6.S891 Lecture 18: Log-Concave Polynomials, Matroids, and the "Trickle-Down" Method

Kuikui Liu

November 14, 2023

In this lecture, we build an FPRAS for counting bases of a matroid. We study spectral independence for the uniform distribution of matroid bases, and establish nearly-linear mixing time bounds for a natural Markov chain. We also connect this theory to log-concavity of the bases generating polynomial.

1 Matroids

We will use just one of the many *cryptomorphic* definitions of a matroid. For a more comprehensive treatment of matroids, see [Oxl11].

Definition 1 (Matroid; Independent Set Definition)**.** *A* matroid $\mathcal{M}$ *is a pair* $(\mathcal{U}, \mathcal{X})$*, where* $\mathcal{U}$ *is a finite ground set, and* $\mathcal{X} \subseteq 2^{\mathcal{U}}$ *is a family of subsets of* $\mathcal{U}$ *satisfying the following properties:*

- ***Downwards Closure:*** *If* $T \in \mathcal{X}$ *and* $S \subseteq T$*, then* $S \in \mathcal{X}$ *as well.*[1]
- ***Exchange Property:*** *If* $S, T \in \mathcal{X}$ *and* $|T| > |S|$*, then there exists* $u \in T \setminus S$ *such that* $S \cup \{u\} \in \mathcal{X}$*.*

An element of $\mathcal{X}$ *is called an* independent set*, and a maximal independent set is called a* basis*. It is well-known that all bases have the same cardinality.*[2] *This common cardinality is called the* rank *of the matroid.*

Matroids were initially introduced in the 1930s [Whi35] as a combinatorial abstraction of the idea of *linear independence* in linear algebra (hence, the name "independent sets"). They have been intensely studied for decades in combinatorial optimization, polyhedral and topological combinatorics, discrete mathematics, mathematical economics, mathematical logic, and more [Oxl11]. There is an entire theory of discrete convexity built around them [Mur03]. Here are some of the prototypical examples.

Example 1 (Graphic Matroids). Let $G = (V, E)$ be a graph. The *graphic matroid* $\mathcal{M} = (\mathcal{U}, \mathcal{X})$ associated to G is given by $\mathcal{U} = E$ and $\mathcal{X} = \{\text{Acyclic } F \subseteq E\}$. If G is connected, then the bases of $\mathcal{M}$ are precisely the *spanning trees* in G, and the rank is $|V| - 1$.

Example 2 (Linear Matroids). Let $\{v_1, \dots, v_m\} \subseteq V$ be a collection of vectors in some abstract vector space V (e.g. $\mathbb{R}^n$). The *linear matroid* $\mathcal{M} = (\mathcal{U}, \mathcal{X})$ associated to $\{v_1, \dots, v_m\}$ is given by $\mathcal{U} = [m]$ and $\mathcal{X} = \{S \subseteq [m] : \{v_i : i \in S\} \text{ is linearly independent}\}$. In other words, the independent sets of $\mathcal{M}$ are precisely the linearly independent subsets of $\{v_1, \dots, v_m\}$. In this case, the rank of $\mathcal{M}$ is precisely the dimension of the linear space of the vectors $\{v_1, \dots, v_m\}$.

Matroids enjoy a number of useful closure properties, which can be directly checked from the definition.

Fact 1.1 (Closure Properties)**.** *Let* $\mathcal{M} = (\mathcal{U}, \mathcal{X})$ *be a rank-*n *matroid. Then the following are all matroids as well.*

- ***Contraction:*** *For any independent set* $S \in \mathcal{X}$*, define the* contraction $\mathcal{M}/S$ *as the matroid with ground set* $\mathcal{U} \setminus S$ *and independent sets* $\{T \setminus S : S \subseteq T \in \mathcal{X}\}$*.*

[1] In the language of algebraic topology, $\mathcal{X}$ is an *abstract simplicial complex.*

[2] In the language of algebraic topology, the simplicial complex $\mathcal{X}$ is *pure*

- ***Restriction:*** *For any $S \subseteq \mathcal{U}$, define the* restriction $\mathcal{M} \mid S$ *as the matroid with ground set S and independent sets $\{T \in \mathcal{X} : T \subseteq S\}$.*
- ***Truncation:*** *For any $0 \le k \le n$, define the* truncation $\mathcal{M}_{\le k}$ *as the matroid with ground set $\mathcal{U}$ and independent sets $\{T \in \mathcal{X} : |T| \le k\}$.*

Our goal is to sample a uniformly random basis in an arbitrary, or equivalently, count the number of bases. By Fact 1.1, such an algorithm would also enable us to count and sample independent sets of size-k for any k. It can also be converted into an algorithm for counting and sampling all independent sets. For applications of solving these problems (e.g. to reliability theory), see [Ana+19] and reference therein.

To sample a uniformly random basis, we will simulate a simple Markov chain called the *basis exchange walk*. It is the natural analog of Glauber dynamics for this problem. If we are currently at a basis B, then we randomly transition to the next basis B' as follows:

1. Select a uniformly random element $e \in B$ and remove e from B.
2. Out of all elements $f \in \mathcal{U}$ such that $B - e + f$ is again a basis, choose one such f uniformly at random and transition to $B' = B - e + f$. Note that we can pick $f = e$.

Let P_{Ex} denote the transition matrix of this Markov chain.

