Down-up walks and spectral independence

USTC, Hefei, China, Jan 17, 2025
Lecturer: Weiming Feng (ETH Zürich)

In this lecture, we will discuss some basic knowledge about down-up walks [ALOV19] and spectral independence [ALO20]. The spectral independence is a new technique for analyzing the mixing time of Glauber dynamics and, more generally, down-up walks. We cannot cover all the details in this lecture, but one can refer to the resources on the course webpage for more details.

1 Glauber dynamics and down-up walks

Let π be a distribution over $[q]^V$ with support $\Omega \subseteq [q]^V$. We can think π as the uniform distribution over proper q-colorings of a graph $G = (V, E)$. In each step, given the current state $X \in \Omega$, the Glauber dynamics picks a vertex $v \in V$ uniformly at random and resamples X_v from the conditional marginal distribution $\pi_v^{X_{V-v}}$, where the notation $\pi_v^{X_{V-v}}$ means $\pi_v(\cdot \mid X_{V-v})$, which is the marginal distribution on v conditional on X_{V-v}. We use $P \in \Omega \times \Omega \to \mathbb{R}$ to denote the transition matrix of Glauber dynamics. One transition of Glauber dynamics can be decomposed into the following down-walk step and up-walk step. Let $X \in \Omega$ denote the current state.

- **Down-walk step:** Pick a vertex $v \in V$ uniformly at random and remove the value of X_v to obtain a *partial* configuration X_{V-v}. Formally, define the set of partial configurations

$$\Omega_{n-1} = \{\sigma \in [q]^S \mid S \subseteq V \land |S| = n-1 \land \pi_S(\sigma) > 0\},$$

where π_S denote the marginal distribution on S projected from π. Here, $\pi_S(\sigma) > 0$ means σ can be extended to a valid full configuration in Ω. In Ω_{n-1}, we use the index $n-1$ to emphasize the size of the partial configuration is $n-1$. Also define

$$\Omega_n = \Omega.$$

The down walk transition matrix $D : \Omega_n \times \Omega_{n-1} \to \mathbb{R}$ is defined by for any $\sigma \in \Omega_n$, any $\tau \in \Omega_{n-1}$ such that τ is a partial configuration on $S \subseteq V$,

$$D(\sigma, \tau) = \begin{cases} \frac{1}{n} & \text{if } \sigma_S = \tau, \\ 0 & \text{otherwise.} \end{cases}$$

- **Up-walk step:** In the up-walk steps, we go from Ω_{n-1} back to Ω_n. Given a partial configuration $\tau \in [q]^S$ such that $\tau \in \Omega_{n-1}$, we can uniquely identify the missing vertex $v = V - S$ and obtain a random $\sigma \in \Omega_n$ such that $\sigma_S = \tau$ and $\sigma_v \sim \pi_v^\tau$. Formally, the transition matrix $U : \Omega_{n-1} \times \Omega_n \to \mathbb{R}_{\geq 0}$ can be written as

$$U(\tau, \sigma) = \begin{cases} \pi_v^\tau(\sigma_v) & \text{if } \sigma_S = \tau, \\ 0 & \text{otherwise.} \end{cases}$$

Now, the transition matrix P can be written as

$$P = DU.$$

Next, we let $\pi_n = \pi$ and $\pi_{n-1} = \pi_n D$. Intuitively, to draw a random sample $X \sim \pi_{n-1}$, one can first sample $X \sim \pi$ and drop the value at a uniform random index $v \in V$. By definition, it is easy to verify that $\pi_{n-1} U = \pi_n = \pi$, so that $\pi P = \pi$.

By decomposing P as a pair of down-walk and up-walk, we can proved some property of the transition matrix P. Let $\langle f, g\rangle_{\pi_n} = \sum_{x \in \Omega_n} \pi_n(x) f(x) g(x)$ be the weighted inner product defined over Ω_n. Similarly, let $\langle f', g'\rangle_{\pi_{n-1}} = \sum_{\tau \in \Omega_{n-1}} \pi_{n-1}(\tau) f'(\tau) g'(\tau)$ be the weighted inner product defined over Ω_{n-1}. The following observation holds.

