Approximate Counting and Mixing Time of Markov Chains	**Fall 2024**

Lecture 9: Spectral Independence

Lecturer: Shayan Oveis Gharan

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

For the next few lectures we study the Glauber dynamics on two state spin systems. In other words, suppose we have a graph $G = (V, E)$ we want sample from the state space $\{\pm 1\}^V$. We are going to see that if the underlying distribution π on $\{\pm 1\}^V$ exhibits limited "spectral" correlations then the Glauber dynamics mies in almost linear time (when G is bounded degree).

For $\sigma \in \{\pm 1\}^V$ let

$$\sigma^{\oplus i}(j) = \begin{cases} -\sigma_j & \text{if } j = i \\ \sigma_j & \text{otherwise} \end{cases}$$

Recall that the Glauber dynamics works as follows: First we choose a u.r. vertex i then, with probability $\frac{\pi(\sigma^{\oplus i})}{\pi(\sigma)+\pi(\sigma^{\oplus i})}$ we move to $\sigma^{\oplus i}$ and otherwise we stay.

It is also instructive to write down the Dirichlet form

$$\mathcal{E}_K(f, f) = \frac{1}{2} \sum_{\sigma \in \{\pm 1\}^V} \pi(\sigma) \frac{1}{n} \sum_{i=1}^{n} \frac{\pi(\sigma^{\oplus i})}{\pi(\sigma) + \pi(\sigma^{\oplus i})} (f(\sigma) - f(\sigma^{\oplus i}))^2$$

Pinning. Let π be a distribution on $\{\pm 1\}^V$. For a any set of vertices $i_1, \dots, i_k$ (for any $1 \le k < n$) and signs $s_1, \dots, s_k$ we let

$$\pi_{(i_1,s_1),\dots,(i_k,s_k))} := \pi | \sigma_{i_1} = s_1, \dots, \sigma_{i_k} = s_k.$$

In other words, this is the conditional measure on all vertices in $V - \{i_1, \dots, i_k\}$ when we pin $i_1, \dots, i_k$ to signs $s_1, \dots, s_k$ respectively.

Averaging / Projection. Conversely, given π and a set $S \subseteq V$ we let π^S to be distribution π projected onto the set S when we average out all vertices outside of S. In other words,

$$\pi^S(\tau \in \{\pm 1\}^S) = \sum_{\sigma : \sigma_S = \tau} \pi(\sigma).$$

Having that we can re-write the Dirichlet form as follows:

$$\begin{aligned}
\mathcal{E}_K(f, f) &= \frac{1}{2} \mathbb{E}_i \sum_{\sigma \in \{\pm 1\}^V} (\pi(\sigma) + \pi(\sigma^{\oplus i})) \cdot \frac{\pi(\sigma^{\oplus i})}{\pi(\sigma) + \pi(\sigma^{\oplus i})} \cdot \frac{\pi(\sigma)}{\pi(\sigma) + \pi(\sigma^{\oplus i})} \cdot (f(\sigma) - f(\sigma^{\oplus i}))^2 \\
&= \mathbb{E}_i \sum_{\sigma \in \{\pm 1\}^V} (\pi(\sigma) + \pi(\sigma^{\oplus i})) \cdot \mathrm{Var}_{\pi_{\sigma - i}}(f) \\
&= \mathbb{E}_i \mathbb{E}_{\tau \sim \pi^{V-i}} \mathrm{Var}_{\pi_\tau}(f).
\end{aligned}$$

Note that $\pi_{\sigma - i}$ is the pinning of π on all vertices in $V - i$ according to σ. It follows that to bound the Poincare constant it is enough to relate the local variance to the global variance of f.

11.1 Spectral Independence

Definition 11.1 (Spectral Independence)**.** *Let π be a probability distribution over $\{\pm 1\}^V$. Define the influence matrix $\Psi_\pi \in \mathbb{R}^{V\times V}$,*

$$\Psi_\pi(i,j) = \mathbb{P}_\pi\left[\sigma_j = +1 | \sigma_i = +1\right] - \mathbb{P}_\pi\left[\sigma_j = +1 | \sigma_i = -1\right].$$

If $i = j$ we simply let $\Psi_\pi(i,i) = 1$. We say π is η-spectrally independent if $\lambda_{\max}(\Psi_\pi) \leq 1 + \eta$.

