Lecture 16: Correlation Decay from Spectral Independence

November 7, 2023

In this lecture, we return to correlation decay and show that it implies fast mixing of Glauber dynamics. Our main case study is again the hardcore model.

1 Optimal Mixing for the Hardcore Model in Uniqueness

Let $G = (V, E)$ be a graph of maximum degree Δ, and let $\lambda \geq 0$ be a parameter. Recall the associated hardcore Gibbs measure is described by

$$\mu_{G,\lambda}(I) \propto \lambda^{|I|}, \qquad \forall I \subseteq V \text{ independent},$$

with partition function

$$Z_G(\lambda) \stackrel{\mathsf{def}}{=} \sum_{I \subseteq V \text{ independent}} \lambda^{|I|}.$$

We previously saw that when $\lambda < \lambda_c(\Delta)$, where $\lambda_c(\Delta) = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$ is the *uniqueness threshold* (w.r.t. the infinite Δ-regular tree), the measure $\mu_{G,\lambda}$ exhibits correlation decay (more precisely, *strong spatial mixing*). Furthermore, in this regime, we have a FPTAS for estimating $Z_G(\lambda)$ based on Weitz's algorithm [Wei06],[1] and this counting problem becomes NP-hard for $\lambda > \lambda_c(\Delta)$ [Sly10].

Theorem 1.1 ([CLV21]; building on [ALO21; CLV20])**.** *Suppose $\lambda \leq (1-\delta)\lambda_c(\Delta)$. Then for every graph $G = (V, E)$ of maximum degree Δ, the hardcore Gibbs measure $\mu_{G,\lambda}$ is $O(1/\delta)$-spectrally independent.*[2] *In particular, Glauber dynamics for $\mu_{G,\lambda}$ mixes in $O_{\Delta,\delta}(n \log n)$-steps.*

Remark 1. Note that the spectral independence result does *not* require the degree Δ to be bounded by a universal constant. However, the optimal mixing time does. With the techniques we have seen so far, we only get mixing time $n^{O(1/\delta)}$ if we don't have the bounded-degree assumption. Later, we'll see a sharper technique which will completely remove this dependence on Δ.

Throughout, we'll write μ instead of $\mu_{G,\lambda}$ since G, λ will be clear from context. As is typical in many proofs of spectral independence, we will bound $\|\Psi_\mu\|_{\ell_\infty \to \ell_\infty} \leq O(1/\delta)$ instead. In particular, we aim to show that for all $r \in V$,

$$\sum_{v \in V} |\Psi_\mu(r \to v)| \leq O(1/\delta).$$

As a warm-up, let us see how this follows directly from correlation decay when the graph $G = (V, E)$ is *amenable*, i.e. the balls in G w.r.t. graph distance grow subexponentially in the radius. For

[1] [PR19] also showed that Barvinok's polynomial interpolation algorithm works in this regime.

[2] Note that one convenient closure property of the hardcore model is that every conditional measure of $\mu_{G,\lambda}$ induced via pinning is another hardcore model on a subgraph of G. Hence, we can simplify the statement to only address spectral independence for $\mu_{G,\lambda}$ itself.

instance, if $G = \mathbb{Z}^d$, then

$$\begin{aligned}
\sum_{v \in V} |\Psi_\mu(r \to v)| &= \sum_{\ell=0}^{\infty} \sum_{v:\mathrm{dist}(r,v)=\ell} |\Psi_\mu(r \to v)| \\
&\leq \sum_{\ell=0}^{\infty} (1 - O(\delta))^\ell \cdot \#\{v \in V : \mathrm{dist}(r,v) = \ell\} && \text{(Correlation Decay)} \\
&\leq \sum_{\ell=0}^{\infty} (1 - O(\delta))^\ell \cdot O\left(\ell^d\right) && \text{(Growth of balls in } G = \mathbb{Z}^d\text{)} \\
&\leq O_d(1/\delta). && \text{(Assuming the dimension-}d\text{ is fixed)}
\end{aligned}$$

This recovers classical coupling-based results connecting temporal mixing of local Markov chains like Glauber dynamics, and spatial mixing of μ [Dye+04; Wei04]. However, even putting aside the fact that the bound additionally depends on the degree Δ, this argument definitely breaks down for e.g. expanders, which constitute most graphs. Note that here, we have *not* used the full power of strong spatial mixing, which concerns the influence of a (potentially large) set of vertices; we have only used the fact that vertex-to-vertex correlations are exponentially small in their distance. While we do not know how to use strong spatial mixing directly to prove Theorem 1.1 for general bounded-degree graphs, we can still open up the proof of strong spatial mixing to establish spectral independence. In particular, we will use contraction of the tree recursion to remove the extra factor of $\#\{v \in V : \mathrm{dist}(r,v) = \ell\}$.

