Local-to-Global Functional Inequalities in Simplicial Complexes

Giorgos Mousa

Doctor of Philosophy
Laboratory for Foundations of Computer Science
School of Informatics
University of Edinburgh
2022

Abstract

A study of random walks over simplicial complexes with a particular emphasis on matroids. A framework is developed that yields results on the entropy contraction and modified log-Sobolev constant of the exchange walks over the levels of a simplicial complex, on the basis of entropy contraction properties of some local walks. This provides a general method for analyzing a variety of Markov chains by analyzing some of their lower-dimensional instances.

Keywords: simplicial complexes, Markov chains, sampling and counting, functional inequalities, mixing time, matroids, broken circuit complex.

Acknowledgements

I would like to express my gratitude towards my supervisor Heng Guo, who guided and supported me greatly throughout the four years of my studies, and who is an important contributor to the work herein. It has been very pleasant working with him and I have gained a lot from our interaction.

We began this project with the aim of proving the rapid mixing of the basis-exchange walk of matroids. Although we did not achieve this ourselves, its resolution happened early on in the course of my studies, and it served as an inspiration for our work. The methods involved in this were very much to my taste and I am glad to have been working in this area in such a period of exciting developments.

I would also like to thank my dear friend Rafail Chionatos, from my hometown in Greece, who – back in our high school years – persuaded me that academic pursuit is something highly valuable and worthwhile.

Lastly, I feel very fortunate to have been given the opportunity to pursue my graduate studies, especially in the wonderful city of Edinburgh. I will remember my time here as having been very special.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Giorgos Mousa*)

Table of Contents

Chapter 1

Introduction

In this thesis we will showcase a recently developed framework for analyzing a variety of Markov chains over discrete spaces which makes use of simplicial complexes. In general, such analysis is often useful for approximately *sampling* from a distribution or for approximating its partition function (e.g., *counting* the number of possible states in the case of a uniform distribution). Specifically, functional inequalities are a good tool towards this purpose, as they provide bounds on the mixing time of a Markov chain, and they can even be used to prove expansion properties and concentration of measure.

The use of simplicial complexes has proved to be versatile, as it turns out that various discrete spaces can be encoded as the faces of a simplicial complex. Moreover, a simplicial complex is a graded structure which lends itself nicely for a multilevel analysis and for inductive arguments. The usual route of a proof using this framework is the following: one encodes the space and distribution of interest at the top level of a simplicial complex. Through a canonical construction, this implies a series of exchange walks over all levels (*global* walks), as well as certain *local* exchange walks. Then, in an inductive manner, the argument builds up from the bottom level walk towards the top level by piecing together information from the local walks. In other words, one can define multiple exchange walks (local and global) which stand in a certain relation to each other. The nature of this relation is the main theory that we will develop. The strength of this theory lies in the ability that, by analyzing some of the easier instances of such walks (the lower-dimensional local walks) and exploiting their relation to the global walks, one can prove properties of the latter harder-to-analyze walks.

Another way of understanding this framework is through the fact that the global exchange walks can be decomposed into local walks. In this decomposition there is overlap, meaning that the same transition may appear in multiple local walks, and thus this analysis departs, for example, from a "projection/restriction chains"-type decomposition [JSTV04].

The achievements of this approach are numerous. The first major result – by applying this theory to the *independence complex* of a matroid – was the proof of rapid mixing of the basis-exchange walk and the resolution of the Mihail-Vazirani conjecture [MV89, Mih92] on the expansion of the basis-exchange graph [ALOV19]. In fact, this theory works wonderfully for matroids, with quantitatively neat expressions and without any ad hoc arguments. After this breakthrough, similar ideas were used to prove other strong and long sought-after results, such as optimal mixing of the Glauber dynamics over spin systems; for example, using this framework, optimal mixing for the hardcore model has been achieved [ALO20, CLV21a].

The goal that served as our inspiration and the one that we shall focus on is the analysis of the basis-exchange walk of matroids. The proof of sharp functional inequalities and mixing time bounds will serve as the main motivation for the general simplicial complex framework that we will develop.

The highlights of our presentation are the following. In terms of general theory, we present some simple proofs of Oppenheim's Trickling Down Theorem [Opp18] and Alev-Lau's Local-to-Global Theorem [AL20], as well as the Entropy Contraction Local-to-Global Theorem (our contribution). In terms of applications, we show sharp functional inequalities/modified log-Sobolev constant for the basis-exchange walk of matroids. After this, we examine a generalization of the independence complex of matroids, the *broken circuit complex*, for which we provide expansion results for rank 3 matroids.

