\min}(\ell))$$
$$K_r^{\max}(\ell) \stackrel{\text{def}}{=} \max_{p : L_r(\ell) \to [0,1]} K_r^p = \varphi(R_r^{\max}(\ell))$$

To control $\sum_{v \in L_r(\ell)} \mathcal{R}_r^v$, it turns out it suffices to control the decay of $\sum_{v \in L_r(\ell)} \mathcal{K}_r^v$ as ℓ increases.

Proposition V.2 (φ-Pseudoinfluence Decay Rate Bound). *Assume λ is up-to-Δ unique with gap $0 < \delta < 1$ (see*

Definition II.3). For $\ell \geq 2$, assume that there exists $\eta \leq 1/2$ such that for all $u \in L_r(1)$, we have the inequality $\left|R_u^{\min}(\ell-1) - R_u^{\max}(\ell-1)\right| \leq \eta$. Then,

$$\frac{\sum_{v \in L_r(\ell)} \mathcal{K}_r^v}{\max_{u \in L_r(1)} \left\{\sum_{v \in L_u(\ell-1)} \mathcal{K}_u^v\right\}} \leq (1+2\eta)^{\Delta+1} \sqrt{1-\delta}.$$

Unfortunately, due to the additional error factor of $(1+2\eta)^{\Delta+1}$, we must control $\eta = \eta(\ell)$. To do this, we leverage a "precise" version of the strong spatial mixing result proved in [Wei06], where the constant is explicit.

Definition V.3. *Define $\eta^* = \frac{R^{\max}(2)}{R^{\min}(2)} \cdot \left|R^{\min}(2) - R^{\max}(2)\right|$. Note by Fact II.8 and the fact that $\lambda \leq O(1/\Delta)$, we have*

$$\eta^* \leq \frac{\lambda}{\frac{\lambda}{(1+\lambda)^\Delta}} \cdot \left|\lambda - \frac{\lambda}{(1+\lambda)^\Delta}\right| \leq O(1/\Delta). \tag{3}$$

Proposition V.4 (Strong Spatial Mixing [Wei06])**.** *Assume that λ is up-to-Δ unique with gap $0 < \delta < 1$. Then for all trees T rooted at r of maximum degree $\leq \Delta$, we have*

$$\left|R_{T,r}^{\min}(\ell) - R_{T,r}^{\max}(\ell)\right| \leq \sqrt{1-\delta}^{\ell-2} \cdot \eta^*$$

A proof of this specific bound is provided in the full version of the paper. With these results in hand, we may deduce Proposition IV.5.

A. Proof Sketch of Proposition IV.5

In order to apply Proposition V.2 and Proposition V.4, we must relate $\mathcal{R}_r^v$ to $\mathcal{K}_r^v$. This is done in the following lemma.

Lemma V.5 (Relating R-Pseudoinfluence to φ-Pseudoinfluence)**.** *Let T be a tree rooted at r. For any $\ell \geq 1$ and any vertex $v \in L_r(\ell)$, we have the bound*

$$\Phi(R^{\max}(\ell)) \cdot \mathcal{R}_r^v \leq \mathcal{K}_r^v \leq \Phi(R^{\min}(\ell)) \cdot \mathcal{R}_r^v.$$

The main idea in the proof of Proposition IV.5 is then to first map the R-pseudoinfluence to the φ-pseudoinfluence (with a loss of $1/\Phi(R^{\max}(\ell))$ using Lemma V.5, apply the φ-pseudoinfluence decay Proposition V.2, and then map back to the R-pseudoinfluence (with a loss of $\Phi(R^{\min}(\ell))$) using Lemma V.5 again. The details are left to the full version of the paper [ALO20].

At this point, all that is left is to prove Proposition V.2 and Proposition V.4. We prove Proposition V.2 and Proposition V.4 in the following subsections.

