

Log Concave Polynomials II: High-Dimensional Walks and an FPRAS for Counting Bases of a Matroid

Nima Anari, Kuikui Liu, Shayan Oveis Gharan, Cynthia Vinzant

Presented by Alex Albors Juez and Rohan Mukherjee

CSE 521, University of Washington

December 14th, 2023

Table of contents

1. Motivation

Counting bases of Matroids

2. Preliminaries

Simplicial Complexes

Log-concave polynomials

Markov Chains and Random Walks

3. Walks on Simplicial Complexes

Upper and lower walks

0-local-spectral-expanders

$$\lambda^*(P_d^\wedge) = \lambda^*(P_{d-1}^\vee)$$

4. From log-concavity to Local Spectral Expanders

Strong Log-Concavity implies 0-local-spectral-expander

5. Putting everything together: Proof of Theorem 1.1

6. References

Matroids

A *matroid* M is a pair $M = (X, \mathcal{I})$ where X is a finite set and $\mathcal{I} \subseteq 2^X$ so that the following holds:

- (i) *Non-emptiness*: $\emptyset \in \mathcal{I}$
- (ii) *Monotonicity*: If $Y \in \mathcal{I}$ and $Z \subseteq Y$ then $Z \in \mathcal{I}$.
- (iii) *Exchange property*: If $Y, Z \in \mathcal{I}$ and $|Y| < |Z|$, then for some $x \in Z \setminus Y$ we have $Y \cup \{x\} \in \mathcal{I}$

Definition (basis)

Let $M = (X, \mathcal{I})$ be a matroid. A maximal independent set $B \in \mathcal{I}$ is called a *basis* of X . All basis elements have the same size, and their size is called the *rank* of the matroid.

Example: The Acyclic subsets of a graph (forests) form a matroid, called a *graphic matroid*.

Bases exchange walk

Procedure:

1. Start with a basis element B .
2. Drop a random element i from B . Pick j uniformly at random from $\{1, \dots, n\}$, and try adding it to $B \setminus \{i\}$. Do it until we can.
3. Repeat step 2.

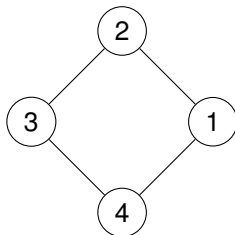


Figure 1: Graph C_4 corresponding to a rank 3 graphic matroid

History

- ▶ 30 years ago, Mihail and Vazirani conjectured that the bases exchange walk mixes in polynomial time.
- ▶ Polynomial mixing time corresponds to being able to count bases in polynomial time (Approximate sampling and approximate counting are equivalent in this scenario [2, JVV86]).
- ▶ Barvinok and Samorodnitsky designed a randomized algorithm that gives a $\log(n)^r$ approx. factor for a matroid with n elements and rank r [4, BS07].
- ▶ In Log-concave polynomials I, Gharan et al. give a deterministic algorithm that returns an e^r approximation factor.[3, AKOV18]
- ▶ In this paper, Gharan et al. give a randomized algorithm yielding a $1 \pm \epsilon$ approximation factor in polynomial time.

Main theorem

Theorem (1.1)

Let $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$ be a d -homogeneous strongly log concave probability distribution. If P_μ denotes the transition probability matrix of M_μ and $X(k)$ denotes the set of size- k subsets of $[n]$ which are contained in some element of $\text{supp}(\mu)$, then for every $0 \leq k \leq d-1$, P_μ has at most $|X(k)| \leq \binom{n}{k}$ eigenvalues of value $> 1 - \frac{k+1}{d}$. In particular, M_μ has spectral gap at least $1/d$, and if $\tau \in \text{supp}(\mu)$ and $0 < \epsilon < 1$, the total variation mixing time of the Markov chain M_μ started at τ is at most $t_\tau(\epsilon) \leq d \log(\frac{1}{\epsilon \mu(\tau)})$.

Simplicial Complexes

Definition

- ▶ A set $X \subseteq 2^{[n]}$ is called a simplicial complex if whenever $\sigma \in X$ and $\tau \subset \sigma$, we have $\tau \in X$.
- ▶ Elements of X are called faces, and the dimension of a face $\tau \in X$ is defined as $\dim(\tau) = |\tau|$.
- ▶ A face of dimension 1 is called a *vertex*, and a face of dimension 2 is called an *edge*.
- ▶ Define $X(k) = \{ \tau \in X \mid \dim(\tau) = k \}$ to be the collection of degree- k faces of X .