1.1 Content Layout

The remainder of the thesis is organized as follows:

- In Chapter 2, we present some basic preliminaries (Linear Algebra, Markov Chains) that can be skipped and consulted when needed.
- In Chapter 3, we introduce the notion of the simplicial complex. We also define weights and distributions, as well as the exchange walks (global and local). This chapter sets the foundations for what follows – it is a more specialized "preliminaries chapter".

- In Chapter 4, we present some general theorems about variance contraction. This is equivalent to theorems about the spectral gap of the exchange walks and the relationship among them. We present some simple proofs – using Dirichlet forms – of the Trickling-Down Theorem [Opp18] and the main theorem of [AL20].

- In Chapter 5, we prove a general local-to-global theorem of entropy contraction. This can be thought as a theorem of the same nature as that of [AL20], but within the quantitatively sharper domain of entropy (instead of variance). This theorem can be used to prove bounds for the modified log-Sobolev constant instead of the spectral gap, and it can also provide, in various cases, tighter bounds on the mixing time of the exchange walks. This chapter contains our main contribution.

- In Chapter 6, after presenting some basic facts about matroids and the Mihail-Vazirani Conjecture (6.1.1), we show how this theory can be applied to matroids to prove sharp functional inequalities, which imply tight mixing time bounds for the basis-exchange walk (see Corollary 6.4.3) and concentration of measure results (see Corollary 6.4.5).

- In Chapter 7, we discuss the case of the broken circuit complex, which is a generalization of the independence complex of matroids. This chapter is independent of the results presented in Chapter 4 and Chapter 5, but it is a natural continuation of Chapter 6. We prove a constant edge expansion result for the graph defined by the reduced broken circuit complex of a rank 3 matroid (see Corollary 7.2.3). The ultimate goal in this direction is to prove rapid mixing of the exchange walks over all levels of the broken circuit complex. This would allow us to efficiently approximate the coefficients of the characteristic polynomial of a matroid. Furthermore, this approach is a candidate for a first FPRAS for approximately counting the number of acyclic orientations of a graph.

- In Chapter 8, we end with a few open problems.

Chapter 2

Preliminaries

2.1 Linear Algebra

In this section we include a few classical results from Linear Algebra. For reference on basic definitions and for the omitted proofs we direct the reader to the book by Horn and Johnson [HJ12]. We choose not to include some results in their full generality, and we restrict ourselves only to what we need.

Notation-wise, throughout the thesis we will regard vectors as column vectors, except for those that correspond to distributions which we view as row vectors. We denote by $\mathbb{1}$ the all-ones vector, and by I the identity matrix. For a vector v, we use $\mathrm{diag}(v)$ to denote the matrix whose diagonal entries are those of v and whose off-diagonal entries are zero; e.g. $I = \mathrm{diag}(\mathbb{1})$.

Proposition 2.1.1. [HJ12, Theorem 1.3.22]. *Let $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{m\times n}$, with $n \le m$. Then, the m eigenvalues of BA are the n eigenvalues of AB together with $m-n$ zeros.*

Theorem 2.1.2 (Spectral Theorem). [HJ12, Theorem 2.5.6]. *A symmetric matrix $A \in \mathbb{R}^{n\times n}$ has real eigenvalues $\{\lambda_k\}_{k=1}^n$. There exists a choice of corresponding real eigenvectors, $\{v_k\}_{k=1}^n$, such that*

$$A = \sum_{k=1}^{n} \lambda_k v_k v_k^{\mathrm{T}},$$

with $\|v_k\| = 1$ and, for $k \neq l$, $\langle v_k, v_l\rangle = 0$; i.e. $\{v_k\}_{k=1}^n$ is an orthonormal basis.

From the Spectral Theorem (2.1.2) it follows that we can order the eigenvalues of a symmetric matrix $A \in \mathbb{R}^{n\times n}$, and we will consistently do so as $\lambda_1(A) \ge \ldots \ge \lambda_n(A)$.