Fact 1.2 (Ergodicity of P_{Ex})**.** P_{Ex} *is ergodic, and is reversible w.r.t. the uniform distribution over bases of $\mathcal{M}$.*

Reversibility can be directly checked. Irreducibility follows from the exchange property, since between any pairs of bases B, B', one can inductively find a sequence of exchanges to transform B into B'. Aperiodicity follows from the fact that we can choose $f = e$. In this lecture, we establish fast mixing of P_{Ex}.

Theorem 1.3. *For every matroid $\mathcal{M}$ of rank-n, the modified log-Sobolev constant of P_{Ex} satisfies $\varrho(\mathsf{P}_{\mathsf{Ex}}) \ge 1/n$. In particular, $\mathrm{T}_{\mathsf{mix}}(\epsilon; \mathsf{P}_{\mathsf{Ex}}) \le O(n \log n + n \log\log m)$ where $m = |\mathcal{U}|$.*

Remark 1. This bound is sharp by considering a *partition matroid* on $2n$ elements: We have n disjoint sets $B_1, \dots, B_n$ each of size-2, and a subset of $\mathcal{U} = \bigsqcup_{i=1}^{n} B_i$ is independent if and only if it has at most element from each of $B_1, \dots, B_n$. In this case, the distribution μ can be identified with the uniform measure over $\{0,1\}^n$, and P_{Ex} is exactly Glauber dynamics.

A spectral gap of $1/n$ along with polynomial-time mixing was first established in [Ana+19] building on prior works on log-concave polynomials [AOV21] and the recently emerging theory of high-dimensional expanders, in particular [KO18]. Building on this, the modified log-Sobolev bound stated above was subsequently obtained by [CGM21] using local-to-global entropy contraction. More recently, it was shown that the mixing time is $O(n \log n)$, independent of m [Ana+21].

2 Spectral and Entropic Independence in the Homogeneous Setting

Since are restricting attention to distributions over the slice $\binom{[m]}{n}$, let us set up some notation specific to this setting. For $S \subseteq [m]$, we write μ^S for the conditional measure over $\{T \in \mathrm{supp}(\mu) : T \supseteq S\}$, given by $\mu^S(T) = \frac{\mu(T)}{\sum_{T' \supseteq S} \mu(T')}$; note this definition only makes sense if there exists at least one $T \in \mathrm{supp}(\mu)$ such that $T \supseteq S$. For every $0 \le k \le n$, define $\mathcal{D}^{n \searrow k}$ as the linear operator which acts on distributions as follows. For any distribution μ on $\binom{[m]}{n}$, we obtain a new distribution $\mu_k \overset{\mathsf{def}}{=} \mu \mathcal{D}^{n \searrow k}$ on $\binom{[m]}{k}$ which is the law of the output of the following process:

1. First sample $T \sim \mu$, a set in $\binom{[m]}{n}$.
2. Output a uniformly random size-k subset $S \subseteq T$.

This "down" operator is agnostic to μ, and the distribution $\mu_0, \dots, \mu_n$ are (mixtures of) various marginal distributions of μ itself. We saw them previously when we discussed local-to-global entropy contraction.

Let us now define an "up" operator $\mathcal{U}_{\mu}^{k \nearrow n}$ which is in some sense "dual" to $\mathcal{D}^{n \searrow k}$, and does depend on the reference measure μ. For a distribution ν_k on $\binom{[m]}{k}$, we obtain a new distribution $\nu = \nu_k \mathcal{U}_{\mu}^{k \nearrow n}$ on $\binom{[m]}{n}$ as the law of the output of the following process:

1. First sample $S \sim \nu_k$, a set in $\binom{[m]}{k}$.
2. Output a random $T \in \binom{[m]}{n}$ drawn according to the conditional measure μ^S.

The composition of these two operators yield Markov chains on various "levels" $\binom{[m]}{k}$. For instance, the *down-up walk* $\mathcal{D}^{n \searrow n-1} \mathcal{U}_{\mu}^{n-1 \nearrow n}$ is a Markov chain on $\binom{[m]}{n}$, which is reversible to μ, given by the following two-step process: Starting from $S \in \binom{[m]}{n}$, we

1. remove a uniformly random element $i \in S$ from S, and
2. add a random element $j \in [m]$ such that $S - i + j \in \operatorname{supp}(\mu)$ with probability proportional to $\mu(S - i + j)$.

When μ is the uniform measure over bases of a matroid, this down-up Markov chain $\mathsf{P}_{\mu} \stackrel{\mathsf{def}}{=} \mathcal{D}^{n \searrow n-1} \mathcal{U}_{\mu}^{n-1 \nearrow n}$ is exactly the basis exchange walk we described earlier.

Let us now relate the mixing properties of this Markov chain P_{μ} to the correlations within μ. We modify the notions of spectral independence and entropic independence slightly.

Definition 2 (Spectral Independence)**.** *Let μ be a probability measure over the slice $\binom{[m]}{n}$. We define the* influence matrix *$\mathcal{I}_{\mu} \in \mathbb{R}^{m \times m}$ by*

$$\mathcal{I}_{\mu}(i \to j) \stackrel{\mathsf{def}}{=} \Pr_{S \sim \mu}[j \in S \mid i \in S] - \Pr_{S \sim \mu}[j \in S], \qquad \forall i, j \in [m].$$

For $\eta \geq 0$, we say μ is η-spectrally independent *if $\lambda_{\max}(\mathcal{I}_{\mu}) \leq 1 + \eta$.*

For further discussion of the relationship between this version of spectral independence and the one previously defined, see Appendix A.