Fact 11.2. *Let D_π be the diagonal matrix with $D_\pi(i,i) = \mathrm{Var}(\sigma_i)$. Then,*

$$\Psi_\pi = D_\pi^{-1} Cov(\pi)$$

In particular π is η-spectrally independent iff $Cov(\pi) \preceq (1+\eta)D_\pi$

Note that this fact in particular implies that Ψ_π has real eigenvalues, as eigenvalues of Ψ is the same as eigenvalues of the symmetric matrix $D_\pi^{-1/2} Cov D_\pi^{-1/2}$.

Proof. The first observation is that

$$\mathrm{Var}(\sigma_i) = \mathbb{P}\left[\sigma_i = +1\right]\mathbb{P}\left[\sigma_i = -1\right](1 - (-1))^2 = 4\mathbb{P}\left[\sigma_i = +1\right]\mathbb{P}\left[\sigma_i = -1\right]$$

On the other hand, let $p_i = \mathbb{P}\left[\sigma_i = 1\right], p_j = \mathbb{P}\left[\sigma_j = 1\right]$. Then,

$$\begin{aligned}
\mathrm{Cov}(i,j) &= \mathbb{E}\sigma_i\sigma_j - \mathbb{E}\sigma_i\mathbb{E}\sigma_j \\
&= \mathbb{E}\sigma_i\sigma_j - (2p_i - 1)(2p_j - 1) \\
&= 4(\mathbb{P}\left[\sigma_i = \sigma_j = +1\right] - \mathbb{P}\left[\sigma_i = +1\right]\mathbb{P}\left[\sigma_j = +1\right]).
\end{aligned}$$

To see the last line observe that

$$\begin{aligned}
-\mathbb{P}\left[\sigma_i = 1, \sigma_j = -1\right] + p_i &= \mathbb{P}\left[\sigma_i = \sigma_j = 1\right], \\
-\mathbb{P}\left[\sigma_i = -1, \sigma_j = 1\right] + p_j &= \mathbb{P}\left[\sigma_i = \sigma_j = 1\right] \\
\mathbb{P}\left[\sigma_i = \sigma_j = -1\right] + p_i + p_j - 1 &= \mathbb{P}\left[\sigma_i = \sigma_j = 1\right]
\end{aligned}$$

Having this, we can write

$$\begin{aligned}
D_\pi^{-1} Cov(i,j) &= \frac{\mathbb{P}\left[\sigma_i = \sigma_j = +1\right] - \mathbb{P}\left[\sigma_i = +1\right]\mathbb{P}\left[\sigma_j = +1\right]}{\mathbb{P}\left[\sigma_i = +1\right]\mathbb{P}\left[\sigma_i = -1\right]} \\
&= \frac{\mathbb{P}\left[\sigma_j = 1 | \sigma_i = 1\right] - \mathbb{P}\left[\sigma_j = 1\right]}{\mathbb{P}\left[\sigma_i = -1\right]} \\
&= \frac{\mathbb{P}\left[\sigma_j = 1 | \sigma_i = 1\right] - \mathbb{P}\left[\sigma_j = 1, \sigma_i = -1\right] - \mathbb{P}\left[\sigma_j = 1, \sigma_i = 1\right]}{\mathbb{P}\left[\sigma_i = -1\right]} \\
&= \frac{\mathbb{P}\left[\sigma_j = 1 | \sigma_i = 1\right](1 - \mathbb{P}\left[\sigma_i = 1\right]) - \mathbb{P}\left[\sigma_j = 1, \sigma_i = -1\right]}{\mathbb{P}\left[\sigma_i = -1\right]} = \Psi_\pi(i,j).
\end{aligned}$$

□

Ex1: Independent Case. Suppose π is a product distribution. In that case for any $i, j \in U$, $\sigma_j | \sigma_i$ is distributed the same as σ_j. Therefore, all off-diagonal entries of Ψ_π are zero. So, $\Psi_\pi = I$ and π is 0-spectrally independent.

Extreme Positively Correlation. Suppose there are only two sets in the support of $\pi(\{+1,\ldots,+1\})) = \pi(\{-1,\ldots,-1\} = 1/2$. In this case the distribution is very positively correlated. It follows that $\Psi_\pi = J_V$, where J is the all-ones matrix. So, π is $n-1$-spectrally independent. Note that in this case any local chain which only flips the state of one particle is not irreducible (the chain has only two states and to mix one has to change the state of all particles simultaneously).