Theorem 2.1.3 (Courant-Fischer Theorem)**.** [HJ12, Theorem 4.2.6]. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix and let S be a subspace of $\mathbb{R}^n$. Then, for $1 \le k \le n$,*

$$\lambda_k(A) = \max_{\substack{S\subseteq\mathbb{R}^n \\ \dim S = k}} \min_{\substack{f\in S \\ \|f\|=1}} \langle f, Af\rangle\,,$$

and,

$$\lambda_k(A) = \min_{\substack{S\subseteq\mathbb{R}^n \\ \dim S = n-k+1}} \max_{\substack{f\in S \\ \|f\|=1}} \langle f, Af\rangle\,.$$

Corollary 2.1.4. *Let $A, B \in \mathbb{R}^{n\times n}$ be symmetric matrices such that $\forall f \in \mathbb{R}^n, \langle f, Af\rangle \le \langle f, Bf\rangle$. We denote this by $A \preceq B$. Then, for $1 \le k \le n$, $\lambda_k(A) \le \lambda_k(B)$.*

Proof. For $1 \le k \le n$, let

$$S_{\max} \in \arg\max_{\substack{S\subseteq\mathbb{R}^n \\ \dim S = k}} \min_{\substack{f\in S \\ \|f\|=1}} \langle f, Af\rangle\,.$$

Then, by the Courant-Fischer Theorem (2.1.3),

$$\begin{aligned}\lambda_k(A) &= \max_{\substack{S\subseteq\mathbb{R}^n \\ \dim S = k}} \min_{\substack{f\in S \\ \|f\|=1}} \langle f, Af\rangle = \min_{\substack{f\in S_{\max} \\ \|f\|=1}} \langle f, Af\rangle \\ &\le \min_{\substack{f\in S_{\max} \\ \|f\|=1}} \langle f, Bf\rangle \le \max_{\substack{S\subseteq\mathbb{R}^n \\ \dim S = k}} \min_{\substack{f\in S \\ \|f\|=1}} \langle f, Bf\rangle = \lambda_k(B).\end{aligned}$$ □

Proposition 2.1.5. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix and let $D \in \mathbb{R}^{n\times n}$ be a diagonal matrix with non-negative entries. If A has at most one positive eigenvalue, then so does DA.*

Proof. By Proposition 2.1.1, the spectrum of DA is the same as that of the symmetric matrix $D^{1/2}AD^{1/2}$. Then, due to the Spectral Theorem (2.1.2) applied to A,

$$\begin{aligned} D^{1/2}AD^{1/2} &= D^{1/2}\left(\lambda_1 v_1 v_1^{\mathsf{T}} + \sum_{k=2}^{n} \lambda_k v_k v_k^{\mathsf{T}}\right) D^{1/2} \\ &= \lambda_1 (D^{1/2}v_1)(D^{1/2}v_1)^{\mathsf{T}} + \sum_{k=2}^{n} \lambda_k (D^{1/2}v_k)(D^{1/2}v_k)^{\mathsf{T}} \\ &\preceq \lambda_1 (D^{1/2}v_1)(D^{1/2}v_1)^{\mathsf{T}}. \qquad \text{(because, for } k \ge 2,\ \lambda_k \le 0\text{)} \end{aligned}$$

The matrix $\lambda_1 (D^{1/2}v_1)(D^{1/2}v_1)^{\mathsf{T}}$ has at most one positive eigenvalue depending on the sign of λ_1 (follows from Proposition 2.1.1). Thus, by Corollary 2.1.4, we get that each of the matrices $D^{1/2}AD^{1/2}$ and DA has at most one positive eigenvalue. □

2.2 Markov Chains

Let Ω be a finite state space and let $\pi \in \mathbb{R}^{\Omega}_{>0}$ be a distribution over Ω, so that $\sum_{x\in\Omega} \pi(x) = 1$. Let $P \in \mathbb{R}^{\Omega\times\Omega}_{\geq 0}$ be the transition matrix of a Markov chain, so that for any $x \in \Omega$ we have $\sum_{y\in\Omega} P(x,y) = 1$. If P corresponds to an irreducible and aperiodic Markov chain, then it has a unique stationary distribution. Such a chain converges to its unique stationary distribution π, meaning that $\mu P^t \to \pi$, as $t \to \infty$, for any initial distribution μ. One way to measure the distance from stationarity is the total variation distance, which for $\mu, \nu \in \mathbb{R}^{\Omega}$ is defined as

$$\|\mu - \nu\|_{\mathrm{TV}} := \max_{A\subseteq\Omega} |\mu(A) - \nu(A)| .$$

A quantity of interest is the time required for the random walk to reach within a distance ε from its stationary distribution, which is given by the mixing time,

$$t_{mix}(P,\varepsilon) := \min\left\{t \;\middle|\; \forall x \in \Omega,\ \|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} \leq \varepsilon\right\}.$$

We say P is *reversible* with respect to π if the detailed balance equations hold, i.e.

$$\pi(x)P(x,y) = \pi(y)P(y,x), \qquad \forall x,y \in \Omega. \tag{2.1}$$

We endow $\mathbb{R}^{\Omega}$ with the inner product

$$\langle f, g\rangle_{\pi} := \sum_{x\in\Omega} \pi(x) f(x) g(x),$$

where $f, g \in \mathbb{R}^{\Omega}$. The induced norm is $\|f\|_{\pi} := \sqrt{\langle f, f\rangle_{\pi}}$.