Observation 1.1 ([AL20]). *D and U is a pair of adjoint operators such that*

$$\forall f \in \mathbb{R}^{\Omega_n}, g \in \mathbb{R}^{\Omega_{n-1}}, \quad \langle f, Dg\rangle_{\pi_n} = \langle Uf, g\rangle_{\pi_{n-1}}.$$

Proof. We can compute

$$\begin{aligned}
\langle f, Dg\rangle_{\pi_n} &= \sum_{\sigma \in \pi_n} \pi_n(\sigma) f(\sigma) \sum_{\tau \in \Omega_{n-1}} D(\sigma, \tau) g(\tau) = \frac{1}{n} \sum_{\sigma \in \Omega_n} \pi_n(\sigma) f(\sigma) \sum_{v \in V} g(\sigma_{V-v}) \\
&= \sum_{v \in V} \sum_{\sigma \in \Omega_n} \pi_{n-1}(\sigma_{V-v}) \pi_v^{\sigma_{V-v}}(\sigma_v) f(\sigma) g(\sigma_{V-v}) \\
&= \sum_{\tau \in \Omega_{n-1}} \pi_{n-1}(\tau) g(\tau) \sum_{\sigma \in \Omega_n} U(\tau, \sigma) f(\sigma) \\
&= \langle Uf, g\rangle_{\pi_{n-1}}.
\end{aligned}$$

□

Glauber dynamics transition matrix is PSD With this observation, we can quickly prove the transition matrix P of Glauber dynamics is PSD. For any $h \in \mathbb{R}^{\Omega_n}$, we have

$$\langle h, Ph\rangle_{\pi_n} = \langle h, DUh\rangle_{\pi_n} = \langle Uh, Uh\rangle_{\pi_{n-1}} \geq 0.$$

Let h be the eigenvector corresponding to the minimum eigenvalue $\lambda_{\min}$ of P. It holds that

$$\langle h, Ph\rangle_{\pi_n} = \lambda_{\min} \cdot \langle h, h\rangle_{\pi_n} \geq 0 \quad \Longrightarrow \quad \lambda_{\min} \geq 0,$$

where the implication holds because $\langle h, h\rangle_{\pi_n} > 0$.

Alternative proof: spectral gap implies mixing We now reprove the fact that the spectral gap implies mixing time for Glauber dynamics using the decomposition of down-up walks. Let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be a convex function such that $f(1) = 0$. Let π and μ be two distributions over Ω such that μ is absolutely continuous with respect to π (denoted by $\mu \ll \pi$), which means $\pi(x) = 0 \implies \mu(x) = 0$. Define their f-divergence by

$$D_f(\mu \| \pi) = \mathbf{E}_\pi\left[f\left(\frac{\mu}{\pi}\right)\right] - f\left(\mathbf{E}_\pi\left[\frac{\mu}{\pi}\right]\right) = \mathbf{E}_\pi\left[f\left(\frac{\mu}{\pi}\right)\right] - f(1) = \mathbf{E}_\pi\left[f\left(\frac{\mu}{\pi}\right)\right].$$

The f-divergence can be viewed as a "distance" between two distributions. For example, if we set $f(x) = \frac{1}{2}|x-1|$, then $D_f(\mu \| \pi) = d_{\mathrm{TV}}(\mu, \pi)$ is the total variation distance. If we set $f(x) = (x-1)^2$, then D_f is known as χ^2-divergence. If we set $f(x) = x \ln x$, then D_f is known as the relative entropy or KL-divergence. We need the following well-known fact about f-divergence.

Lemma 1.2 (data processing inequality). *Let P be an arbitrary transition matrix from Ω to some space Ω'. For any two distributions μ_1, μ_2 over Ω such that $\mu_1 \ll \mu_2$, it holds that*

$$D_f(\mu_1 P \| \mu_2 P) \leq D_f(\mu_1 \| \mu_2).$$

The data processing inequality follows from the convexity of f. The proof can be found in many textbooks such as [PW25].

We also need to use the following property of down-up walks. For two functions (vectors) f, g, we use $\frac{f}{g}$ to denote the function (vector) such that $\frac{f}{g}(x) = \frac{f(x)}{g(x)}$.