B. The φ-Pseudoinfluence Decay: Proof Sketch of Proposition V.2

Our goal in this subsection is to prove Proposition V.2. While initially this appears to be a more daunting task, it is made feasible by the fact that the tree recurrence F for R induces a corresponding tree recurrence for K given by

$$K_r^\sigma = (\varphi \circ F \circ \varphi^{-1})(K_u^\sigma : u \in L_r(1)).$$

Using this tree recurrence for K_r^σ, we prove Lemma V.6 and Lemma V.8. Chained together with Lemma V.7, which lies at the heart of the results in [LLY13], we immediately obtain Proposition V.2.

Throughout, we will let $\mathbf{K} = (K_u : u \in L_r(1))$, $\mathbf{K}^{\max}(\ell) = (K_u^{\max}(\ell-1) : u \in L_r(1))$ and $\mathbf{K}^{\min}(\ell) = (K_u^{\min}(\ell-1) : u \in L_r(1))$ denote vectors with $|L_r(1)|$ many entries.

Lemma V.6 (True Decay)**.** *For every λ, and every tree T rooted at r, we have the inequality*

$$\frac{\sum_{v \in L_r(\ell)} \mathcal{K}_r^v}{\max_{u \in L_r(1)} \left\{\sum_{v \in L_u(\ell-1)} \mathcal{K}_u^v\right\}} \leq \sum_{u \in L_r(1)} \max_{\mathbf{K}^{\min}(\ell) \leq \mathbf{K} \leq \mathbf{K}^{\max}(\ell)} \left|\partial_{K_u}(\varphi \circ F \circ \varphi^{-1})(\mathbf{K})\right|.$$

Lemma V.7 (Ideal Decay; [LLY13] Lemmas 12, 13, 14)**.** *Assume λ is up-to-Δ unique with gap $0 < \delta < 1$. Let T be any tree of maximum degree $\leq \Delta$ rooted at r. Then we have the bound*

$$\max_{0 \leq \mathbf{K} \leq \infty} \left\|\nabla(\varphi \circ F \circ \varphi^{-1})(\mathbf{K})\right\|_1 \leq \sqrt{1-\delta}$$

Lemma V.8 (Relating True Decay to Ideal Decay)**.** *Assume $\left|R_u^{\max}(\ell-1) - R_u^{\min}(\ell-1)\right| \leq \eta$ for all $u \in L_r(1)$, where $\eta \leq \frac{1}{2}$. Then for every λ, and every tree T with maximum degree $\leq \Delta$ rooted at r, we have the inequality*

$$\sum_{u \in L_r(1)} \max_{\mathbf{K}^{\min}(\ell) \leq \mathbf{K} \leq \mathbf{K}^{\max}(\ell)} \left|\partial_{K_u}(\varphi \circ F \circ \varphi^{-1})(\mathbf{K})\right| \leq (1+2\eta)^{\Delta+1} \left\|\nabla(\varphi \circ F \circ \varphi^{-1})(\mathbf{K}^{\max}(\ell))\right\|_1$$

VI. Conclusion and Open Problems

In this work we have shown that for the hardcore distribution on independent sets of a graph of maximum degree $\leq \Delta$ with parameter $\lambda = (1-\delta)\lambda_c(\Delta)$, there is a constant $C(\delta)$ such that the Glauber dynamics mixes in $O(n^{C(\delta)})$ steps. While this running time does not have an exponential dependence on $\log \Delta$ as in the correlation decay algorithm of [Wei06], its dependence on δ is significantly worse. Specifically, we have that $C(\delta) \leq \exp(O(1/\delta))$, while the correlation decay algorithm of [Wei06] exhibits a dependence of $C(\delta) \leq O(1/\delta)$.

In a follow-up work by [CLV20], they showed how one can bound the total pairwise influence of the root of a tree on all other vertices. This is in contrast to our analysis, which focuses on the total pairwise influence of all other vertices on the root. They achieve an upper bound of $O(1/\delta)$, giving $O(1/\delta)$-spectral independence and $n^{O(1/\delta)}$ mixing. They also generalize to all antiferromagnetic two-state spin systems. We leave it as an open problem to bound the total pairwise influence on the root by $O(1/\delta)$, and to generalize this to other two-state spin systems.