Proposition 2.2.1. *A transition matrix P which is reversible with respect to a distribution π has real eigenvalues. Furthermore, we can use the Courant-Fischer Theorem (2.1.3) with inner product $\langle\cdot,\cdot\rangle_{\pi}$ and norm $\|\cdot\|_{\pi}$ to determine its eigenvalues.*

Proof. Let $D = \mathrm{diag}(\pi)$. Notice that (2.1) can be written in matrix form as $DP = (DP)^{\mathrm{T}}$. This means that the matrix $D^{-1/2}DPD^{-1/2} = D^{1/2}PD^{-1/2}$ is symmetric and by the Spectral Theorem (2.1.2) it has real eigenvalues. By Proposition 2.1.1, P has the same spectrum as $D^{1/2}PD^{-1/2}$ and thus it also has real eigenvalues. The proposition's last assertion follows by applying the Courant-Fischer Theorem (2.1.3) to $D^{1/2}PD^{-1/2}$ and changing the variable $f \mapsto D^{1/2}f$. □

The adjoint operator P^* of P is defined as $P^*(x,y) = \frac{\pi(y)P(y,x)}{\pi(x)}$. This is the unique operator which satisfies $\langle f, Pg\rangle_{\pi} = \langle P^* f, g\rangle_{\pi}$. It is easy to verify that if P satisfies the detailed balance equations (2.1) (so P is reversible), then P is *self-adjoint*, namely $P = P^*$.

The *Dirichlet form* is defined as

$$\mathcal{E}_P(f,g) := \langle (I-P)f, g\rangle_\pi . \tag{2.2}$$

Let the *Laplacian* $\mathcal{L} := I - P$. Then,

$$\begin{aligned}\mathcal{E}_P(f,g) &= \sum_{x,y\in\Omega} \pi(x)g(x)\mathcal{L}(x,y)f(y)\\ &= g^{\mathrm{T}}\operatorname{diag}(\pi)\mathcal{L}f.\end{aligned}$$

In particular, if P is reversible, then $\mathcal{L}^* = \mathcal{L}$ and

$$\begin{aligned}\mathcal{E}_P(f,g) &= \langle \mathcal{L}f, g\rangle_\pi = \langle f, \mathcal{L}^* g\rangle_\pi = \langle f, \mathcal{L}g\rangle_\pi = \mathcal{E}_P(g,f)\\ &= f^{\mathrm{T}}\operatorname{diag}(\pi)\mathcal{L}g. \end{aligned}\tag{2.3}$$

In this thesis all Markov chains are reversible and we will most commonly use the form (2.3). Another common expression of the Dirichlet form for reversible P is

$$\mathcal{E}_P(f,g) = \frac{1}{2}\sum_{x\in\Omega}\sum_{y\in\Omega}\pi(x)P(x,y)(f(x)-f(y))(g(x)-g(y)), \tag{2.4}$$

but we will not need this expression until Section 5.4. It is well known that the spectral gap of P, or equivalently the smallest positive eigenvalue of $\mathcal{L}$, controls the convergence rate of P. It also has a variational characterisation.

Let the expected value and variance of f be

$$\mathbb{E}_\pi f = \sum_{x\in\Omega}\pi(x)f(x), \qquad \operatorname{Var}_\pi(f) := \mathbb{E}_\pi f^2 - (\mathbb{E}_\pi f)^2 .$$

Then, the Poincaré constant is defined as

$$\lambda(P) := \min_{f\in\mathbb{R}^\Omega}\left\{\frac{\mathcal{E}_P(f,f)}{\operatorname{Var}_\pi(f)} \mid \operatorname{Var}_\pi(f)\neq 0\right\}. \tag{2.5}$$

For a reversible transition matrix P, the Poincaré constant $\lambda(P)$ and the spectral gap $1-\lambda_2(P)$ are equal. The usefulness of $\lambda(P)$ is due to the fact that, if, say, all eigenvalues of P are non-negative, then

$$t_{\mathrm{mix}}(P,\varepsilon) \le \frac{1}{\lambda(P)}\left(\frac{1}{2}\log\frac{1}{\pi_{\min}} + \log\frac{1}{2\varepsilon}\right), \tag{2.6}$$

where $\pi_{\min} = \min_{x\in\Omega}\pi(x)$. See, for example, [LP17, Theorem 12.4].