Fact 2.1. *We have the identity $\mathcal{I}_{\mu} = D_{\mu}^{-1} \operatorname{Cov}(\mu)$, where $D_{\mu} = \operatorname{diag}\left(\Pr_{S \sim \mu}[i \in S]\right)_{i \in [n]}$ and $\operatorname{Cov}(\mu)$ is the covariance matrix of μ after identifying each $S \in \operatorname{supp}(\mu)$ with its $\{0,1\}$-indicator vector $\mathbf{1}_S \in \mathbb{R}^m$. In particular, μ is η-spectrally independent in the sense of Definition 2 if and only if $\operatorname{Cov}(\mu) \preceq (1 + \eta) \cdot D_{\mu}$.*

Definition 3 (Entropic Independence)**.** *Let μ be a probability measure over the slice $\binom{[m]}{n}$. We say μ is* η-entropically independent *if*

$$\mathscr{D}\left(\nu \mathcal{D}^{n \searrow 1} \,\|\, \mu \mathcal{D}^{n \searrow 1}\right) \leq \frac{1+\eta}{n} \cdot \mathscr{D}\left(\nu \,\|\, \mu\right), \qquad \forall \text{ distributions } \nu \text{ on } \binom{[m]}{n}.$$

We have the following analogs of the local-to-global theorems we saw previously, with essentially the same proofs.

Theorem 2.2 (Local-to-Global)**.** *Let μ be a probability measure over the slice $\binom{[m]}{n}$.*

- *Suppose there exists $\eta \geq 0$ such that μ and all of its conditional measures are all η-spectrally independent. Then $\gamma(\mathsf{P}_{\mu}) \geq \Omega\left(n^{-(1+\eta)}\right)$. [AL20; ALO21]*
- *Suppose there exists $\eta \geq 0$ such that μ and all of its conditional measures are all η-entropically independent. Then $\varrho(\mathsf{P}_{\mu}) \geq \Omega\left(n^{-(1+\eta)}\right)$. [Ana+22]*

We can also connect spectral independence to entropic independence just as we did previously.

Theorem 2.3 ([Ana+22])**.** *Let μ be a probability measure over the slice $\binom{[m]}{n}$. For $\boldsymbol{\lambda} \in \mathbb{R}_{\geq 0}^m$, define $\mu_{\boldsymbol{\lambda}}$ to be the distribution on $\binom{[m]}{n}$ given by $\mu_{\boldsymbol{\lambda}}(S) \propto \mu(S) \boldsymbol{\lambda}^S$ where recall $\boldsymbol{\lambda}^S \stackrel{\mathsf{def}}{=} \prod_{i \in S} \lambda_i$. The following are equivalent.*

- *For every $\boldsymbol{\lambda} \in \mathbb{R}_{\geq 0}^m$, $\mu_{\boldsymbol{\lambda}}$ is η-spectrally independent.*
- *For every $\boldsymbol{\lambda} \in \mathbb{R}_{\geq 0}^m$, $\mu_{\boldsymbol{\lambda}}$ is η-entropically independent.*

3 The Mixture Method for Spectral Independence

Previously, we showed that spectral independence implies fast mixing of local Markov chains. The high-level idea there was to decompose the probability measure μ, e.g. based on conditioning the coordinates/elements. Here, we will use the same type of strategy to reason about spectral independence itself.

Theme 3.1. *To establish spectral independence for μ, find a decomposition of μ into a mixture ξ of component measures $\{\nu_\alpha : \alpha \in \mathcal{I}\}$ (for some index set $\mathcal{I}$) such that*

- *the mixture measure ξ on $\mathcal{I}$ itself satisfies some kind of mixing condition (e.g. a Poincaré Inequality), and*
- *each component measure ν_α has well-behaved correlations (e.g. is spectrally independent).*

One instantiation of this we have already seen was when we proved the Glauber dynamics having spectral gap $\Omega(1/n)$ implies $O(1)$-spectral independence for μ. The mixture measure was μ itself, and the component measures were Dirac masses. Another example is in the third problem set, where in the application to Ising models, the mixture measure ξ is log-concave (in $\mathbb{R}^d$) and the component measures are product measures on $\{\pm 1\}^n$. In this lecture, we again decompose μ into its pinnings, and relate the spectral independence of the conditionals of μ to the spectral independence of μ itself.

The main result in this section is the following.

Theorem 3.2 ([Ana+19])**.** *Let $\mathcal{M} = (\mathcal{U}, \mathcal{X})$ be a matroid of rank-n, and let μ be the uniform distribution over its bases. Then for every $\boldsymbol{\lambda} \in \mathbb{R}^{|\mathcal{U}|}_{\geq 0}$, $\mu_{\boldsymbol{\lambda}}$ is 0-spectrally independent (and hence, 0-entropically independent).*

Combined with Theorem 2.2, this immediately implies Theorem 1.3. By going through the proof and tightening all parts of the analysis, we can also get exactly $1/n$ as the lower bound instead of $\Omega(1/n)$.

3.1 Trickling Down and 0-Spectral Independence for Matroids

The following seminal result of Oppenheim was first proved in the setting of high-dimensional expanders, and originally stated the language of simplicial complexes. We state and prove a very special case for convenience.

Theorem 3.3 (Oppenheim's Trickle-Down Theorem (Special Case); [Opp18])**.** *Let μ be a probability measure over the slice $\binom{[m]}{n}$. Assume the following conditions:*

- ***Weak Spectral Independence:*** *The distribution μ satisfies $\lambda_{\max}(\mathcal{I}_\mu) < n - 1$.*
- ***Spectral Independence for Conditionals:*** *For every $i \in [m]$, the conditional measure μ^i is 0-spectrally independent.*

Then μ itself is 0-spectrally independent.