Examples

Any (undirected) graph $G = (V, E)$ is an example of a simplicial complex.

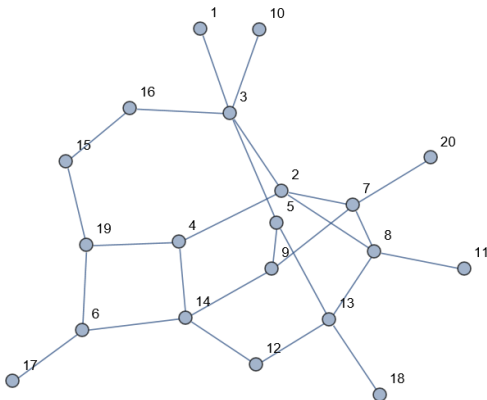


Figure 2: Example of a simplicial complex

Examples (contd.)

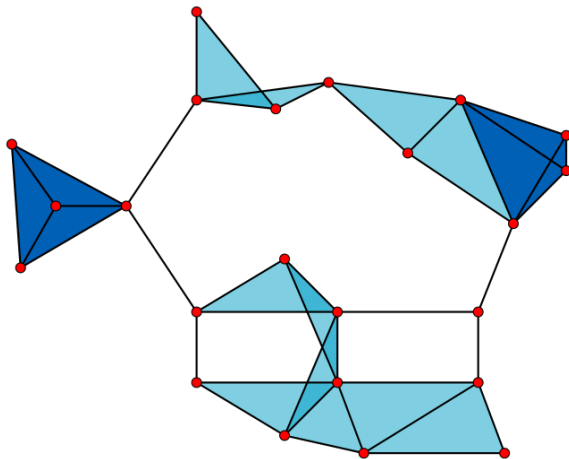


Figure 3: Example of a simplicial complex

Definitions contd.

- ▶ A simplicial complex X is pure if all maximal (w.r.t. inclusion) faces have the same dimension.
- ▶ The link of a face $\tau \in X$ is defined by $X_\tau = \{ \sigma \setminus \tau \mid \sigma \in X, \tau \subset \sigma \}$.
Importantly, if X is pure of dimension d and $\tau \in X(k)$, then X_τ is pure of dimension $d - k$.
- ▶ Can equip a weight function $w : X \rightarrow \mathbb{R}_{>0}$ to X by assigning a positive weight to each face of X . Say a weight function $w : X \rightarrow \mathbb{R}_{>0}$ is balanced if for any $\tau \in X$,

$$w(\tau) = \sum_{\substack{\sigma \in X(k+1) \\ \tau \subset \sigma}} w(\sigma)$$

- ▶ Notice that we can equip X with a (balanced) weight function by assigning its maximal faces weights and then assigning weights to the rest of the faces inductively.

Weights contd.

- ▶ Any (balanced) weight function on X induces a weighted graph on the vertices of X as follows: the 1-skeleton of X is the (weighted) graph $G = (X(1), X(2), w)$ where w has been restricted from X to $X(2)$. In this case $w(v)$ for $v \in X(1)$ is the weighted degree of v .

d -homogeneous polynomials

A polynomial $p \in \mathbb{R}[x_1, \dots, x_n]$ is d -homogeneous if $p(\lambda x_1, \dots, \lambda x_n) = \lambda^d p(x_1, \dots, x_n)$ for every $\lambda \in \mathbb{R}$. Notice in this case that,

$$\sum_{k=1}^n x_k \partial_k p(x) = d \cdot p(x)$$

Example. Consider $p(x, y, z) = xyz^2 + x^2yz$. Then,

$$\begin{aligned} \sum_{k=1}^3 x_k \partial_k p(x) &= (xyz^2 + 2x^2yz) + (xyz^2 + x^2yz) + (2xyz^2 + x^2yz) \\ &= 4xyz^2 + 4x^2yz \end{aligned}$$

Constructing Simplicial Complexes from Polynomials

From a d -homogeneous $p \in \mathbb{R}_{\geq 0}[x_1, \dots, x_n]$ $p(x) = \sum_S c_S x^S$, can construct a (weighted) simplicial complex X^p by doing the following: include a d -dimensional face S with weight $w(S) = c_S$ and include all subsets of these maximal faces inductively.