The (standard) log-Sobolev constant relates $\mathcal{E}_P\left(\sqrt{f},\sqrt{f}\right)$ with the following entropy-like quantity:

$$\operatorname{Ent}_\pi(f) := \mathbb{E}_\pi(f\log f) - \mathbb{E}_\pi f\log\mathbb{E}_\pi f, \tag{2.7}$$

for $f \in \mathbb{R}^{\Omega}_{\geq 0}$. Here, log stands for the natural logarithm with base e and we follow the convention that $0 \log 0 = 0$. The log-Sobolev constant is defined as

$$\alpha(P) := \inf_{f \in \mathbb{R}^{\Omega}_{\geq 0}} \left\{ \frac{\mathcal{E}_P\left(\sqrt{f}, \sqrt{f}\right)}{\mathrm{Ent}_\pi(f)} \mid \mathrm{Ent}_\pi(f) \neq 0 \right\}.$$

The constant $\alpha(P)$ gives a better control of the mixing time of P. As shown by [DSC96], for a continuous-time chain,[1]

$$t_{\mathrm{mix}}(P, \varepsilon) \leq \frac{1}{4\alpha(P)} \left(\log\log \frac{1}{\pi_{\min}} + \log \frac{1}{2\varepsilon^2} \right). \tag{2.8}$$

The saving seems modest comparing to (2.6), but it is quite common that $\pi_{\min}$ is exponentially small in the instance size, in which case the saving is a polynomial factor.

What we are interested in, however, is the following modified log-Sobolev constant introduced by [BT06]:

$$\rho(P) := \inf_{f \in \mathbb{R}^{\Omega}_{\geq 0}} \left\{ \frac{\mathcal{E}_P(f, \log f)}{\mathrm{Ent}_\pi(f)} \mid \mathrm{Ent}_\pi(f) \neq 0 \right\}.$$

Similar to (2.8), for a continuous-time chain,

$$t_{\mathrm{mix}}(P, \varepsilon) \leq \frac{1}{\rho(P)} \left(\log\log \frac{1}{\pi_{\min}} + \log \frac{1}{2\varepsilon^2} \right), \tag{2.9}$$

as shown by [BT06, Corollary 2.8]. (Recall Footnote 1.)

For reversible P, the following relationships among these constants are known,

$$2\lambda(P) \geq \rho(P) \geq 4\alpha(P).$$

See, for example, [BT06, Proposition 3.6].

Thus, lower bounds on these constants (with the same asymptotic behaviour) are increasingly difficult to obtain. However, to get the best asymptotic control of the mixing time, one only needs to lower bound the modified log-Sobolev constant $\rho(P)$ instead of $\alpha(P)$ by comparing (2.8) and (2.9). Indeed, as observed by [HS19], by taking f in the definition of $\alpha(P)$ to be the indicator function $\frac{1}{\pi(x)} \mathbb{1}_x$, for all $x \in \Omega$, we get

$$\alpha(P) \leq \min_{x \in \Omega} \left\{ \frac{1}{-\log \pi(x)} \right\}.$$

[1] [DSC96] and [BT06] showed (2.8) and (2.9), respectively, for continuous-time Markov chains. For the discrete-time chains of interest in this thesis, we will be able to prove the same bound as in (2.9) by a step-wise entropy contraction argument.

In many settings (such as strongly log-concave distributions), we cannot hope for a good uniform bound for $\alpha(P)$, as the right hand side of the above can be arbitrarily small for fixed input size.

By (2.2) and (2.7), it is clear that if we replace f by cf for some constant $c > 0$, then both $\mathscr{E}_P(f, \log f)$ and $\mathrm{Ent}_\pi(f)$ increase by the same factor c. Thus, in order to bound ρ, we may further assume that $\mathbb{E}_\pi f = 1$. This assumption allows the simplification $\mathrm{Ent}_\pi(f) = \mathbb{E}_\pi(f \log f)$. In this case, $\pi(\cdot)f(\cdot)$ is a distribution, and $\mathrm{Ent}_\pi(f)$ is the relative entropy (or Kullback–Leibler divergence) between $\pi(\cdot)f(\cdot)$ and $\pi(\cdot)$.