Proposition 1.2 (Total Influence Decay)**.** *Suppose $\lambda \leq (1-\delta)\lambda_c(\Delta)$. Then there exists an absolute constant $C > 0$ such that for every graph $G = (V,E)$ of maximum degree Δ, every $r \in V$, and every $\ell \in \mathbb{N}$, we have the estimate*

$$\sum_{v:\mathrm{dist}(r,v)=\ell} |\Psi_\mu(r \to v)| \leq C \cdot (1 - O(\delta))^\ell.$$

By the arguments we have already seen, Theorem 1.1 is an immediate consequence of Proposition 1.2. The rest of this note is devoted to establishing Proposition 1.2.

2 Total Influence Decay via Contractive Tree Recursions

What does it mean for the influence of r onto v to be small? It means that the marginal distribution of v is *stable* w.r.t. pinning r; whether or not you pin r to be in/out, the marginal distribution of r roughly stays the same. This stability viewpoint suggests a way to certify that the influence small.

Theme 2.1. *We can* "witness" *that $\Psi_\mu(r \to v)$ is small by exhibiting an algorithm for exactly computing the marginals of v which is* "stable" *w.r.t. its input. Note that this algorithm need not be efficient.*

Ultimately the existence of such a "stable" algorithm must depend on whether or not $\lambda < \lambda_c(\Delta)$. In light of our proof of strong spatial mixing, a natural choice of this algorithm is the tree recursion for the hardcore model, combined with the self-avoiding walk tree gadget to make it applicable to general graphs. Stability is then guaranteed by contraction of the recursion.

To formalize all of this, the key is to interpret influences as *derivatives* of the marginals (although for technical reasons, we will instead consider the marginal *ratios*). This is essentially the same as the classical fact that the second-order derivatives of the log-partition function yield covariances. To see this, it is again best to switch to the multivariate perspective.

Lemma 2.2. *For a vector of vertex-dependent fugacities $\boldsymbol{\lambda} \in \mathbb{R}^V_{\geq 0}$, define*

$$R_{G,r}(\boldsymbol{\lambda}) \overset{\mathsf{def}}{=} \frac{\lambda_r \cdot Z_{G-N[r]}(\boldsymbol{\lambda})}{Z_{G-r}(\boldsymbol{\lambda})} = \frac{\Pr_{I \sim \mu_{G,\boldsymbol{\lambda}}}[r \in I]}{\Pr_{I \sim \mu_{G,\boldsymbol{\lambda}}}[r \notin I]}. \tag{1}$$

Then

$$\Psi_\mu(r \to v; \boldsymbol{\lambda}) = \partial_{\log \lambda_v} \log R_{G,r}(\boldsymbol{\lambda}), \tag{2}$$

where we have highlighted the dependence of $\Psi_\mu(r \to v)$ on $\boldsymbol{\lambda}$.

Since Eq. (2) involves the ratios $R_{G,r}(\boldsymbol{\lambda})$, let us write down the analog of the tree recursion. This will be our algorithm for computing the marginal ratios.

Lemma 2.3 (Tree Recursion for Ratios)**.**

$$F_{d,\lambda}(R_1,\dots,R_d) \stackrel{\mathsf{def}}{=} \lambda \prod_{i=1}^{d} \frac{1}{1+R_i}. \tag{3}$$

To illustrate Theme 2.1, let us first prove a very special but representative case of Proposition 1.2.

Proof of Proposition 1.2 for Trees when $\lambda \le \frac{1-\delta}{\Delta-1}$. Suppose G is a tree T of maximum degree Δ, and let r be an arbitrarily chosen root vertex. Let $u_1,\dots,u_d$ be the children of r, with corresponding rooted subtrees $T_i = T_{u_i}$. Then

$$K_{T,r}(\boldsymbol{\lambda}) = G_{d,\lambda}\left(K_{T_1,u_1}(\boldsymbol{\lambda}),\dots,K_{T_d,u_d}(\boldsymbol{\lambda})\right), \tag{4}$$

where we use the change of variables $K_{T,r}(\boldsymbol{\lambda}) = \log R_{T,r}(\boldsymbol{\lambda})$ in keeping with Lemma 2.2, and

$$G_{d,\lambda}(\boldsymbol{K}) \stackrel{\mathsf{def}}{=} \log F_{d,\lambda}\left(e^{K_1},\dots,e^{K_d}\right)$$

gives the induced tree recursion for these new variables. Since we're assuming the stronger bound $\lambda \le \frac{1-\delta}{\Delta-1}$, this composition $G_{d,\lambda}$ is a contraction in the following sense.