Lemma 1.3. *Note that $\pi_n = \pi$. For any distribution μ over Ω_n, $\frac{\mu D}{\pi_n D} = U \frac{\mu}{\pi_n}$.*

Proof. For any $x \in \Omega_{n-1}$ such that $x \in [q]^S$, on the one hand,

$$\frac{\mu D}{\pi_n D}(x) = \frac{\sum_{y \in \Omega_n : y_S = x} \mu(y)/n}{\sum_{y \in \Omega_n : y_S = x} \pi(y)/n} = \frac{\mu_S(x)}{\pi_S(x)}.$$

One the other hand, let $v = V - S$,

$$U \frac{\mu}{\pi_n} = \sum_{y \in \Omega_n : y_S = x} \pi_v^x(y_v) \frac{\mu(y)}{\pi(y)} = \sum_{y \in \Omega_n : y_S = x} \frac{\pi(y)}{\pi_S(x)} \frac{\mu(y)}{\pi(y)} = \frac{\mu_S(x)}{\pi_S(x)}.$$

□

Now, let us focus on the χ^2-divergence, which is closely related to the spectral gap. Let π be the stationary distribution of Glauber dynamics. Let μ_0 be an arbitrary initial distribution. Let $\mu_t = \mu_0 P^t$ be the distribution after t steps. We show that $D_{\chi^2}(\mu_t \| \pi)$ decays exponentially with t. Define the function $h_t = \frac{\mu_t}{\pi}$[1]. Note that $\mathbf{E}_\pi [h_t] = 1$ and $D_{\chi^2}(\mu_t \| \pi) = \mathbf{Var}_\pi [h_t]$. Then

$$D_{\chi^2}(\mu_{t+1} \| \pi) = D_{\chi^2}(\mu_t P \| \pi P) = D_{\chi^2}(\mu_t DU \| \pi DU) \le D_{\chi^2}(\mu_t D \| \pi D),$$

where the last inequality holds from the data processing inequality. We want to compare $D_{\chi^2}(\mu_{t+1} \| \pi)$ to $D_{\chi^2}(\mu_t \| \pi)$. The above inequality shows that it suffices to compare $D_{\chi^2}(\mu_t D \| \pi D)$ to $D_{\chi^2}(\mu_t \| \pi)$. Note that $D_{\chi^2}(\mu_t \| \pi) = \mathbf{Var}_{\pi_n} [h_t]$ and $D_{\chi^2}(\mu_t D \| \pi D) = \mathbf{Var}_{\pi_{n-1}} \left[\frac{\mu_t D}{\pi_n D} \right] = \mathbf{Var}_{\pi_{n-1}} \left[U \frac{\mu_t}{\pi_n} \right]$, where the last equation holds from Lemma 1.2 Then, we have

$$\begin{aligned}
D_{\chi^2}(\mu_t \| \pi) - D_{\chi^2}(\mu_{t+1} \| \pi) &\ge D_{\chi^2}(\mu_t \| \pi) - D_{\chi^2}(\mu_t D \| \pi D) \\
&= \mathbf{Var}_{\pi_n} [h_t] - \mathbf{Var}_{\pi_{n-1}} [U h_t] \\
\text{(by } \mathbf{E}_{\pi_n} [h_t] = \mathbf{E}_{\pi_{n-1}} [U h_t] = 1) \quad &= \langle h_t, h_t \rangle_{\pi_n} - \langle U h_t, U h_t \rangle_{\pi_{n-1}} \\
\text{(by Observation 1.1)} \quad &= \langle h_t, h_t \rangle_{\pi_n} - \langle h_t, DU h_t \rangle_{\pi_n} \\
&= \langle h_t, (I - P) h_t \rangle_{\pi_n} \\
&\ge \gamma \mathbf{Var}_{\pi_n} [h_t] = \gamma D_{\chi^2}(\mu_t \| \pi).
\end{aligned}$$

where $\gamma = \inf_{f : \mathbf{Var}_\pi[f] > 0} \frac{\mathcal{E}(f,f)}{\mathbf{Var}_\pi[f]}$ is the spectral gap. Rearranging the inequality, we have