Chapter 3

Simplicial Complexes

Definition 3.0.1. *A simplicial complex $\mathscr{C}$ is a nonempty downwards closed collection of sets (called faces) over a finite ground set of elements. It satisfies,*

- $\emptyset \in \mathscr{C}$;
- *if* $S \in \mathscr{C}$ *and* $T \subseteq S$, *then* $T \in \mathscr{C}$.

Remark. *More precisely, this defines an abstract and finite simplicial complex.*

The tuple $(\mathscr{C}, \subseteq)$ forms a poset, which we equip with a rank function $r : \mathscr{C} \to \mathbb{N}$, such that $r(S) = |S|$. We write $r = r(\mathscr{C}) = \max_{S \in \mathscr{C}} |S|$ for the rank of $\mathscr{C}$. We will use this rank function (the cardinality) for the grading of the faces of a simplicial complex. We denote by $\mathscr{C}(k)$ the collection of faces of size k, where $0 \leq k \leq r$. We can thus write $\mathscr{C} = \cup_{k=0}^{r} \mathscr{C}(k)$. We also write v as a shorthand for $\{v\}$, for $\{v\} \in \mathscr{C}(1)$. A *pure* simplicial complex has maximal faces (*facets*) of the same cardinality r.

We note that a simplicial complex (without the adjective "abstract") usually refers to a collection of faces that is geometrically realized as a collection of corresponding simplices that satisfies some simple conditions: (1) every face of a simplex belongs to the simplicial complex, and (2) the intersection of two simplices is a face of both. With this geometric viewpoint in mind, we call the sets in $\mathscr{C}(1)$ points or vertices, in $\mathscr{C}(2)$ edges, and – in general – the sets in $\mathscr{C}(k)$ (k-1)-simplices. We also define a function $\dim : \mathscr{C} \to \{-1\} \cup \mathbb{N}$ such that $\dim(S) = |S| - 1$ and $d = \dim(\mathscr{C}) = \max_{S \in \mathscr{C}} \dim(S)$. Any abstract simplicial complex can be realized geometrically, for example by embedding its points in $\mathbb{R}^{2d+1}$ in general position, i.e. through a map $f : \mathscr{C}(1) \to \mathbb{R}^{2d+1}$, and taking the convex hull $\operatorname{conv}(\{f(v) \mid v \in S\})$ of every face $S \in \mathscr{C}$ to form the corresponding simplices.

As we will only deal with pure, finite, abstract simplicial complexes, we drop these adjectives for conciseness.

3.1 Links and Skeletons

The links and skeletons of a simplicial complex are subcomplexes, i.e. substructures of the initial complex which are also simplicial complexes. These notions are useful for inductive arguments.

Definition 3.1.1. *The link of a simplicial complex $\mathscr{C}$ at a face S is the simplicial complex $\mathscr{C}_S := \{T \mid T \cap S = \emptyset,\ T \cup S \in \mathscr{C}\}$.*

Definition 3.1.2. *The k-skeleton of a simplicial complex $\mathscr{C} = \cup_{i=0}^{r} \mathscr{C}(i)$ is the simplicial complex $\cup_{i=0}^{k+1} \mathscr{C}(i)$, for $k \leq \dim(\mathscr{C})$.*

Notice that, in the above definition, k refers to the dimension of the resulting simplicial complex. Of special interest is the 1-skeleton of the link at a face S, which corresponds to a graph G_S whose vertex set is $\mathscr{C}_S(1)$, and whose edge set is $\mathscr{C}_S(2)$.

3.2 Weight Function and Distributions

We may equip a simplicial complex with a weight function w.

Definition 3.2.1. *A weighted simplicial complex $(\mathscr{C}, w)$ is a simplicial complex $\mathscr{C}$ along with a weight function $w : \mathscr{C} \to \mathbb{R}_{>0}$, such that*

$$w(S) = \sum_{T \supset S,\ |T| = |S|+1} w(T), \qquad \text{if } |S| < r.$$

Thus, it suffices to specify the weight function over $\mathscr{C}(r)$. A weight function that is usually of interest is the one for which $w(S) = 1$, for $S \in \mathscr{C}(r)$. In this case, we have that $w(S)$ is equal to the number of maximal chains in $(\mathscr{C}_S, \subseteq)$, for $S \in \mathscr{C}$.

We can now define the distributions $\{\pi_k\}_{k=0}^{r}$, one over each $\mathscr{C}(k)$, as

$$\pi_k(S) \propto w(S), \qquad S \in \mathscr{C}(k).$$

We view distributions as row vectors, contrary to other vectors which we view as column vectors. We denote by $Z_k := \sum_{S \in \mathscr{C}(k)} w(S)$ the partition function of the weight function w over $\mathscr{C}(k)$. Thus, we can write $\pi_k(S) = \frac{w(S)}{Z_k}$. We wish to extend the scope of the weight function to the links and skeletons of the simplicial complex.