Visuals

This polynomial yields the above (weighted) simplex where each tetrahedral face has weight 1:

$$p(x_1, \dots, x_7) = x_1x_2x_3x_4 + x_3x_5x_6x_7$$

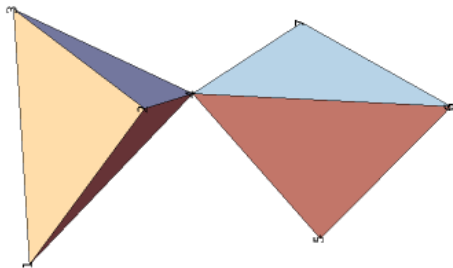


Figure 4: Two Tetrahedrons Glued Together

Roadmap

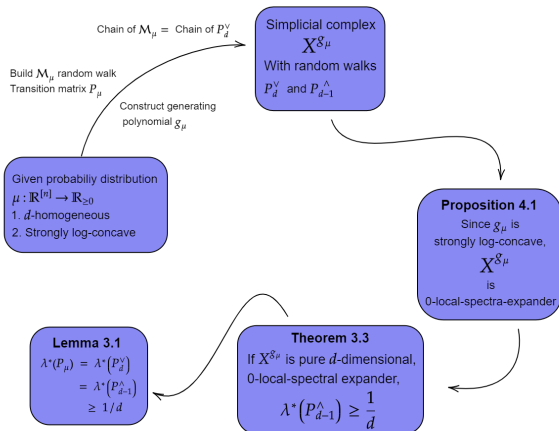


Figure 5: Roadmap

Log-concave polynomial identities

Definition

A polynomial $p \in \mathbb{R}_{\geq 0}[x_1, \dots, x_n]$ is log-concave if $\log p$ is concave, equivalently if

$$\nabla^2 \log p = \frac{p \cdot (\nabla p)^2 - (\nabla p)(\nabla p)^T}{p^2}$$

is NSD. For convenience, define $p(x) \equiv 0$ to be log-concave.

Log-concave properties contd.

- ▶ By Cauchy's interlacing theorem, if p is log-concave then $p \cdot (\nabla^2 p)$ has at most one positive eigenvalue at any $x \in \mathbb{R}_{>0}^n$.
- ▶ Since p has nonnegative coefficients, log-concavity is equivalent to $\nabla^2 p \preceq \frac{(\nabla p)(\nabla p)^T}{p}$, so in this case $\nabla^2 p$ has at most 1 one positive eigenvalue.
- ▶ Turns out converse is true too: if p is a degree d -homogeneous polynomial in $\mathbb{R}[x_1, \dots, x_n]$, and $(\nabla^2 p)(x)$ has at most one positive eigenvalue for all $x \in \mathbb{R}_{>0}^n$, then p is log-concave.

Definition

A polynomial $p \in \mathbb{R}[x_1, \dots, x_n]$ is strongly log concave if for all $k \geq 0$ and all $\alpha \in [n]^k$, we have $\partial^\alpha p$ is log-concave (i.e., all sequences of partials are log-concave).

Markov Chains and Random Walks

- ▶ A Markov Chain is a triple (Ω, P, π) where Ω denotes a finite state space, $P \in \mathbb{R}_{\geq 0}^{\Omega \times \Omega}$ is a transition probability matrix. That is,

$$P(i, j) = P_{ij} = \mathbb{P}(X_{t+1} = j \mid X_t = i)$$

It follows that the matrix is stochastic, such that $P \mathbf{1} = \mathbf{1}$. Finally, $\pi \in \mathbb{R}_{\geq 0}^{\Omega}$ denotes the stationary distribution of the chain ($\pi P = \pi$).

- ▶ The Markov Chain (Ω, P, π) is reversible if

$$\pi(\tau)P(\tau, \sigma) = \pi(\sigma)P(\sigma, \tau)$$

for all $\tau, \sigma \in \Omega$.

Markov Chains and Random Walks continued

- ▶ For any reversible Markov chain (Ω, P, π) , the largest eigenvalue of P is 1 (Perron-Fröbenius Theorem). We let $\lambda^*(P) = \max\{|\lambda_2|, |\lambda_n|\}$. The *spectral gap* of the Markov chain is $1 - \lambda^*(P)$.

Theorem (2.9, (DS))

For any reversible irreducible Markov chain (Ω, P, π) , $\epsilon > 0$, and any starting state

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

where

$$t_\tau(\epsilon) = \min \left\{ t \in \mathbb{N} \mid \|P^t(\tau, \cdot) - \pi\|_1 \leq \epsilon \right\}$$

Setting the stage

- ▶ Consider a pure d -dimensional complex X with a balanced weight function $w : X \rightarrow \mathbb{R}_{>0}$.
- ▶ Going to define a bipartite graph G_k with one side $X(k)$ and the other side $X(k+1)$. Connect $\tau \in X(k)$ to $\sigma \in X(k+1)$ with an edge of weight $w(\sigma)$ iff $\tau \subset \sigma$. Consider simple random walk on G_k : choose a neighbor proportional to the weight of the edge connecting the two vertices.