Remark 2. "Weak spectral independence" indeed is very weak, and we almost always get it "for free". One cannot see this fact by studying $\|\mathcal{I}_\mu\|_{\ell_\infty \to \ell_\infty}$. However, it can be reduced to mere connectivity of a certain "local random walk" Q_μ encoding the correlations within μ, which have seen before:

$$Q_\mu(i \to j) \overset{\mathsf{def}}{=} \frac{1}{n-1} \Pr_{S \sim \mu}[j \in S \mid i \in S], \qquad \forall i \neq j.$$

Q_μ is a Markov chain on $[m]$ which is reversible w.r.t. μ_1. One can show that $\lambda_2(Q_\mu) = \frac{\lambda_{\max}(\mathcal{I}_\mu)}{n-1}$, and so "weak spectral independence" is equivalent to connectivity of Q_μ.

Remark 3. Oppenheim's result generalizes to when the conditionals are η-spectrally independent for positive η. The final conclusion is then $\frac{(n-1)\eta}{(n-2)-\eta}$ spectral independence for μ itself. Notably, there is a *degradation* in the spectral independence parameter, which becomes rather severe if this Trickle-Down Theorem is applied too many times. [ALO22] ameliorated this degradation by using more complicated matrix upper bounds, extending the applicability of the trickle-down method to e.g. edge-colorings; see also [AO23; WZZ23].

Proof of Theorem 3.3. Similar to previous lectures, let us write

$$\boldsymbol{m}(\mu) \stackrel{\mathsf{def}}{=} \mathbb{E}_{S\sim\mu}\left[\mathbf{1}_S\right] = \left[\Pr_{S\sim\mu}[i \in S]\right]_{i\in[m]} \in \mathbb{R}^m$$

for convenience. Then $\mu_1 = \mu\mathcal{D}^{n\searrow 1} = \frac{1}{n}\boldsymbol{m}(\mu)$ and μ may be decomposed as a mixture $\mathbb{E}_{i\sim\mu_1}\left[\mu^i\right]$. By the Law of Total Covariance,

$$\mathrm{Cov}(\mu) = \mathbb{E}_{i\sim\mu_1}\left[\mathrm{Cov}\left(\mu^i\right)\right] + \mathrm{Cov}_{i\sim\mu_1}\left(\boldsymbol{m}\left(\mu^i\right)\right).$$

Here, we are still viewing μ^i as a probability measure over $\binom{[m]}{n}$, although all sets in $\mathrm{supp}\left(\mu^i\right)$ are forced to contain the element i. Hence, $\boldsymbol{m}\left(\mu^i\right)$ is still a vector in $\mathbb{R}^m$, and has entry 1 in coordinate i. By our spectral independence assumption for the conditional measures, we have that the first term above is upper bounded as

$$\mathbb{E}_{i\sim\mu_1}\left[\mathrm{Cov}\left(\mu^i\right)\right] \preceq \mathbb{E}_{i\sim\mu_1}\left[\mathrm{diag}\left(\boldsymbol{m}\left(\mu^i\right)\right) - \mathbf{1}_i\mathbf{1}_i^\top\right] = \frac{n-1}{n}\,\mathrm{diag}\left(\boldsymbol{m}(\mu)\right).$$

Here, the correction $\mathbf{1}_i\mathbf{1}_i^\top$ again comes from including the ith coordinate in $\boldsymbol{m}\left(\mu^i\right)$, which has value 1. For the second term, we use the fact that for each $j \in [m]$, the jth entry of $\boldsymbol{m}\left(\mu^i\right)$ is precisely the conditional probability $\Pr[j \mid i]$ appearing in spectral independence.

Claim 3.4. *We have the identity*

$$\mathrm{Cov}_{i\sim\mu_1}\left(\boldsymbol{m}\left(\mu^i\right)\right) = \frac{1}{n}\mathrm{Cov}\left(\mu\right)\mathrm{diag}\left(\boldsymbol{m}(\mu)\right)^{-1}\mathrm{Cov}(\mu).$$

We prove Claim 3.4 in a moment. We first use it to conclude the proof of the theorem. Claim 3.4 and 0-spectral independence for the conditionals gives us the inequality

$$\mathrm{Cov}(\mu) \preceq \frac{n-1}{n}\,\mathrm{diag}\left(\boldsymbol{m}(\mu)\right) + \frac{1}{n}\mathrm{Cov}\left(\mu\right)\cdot\mathrm{diag}\left(\boldsymbol{m}(\mu)\right)^{-1}\cdot\mathrm{Cov}(\mu).$$

Letting $A_\mu = \mathrm{diag}\left(\boldsymbol{m}(\mu)\right)^{-1/2}\mathrm{Cov}(\mu)\,\mathrm{diag}\left(\boldsymbol{m}(\mu)\right)^{-1/2}$, the above inequality is equivalent to

$$A_\mu \cdot \left(\mathsf{Id} - \frac{1}{n}A_\mu\right) \preceq \frac{n-1}{n}\mathsf{Id}.$$