Negatively Correlated Case. For another example, suppose π is a *negatively correlated* distribution over $\{\pm 1\}^V$ which is k-homogeneous, namely for any $i,j \in [n]$, $\mathbb{P}[\sigma_i = +|\sigma_j = +] \le \mathbb{P}[\sigma_i = +|\sigma_j = -]$ and that every σ in the support of π has exactly k many $+$'s. It is welll-known that for any matrix M,

$$\lambda_{\max}(M) \le \max_i \sum_j |M_{i,j}|$$

In our setting, we get

$$\begin{aligned}
\lambda_{\max}(\Psi_\pi) &\le \max_i 1 + \sum_{j\neq i} |\mathbb{P}[\sigma_j = +|\sigma_i = +] - \mathbb{P}[\sigma_j = +|\sigma_i = -]| \\
&= 1 + \max_i \left| \sum_{j \neq i} \mathbb{P}[\sigma_j = +|\sigma_i = +] - \mathbb{P}[\sigma_j = +|\sigma_i = -] \right| && \text{(negative correlation)} \\
&= 1 + \max_i |\mathbb{E}[|\sigma|_+ - 1|\sigma_i = +] - \mathbb{E}[|\sigma_i|_+ |\sigma_i = -]| = 2
\end{aligned}$$

where $|\sigma|_+$ is the number of $+$'s in σ. So, π is 2-spectrally independent.

By now there are various proofs of this theorem. However, this was first proven following a long line of works on simplicial complexes started with works of [DK17; KO20; AL20; ALO21]

Theorem 11.3 (Mixing for Sparse Graphical Models)**.** *Let π be a probability measure on $\{\pm 1\}^V$. Suppose π satisfies the following properties: (i) Spectral Independence: There exists $\eta \le O(1)$ such that for every $S \subseteq [n]$ and every pinning $\tau : S\{\pm 1\}$, the conditional distribution π_τ is η-spectrally independent. Then Glauber dynamics has spectral gap at least ?.*

11.2 Poincaré Inequality via Spectral Independence

In this section we prove Theorem 11.3. The main tool is the following tensorization of variance.

Lemma 11.4 (Approximate Tensorization of Variance)**.** *Let π be a distribution on $\{\pm 1\}^V$ that is η-spectrally independent. Then, for any function $f : \{\pm 1\}^V \to \mathbb{R}$,*

$$(1 - \frac{1+\eta}{n}) \operatorname{Var}_\pi(f) \le \mathbb{E}_i \mathbb{E}_{\sigma_i \sim \pi^i} \operatorname{Var}_{\pi_{i,\sigma_i}}(f) \tag{11.1}$$

Proof of Theorem 11.3. By repeatedly applying (11.1) we can write

$$\begin{aligned}
\mathrm{Var}(f) &\le \left(1-\frac{1+\eta}{n}\right)^{-1} \mathbb{E}_i \mathbb{E}_{\sigma_i\sim\pi^i} \mathrm{Var}_{\pi_{i,\sigma_i}}(f) && \text{(spect Ind of } \pi) \\
&\le \left(1-\frac{1+\eta}{n}\right)^{-1}\left(1-\frac{1+\eta}{n-1}\right)^{-1} \mathbb{E}_{i,j} \mathbb{E}_{\sigma_i,\sigma_j\sim\pi^{i,j}} \mathrm{Var}_{\pi_{i,\sigma_i,j,\sigma_j}}(f) && \text{(spect Ind of } \pi_{i,\sigma_i}) \\
&\dots \\
&\le \prod_{j=0}^{k-1}\left(1-\frac{1+\eta}{n-j}\right)^{-1} \mathbb{E}_{S\sim\binom{n}{k}} \mathbb{E}_{\sigma\sim\pi^S} \mathrm{Var}_{\pi_{S,\sigma}}(f) \\
&\dots \\
&\le \prod_{j=0}^{n-1}\left(1-\frac{1+\eta}{n-j}\right)^{-1} \mathbb{E}_{S\sim\binom{V}{n-1}} \mathbb{E}_{\sigma\sim\pi^S} \mathrm{Var}_{\pi_{S,\sigma}}(f) \\
&\lesssim \exp\left((1+\eta)\sum_{i=0}^{n-1}\frac{1}{n-i}\right)\mathcal{E}(f,f) \lesssim n^{1+\eta}\mathcal{E}(f,f)
\end{aligned}$$