Examples

- ▶ One on $X(k)$ called P_k^\wedge , where given $\tau \in X(k)$ you take two steps of the walk in G_k to transition to the next k -face w.r.t. the P_k^\wedge matrix.
- ▶ One on $X(k+1)$ called P_{k+1}^\vee , where given $\sigma \in X(k+1)$ you take two steps to transition to the next $k+1$ face w.r.t. P_{k+1}^\vee .

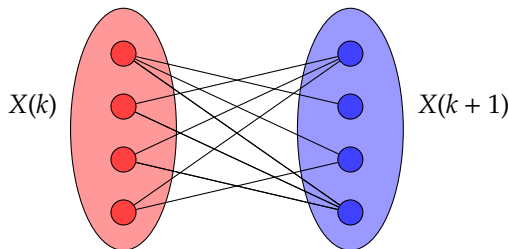


Figure 6: Bipartite graph G_k

Values of the transition matrix

$$P_k^\wedge(\tau, \tau') = \begin{cases} \frac{1}{k+1} & \text{if } \tau = \tau' \\ \frac{w(\tau \cup \tau')}{(k+1)w(\tau)} & \text{if } \tau \cup \tau' = X(k+1) \\ 0, & \text{otherwise} \end{cases}$$

$$P_{k+1}^\vee(\sigma, \sigma') = \begin{cases} \sum_{\tau \in X(k); \tau \subset \sigma} \frac{w(\sigma)}{(k+1)w(\tau)} & \text{if } \sigma = \sigma' \\ \frac{w(\sigma')}{(k+1)w(\sigma \cap \sigma')} & \text{if } \sigma \cap \sigma' = X(k) \\ 0, & \text{otherwise} \end{cases}$$

Note that both random walks are reversible with the same stationary distribution:

$$w(\tau)P_k^\wedge(\tau, \tau') = w(\tau')P_k^\wedge(\tau', \tau) \quad \text{and} \quad w(\sigma)P_{k+1}^\vee(\sigma, \sigma') = w(\sigma')P_{k+1}^\vee(\sigma', \sigma)$$

Proving $\lambda^*(P_d^\wedge) = \lambda^*(P_{d-1}^\vee)$

Fact (Useful)

Let $A \in \mathbb{R}^{n \times k}$ and $B \in \mathbb{R}^{k \times n}$ be arbitrary matrices. Then, non-zero eigenvalues of AB are equal to non-zero eigenvalues of BA with the same multiplicity.

Lemma (3.1)

For any $1 \leq k \leq d$, P_k^\wedge and P_{k+1}^\vee are stochastic, self-adjoint w.r.t. the ω -induced inner product, PSD, and have the same (with multiplicity) non-zero eigenvalues.

Proving $\lambda^*(P_d^\wedge) = \lambda^*(P_{d-1}^\vee)$

Proof.

Since G_k is bipartite, we may write the transition of the random walk on G_k as

$$P_k = \begin{bmatrix} 0 & P_k^\downarrow \\ P_k^\uparrow & 0 \end{bmatrix}$$

Note that P_k^\uparrow and P_k^\downarrow are stochastic matrices. Then we see that

$$P_k^2 = \begin{bmatrix} P_k^\downarrow P_k^\uparrow & \\ & P_k^\uparrow P_k^\downarrow \end{bmatrix}$$

It is easy to see P_k^2 is PSD and stochastic. But now we note that

$P_k^\wedge = P_k^\downarrow P_k^\uparrow$ and $P_{k+1}^\vee = P_k^\uparrow P_k^\downarrow$ and we're done. □

Looking at P_1^\wedge

- ▶ P_1^\wedge is the transition probability matrix of the simple $(1/2)$ -lazy random walk on the weighted 1-skeleton of X where the weight of each edge $e \in X(2)$ is $w(e)$.
- ▶ Also consider the non-lazy variant of this random walk, given by the transition matrix $\widetilde{P}_1^\wedge = 2(P_1^\wedge - I/2)$
- ▶ Similarly, for any face $\tau \in X(k)$, we define the upper random walk on the faces of the link X_τ . Specifically, let $P_{\tau,1}^\wedge$ denote the transition matrix of the upper walk, as above, on the 1-dimensional faces of X_τ , and $\widetilde{P}_{\tau,1}^\wedge = 2(P_{\tau,1}^\wedge - I/2)$ be the transition matrix for the non-lazy version.