Our goal is to deduce $A_\mu \preceq \mathsf{Id}$, which is exactly 0-spectral independence for μ itself. Let $\boldsymbol{\lambda}$ be any eigenvalue of A_μ. We wish to show that $\boldsymbol{\lambda} \le 1$, and the above inequality asserts that $\boldsymbol{\lambda}\left(1 - \frac{\boldsymbol{\lambda}}{n}\right) \le 1 - \frac{1}{n}$, or equivalently, $(1-\boldsymbol{\lambda})\left(\frac{1+\boldsymbol{\lambda}}{n} - 1\right) \le 0$. The quadratic on the left-hand side has negative leading coefficient, and roots at 1 and $n-1$. Hence, to ensure $\boldsymbol{\lambda} \le 1$, we just need to rule out the possibility that $\boldsymbol{\lambda} \ge n-1$. This is guaranteed by our "weak spectral independence" assumption. □

Proof of Claim 3.4. The left-hand side is given by

$$\begin{aligned}
&\frac{1}{n}\sum_{i\in[m]} m_i(\mu)\cdot\boldsymbol{m}\left(\mu^i\right)^{\otimes 2} - \left(\frac{1}{n}\sum_{i\in[m]}\boldsymbol{m}\left(\mu^i\right)\right)^{\otimes 2} \\
&= \frac{1}{n}\sum_{i\in[m]} m_i(\mu)^{-1}\cdot\left(m_i(\mu)\cdot\boldsymbol{m}\left(\mu^i\right)\right)^{\otimes 2} - \boldsymbol{m}(\mu)^{\otimes 2} \\
&= \frac{1}{n}\mathbb{E}_{S\sim\mu}\left[\mathbf{1}_S\mathbf{1}_S^\top\right]\cdot\mathrm{diag}\left(\boldsymbol{m}(\mu)\right)^{-1}\cdot\mathbb{E}_{S\sim\mu}\left[\mathbf{1}_S\mathbf{1}_S^\top\right] - \boldsymbol{m}(\mu)^{\otimes 2},
\end{aligned}$$

where in the final step, we used the fact that the ith row and column of $\mathbb{E}_{S\sim\mu}\left[\mathbf{1}_S\mathbf{1}_S^\top\right]$ is precisely the vector $m_i(\mu)\cdot\boldsymbol{m}\left(\mu^i\right)$. Expanding the right-hand side yields the same expression after using the fact that $\mathbf{1}^\top\boldsymbol{m}(\mu) = \sum_{i\in[m]}\Pr_{S\sim\mu}[i \in S] = n$. □

Proof of Theorem 3.2. We prove the claim for the special case $\boldsymbol{\lambda} = \mathbf{1}$; the case of general $\boldsymbol{\lambda} \in \mathbb{R}^{\mathcal{U}}_{\ge 0}$ is similar. By Theorem 3.3 and induction, it suffices to prove that for every independent set $S \in \mathcal{X}$ of size-$(n-2)$, the conditional distribution μ^S is 0-spectrally independent. Note that μ^S is just the uniform distribution over the bases of the contraction $\mathcal{M}_S$, which has rank-2 since S is independent and has cardinality $n-2$. Hence, we just need to establish 0-spectral independence for all rank-2 matroids.

Let $\mathcal{M}$ be a rank-2 matroid, and let μ denote the uniform distribution over its bases. Our goal is to establish 0-spectral independence. By Fact 2.1, this is equivalent to establishing that

$$\mathbb{E}_{S\sim\mu}\left[\mathbf{1}_S\mathbf{1}_S^\top\right] - \operatorname{diag}\left(\boldsymbol{m}(\mu)\right) \preceq \boldsymbol{m}(\mu)\boldsymbol{m}(\mu)^\top. \tag{1}$$

We use the following purely linear algebraic fact, whose proof is provided in Appendix B.

Lemma 3.5. *Let $A \in \mathbb{R}_{\geq 0}^{n\times n}$ be a nonnegative symmetric matrix which is not all zero. Then the following are equivalent.*

1. *For every $v \in \mathbb{R}_{>0}^n$ and $x \in \mathbb{R}^n$, $\left(x^\top Ax\right)\cdot\left(v^\top Av\right) \leq \left(x^\top Av\right)^2$.*
2. *There exists $v \in \mathbb{R}_{>0}^n$ such that for every $x \in \mathbb{R}^n$, $\left(x^\top Ax\right)\cdot\left(v^\top Av\right) \leq \left(x^\top Av\right)^2$.*
3. *A has exactly one positive eigenvalue.*

We prove Lemma 3.5 in a moment. We first use it to establish Eq. (1). We instantiate Lemma 3.5 with

$$A = \mathbb{E}_{S\sim\mu}\left[\mathbf{1}_S\mathbf{1}_S^\top\right] - \operatorname{diag}\left(\boldsymbol{m}(\mu)\right), \tag{2}$$

the left-hand side in Eq. (1), and $v = \mathbf{1}$. Note that $Av = \boldsymbol{m}(\mu)$ since all sets $S \in \operatorname{supp}(\mu)$ have size-2. Similarly, $\mathbf{1}^\top A\mathbf{1} = 2$. We will establish that $2x^\top Ax \leq \left(x^\top A\mathbf{1}\right)^2$ for all $x \in \mathbb{R}^n$, which is stronger than Eq. (1). By Lemma 3.5, to prove the former, it suffices to certify that A has at most one positive eigenvalue. This is where we will use the fact that we are looking at a matroid.