We show in the full version of the paper that in general, one cannot bound $\lambda_{\max}(\Psi_\mu)$ asymptotically better than $O(1/\delta)$, even for the special case of trees. We do this by showing for the infinite Δ-regular tree that the total pairwise influence on a vertex is $\Theta(1/\delta)$. This shows that in general the best bound on the mixing time of the Glauber dynamics one can hope to achieve by bounding the spectral independence and applying the local-to-global theorem of [AL20] is $n^{O(1/\delta)}$. However, prior results [Vig01; Eft+16] for this problem appear to suggest that $O(C(\delta)n\log n)$ should be possible, which illustrates a key limitation of the current local-to-global results.

Acknowledgements

We thank Vedat Levi Alev and Lap Chi Lau for stimulating discussions, and for sharing with us a preliminary draft of their results [AL20]. We also thank Dorna Abdolazimi, Zongchen Chen, and Chihao Zhang for helpful comments, and for pointing out bugs in earlier versions of this paper. Finally, we thank Eric Vigoda, Lap Chi Lau and the anonymous referees for delivering helpful feedback on a preliminary draft of this paper.

Shayan Oveis Gharan and Kuikui Liu are supported by the NSF grants CCF-1552097 and CCF-1907845, and ONR-YIP grant N00014-17-1-2429.

References

[AGR16] Nima Anari, Shayan Oveis Gharan, and Alireza Rezaei. "Monte Carlo Markov Chain Algorithms for Sampling Strongly Rayleigh Distributions and Determinantal Point Processes". In: ed. by Vitaly Feldman, Alexander Rakhlin, and Ohad Shamir. Vol. 49. Proceedings of Machine Learning Research. Columbia University, New York, New York, USA: PMLR, June 2016, pp. 103–115.

[AL20] Vedat Levi Alev and Lap Chi Lau. "Improved Analysis of Higher Order Random Walks and Applications". In: *Proceedings of the 52nd Annual ACM Symposium on Theory of Computing (STOC)*. 2020.

[ALO20] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. "Spectral Independence in High-Dimensional Expanders and Applications to the Hardcore Model". In: *arXiv preprint arXiv:2001.00303* (2020).

[Ana+19] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. "Log-Concave Polynomials II: High-Dimensional Walks and an FPRAS for Counting Bases of a Matroid". In: *Proceedings of the 51st Annual ACM SIGACT Symposium on the Theory Computing*. STOC 2019. Phoenix, AZ, USA: Association for Computing Machinery, 2019, pp. 1–12. ISBN: 9781450367059. DOI: 10.1145/3313276.3316385.

[Bar16a] Alexander Barvinok. *Combinatorics and Complexity of Partition Functions*. Vol. 30. Springer Algorithms and Combinatorics, 2016.

[Bar16b] Alexander Barvinok. "Computing the Permanent of (Some) Complex Matrices". In: *Found. Comput. Math.* 16.2 (Apr. 2016), pp. 329–342. ISSN: 1615-3375.

[BCV20] Antonio Blanca, Zongchen Chen, and Eric Vigoda. "Swendsen-Wang dynamics for general graphs in the tree uniqueness region". In: *Random Structures and Algorithms* 56 (2020), pp. 373–400.

[Bla+13] Antonio Blanca, David Galvin, Dana Randall, and Prasad Tetali. "Phase Coexistence and Slow Mixing for the Hard-Core Model on $\mathbb{Z}^2$". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*. Ed. by Prasad Raghavendra, Sofya Raskhodnikova, Klaus Jansen, and José D. P. Rolim. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 379–394.

[Bla+19] Antonio Blanca, Yuxuan Chen, David Galvin, Dana Randall, and Prasad Tetali. "Phase Coexistence for the Hard-Core Model on $\mathbb{Z}^2$". In: *Combinatorics, Probability & Computing* 28 (2019), pp. 1–22.