$$D_{\chi^2}(\mu_{t+1} \| \pi) \le (1 - \gamma) D_{\chi^2}(\mu_t \| \pi).$$

Hence, the χ^2 divergence decays exponentially with rate $(1 - \gamma)$. We have $D_{\chi^2}(\mu_t \| \pi) \le (1 - \gamma)^t D_{\chi^2}(\mu_0 \| \pi)$. The maximum χ^2-divergence between μ_0 and π is achieved when $\mu_0(x) = 1$ such that $\pi(x) = \pi_{\min}$. Hence, $D_{\chi^2}(\mu_t \| \pi) \le (1 - \gamma)^t (\frac{1}{\pi_{\min}})^2$.

Exercise 1.4. Find a relation between χ^2-divergence and total variation distance and use it to bound the mixing time of Glauber dynamics.

[1] The definition means $h_t(x) = \frac{\mu_t(x)}{\pi(x)}$ for all $x \in \Omega$

2 Multi-level down-up walks

We give a generalized version of down-up walks. Let $\pi = \pi_n$ be a distribution over $\Omega = \Omega_n \subseteq [q]^V$, where $|V| = n$. We have defined Ω_{n-1} in previous section. We can generalize this to any $k \le n$. For any $0 \le k < n$, define Ω_k as follows

$$\Omega_k = \{\sigma \in [q]^S : S \subseteq V \wedge |S| = k \wedge \pi_S(\sigma) > 0\},$$

which contains all feasible partial configurations of size k. We can define the down-walk between every two adjacent levels. Define the $k \to (k-1)$ down-walk $D_{k\to(k-1)}$ as follows. Given a state $X \in \Omega_k$, say $X \in [q]^S$ for some $S \subseteq V$ with $|S| = k$, one step of the chain does as follows:

- pick a vertex $v \in S$ uniformly at random;
- remove the configuration of X on v to obtain $X_{S-v} \in \Omega_{k-1}$.

For any two levels $\ell > k$, we can define the down-walk $D_{\ell\to k}$ such that

$$D_{\ell\to k} = D_{\ell\to(\ell-1)} D_{(\ell-1)\to(\ell-2)} \cdots D_{(k+1)\to k}.$$

The distribution π can also be generalized to every level k such that

$$\pi_k = \pi_n D_{n\to k}.$$

Similarly, we can define the up-walk from Ω_k to Ω_{k+1} as follows. Given a state $X \in \Omega_k$ where $X \in [q]^S$, one step of the chain does as follows:

- pick a vertex $v \in V - S$ uniformly at random;
- sample X_v from the distribution π_v^X and then extend the configuration X further at the position v to obtain a new configuration $X \in \Omega_{k+1}$.

Alternatively, the up-walk can be interpreted as follows. We say $X \in [q]^S$ is a sub-configuration of $Y \in [q]^{S'}$ if $S \subseteq S'$. Given a state $X \in \Omega_k$, one step of the chain moves X to a new configuration $\sigma \in \Omega_{k+1}$ with probability

$$\begin{aligned} U_{k\to(k+1)}(X,\sigma) &= \mathbf{Pr}_{Y\sim\pi_{k+1}}[Y = \sigma \mid X \text{ is a sub-configuration of } Y] \\ &\propto \mathbf{1}[X \text{ is a sub-configuration of } \sigma] \cdot \pi_{k+1}(\sigma). \end{aligned}$$

Again, for different levels $\ell < k$, we can define the up-walk $U_{\ell\to k}$ by

$$U_{\ell\to k} = U_{\ell\to(\ell+1)} U_{(\ell+1)\to(\ell+2)} \cdots U_{(k-1)\to k}.$$

3 Spectral independence

We are ready to give the definition of spectral independence, a powerful tool to analyze the mixing time of Glauber dynamics. The Glauber dynamics for π is the down-up walk $D_{n\to(n-1)}U_{(n-1)\to n}$ between levels n and $n-1$. To prove the mixing of Glauber dynamics, the spectral independence considers the down-up walk $P^{\vee}_{1,n} = D_{n\to 1}U_{1\to n}$ between level n and level 1.