Definition (Local Spectral Expanders, KO18)

For $\lambda > 0$, a pure d -dimensional weighted complex (X, w) is a λ -local-spectral-expander if for every $0 \leq k < d - 1$, and for every $\tau \in X(k)$, we have $\lambda_2(\widetilde{P}_{\tau,1}^\wedge) \leq \lambda$.

Theorem 3.3

Theorem

Let (X, w) be a pure d -dimensional weighted 0-local spectral expander and let $0 \leq k < d$. Then for all $-1 \leq i \leq k$, P_k^\wedge has at most $|X(i)| \leq \binom{n}{i}$ eigenvalues of value $> 1 - \frac{i+1}{k+1}$, where (by convention) $X(-1) = \emptyset$ and $\binom{n}{-1} = 0$. In particular, the second largest eigenvalue of P_k^\wedge is at most $\frac{k}{k+1}$.

Lemma

$P_k^\wedge \leq \frac{k}{k+1} P_k^\vee + \frac{1}{k+1} I$ for all $0 \leq k < d$.

From log-concavity to Local Spectral Expanders

Theorem (Proposition 4.1)

Let $p \in \mathbb{R}[x_1, \dots, x_n]$ be a multiaffine homogeneous polynomial with nonnegative coefficients. If p is strongly log-concave, then (X^p, w) is a 0-local-spectral-expander, where $w(S) = c_S$ for every maximal face $S \in X^p$

► Let $p_\tau = (\prod_{i \in \tau} \partial_i) p$

Lemma (4.2)

For any $0 \leq k \leq d$, and any simplex $\tau \in X^p(k)$, $w(\tau) = (d - k)! p_\tau(\mathbf{1})$.

From log-concavity to Local Spectral Expanders

Lemma (Lemma 4.2)

For any $0 \leq k \leq d$, and any simplex $\tau \in X^p(k)$, $w(\tau) = (d - k)!p_\tau(\mathbf{1})$.

Proof of Lemma.

Induction on $d - k$. If $\dim(\tau) = d$, then $p_\tau = c_\tau$, and done. So suppose statement holds for $\sigma \in X^p(k + 1)$ and fix simplex $\tau \in X^p(k)$. Then,

$$w(\tau) = \sum_{\substack{\sigma \in X^p(k+1) \\ \tau \subset \sigma}} w(\sigma) = \sum_{i \in X_\tau^p(1)} w(\tau \cup \{i\})$$

Since $\partial_i p_\tau = 0$ for $i \notin X_\tau^p(1)$, we have

$$w(\tau) = (d - k - 1)! \sum_{i \in X_\tau^p(1)} p_{\tau \cup \{i\}}(\mathbf{1}) = (d - k - 1)! \sum_{i=1}^n \partial_i p_\tau(\mathbf{1}) = (d - k)!p_\tau(\mathbf{1})$$

Where the last equality holds by Euler's identity. □

Proof of Proposition 4.1

Since p is strongly log-concave, $\nabla^2 p(\mathbf{1})$ has at most one positive eigenvalue. Let

$$\tilde{\nabla}^2 p = \frac{1}{d-k-1} (\text{diag}(\nabla p))^{-1} \nabla^2 p(\mathbf{1})$$

Claim: $\tilde{\nabla}^2 p = \tilde{P}_{\tau,1}^\wedge$. Note that

$$\tilde{P}_{\tau,1}^\wedge(i,j) = \frac{w_\tau(\{i,j\})}{w_\tau(\{i\})} = \frac{w(\tau \cup \{i,j\})}{w(\tau \cup \{i\})}$$

While,

$$(\tilde{\nabla}^2 p)(i,j) = \frac{(\partial_i \partial_j p)(\mathbf{1})}{(d-k-1)(\partial_i p)(\mathbf{1})}$$

By lemma, equal.

Proof contd.