[Che+20] Zongchen Chen, Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. "Rapid Mixing for Colorings via Spectral Independence". In: *arXiv preprint arXiv:2007.08058* (2020).

[CLV20] Zongchen Chen, Kuikui Liu, and Eric Vigoda. "Rapid Mixing of Glauber Dynamics up to Uniqueness via Contraction". In: *arXiv preprint arXiv:2004.09083* (2020).

[DFJ02] Martin Dyer, Alan Frieze, and Mark Jerrum. "On Counting Independent Sets in Sparse Graphs". In: *SIAM J. Comput.* 31.5 (May 2002), pp. 1527–1541. ISSN: 0097-5397.

[DG00] Martin Dyer and Catherine Greenhill. "On Markov Chains for Independent Sets". In: *Journal of Algorithms* 35.1 (2000), pp. 17–49. ISSN: 0196-6774.

[DGJ09] Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. "Matrix Norms and Rapid Mixing for Spin Systems". In: *The Annals of Applied Probability* 19.1 (Feb. 2009), pp. 71–107.

[DK17] I. Dinur and T. Kaufman. "High Dimensional Expanders Imply Agreement Expanders". In: *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*. 2017, pp. 974–985.

[Dob70] Roland Lvovich Dobrushin. "Prescribing a System of Random Variables by Conditional Distributions". In: *Theory of Probability and its Applications* 15.3 (1970), pp. 458–486.

[DS85a] R. L. Dobrushin and S. B. Shlosman. "Completely Analytical Gibbs Fields". In: *Statistical Physics and Dynamical Systems: Rigorous Results*. Ed. by J. Fritz, A. Jaffe, and D. Szász. Boston, MA: Birkhäuser Boston, 1985, pp. 371–403. ISBN: 978-1-4899-6653-7.

[DS85b] R. L. Dobrushin and S. B. Shlosman. "Constructive Criterion for the Uniqueness of Gibbs Field". In: *Statistical Physics and Dynamical Systems: Rigorous Results*. Ed. by J. Fritz, A. Jaffe, and D. Szász. Boston, MA: Birkhäuser Boston, 1985, pp. 347–370. ISBN: 978-1-4899-6653-7.

[DS87] R. L. Dobrushin and S. B. Shlosman. "Completely analytical interactions: Constructive description". In: *Journal of Statistical Physics* 46.5 (Mar. 1987), pp. 983–1014.

[DS91] Persi Diaconis and Daniel Stroock. "Geometric bounds for eigenvalues of Markov chains". In: *The Annals of Applied Probability* (1991), pp. 36–61.

[Dye+02] Martin Dyer, Alistair Sinclair, Eric Vigoda, and Dror Weitz. "Mixing in Time and Space for Lattice Spin Systems: A Combinatorial View". In: *Randomization and Approximation Techniques in Computer Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 149–163.

[Eft+16] Charilaos Efthymiou, Thomas P. Hayes, Daniel Štefankovič, Eric Vigoda, and Yitong Yin. "Convergence of MCMC and Loopy BP in the Tree Uniqueness Region for the Hard-Core Model". In: *FOCS*. 2016, pp. 704–713.

[Fen+20] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. "Rapid mixing from spectral independence beyond the Boolean domain". In: *arXiv preprint arXiv:2007.08091* (2020).

[FM92] Tomás Feder and Milena Mihail. "Balanced matroids". In: *STOC*. 1992, pp. 26–38.

[Gal+14] Andreas Galanis, Qi Ge, Daniel Štefankovič, Eric Vigoda, and Linji Yang. "Improved Inapproximability Results for Counting Independent Sets in the Hard-core Model". In: *Random Struct. Algorithms* 45.1 (Aug. 2014), pp. 78–110. ISSN: 1042-9832.

[Gre00] Catherine S. Greenhill. "The complexity of counting colourings and independent sets in sparse graphs and hypergraphs". In: *computational complexity* 9 (2000), pp. 52–72.