Definition 3.2.2. *The link of a weighted simplicial complex $(\mathscr{C}, w)$ at a face $S \in \mathscr{C}$ is the weighted simplicial complex $(\mathscr{C}_S, w_S)$, where $w_S(T) := w(S \cup T)$.*

Definition 3.2.3. *The k-skeleton of a weighted simplicial complex $(\mathscr{C}, w)$ is the weighted simplicial complex $\left(\cup_{i=0}^{k+1}\mathscr{C}(i), w\right)$, where with slight abuse of notation we also denote by w the restriction of the weight function to $\cup_{i=0}^{k+1}\mathscr{C}(i)$.*

In a similar fashion to the above definitions and for all appropriate objects, for the links we shall use a subscript indicating the face over which we have taken the link, while for the skeletons we will simply consider the restriction.

For example, we similarly define distribution $\pi_{S,k}$ over $\mathscr{C}_S(k)$ simply as

$$\pi_{S,k}(T) \propto w_S(T), \qquad\qquad T \in \mathscr{C}_S(k).$$

These distributions (and behind those the weight function) are a natural choice because they appear as the stationary distributions of the exchange walks that we are interested in studying.

3.3 Global walks

There are two natural exchange walks on $\mathscr{C}(k)$, $P_k^{\vee}$ and $P_k^{\wedge}$, which start by removing or adding an element respectively, and coming back to $\mathscr{C}(k)$. We call these walks "global" as they are defined over the whole of $\mathscr{C}(k)$. They are comprised by two parts:

- "Going-down", $P_k^{\downarrow}$; starting from a set $S \in \mathscr{C}(k)$, we remove an element $i \in S$ uniformly at random.
- "Going-up", $P_k^{\uparrow}$; starting from a set $S \in \mathscr{C}(k)$, we add an element $i \in \mathscr{C}_S(1)$ with probability $\pi_{S,1}(i)$.

We can now write

$$P_k^{\vee} = P_k^{\downarrow} P_{k-1}^{\uparrow}, \qquad\qquad P_k^{\wedge} = P_k^{\uparrow} P_{k+1}^{\downarrow}. \tag{3.1}$$

The down-up walk, $P_k^{\vee}$, is defined for $1 \le k \le r$, and up-down walk, $P_k^{\wedge}$, for $0 \le k \le r-1$.

Walks similar or equivalent to these can be traced back to [KM17, DK17] and [Opp18].

More explicitly, for $1 \leq k \leq r$ and $S, T \in \mathscr{C}(k)$, we have that

$$P_k^{\vee}(S,T) = \begin{cases} \sum_{S' \in \mathscr{C}(k-1), S' \subset S} \frac{w(S)}{kw(S')} & \text{if } S = T; \\ \frac{w(T)}{kw(S \cap T)} & \text{if } S \cap T \in \mathscr{C}(k-1); \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

This can be noticed by the fact that for two different sets $S, T \in \mathscr{C}(k)$, there can only be a transition from S to T if their intersection, $S \cap T$, belongs to $\mathscr{C}(k-1)$. Then, the chain can move from S to T if it first drops the unique element in $S \setminus T$ (which happens with probability $\frac{1}{k}$), and if it then adds the unique element in $T \setminus S$ (which happens with probability $\pi_{S\cap T,1}(T \setminus S) = \frac{w_{S\cap T}(T\setminus S)}{w_{S\cap T}(\emptyset)} = \frac{w(T)}{w(S\cap T)}$).

Likewise, for $0 \leq k \leq r-1$,

$$P_k^{\wedge}(S,T) = \begin{cases} \frac{1}{k+1} & \text{if } S = T; \\ \frac{w(S\cup T)}{(k+1)w(S)} & \text{if } S \cup T \in \mathscr{C}(k+1); \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