Claim 2.4. *Suppose $\lambda \le \frac{1-\delta}{\Delta-1}$. Then for every $1 \le d \le \Delta-1$, and every $\boldsymbol{K} \in \mathbb{R}^d$ such that $K_i \le \log\lambda$ for all $i \in [d]$,*

$$\left\|\nabla G_{d,\lambda}(\boldsymbol{K})\right\|_1 \le 1 - O(\delta). \tag{5}$$

This is our stability property; for completeness, we provide a proof of this contraction at the end of this section. The restriction $K_i \le \log\lambda$ is not an issue, because a direct corollary of Lemma 2.3 and monotonicity of $F_{d,\lambda}$ is that $R_{G,r}(\boldsymbol{\lambda}) \le \lambda_r$ for every *unpinned* vertex r.

Now fix a vertex $v \in T$, and let u_i denote the unique child of r such that v is in the subtree $T_i = T_{u_i}$. Then a direct consequence of the Chain Rule and Lemma 2.2 is that

$$\begin{aligned}
\Psi_T(r \to v) &= \partial_{\log\lambda_v} K_{T,r}(\boldsymbol{\lambda}) \\
&= \sum_{j=1}^{d} \partial_{K_j} G_{d,\lambda}(\boldsymbol{K}) \cdot \partial_{\log\lambda_v} K_{T_j,u_j}(\boldsymbol{\lambda}) \\
&= \partial_{K_i} G_{d,\lambda}(\boldsymbol{K}) \cdot \Psi_{T_i}(u_i \to v).
\end{aligned}$$

It follows that

$$\begin{aligned}
\sum_{v\in T:\mathrm{dist}(r,v)=\ell} \left|\Psi_T(r\to v)\right| &= \sum_{j=1}^{d} \left|\partial_{K_j} G_{d,\lambda}(\boldsymbol{K})\right| \sum_{v\in T_j:\mathrm{dist}(u_j,v)=\ell-1} \left|\Psi_{T_j}(u_j \to v)\right| \\
&\le \left\|\nabla G_{d,\lambda}(\boldsymbol{K})\right\|_1 \cdot \max_{j=1,\dots,d} \left\{ \sum_{v\in T_j:\mathrm{dist}(u_j,v)=\ell-1} \left|\Psi_{T_j}(u_j\to v)\right| \right\} \\
&\hspace{20em}\text{(Hölder's Inequality)} \\
&\le \cdots \qquad\qquad \text{(Induct on subtrees } T_1,\dots,T_d\text{)} \\
&\le \left(\max_{1\le d\le \Delta-1} \sup_{\boldsymbol{K}} \left\|\nabla G_{d,\lambda}(\boldsymbol{K})\right\|_1 \right)^{\ell} \\
&\le (1-O(\delta))^{\ell}. \qquad\qquad \text{(Contraction of } G_{d,\lambda}\text{)}
\end{aligned}$$

Again, in the final step, we are justified in using contraction of $G_{d,\lambda}$ since we only need to consider unpinned vertices in T. ☐

Proof of Claim 2.4. We have

$$G_{d,\lambda}(\boldsymbol{K}) = \log\lambda - \sum_{i=1}^{d} \log\left(1 - e^{K_i}\right),$$

and so

$$\|\nabla G_{d,\lambda}(\boldsymbol{K})\|_1 = \sum_{i=1}^{d} \frac{e^{K_i}}{1+e^{K_i}}.$$

Since $K_i \le \log\lambda$ and $\lambda \le \frac{1-\delta}{\Delta-1}$, the claim follows. $\square$

3 The Computation Tree and the Potential Method

Now let us consider the general case. For a general graph, we can still use the tree recursion, except we need to be careful about the inputs to $F_{d,\lambda}$. This was handled by the computation tree (or self-avoiding walk tree) we saw previously.