[GŠV15] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. "Inapproximability for Antiferromagnetic Spin Systems in the Tree Nonuniqueness Region". In: *J. ACM* 62.6 (Dec. 2015), 50:1–50:60. ISSN: 0004-5411.

[GŠV16] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. "Inapproximability of the Partition Function for the Antiferromagnetic Ising and Hard-Core Models". In: *Combinatorics, Probability and Computing* 25.4 (2016), pp. 500–559.

[Hay06] Thomas P. Hayes. "A simple condition implying rapid mixing of single-site dynamics on spin systems". In: *FOCS* (2006), pp. 39–46.

[HV05] Tom Hayes and Eric Vigoda. "Coupling with the Stationary Distribution and Improved Sampling for Colorings and Independent Sets". In: *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '05. Vancouver, British Columbia: Society for Industrial and Applied Mathematics, 2005, pp. 971–979. ISBN: 0-89871-585-7.

[JS89] M. Jerrum and Alistair Sinclair. "Approximating the Permanent". In: *SIAM J. Comput.* 18.6 (Dec. 1989), pp. 1149–1178. ISSN: 0097-5397.

[Kel85] F. P. Kelly. "Stochastic Models of Computer Communication Systems". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 47.3 (1985), pp. 379–395.

[KM17] Tali Kaufman and David Mass. "High Dimensional Random Walks and Colorful Expansion". In: *ITCS*. 2017, 4:1–4:27.

[KO18] Tali Kaufman and Izhar Oppenheim. "High Order Random Walks: Beyond Spectral Gap". In: *APPROX/RANDOM*. 2018, 47:1–47:17.

[LLY12] Liang Li, Pinyan Lu, and Yitong Yin. "Approximate Counting via Correlation Decay in Spin Systems". In: *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2012, pp. 922–940.

[LLY13] Liang Li, Pinyan Lu, and Yitong Yin. "Correlation Decay Up to Uniqueness in Spin Systems". In: *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2013, pp. 67–84.

[LV97] Michael Luby and Eric Vigoda. "Approximately Counting Up to Four (Extended Abstract)". In: *Proceedings of the Twenty-ninth Annual ACM Symposium on Theory of Computing*. STOC '97. El Paso, Texas, USA: ACM, 1997, pp. 682–687. ISBN: 0-89791-888-6.

[LV99] Michael Luby and Eric Vigoda. "Fast convergence of the Glauber dynamics for sampling independent sets". In: *Random Struct. Algorithms* 15 (1999), pp. 229–241.

[MS08] Elchanan Mossel and Allan Sly. "Rapid Mixing of Gibbs Sampling on Graphs That Are Sparse on Average". In: *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '08. San Francisco, California: Society for Industrial and Applied Mathematics, 2008, pp. 238–247.

[MS13] Elchanan Mossel and Allan Sly. "Exact thresholds for Ising–Gibbs samplers on general graphs". In: *Annals of Probability* 41.1 (2013), pp. 294–328.

[MSW03] F. Martinelli, A. Sinclair, and D. Weitz. "The Ising model on trees: boundary conditions and mixing time". In: *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.* Oct. 2003, pp. 628–639.

[MSW04] Fabio Martinelli, Alistair Sinclair, and Dror Weitz. "Fast Mixing for Independent Sets, Colorings and Other Models on Trees". In: *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '04. New Orleans, Louisiana, 2004, pp. 456–465. ISBN: 0-89871-558-X.

[MWW07] Elchanan Mossel, Dror Weitz, and Nicholas C. Wormald. "On the hardness of sampling independent sets beyond the tree threshold". In: *Probability Theory and Related Fields* 143 (2007), pp. 401–439.

[Opp18] Izhar Oppenheim. "Local spectral expansion approach to high dimensional expanders part I: Descent of spectral gaps". In: *Discrete and Computational Geometry* 59.2 (2018), pp. 293–330.