as desired. □

Remark 11.5. *Note that if $\eta > 1$, the quantity $1-\frac{1+\eta}{n-k} < 0$ for values of k very close to n. So, instead, one needs a slightly different bound on spectral independence for such values of k. Typically just the connectivity of the support of the Ψ is enough to show prove that $\lambda(\Psi_{\pi_\tau}) \le C(n-k)$ when we pin $n-k$ coordinates in τ. We ignore those details as they simply change the final bounds by constants.*

Note that if we have a perfect independent distribution then we would have

$$\mathrm{Var}(f) = \sum_i \mathbb{E}_{\sigma\sim\pi^{V-i}} \mathrm{Var}_{\pi_{V-i,\sigma}}(f) = n\cdot\mathcal{E}(f,f).$$

In other words, the extra "small" loss n^η is due to small correlation/dependence between the particles in the system.

Let me also formalize the following tensorization of variance for disconnected graphs:

Lemma 11.6. *Suppose G is a disconnected graph with components $S, \overline{S}$. So, π is a product measure where $\sigma(\sigma_S, \sigma_{\overline{S}}) = \pi_{G[S]}(\sigma_S)\cdot\pi_{G[\overline{S}]}(\sigma_{\overline{S}})$. Similarly, for a function $f \in \{\pm1\}^n \to \mathbb{R}$ let f_S and $f_{\overline{S}}$ be the specialization of f to vertices in S, $\overline{S}$ respectively.*

$$\mathrm{Var}_\pi(f) = \mathrm{Var}_{\pi_{G[S]}}(f_S) + \mathrm{Var}_{\pi_{G[\overline{S}]}}(f_{\overline{S}})$$

In the rest of this section we prove Lemma 11.4.

There is a well-known fact in probability called the law of total variance.

Lemma 11.7 (Law of Total Variance)**.** *For random variables X, Y jointly distributed we have*

$$\mathrm{Var}(Y) = \mathbb{E}\,\mathrm{Var}(Y|X) + \mathrm{Var}(\mathbb{E}[Y|X])$$

We can use this lemma to write

Lemma 11.8. *Let $\pi^1 : V \times \{\pm1\} \to \mathbb{R}_{\ge 0}$ be the average of π^i measures. Namely, $\pi^1(i,s) = \frac{1}{n}\mathbb{P}[\sigma_i = s]$. and let $f^1(i,s) = \mathbb{E}_\sigma[f(\sigma)|\sigma_i = s]$. Then, by law of total variance,*

$$\mathrm{Var}(f) = \mathbb{E}_i \mathbb{E}_{\sigma_i\sim\pi^i} \mathrm{Var}_{\pi_{i,\sigma_i}}(f) + \mathrm{Var}_{\pi^1}(f_1)$$

Lemma 11.9.

$$\frac{\mathrm{Var}_{\pi^1}(f^1)}{\mathrm{Var}_\pi(f)} \leq \frac{1+\eta}{n}$$

Lemma 11.4 simply follows from this and the law of total vriance.

Proof. For any $i \in [n]$ and $s \in \{\pm 1\}$, think of $\pi_{i,s}$ as a vector in $\mathbb{R}^{2^n}$ where

$$\pi_{i,s}(\sigma) = \begin{cases} \frac{\pi(\sigma)}{\mathbb{P}[\sigma_i = s]} & \text{if } \sigma_i = s \\ 0 & \text{otherwise} \end{cases}.$$

Similarly, let $1_{i,s} = \mathbb{I}[\sigma_i = s]$. Then, observe that

$$\begin{aligned}\mathrm{Var}_{\pi^1}(f^1) &= \sum_{i,s} \pi^1(i,s) f^1(i,s)^2 - \langle f, 1\rangle_\pi \\ &= f^\top \left(\sum_{i,s} \pi^1(i,s)\pi_{i,s}\pi_{i,s}^\top \right) f - \langle f, 1\rangle_\pi =: \langle Pf, f\rangle_\pi,\end{aligned}$$

where $P = \sum_{i,s} \frac{1}{n} 1_{i,s}\pi_{i,s}^\top$. Note that P is a stochastic matrix.