Since p has nonnegative coefficients, the vector $(\nabla p)(\mathbf{1})$ has nonnegative entries which implies $\text{diag}(\nabla p)(\mathbf{1}) \geq 0$. Fact: if $B \geq 0$ and A has (at most) k positive eigenvalues then BA has at most k positive eigenvalues. Since $(\nabla^2 p)(\mathbf{1})$ has at most 1 positive eigenvalue, $\nabla^2 p$ has at most 1 positive eigenvalue by the fact. Thus, $\tilde{\nabla}^2 p = \tilde{P}_{\tau,1}^\wedge$ has at most one positive eigenvalue, so $\lambda_2(\tilde{P}_{\tau,1}^\wedge) \leq 0$. □

Generating polynomial of μ and \mathcal{M}_μ

- ▶ Let $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$ be a probability distribution. Assigning a multiaffine polynomial with variables x_1, \dots, x_n to μ :

$$g_\mu(x) = \sum_{S \subseteq [n]} \mu(S) \cdot \prod_{i \in S} x_i$$

- ▶ Say μ is d -homogeneous if g_μ is d -homogeneous, and (strongly) log-concave if g_μ is.
- ▶ We can define a random walk \mathcal{M}_μ by the following: We take the state space of \mathcal{M}_μ to be the support of μ , namely $\text{supp}(\mu) = \{S \subseteq [n] \mid \mu(S) \neq 0\}$. For $\tau \in \text{supp}(\mu)$, first we drop an element $i \in \tau$, chosen uniformly at random from τ . Then, among all sets $\sigma \supseteq \tau \setminus \{i\}$ in the support of μ , we choose one with probability proportional to $\mu(\sigma)$.

Proof of Theorem 1.1

Theorem (1.1)

Let $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$ be a d -homogeneous strongly log concave probability distribution. If P_μ denotes the transition probability matrix of M_μ and $X(k)$ denotes the set of size- k subsets of $[n]$ which are contained in some element of $\text{supp}(\mu)$, then for every $0 \leq k \leq d-1$, P_μ has at most $|X(k)| \leq \binom{n}{k}$ eigenvalues of value $> 1 - \frac{k+1}{d}$. In particular, M_μ has spectral gap at least $1/d$, and if $\tau \in \text{supp}(\mu)$ and $0 < \epsilon < 1$, the total variation mixing time of the Markov chain M_μ started at τ is at most $t_\tau(\epsilon) \leq d \log(\frac{1}{\epsilon \mu(t)})$.

Proof.

Let μ be a d -homogeneous strongly log-concave distribution, and let P_μ be the transition probability matrix of the chain M_μ . By Theorem 2.9, it is enough to show that $\lambda^*(P_\mu) \leq 1 - \frac{1}{d}$. Observe that the chain M_μ is exactly the same as the chain P_d^\vee for the simplicial complex X^{g_μ} defined above. Therefore, $\lambda^*(P_\mu) = \lambda^*(P_d^\vee) = \lambda^*(P_{d-1}^\wedge)$, where the last equality follows by Lemma 3.1. Since g_μ is strongly log-concave, by Proposition 4.1, X^{g_μ} is a 0-local-spectral-expander. Therefore, by Theorem 3.3,

$$\lambda^*(P_{d-1}^\wedge) \leq 1 - \frac{1}{(d-1)+1} = 1 - \frac{1}{d}.$$

□

Roadmap

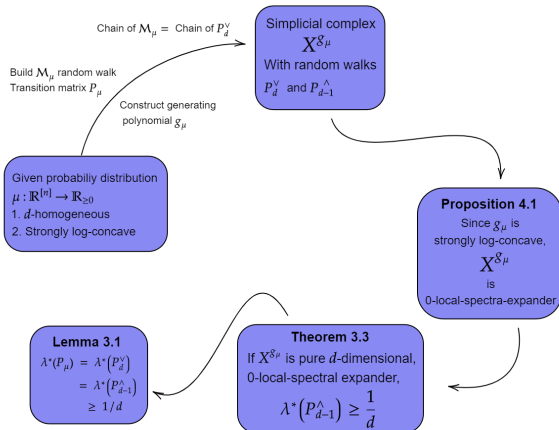


Figure 7: Roadmap

References



Anari, N., Liu, K., Gharan, S. O., Vinzant, C. (2019, June). *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 (pp. 1-16).



Mark Jerrum, Leslie Valiant, and Vijay Vazirani. “Random Generation of Combinatorial Structures from a Uniform Distribution”. In: Theoretical Computer Science 43 (1986), pp. 169–188.



Nima Anari, Shayan Oveis Gharan, and Cynthia Vinzant. “Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids”. In: *FOCS*. to appear. 2018.



Alexander Barvinok and Alex Samorodnitsky. “Random weighting, asymptotic counting, and inverse isoperimetry”. In: Israel Journal of Mathematics 158.1 (Mar. 2007), pp. 159–191. issn: 1565-8511.