Theorem 3.1 (Computation Tree for Ratios; [Wei06])**.** *Let $G = (V, E)$ be an arbitrary graph, and let $r \in V$ be an arbitrary vertex with neighbors $u_1, \dots, u_d$ ordered arbitrarily. Then*

$$R_{G,r} = F_{d,\lambda}(R_{G_1,u_1}, \dots, R_{G_d,u_d}),$$

where for all $k = 1, \dots, d$, the graph G_k is obtained from G by deleting r and $u_1, \dots, u_{k-1}$. In particular, there exists a tree $T = T_{\mathsf{SAW}}(G, r)$ of maximum degree Δ rooted at r, whose vertices correspond to pairs (H, u) where H is a subgraph of G and $u \in V_G$, such that $R_{G,r} = R_{T,r}$.

Remark 2. Using the Chain Rule, one can also deduce from this that for every vertex $r \in G$, if $T = T_{\mathsf{SAW}}(G, r)$ denotes the corresponding computation tree/self-avoiding walk tree, then

$$\sum_{v\in V_G} |\Psi_G(r \to v)| \le \sum_{\hat{v}\in V_T} |\Psi_T(r \to \hat{v})|.$$

The remaining issue is getting up to the true critical threshold $\lambda_c(\Delta)$. For this, the contraction from Claim 2.4 isn't strong enough. But as we saw previously, we can use a different potential function.

Proposition 3.2 (Contraction; [LLY13])**.** *Define the potential function $\varphi : \mathbb{R}_{\ge 0} \cup \{+\infty\} \to \mathbb{R}_{\ge 0} \cup \{+\infty\}$ implicitly by its derivative $\varphi'(R) = \Phi(R) \overset{\text{def}}{=} \frac{1}{\sqrt{R(1+R)}}$,*[3] *and consider the modified tree recursion*

$$G_{d,\lambda}(\boldsymbol{K}) \overset{\text{def}}{=} \varphi\left(F_{d,\lambda}\left(\varphi^{-1}(K_1), \dots, \varphi^{-1}(K_d)\right)\right)$$

in the variables $\boldsymbol{K} = \varphi(\boldsymbol{R})$. If $\lambda \le (1-\delta)\lambda_c(\Delta)$, then for every $1 \le d \le \Delta - 1$, $G_{d,\lambda}$ is a contraction in the sense that

$$\|\nabla G_{d,\lambda}(\boldsymbol{K})\|_1 \le 1 - O(\delta), \qquad \forall \boldsymbol{K} \in (\mathbb{R}_{\ge 0} \cup \{+\infty\})^d.$$

Even though Theorem 3.1 and Proposition 3.2 are now written using the marginal ratios, their proofs are exactly the same as what we did previously. Hence, we omit them for brevity.

Proof of Proposition 1.2. We follow the same proof as what we did previously, replacing the logarithmic potential with the new potential φ, and making use of the computation tree/self-avoiding

[3] For concreteness, we get that $\varphi(R) = 2\,\mathrm{arcsinh}\left(\sqrt{R}\right)$, although we will not need the true form of φ here.

walk tree from Theorem 3.1.

$$
\begin{aligned}
&\sum_{v:\mathrm{dist}_G(r,v)=\ell} |\Psi_G(r \to v)| \\
&= \sum_{v:\mathrm{dist}_G(r,v)=\ell} |\partial_{\log \lambda_v} R_{G,r}(\boldsymbol{\lambda})| && \text{(Lemma 2.2)} \\
&= \frac{1}{|\Phi(R_{G,r}(\boldsymbol{\lambda}))|} \sum_{v:\mathrm{dist}_G(r,v)=\ell} |\partial_{\log \lambda_v} K_{G,r}(\boldsymbol{\lambda})| && \text{(Chain Rule)} \\
&= \frac{1}{|\Phi(R_{G,r}(\boldsymbol{\lambda}))|} \sum_{j=1}^{d} |\partial_j G_{d,\lambda}(\boldsymbol{K})| \sum_{v} \left|\partial_{\log \lambda_v} K_{G_j,u_j}(\boldsymbol{\lambda})\right| && \text{(Theorem 3.1)} \\
&\leq \frac{1}{|\Phi(R_{G,r}(\boldsymbol{\lambda}))|} \cdot \|\nabla G_{d,\lambda}(\boldsymbol{K})\|_1 \cdot \max_{j=1,\dots,d} \left\{ \sum_{v} \left|\partial_{\log \lambda_v} K_{G_j,u_j}(\boldsymbol{\lambda})\right| \right\} && \text{(Hölder's Inequality)} \\
&\leq \cdots && \text{(Induction)} \\
&\leq \frac{1}{|\Phi(R_{G,r}(\boldsymbol{\lambda}))|} \cdot \left(\max_{1 \leq d \leq \Delta-1} \sup_{\boldsymbol{K}} \|\nabla G_{d,\lambda}(\boldsymbol{K})\|_1 \right)^{\ell} \cdot \max_{(H,v):\mathrm{dist}_G(r,v)=\ell} \underbrace{\left|\partial_{\log \lambda_v} K_{H,v}(\boldsymbol{\lambda})\right|}_{=|\Phi(R_{H,v}(\boldsymbol{\lambda}))|\cdot|\Psi_H(v\to v)|} \\
&= \underbrace{\frac{\max_{(H,v):\mathrm{dist}_G(r,v)=\ell} \Phi(R_{H,v}(\boldsymbol{\lambda}))}{|\Phi(R_{G,r}(\boldsymbol{\lambda}))|}}_{(A)} \cdot \left(\max_{1 \leq d \leq \Delta-1} \sup_{\boldsymbol{K}} \|\nabla G_{d,\lambda}(\boldsymbol{K})\|_1 \right)^{\ell} \\
&\leq (A) \cdot (1 - O(\delta))^{\ell}. && \text{(Contraction (see Proposition 3.2))}
\end{aligned}
$$