[PR17] Viresh Patel and Guus Regts. "Deterministic Polynomial-Time Approximation Algorithms for Partition Functions and Graph Polynomials". In: *SIAM Journal on Computing* 46 (2017), pp. 1893–1919.

[PR19] Han Peters and Guus Regts. "On a conjecture of Sokal concerning roots of the independence polynomial". In: *The Michigan Mathematical Journal* 68.1 (2019), pp. 33–55.

[Res+13] Ricardo Restrepo, Jinwoo Shin, Prasad Tetali, Eric Vigoda, and Linji Yang. "Improved mixing condition on the grid for counting and sampling independent sets". In: *Probability Theory and Related Fields* 156.1 (June 2013), pp. 75–99.

[Sin+15] Alistair Sinclair, Piyush Srivastava, Daniel Štefankovič, and Yitong Yin. "Spatial Mixing and the Connective Constant: Optimal Bounds". In: *Proceedings of the Twenty-sixth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '15. San Diego, California: Society for Industrial and Applied Mathematics, 2015, pp. 1549–1563.

[Sly10] Allan Sly. "Computational Transition at the Uniqueness Threshold". In: *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 2010, pp. 287–296.

[Sri14] Piyush Srivastava. "Counting and Correlation Decay in Spin Systems". PhD thesis. EECS Department, University of California, Berkeley, Aug. 2014.

[SS05] Alexander D. Scott and Alan D. Sokal. "The Repulsive Lattice Gas, the Independent-Set Polynomial, and the Lovász Local Lemma". In: *Journal of Statistical Physics* 118.5 (2005), pp. 1151–1261.

[SS14] Allan Sly and Nike Sun. "The Computational Hardness of Counting in Two-Spin Models on d-Regular Graphs". In: *The Annals of Probability* 42.6 (2014), pp. 2383–2416.

[SST14] Alistair Sinclair, Piyush Srivastava, and Marc Thurley. "Approximation Algorithms for Two-State Anti-Ferromagnetic Spin Systems on Bounded Degree Graphs". In: *Journal of Statistical Physics* 155.4 (2014), pp. 666–686.

[SSY13] Alistair Sinclair, Piyush Srivastava, and Yitong Yin. "Spatial Mixing and Approximation Algorithms for Graphs with Bounded Connective Constant". In: *Proceedings of the 2013 IEEE 54th Annual Symposium on Foundations of Computer Science*. FOCS '13. Washington, DC, USA: IEEE Computer Society, 2013, pp. 300–309. ISBN: 978-0-7695-5135-7.

[Vad02] Salil P. Vadhan. "The Complexity of Counting in Sparse, Regular, and Planar Graphs". In: *SIAM J. Comput.* 31.2 (Feb. 2002), pp. 398–427. ISSN: 0097-5397.

[Vad95] Salil P. Vadhan. *The Complexity of Counting*. Undergraduate thesis, Harvard University. Available from https://people.seas.harvard.edu/ salil/research/ugthesis-abs.html. 1995.

[Val79] Leslie G. Valiant. "The Complexity of Enumeration and Reliability Problems". In: *SIAM J. Comput.* 8 (1979), pp. 410–421.

[Vig01] Eric Vigoda. "A Note on the Glauber Dynamics for Sampling Independent Sets". In: *Electr. J. Comb.* 8 (2001).

[VVY13] Juan C. Vera, Eric Vigoda, and Linji Yang. "Improved Bounds on the Phase Transition for the Hard-Core Model in 2-Dimensions". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*. Ed. by Prasad Raghavendra, Sofya Raskhodnikova, Klaus Jansen, and José D. P. Rolim. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 699–713.

[Wei04] Dror Weitz. "Mixing in Time and Space for Discrete Spin Systems". AAI3147046. PhD thesis. 2004. ISBN: 0-496-05573-9.

[Wei06] Dror Weitz. "Counting Independent Sets Up to the Tree Threshold". In: *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)*. 2006, pp. 140–149.