Definition 3.1 (SI: down-up-walk-based definition)**.** Let $C \ge 1$ be a constant. A distribution π over $[q]^V$ is said to be C-spectral independence (C-SI) if the second largest eigenvalue of $1 \leftrightarrow n$ down-up walk $P^{\vee}_{1,n} = D_{n\to 1}U_{1\to n}$ at most $\frac{C}{n}$, where $n = |V|$ denote the size of V.

There is an equivalent and more popular definition of spectral independence. Consider the $1 \leftrightarrow n$ up-down walk $P^{\wedge}_{1,n} = U_{1\to n}D_{n\to 1}$ over Ω_1. Note that two matrices AB and BA have the same eigenvalues except for some zeros. One can verify that $P^{\vee}_{1,n} = D_{n\to 1}U_{1\to n}$ is PSD (using the same proof as that for Glauber dynamics). The up-down walk $P^{\wedge}_{1,n}$ and down-up walk $P^{\vee}_{1,n}$ have the same second largest eigenvalue. Compared to the down-up walk, the up-down walk $P^{\wedge}_{1,n}$ is very simple. Every state in Ω_1 is a value at a single vertex. We can use a pair (v, c) to denote a state. Hence

$$\Omega_1 = \{vc \mid \pi_v(c) > 0\}.$$

The transition matrix $P^{\wedge}_{1,n}$ has a simple form such that for any $(v, a), (u, b) \in \Omega_1$,

$$P^{\wedge}_{1,n}(va, ub) = \frac{1}{n}\pi_u^{v\leftarrow a}(b) = \frac{1}{n}\mathbf{Pr}_{X\sim\pi}\left[X_u = b \mid X_v = a\right].$$

It is easy to verify that $P^{\wedge}_{1,n}$ is reversible with respect to π_1 and for any $(v, c) \in \Omega_1$,

$$\pi_1(vc) = \frac{\pi_v(c)}{n}.$$

Lemma 3.2. *Show that the λ_2 of up-down walk $P^{\wedge}_{1,n}$ is at least $\frac{1}{n}$, which is the reason why we assume $C \ge 1$ in the definition of spectral independence.*

Proof. Fix a variable v and color c such that $\pi_v(c) \in (0, 1)$. Define t such that $\frac{t}{n} = \pi_v(c)$. Let $A = \{vc\}$, $B = \{va \mid a \ne c\}$, and $C = \Omega_1 - A - B$. Define f such that for any $a \in A$, $f(a) = \frac{n}{t}$, for any $b \in B$, $f(b) = -\frac{n}{1-t}$, and for any $c \in C$, $f(c) = 0$. We can verify that

$$\mathbf{E}_{\pi_1}[f] = \pi_v(c)f(a) + \frac{1-t}{n}f(b) = 0.$$

Note that the up-down walk cannot make transitions between A and B. Hence, to compute the Dirichlet form, we only need to consider transitions between A anc C and transitions between B and C. We have

$$\begin{aligned}
\mathcal{E}(f, f) &= \sum_{a\in A}\pi_1(a)\sum_{c\in C}P^{\wedge}_{1,n}(a, c)\frac{n^2}{t^2} + \sum_{b\in B}\pi_1(b)\sum_{c\in C}P^{\wedge}_{1,n}(b, c)\frac{n^2}{(1-t)^2}\\
&= \frac{t}{n}\left(1 - \frac{1}{n}\right)\frac{n^2}{t^2} + \frac{1-t}{n}\left(1 - \frac{1}{n}\right)\frac{n^2}{(1-t)^2}\\
&= \left(1 - \frac{1}{n}\right)\left(\frac{n}{t} + \frac{n}{1-t}\right).
\end{aligned}$$

In the first equation, we only enumerate $a \in A$ and $c \in C$ to sum over all possible transitions from A to C. We ignore transitions from C to A due to the reversibility property of $P^{\wedge}_{1,n}$ (also, we remove the factor $\frac{1}{2}$). The same argument applies to transitions from B to C. A simple calculation shows that

$$\langle f, f\rangle_{\pi_1} = \frac{t}{n}\frac{n^2}{t^2} + \frac{1-t}{n}\frac{n^2}{(1-t)^2} = \frac{n}{t} + \frac{n}{1-t}.$$