All that remains is to bound (A) by a universal constant, which we do using marginal bounds.

Claim 3.3. *Suppose $\lambda < \lambda_c(\Delta)$. Then for every graph H of maximum degree Δ and every vertex $v \in H$, we have $C \cdot \lambda \leq R_{H,v}(\lambda \mathbf{1}) \leq \lambda$ for some universal constant $C > 0$. In particular, $(A) \leq O(1)$ independent of Δ, λ.*

Proof. The upper bound $R_{H,v}(\boldsymbol{\lambda}) \leq \lambda$ is immediate, e.g. by using monotonicity of the tree recursion Eq. (3) and Theorem 3.1. Now let $u_1, \dots, u_d$ be the neighbors of v of which there are $d \leq \Delta$ many. Applying this upper bound to the marginal ratios for these neighbors yields the lower bound $R_{H,v}(\boldsymbol{\lambda}) \geq \lambda(1+\lambda)^{-d}$. If we assume $\lambda \leq \lambda_c(\Delta)$, which is at most $O(1/\Delta)$, we get that $(1+\lambda)^{-d} \geq \Omega(1)$. The claim follows. □

□

References

[ALO21] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. "Spectral Independence in High-Dimensional Expanders and Applications to the Hardcore Model". In: *SIAM Journal on Computing* 0.0 (2021). Special Section FOCS 2020, FOCS20-1-FOCS20–37. DOI: 10.1137/20M1367696. eprint: https://doi.org/10.1137/20M1367696 (cit. on p. 1).

[CLV20] Zongchen Chen, Kuikui Liu, and Eric Vigoda. "Rapid Mixing of Glauber Dynamics up to Uniqueness via Contraction". In: *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*. FOCS 2020. 2020, pp. 1307–1318 (cit. on p. 1).

[CLV21] Zongchen Chen, Kuikui Liu, and Eric Vigoda. "Optimal Mixing of Glauber Dynamics: Entropy Factorization via High-Dimensional Expansion". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2021. Virtual, Italy: Association for Computing Machinery, 2021, pp. 1537–1550 (cit. on p. 1).

[Dye+04] Martin Dyer, Alistair Sinclair, Eric Vigoda, and Dror Weitz. "Mixing in time and space for lattice spin systems: A combinatorial view". In: *Random Structures & Algorithms* 24 (4 2004), pp. 461–479 (cit. on p. 2).

[LLY13] Liang Li, Pinyan Lu, and Yitong Yin. "Correlation Decay Up to Uniqueness in Spin Systems". In: *Proceedings of the Twenty-fourth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '13. New Orleans, Louisiana: Society for Industrial and Applied Mathematics, 2013, pp. 67–84. ISBN: 978-1-611972-51-1 (cit. on p. 4).

[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 (cit. on p. 1).

[Sly10] Allan Sly. "Computational Transition at the Uniqueness Threshold". In: *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*. 2010, pp. 287–296. DOI: 10.1109/FOCS.2010.34 (cit. on p. 1).

[Wei04] Dror Weitz. "Mixing in Time and Space for Discrete Spin Systems". AAI3147046. PhD thesis. University of California, Berkeley, 2004. ISBN: 0-496-05573-9 (cit. on p. 2).

[Wei06] Dror Weitz. "Counting Independent Sets Up to the Tree Threshold". In: *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)*. Seattle, WA, USA, 2006, pp. 140–149 (cit. on pp. 1, 4).