Since μ is uniform over the bases of a rank-2 matroid, up to scaling by some positive constant (namely $\#\{\text{Bases}\}^{-1}$), A is just the $\{0,1\}$-adjacency matrix of the following graph $G_{\mathcal{M}}$ on vertex set $\mathcal{U}$: Two distinct elements $i, j \in \mathcal{U}$ are connected by an edge if and only if $\{i, j\}$ is a basis in $\mathcal{M}$. In other words, $G_{\mathcal{M}}$ is the complement of the dependency graph of G.[3]

Claim 3.6. *The graph $G_{\mathcal{M}}$ is a* complete multipartite *graph, i.e. there exists a partition $A_1 \sqcup \cdots \sqcup A_k$ of $\mathcal{U}$ such that $i \sim j$ if and only if i, j belong in different parts of the partition.*

Proof. We claim that the relation $i \not\sim j$ in $G_{\mathcal{M}}$ is an *equivalence relation* on $\mathcal{U}$. Indeed if $i, j, k \in \mathcal{U}$ are distinct elements and $i \sim k$, then applying the exchange axiom to $S = \{j\}$ and $T = \{i, k\}$ implies that $i \sim j$ or $j \sim k$. The contrapositive says that $i \not\sim j$ and $j \not\sim k$ implies $i \not\sim k$. Now that we know $\not\sim$ is an equivalence relation, we may take $A_1, \ldots, A_k$ to be the equivalence classes, concluding the proof. $\square$

We now use Claim 3.6 to show that A has at most one positive eigenvalue. Observe that

$$A \propto \mathsf{Adj}(G_{\mathcal{M}}) = \mathbf{1}\mathbf{1}^\top - \sum_{i=1}^{k}\mathbf{1}_{A_i}\mathbf{1}_{A_i}^\top.$$

Since A is equal to a rank-1 matrix minus a positive semidefinite matrix, it has exactly one positive eigenvalue as desired. $\square$

4 Connections with Log-Concave Polynomials

While none of the above results or proofs invoked polynomials, we can capture them again using the generating polynomial for the distribution μ, which recall is given by $g_\mu(\boldsymbol{z}) = \sum_{S\subseteq[m]}\mu(S)\boldsymbol{z}^S$. If μ is supported on the slice $\binom{[m]}{n}$, then g_μ is a *homogeneous* multiaffine polynomial of degree-m. We have the following equivalence between log-concavity of g_μ and 0-spectral independence.

Lemma 4.1. *Let μ be a probability measure over the slice $\binom{[m]}{n}$. For every $\boldsymbol{\lambda} \in \mathbb{R}_{\geq 0}^m$, the following are equivalent:*

[3] In the language of algebraic topology, where we view $\mathcal{M}$ as a simplicial complex, $G_{\mathcal{M}}$ is the 1-*skeleton* graph of $\mathcal{M}$.

- 0-***Spectral Independence:*** *The distribution* $\mu_{\boldsymbol{\lambda}}$ *is* 0*-spectrally independent.*
- ***Log-Concavity of*** g_μ***:*** *The polynomial* g_μ *is log-concave at* $\boldsymbol{\lambda}$*, i.e.* $\nabla^2 \log g_\mu(\boldsymbol{\lambda}) \preceq 0$.

Proof. A direct calculation reveals that

$$\mathrm{Cov}\left(\mu_{\boldsymbol{\lambda}}\right) - \mathrm{diag}\left(\boldsymbol{m}\left(\mu_{\boldsymbol{\lambda}}\right)\right) = \mathrm{diag}(\boldsymbol{\lambda}) \cdot \nabla^2 \log g_\mu(\boldsymbol{\lambda}) \cdot \mathrm{diag}(\boldsymbol{\lambda}).$$

Since $\boldsymbol{\lambda} \in \mathbb{R}^m_{\geq 0}$, the former is negative semidefinite (i.e. μ is 0-spectrally independent) if and only if g_μ is log-concave at $\boldsymbol{\lambda}$. □

A straightforward generalization of the proof also shows that η-spectral independence for $\mu_{\boldsymbol{\lambda}}$ is equivalent to concavity of the function $\boldsymbol{z} \mapsto \log g_\mu\left(\boldsymbol{z}^{\frac{1}{1+\eta}}\right)$ at $\boldsymbol{z} = \boldsymbol{\lambda}$ [Ali+21].

The following beautiful theorem was first proved in [AOV21], building on seminal work of Adiprasito–Huh–Katz [AHK18]. Using essentially Lemma 4.1 and Theorem 3.2, an elementary proof was given in [Ana+19].

Theorem 4.2 ([AOV21])**.** *The generating polynomial for the uniform measure over bases of any matroid is log-concave on all of* $\mathbb{R}^{\mathcal{U}}_{\geq 0}$.

The theory of log-concave polynomials was first studied by Gurvits [Gur06] towards proving combinatorial log-concavity inequalities between sequences of numbers. In a momumental breakthrough, Adiprasito–Huh–Katz built a "combinatorial Hodge theory" for proving such inequalities, resolving many decades-old conjectures on log-concavity of the coefficients of the chromatic polynomial, of the number of independent sets of a fixed size in any matroid, and more [AHK18]. Many of these conjectures now have elementary proofs using the theory of log-concave polynomials [Ana+18; BH20; BL23]. Log-concave polynomials are often called *Lorentzian polynomials* [BH20], particularly in the algebraic combinatorics community where there are many applications.

5 Open Problems

We conclude this lecture with some open problems.