Hence, the spectral gap of $P^{\wedge}_{1,n}$ satisfies

$$\inf_{h:h\perp_{\pi_1}\mathbf{1}}\frac{\mathcal{E}(h, h)}{\langle h, h\rangle_{\pi_1}} \le \frac{\mathcal{E}(f, f)}{\langle f, f\rangle_{\pi_1}} = 1 - \frac{1}{n}. \qquad \square$$

In the last lecture, we show that any reversible Markov chain P admits a spectral decomposition $P = \sum_i \lambda_i v_i v_i^T D$, where λ_i is the eigenvalue, v_i is the eigenvector and D is the diagonal matrix generated by the stationary distribution. We can apply this result to the up-down walk. Note that $\lambda_1 = 1$ and $v_1 = \mathbf{1}$. We can remove the term contributed by the top eigenvalue to obtain

$$A = P^{\wedge}_{1,n} - \mathbf{1}\mathbf{1}^T \mathrm{diag}(\pi_1) \quad \Leftrightarrow \quad A(va, ub) = \frac{1}{n}\left(\pi_u^{v \leftarrow a}(b) - \pi_u(b)\right),$$

which motivates the following definition of influence matrix $\Psi : \Omega_1 \times \Omega_1 \to \mathbb{R}$ such that

$$\Psi_\pi(va, ub) = \pi_u^{v \leftarrow a}(b) - \pi_u(b).$$

The definition of the influence matrix is very intuitive. The difference $\pi_u^{v \leftarrow a}(b) - \pi_u(b)$ can be viewed as the influence on u taking the value b from the event that v taking the value a. The following definition of spectral independence is well-known.

Definition 3.3 (SI: influence-matrix-based definition)**.** Let $C \geq 1$ be a constant. A distribution π is said to be C-SI if the largest eigenvalue of influence matrix Ψ_π is at most C.

The above definition of spectral independence was discovered in [CGŠV21]. However, the notion of spectral independence was first introduced in [ALO20] for Boolean distributions ($q = 2$). The definition of influence matrix in [ALO20] is different from the definition here and the definition only works for Boolean distributions.

3.1 Spectral independence implies rapid mixing

We show the spectral independence implies the mixing of down-up walks. Then, we add a minor additional assumption to prove the mixing of Glauber dynamics.

Let $k \geq 1$. Define the k-block dynamics as the $n \leftrightarrow (n-k)$ down-up walks such that

$$P_k^B = P^{\vee}_{n,n-k} = D_{n \to (n-k)} U_{(n-k) \to n}.$$

In words, at every step, given $X \in \Omega$, P_k^B samples a subset $S \subseteq V$ with $|S| = k$ uniformly at random, and then resamples X_S from the conditional distribution $\pi_S^{X_{V-S}}$.

Observation 3.4. *The Glauber dynamics is* 1*-block dynamics.*

To introduce the mixing result for general k-block dynamics, we first introduce the notation of conditional distributions. Let $\Lambda \subseteq V$ be a subset of variables. Let $\sigma \in [q]^\Lambda$ be a feasible partial configuration on Λ. We use $\pi^\sigma_{V-\Lambda}$ to denote the marginal distribution on $V - \Lambda$ conditional on σ. For example, in graph coloring, given the partial coloring σ, for any vertex $v \in V - \Lambda$, v can only take colors in a list $L_v^\sigma = [q] - \{\sigma_u \mid u \in \Lambda \text{ is a neighbor of } v\}$.

Theorem 3.5. *Let $C \geq 1$ be a constant. Let $\ell > C$. Let π be a distribution over $[q]^V$ with $|V| = n$. If for any feasible partial configuration $\sigma \in [q]^{V-S}$ where $S \subseteq V$ and $|S| > \ell$, the conditional distribution π_S^σ is C-SI, then the spectral gap of ℓ-block dynamics is at least*

$$\prod_{i=\ell+1}^{n} \left(1 - \frac{C}{i}\right) \approx \exp\left(-\sum_{i=\ell+1}^{n} \frac{C}{n}\right) \approx \left(\frac{\ell}{n}\right)^C.$$

Remark 3.6. For a conditional distribution $\pi^\sigma_{V-\Lambda}$ with $|V - \Lambda| = k$, Definition 3.1 holds for $\pi^\sigma_{V-\Lambda}$ means the second largest eigenvalue of $k \leftrightarrow 1$ down-up walk transition matrix at most $\frac{C}{k}$, which is equivalent to Definition 3.3 such that $\lambda_{\max}(\Psi_{\pi^\sigma_{V-\Lambda}}) \leq C$.