$$\mathrm{Var}_\pi(f) = \langle f, f\rangle_\pi - \langle f, 1\rangle_\pi.$$

It follows that

$$\max_f \frac{\mathrm{Var}_{\pi^1}(f^1)}{\mathrm{Var}(f)} = \max_f \frac{\langle Pf, f\rangle_\pi - \langle f, 1\rangle_\pi}{\langle f, f\rangle_\pi - \langle f, 1\rangle_\pi} = \lambda_2(P)$$

So, to prove the lemma it is enough to show that $\lambda_2(P) \leq \frac{1+\eta}{n}$. The observation is that P is a low-rank matrix. In particular, let U be the matrix with columns $\frac{1}{n}1_{i,s}$ and R be the matrix with rows $\pi_{i,s}$. Then, $P = UR$, and $\lambda_2(P) = \lambda_2(UR) = \lambda_2(RU)$. Let $M = RU \in \mathbb{R}^{2n \times 2n}$. In particular,

$$M(i,s_1,j,s_2) = \frac{1}{n}\sum_\sigma \frac{\pi(\sigma)}{\mathbb{P}[\sigma_i = s_1]}\mathbb{I}[\sigma_i = s_1, \sigma_j = s_2] = \frac{\mathbb{P}[\sigma_i = s_1, \sigma_j = s_2]}{n\mathbb{P}[\sigma_i = s_1]} = \frac{1}{n}\mathbb{P}[\sigma_j = s_2 | \sigma_i = s_1]$$

In fact if we take off the projection on the all-ones vector, we have

$$(nM - 1\pi^\top)(i,s_1,j,s_2) = (\mathbb{P}[\sigma_j = s_2|\sigma_i = s_1] - \mathbb{P}[\sigma_j = s_2])$$

The above matrix is very similar to the co-variance/influence matrix. It turns out that

$$nM - 1\pi^\top - I = \begin{bmatrix} A_\pi & -A_\pi \\ B_\pi & -B_\pi \end{bmatrix}$$

where $\Psi_\pi - I = A_\pi - B_\pi$. It then follows that $\Psi_\pi - I$ has the same non-zero eigenvalues as $nM - 1\pi^\top - I$. This implies that $\lambda_2(M) = \frac{1}{n}\lambda_{\max}(nM - 1\pi^\top) = \frac{1}{n}(1 + \lambda_{\max}(\Psi_\pi - I)) \leq \frac{1+\eta}{n}$ as desired. □

11.3 Shattering Lemma and Optimal Poincaré Constant

In this section, we sharpen our previous analysis in the setting of graphical models on bounded-degree graphs. In particular, we assume that G is a graph with maximum degree Δ. and we prove an $\Omega_{\eta,\Delta}(1/n)$

lower bound on the spectral gap of Glauber dynamics without assuming marginal boundedness. This implies $O(n^2)$-mixing. In future lectures we will bound the MLS constant by $1/n$ which gives the optimal $O(n \log n)$ mixing.

First, recall that

Lemma 11.10. *Let G be a graph with n vertices and maximum degree Δ. Then, for any any positive integer $\ell \geq 1$, and for any vertex $v \in V$,*

$$\mathbb{P}_{S \in \binom{n}{\theta n}}\left[|S_v| = \ell\right] \leq (2e\Delta\theta)^{\ell-1}$$

Here S_v is the unique maximal connected component of the induced subgraph $G[S]$ that contains v.

The main observation is the following lemma of Borgs-CHayes-Kahn-Lovász

Lemma 11.11. *The number of connected induced subgraph of G that contains v is at most $(e\Delta)^{\ell-1}$.*

Proof Sketch. Let T be the complete Δ-ary tree of depth ℓ rooted at r. The observation is that the number of subtrees of T rooted at r with exactly ℓ vertices is exactly

$$\frac{1}{\ell}\binom{\ell\Delta}{\ell-1} \leq \frac{(e\Delta)^{\ell-1}}{2}$$

(see e.g., Stanley's book for a proof). Now, the number of subtrees of G with ℓ nodes containing v is at most that of T (one can give a one-to-one mapping of such subtrees of G to subtrees of T). □

Proof. Now, we are read to finish the proof. Let $k = \theta n$.