5.1 Deterministic Algorithms

The basis exchange walk only assumes that we have an *independence oracle* for matroid. This is some blackbox function which outputs whether or not a subset $S \subseteq \mathcal{U}$ is independent or not; we only get to query it with subsets of $\mathcal{U}$, without knowing how this function works. It turns out that in this restrictive model, there can be no FPTAS for counting bases of a matroid [ABF94]. In this oracle setting, there do exist deterministic algorithms which can approximate the number of bases of an arbitrary matroid up to a multiplicative factor of $2^{O(r)}$ [AOV21]. Note that the naïve approximation factor is $\binom{n}{r} \approx n^{O(r)}$. However, in many applications, we have much more than just access to an independence oracle. Hence, the following question is very natural.

Question 1. *Does there exist an* FPTAS *for counting bases of an "explicitly given" matroid? For instance, given a collection of vectors* $v_1, \ldots, v_n \in \mathbb{R}^d$*, does there exist an* FPTAS *for counting the bases of the linear matroid induced by* $\{v_1, \ldots, v_n\}$*?*

An example of a positive result in this direction is the classical Kirchhoff Matrix Tree Theorem [Kir47]. It says that we can express the number of spanning trees in a graph as the determinant of an $n \times n$ matrix, which can then be computed exactly and deterministically in polynomial-time.

5.2 Matroid Intersection

Now that we can approximately count bases of a single matroid, a natural extension is to count the common bases of two matroids. This simultaneously generalizes the single matroid case, as well as many other important combinatorial problems (e.g. perfect matchings in bipartite graphs, arborescences, rainbow spanning trees, etc.). This problem remains wide open.

Question 2. *Does there exist an* FPRAS *for counting the common bases of two matroids* $\mathcal{M}, \mathcal{N}$ *on the same ground set?*

[AOV21] gave a deterministic $2^{O(r)}$-approximation for this problem as well. Note that finding a common basis of *three* matroids is already an NP-hard problem, since it can encode the problem of finding Hamiltonian path.

References

[ABF94] Y. Azar, A.Z. Broder, and A.M. Frieze. "On the problem of approximating the number of bases of a matroid". In: *Information Processing Letters* 50.1 (1994), pp. 9–11. ISSN: 0020-0190 (cit. on p. 7).

[AHK18] Karim Adiprasito, June Huh, and Eric Katz. "Hodge theory for combinatorial geometries". In: *Annals of Mathematics* 188.2 (2018), pp. 381–452 (cit. on p. 7).

[AL20] Vedat Levi Alev and Lap Chi Lau. "Improved Analysis of Higher Order Random Walks and Applications". In: *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2020. Chicago, IL, USA: Association for Computing Machinery, 2020, pp. 1198–1211 (cit. on p. 3).

[Ali+21] Yeganeh Alimohammadi, Nima Anari, Kirankumar Shiragur, and Thuy-Duong Vuong. "Fractionally Log-Concave and Sector-Stable Polynomials: Counting Planar Matchings and More". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2021. Virtual, Italy: Association for Computing Machinery, 2021, pp. 433–446 (cit. on p. 7).

[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. 3).

[ALO22] Dorna Abdolazimi, Kuikui Liu, and Shayan Oveis Gharan. "A Matrix Trickle-Down Theorem on Simplicial Complexes and Applications to Sampling Colorings". In: *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*. 2022, pp. 161–172 (cit. on p. 4).

[Ana+18] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. "Log-Concave Polynomials III: Mason's Ultra-Log-Concavity Conjecture for Independent Sets of Matroids". In: *arXiv preprint arXiv:1811.01600* (2018) (cit. on p. 7).

[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 Theory of Computing*. STOC 2019. Pheonix, Arizona, USA: Association for Computing Machinery, 2019, pp. 1–12 (cit. on pp. 2, 4, 7).

[Ana+21] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, Cynthia Vinzant, and Thuy-Duong Vuong. "Log-Concave Polynomials IV: Approximate Exchange, Tight Mixing Times, and near-Optimal Sampling of Forests". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2021. Virtual, Italy: Association for Computing Machinery, 2021, pp. 408–420 (cit. on p. 2).

[Ana+22] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. "Entropic Independence: Optimal Mixing of down-up Random Walks". In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2022. Rome, Italy: Association for Computing Machinery, 2022, pp. 1418–1430 (cit. on p. 3).

[AO23] Dorna Abdolazimi and Shayan Oveis Gharan. "An Improved Trickle down Theorem for Partite Complexes". In: *38th Computational Complexity Conference (CCC 2023)*. Ed. by Amnon Ta-Shma. Vol. 264. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 10:1–10:16. ISBN: 978-3-95977-282-2. DOI: 10.4230/LIPIcs.CCC.2023.10 (cit. on p. 4).

[AOV21] Nima Anari, Shayan Oveis Gharan, and Cynthia Vinzant. "Log-concave polynomials, I: Entropy and a deterministic approximation algorithm for counting bases of matroids". In: *Duke Mathematical Journal* (2021), pp. 1–46 (cit. on pp. 2, 7, 8).

[BH20] Petter Brändén and June Huh. “Lorentzian polynomials”. In: *Annals of Mathematics* 192 (3 2020), pp. 821–891 (cit. on p. 7).

[BL23] Petter Brändén and Jonathan Leake. “Lorentzian polynomials on cones”. In: *arXiv preprint arXiv:2304.13203* (2023) (cit. on p. 7).

[CGM21] Mary Cryan, Heng Guo, and Giorgos Mousa. “Modified log-Sobolev inequalities for strongly log-concave distributions”. In: *Annals of Probability* 49.1 (2021). A preliminary version of this paper appeared in Proceedings of the 2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)., pp. 506–525 (cit. on p. 2).

[Gur06] Leonid Gurvits. “Hyperbolic polynomials approach to Van der Waerden/Schrijver-Valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications”. In: *STOC*. 2006, pp. 417–426 (cit. on p. 7).