Remark 3.7. The above theorem requires that the spectral independence holds not only for π but also for all conditional distributions induced by π. In many papers, the spectral independence is also directed defined as Definition 3.1 and Definition 3.3 hold for π and all conditional distributions induced by π.

Proof. For any $f : \Omega \to \mathbb{R}$, let $\mathcal{E}_k(f,f)$ denote the Dirichlet form for P_k^B. For any subset Λ, let Ω_Λ denote the set of all feasible configurations on Λ and let $\Lambda^c = V - \Lambda$. By definition,

$$\begin{aligned}\mathcal{E}_k(f,f) &= \frac{1}{2}\sum_{x\in\Omega}\pi(x)\sum_{y\in\Omega}P_k^B(x,y)(f(x)-f(y))^2\\ &= \frac{1}{2\binom{n}{k}}\sum_{S\subseteq V:|S|=k}\ \sum_{\substack{\sigma\in\Omega_{S^c}\\ y_{S^c}=x_{S^c}=\sigma}}\pi_{S^c}(\sigma)\sum_{x_S\in[q]^S,y_S\in[q]^S}\pi_S^\sigma(x_S)\pi_S^\sigma(y_S)(f(x)-f(y))^2.\end{aligned}$$

Note that $\mathbf{Var}_\mu[f] = \frac{1}{2}\sum_{x,y}\mu(x)\mu(y)(f(x)-f(y))^2$ for any distribution μ. Let Ω_S^σ denote the support of π_S^σ. Let f^σ denote a function from Ω_S^σ to $\mathbb{R}$ such that for any $\tau\in\Omega_S^\sigma$, $f^\sigma(\tau) = f(\tau\sigma)$, where $\tau\sigma$ is a full configuration obtained by concatenating τ and σ. Let $\binom{V}{k}$ denote all subsets $\Lambda\subseteq V$ such that $|\Lambda| = k$. Let $S\sim\binom{V}{k}$ denote a uniform random element in $\binom{V}{k}$. Then

$$\mathcal{E}_k(f,f) = \mathbf{E}_{S\sim\binom{V}{k}}\left[\mathbf{E}_{\sigma\sim\pi_{S^c}}\left[\mathbf{Var}_{\pi_S^\sigma}[f^\sigma]\right]\right].$$

Our starting point is Definition 3.1, which shows that P_{n-1}^B for π has spectral gap $1-\frac{C}{n}$:

$$\mathbf{Var}_\pi[f] \le \left(1-\frac{C}{n}\right)^{-1}\mathbf{E}_{S\sim\binom{V}{n-1}}\left[\mathbf{E}_{\sigma\sim\pi_{S^c}}\left[\mathbf{Var}_{\pi_S^\sigma}[f^\sigma]\right]\right].$$

Next, we apply Definition 3.1 for conditional distribution π_S^σ on the inner variance term. Since π_S^σ is a joint distribution over $n-1$ variables in S, we have

$$\mathbf{Var}_{\pi_S^\sigma}[f^\sigma] \le \left(1-\frac{C}{n-1}\right)^{-1}\mathbf{E}_{T\sim\binom{S}{n-2}}\left[\mathbf{E}_{\tau\sim\pi_{S-T}^\sigma}\left[\mathbf{Var}_{\pi_T^{\sigma\tau}}[f^{\sigma\tau}]\right]\right].$$

Combining the two inequalities, we have

$$\mathbf{Var}_\pi[f] \le \left(1-\frac{C}{n}\right)^{-1}\left(1-\frac{C}{n-1}\right)^{-1}\mathbf{E}_{S\sim\binom{V}{n-2}}\left[\mathbf{E}_{\sigma\sim\pi_{S^c}}\left[\mathbf{Var}_{\pi_S^\sigma}[f^\sigma]\right]\right].$$