$$\begin{aligned}
\mathbb{P}_{S \sim \binom{n}{k}}\left[|S_v| = \ell\right] &\leq \sum_{U \in \binom{V}{\ell}, v \in U, G[U] \text{ conn}} \mathbb{P}_S\left[U \subseteq S\right] \quad \leq \left|\left\{U \in \binom{V}{\ell} : v \in U, G[U] \text{ conn}\right\}\right| \cdot \mathbb{P}\left[U \subseteq S\right] \\
&\leq (e\Delta)^{\ell-1}\theta^{\ell} && \text{(Lemma 11.11)} \\
&\leq (e\Delta\theta)^{\ell-1} && (k = \theta n)
\end{aligned}$$

To see the second to the last line say $U = \{u_1, \ldots, u_\ell\}$; then

$$\mathbb{P}\left[U \subseteq S\right] = \mathbb{P}\left[u_1 \in S\right]\mathbb{P}\left[u_2 \in S | u_1 \in S\right] \ldots \mathbb{P}\left[u_\ell \in S | u_1, \ldots, u_{\ell-1} \in S\right] \leq \theta^{\ell}.$$

□

Theorem 11.12. *If G has maximum degree Δ (and $\pi(\sigma) \geq 1/C^n$ for a constant C and any σ then, the Poincaré constant of the Glauber dynamics is at most $O_{\eta,\Delta}(1/n)$.*

Proof. Let $k = (1-\theta)n$ for a value of $\theta \ll 1/\Delta$ that we choose later. Similar to before, we can write

$$\begin{aligned}
\mathrm{Var}(f) &\leq \theta^{-(1+\eta)}\mathbb{E}_{S \sim \binom{n}{k}}\mathbb{E}_{\sigma \sim \pi^S} \mathrm{Var}_{\pi_{S,\sigma}}(f) \\
&= \theta^{-(1+\eta)}\mathbb{E}_{S \sim \binom{n}{k}} \sum_{U \text{ comp of } G \setminus S} \mathrm{Var}_{\pi_{G[U]}}(f_U) && \text{(Independence of components)} \\
&\leq \theta^{-(1+\eta)}\mathbb{E}_{S \sim \binom{n}{n-k}}\mathbb{E}_{\tau \sim \pi^S} \sum_{v \notin S} \mathbb{E}_{\sigma \sim \pi_\tau^{V-v}} \lambda_{|S_v|} \mathrm{Var}_{\pi_{V-v,\tau,\sigma}}(f) \\
&\leq \theta^{-(1+\eta)} \sum_v \mathbb{E}_{\sigma \sim \pi^{V-v}} \mathrm{Var}_{\pi_{V-v,\sigma}}(f) \cdot \sum_\ell \lambda_\ell (2e\Delta\theta)^{\ell-1} \\
&\leq O_{\eta,\Delta}(n)\mathbb{E}_v m E_{\sigma \sim \pi^{V-v}} \mathrm{Var}_{\pi_{V-v,\sigma}}(f)
\end{aligned}$$

In third equation α_ℓ is the worst Poincaré constant over (connected) graphs of size ℓ constaining i. The second to the last line uses that the probability that $|S_v|$ has size ℓ for S of size $|S| = n - k = \theta n$ is at most $2e\Delta\theta)^{\ell-1}$. The last line simply uses that connectivity and that the Poincaré constant, λ_ℓ is at most exponentially large in ℓ. ☐

References

[AL20] A. Alev and L. C. Lau. "Improved analysis of higher order random walks and applications". In: *Combinatorica* 40.5 (2020), pp. 703–731 (cit. on p. 11-3).

[ALO21] N. Anari, K. Liu, and S. Oveis Gharan. "Spectral Independence in High-Dimensional Expanders and Applications to the Hardcore Model". In: *Siam Journal of Computing* (July 2021). Special Issue of FOCS 2020, FOCS20–1–FOCS20–37 (cit. on p. 11-3).

[DK17] I. Dinur and T. Kaufman. "High Dimensional Expanders Imply Agreement Expanders". In: *FOCS*. 2017, pp. 974–985 (cit. on p. 11-3).

[KO20] T. Kaufman and I. Oppenheim. "High Order Random Walks: Beyond Spectral Gap". In: *Combinatorica* 40 (2020), 245–281 (cit. on p. 11-3).