[Kir47] G. Kirchhoff. “Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird”. In: *Annalen der Physik* 148 (12 1847), pp. 497–508 (cit. on p. 7).

[KO18] Tali Kaufman and Izhar Oppenheim. “Construction of New Local Spectral High Dimensional Expanders”. In: *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2018. Los Angeles, CA, USA: Association for Computing Machinery, 2018, pp. 773–786. ISBN: 9781450355599. DOI: 10.1145/3188745.3188782 (cit. on p. 2).

[Mur03] Kazuo Murota. *Discrete Convex Analysis*. Monographs on Discrete Mathematics and Applications. Society for Industrial and Applied Mathematics, 2003, p. 389 (cit. on p. 1).

[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 (cit. on p. 4).

[Oxl11] James Oxley. *Matroid Theory*. 2nd ed. Oxford Graduate Texts in Mathematics. Oxford University Press, 2011 (cit. on p. 1).

[Whi35] Hassler Whitney. “On the Abstract Properties of Linear Dependence”. In: *American Journal of Mathematics* 57.3 (1935), pp. 509–533 (cit. on p. 1).

[WZZ23] Yulin Wang, Chihao Zhang, and Zihan Zhang. “Sampling Proper Colorings on Line Graphs Using $(1+o(1))\Delta$ Colors”. In: *arXiv preprint arXiv:2307.08080* (2023) (cit. on p. 4).

A A Remark on Homogeneity

We used a different notation $\mathcal{I}_\mu$ for the influence matrix in the homogeneous setting to distinguish it from the previous one Ψ_μ, where $\Psi_\mu(i \to j) = \Pr_{S\sim\mu}[j \in S \mid i \in S] - \Pr_{S\sim\mu}[j \in S \mid i \notin S]$. Both notion of influence make sense, but requiring $\lambda_{\max}(\mathcal{I}_\mu) \le 1+\eta$ is less stringent than requiring $\lambda_{\max}(\Psi_\mu) \le 1+\eta$.

Homogeneity is actually not a restrictive assumption, since any distribution (e.g. over $\{\pm 1\}^n$ or $[q]^n$) can always be *homogenized*. For instance, let μ be a distribution over $\{\pm 1\}^n$ (e.g. all independent sets in some graph), which a priori is not homogeneous. We define a homogenized version μ^{hom} over the slice $\binom{[n]\times\{\pm 1\}}{n}$, which is supported on sets of the form $S_\sigma \stackrel{\mathsf{def}}{=} \{(i,\sigma(i)) : i \in [n]\} \in \binom{[n]\times\{\pm 1\}}{n}$ and given by $\mu^{\mathsf{hom}}(S_\sigma) = \mu(\sigma)$.

This homogenized distribution μ^{hom} possesses all of the same essential features as that of μ. Glauber dynamics for μ is exactly the down-up walk for the homogenized distribution μ^{hom}. Furthermore, one can show that Ψ_μ and $\mathcal{I}_{\mu^{\mathsf{hom}}}$ have the same eigenvalues (up to multiplicity of the zero eigenvalue). In particular, η-spectral independence for μ in the sense we previously discussed is the same as η-spectral independence for μ^{hom} in the sense of Definition 2.

B Unfinished Proofs

Proof of Lemma 3.5. Item 1 clearly implies Item 2. Suppose Item 2 holds. Then the matrix

$$A - \frac{(Av)(Av)^\top}{v^\top Av}$$

is negative semidefinite. Note that $\mathbf{1}^\top A\mathbf{1} > 0$ and so A has at least one positive eigenvalue. If A has more than one positive eigenvalue, then there exist a subspace $W \subseteq \mathbb{R}^n$ of dimension ≥ 2 on which A is a positive definite quadratic form. But this implies $A - \frac{(Av)(Av)^\top}{v^\top Av}$ is positive definite on $W \cap \operatorname{span}\{Av\}^\perp$, which has dimension ≥ 1, contradicting $A - \frac{(Av)(Av)^\top}{v^\top Av}$ being negative semidefinite. Hence, Item 3 must hold.[4]

Finally, suppose Item 3 holds. Then we can express A as $A = B + ww^\top$, where B is negative semidefinite and $w \in \mathbb{R}^n$. Now let $v \in \mathbb{R}^n_{>0}$ and $x \in \mathbb{R}^n$ be arbitrary, and let $P \in \mathbb{R}^{2\times n}$ have v, x as its rows. Then $PAP^\top = PBP^\top + (Pw)(Pw)^\top \in \mathbb{R}^{2\times 2}$ has at most one positive eigenvalue (e.g. by the same argument as we did above). Furthermore

$$PAP^\top = \begin{bmatrix} v^\top Av & x^\top Av \\ x^\top Av & x^\top Ax \end{bmatrix}.$$

Since $v \in \mathbb{R}^n_{>0}$, $v^\top Av > 0$ and so $PAP^\top$ has at least one positive eigenvalue. It follows that $PAP^\top$ has exactly one positive eigenvalue and one negative eigenvalue, and so

$$\left(v^\top Av\right) \cdot \left(x^\top Ax\right) - \left(x^\top Av\right)^2 = \det\left(PAP^\top\right) \leq 0.$$

Since $v \in \mathbb{R}^n_{>0}$ and $x \in \mathbb{R}^n$ were arbitrary, we have Item 1. □

[4]A more streamlined approach would have been to just apply the Cauchy Interlacing Theorem.