Repeating this process until the size of S becomes ℓ, we have

$$\mathbf{Var}_\pi[f] \le \prod_{i=\ell+1}^{n}\left(1-\frac{C}{i}\right)^{-1}\mathbf{E}_{S\sim\binom{V}{\ell}}\left[\mathbf{E}_{\sigma\sim\pi_{S^c}}\left[\mathbf{Var}_{\pi_S^\sigma}[f^\sigma]\right]\right]. \tag{1}$$

This proves the theorem. □

To prove the mixing of Glauber dynamics, we cannot simply set $\ell = 1$ regardless of C. This is because $1-\frac{C}{i}$ can be negative if $i < C$. However, we can further assume that for any π_S^σ, where $\sigma\in S^c$ and $|S| = k$, the largest eigenvalue of the influence matrix Ψ for π_S^σ is at most $k\eta$ for some constant $\eta < 1$, or equivalently, the second largest eigenvalue of the $k \leftrightarrow 1$ down-up walk is at most $\eta < 1$. Then, we can keep doing the proof until $\ell = 1$.

The above proof only shows a (large) polynomial mixing time of Glauber dynamics. For graphical models (e.g. graph coloring), where the underlying graph has bounded degree, see [CLV21] for a proof of $O(n\log n)$ optimal mixing from the spectral independence. Also, see [AJK+21, CFYZ21, CFYZ22, CE22] for the optimal mixing for certain graphical models on general graphs with unbounded degree.

3.2 Establish spectral independence from total influence

To verify the spectral independence condition, Definition 3.3 requires us to compute the max eigenvalue of the influence matrix. This is usually not easy. Instead, we can bound the total influence, which is the sum of influences.

$$\lambda_{\max}(\Psi) \leq \|\Psi\|_\infty = \max_{va} \sum_{ub} |\Psi(va, ub)|.$$

Sometimes, it is easier to work with the following n by n influence matrix $\Psi' : V \times V \to \mathbb{R}$. For any two variables $v, u \in V$, define the influence from u to v by

$$\Psi'(v, u) = \max_{a,b} d_{\mathrm{TV}}\left(\pi_u^{v \leftarrow a}, \pi_u^{v \leftarrow b}\right).$$

To upper bound the top eigenvalue of Ψ, the task can be reduced to upper bound the total influence of Ψ' such that

$$\|\Psi'\|_\infty = \max_{v} \sum_{u} |\Psi'(v, u)|.$$

For graph q-coloring on graph G, one can obtain a constant the total influence bound if

- G has constant degree and $q \geq 1.809\Delta$ [CV24];
- G is triangle-free and $q \geq 1.763\Delta$ [FGYZ21, CGŠV21];
- G has constant degree and constant large girth, and $q \geq \Delta + 3$ [CLMM23].

The spectral independence shows $O(n \log n)$ mixing time in bounded degree graphs [CLV21].

References

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

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

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

[ALOV19] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials II: high-dimensional walks and an FPRAS for counting bases of a matroid. In *STOC*, pages 1–12, 2019.

[CE22] Yuansi Chen and Ronen Eldan. Localization schemes: A framework for proving mixing bounds for markov chains. In *FOCS*, pages 110–122, 2022.

[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, 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, 2022.

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

[CLMM23] Zongchen Chen, Kuikui Liu, Nitya Mani, and Ankur Moitra. Strong spatial mixing for colorings on trees and its algorithmic applications. In *FOCS*, pages 810–845, 2023.

[CLV21] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of Glauber dynamics: entropy factorization via high-dimensional expansion. In *STOC*, pages 1537–1550, 2021.

[CV24] Charlie Carlson and Eric Vigoda. Flip dynamics for sampling colorings: Improving $(11/6\text{-}\varepsilon)$ using a simple metric. *CoRR*, abs/2407.04870, 2024.

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

[PW25] Yury Polyanskiy and Yihong Wu. *Information Theory: From Coding to Learning*. Cambridge University Press, 2025. https://people.lids.mit.edu/yp/homepage/data/itbook-export.pdf.