We first give some basic decompositions of $P_k^{\vee}$ and $P_k^{\wedge}$. Let A_k be a matrix whose rows are indexed by $\mathscr{C}(k)$ and columns by $\mathscr{C}(k+1)$ such that

$$A_k(S,T) := \begin{cases} 1 & \text{if } S \subset T; \\ 0 & \text{otherwise,} \end{cases}$$

and let w_k be the vector of $\{w(S)\}_{S \in \mathscr{C}(k)}$. Then,

$$P_k^{\uparrow} = \operatorname{diag}(w_k)^{-1} A_k \operatorname{diag}(w_{k+1}), \tag{3.4}$$

$$P_{k+1}^{\downarrow} = \frac{1}{k+1} A_k^{\mathtt{T}}. \tag{3.5}$$

Let $D_k = \operatorname{diag}(\pi_k)$. Using (3.4) and (3.5), we get that

$$D_{k+1} P_{k+1}^{\downarrow} = (P_k^{\uparrow})^{\mathtt{T}} D_k. \tag{3.6}$$

By multiplying equation (3.6) by the all-ones vector, we also get that

$$\pi_{k+1} P_{k+1}^{\downarrow} = \pi_k, \tag{3.7}$$

$$\pi_k P_k^{\uparrow} = \pi_{k+1}. \tag{3.8}$$

Proposition 3.3.1. *The walks $P_k^{\vee}$ and $P_k^{\wedge}$ are reversible with respect to π_k and they have non-negative eigenvalues.*

Proof. For the reversibility, one can check by direct computation that the detailed balance equations hold, namely (2.1). From (3.6) we can deduce that the matrices $P_k^\vee$ and $P_k^\wedge$ have non-negative eigenvalues as follows: for $P_k^\vee$ (and similarly for $P_k^\wedge$), and any $f \in \mathbb{R}^{\mathscr{C}(k)}$, we have

$$\langle f, P_k^\vee f\rangle_{\pi_k} = f^{\mathrm{T}} D_k P_k^\vee f \overset{(3.6)}{=} f^{\mathrm{T}} (P_{k-1}^\uparrow)^{\mathrm{T}} D_{k-1} P_{k-1}^\uparrow f = \left\| P_{k-1}^\uparrow f \right\|_{\pi_{k-1}}^2 \geq 0.$$

The proposition now follows from Proposition 2.2.1. □

Corollary 3.3.2. *The spectral gaps of $P_k^\vee$ and $P_{k-1}^\wedge$ are equal, i.e., for $1 \leq k \leq r$, $\lambda(P_k^\vee) = \lambda(P_{k-1}^\wedge)$.*

Proof. Notice that the matrices $P_k^\vee = P_k^\downarrow P_{k-1}^\uparrow$ and $P_{k-1}^\wedge = P_{k-1}^\uparrow P_k^\downarrow$ are of the form AB and BA, and thus they have the same nonzero eigenvalues (with multiplicity, by Proposition 2.1.1). Given also that they have non-negative eigenvalues (Proposition 3.3.1), it follows that they have equal spectral gaps. Note that for the cases where the state space Ω of a walk P has size 1, we adopt the convention that $\lambda(P) = 1$, which is consistent with the desired equality. □

3.4 Local Walks

We shall also use the notation $P_{S,k}^\wedge$ and $P_{S,k}^\vee$, where $S \in \mathscr{C}$, to denote the walks over $\mathscr{C}_S(k)$. These walks are local in the sense that they operate over $\mathscr{C}_S(k)$ instead of the whole of $\mathscr{C}(|S|+k)$. However, it is good to note that – in another sense – they are not completely "local" as they have weights that perform counting of higher level supersets. The local walks that will be of particular interest are the walks $P_{S,1}^\wedge$, for every face $S \in \mathscr{C}$ with $|S| \leq r-2$, as well as their non-lazy counterparts (i.e., those without self-loops), $G_S := 2P_{S,1}^\wedge - I$. These walks operate over the 1-skeleton of the link at S. In particular, G_S is the simple random walk over the weighted graph that corresponds to this weighted simplicial complex.

3.5 Some Properties

From Definition 3.2.1, it follows that

$$w(S) = r_S! \cdot w(\mathscr{C}_S(r_S)), \qquad \forall S \in \mathscr{C}.$$

where $r_S = r - |S|$ is the rank of $\mathscr{C}_S$. Thus, for any $S \in \mathscr{C}(k)$, $\pi_k(S)$ is proportional to the total weight of the maximal supersets of S. Recalling that $Z_k = \sum_{S\in\mathscr{C}(k)} w(S)$ is the normalisation constant of π_k, we have that, for any $0 \le k \le r$,

$$\begin{aligned} Z_k &= \sum_{S\in\mathscr{C}(k)} r_S! \cdot w(\mathscr{C}_S(r_S)) = (r-k)! \sum_{S\in\mathscr{C}(k)} w(\mathscr{C}_S(r_S)) \\ &= (r-k)! \binom{r}{k} w(\mathscr{C}(r)) = \frac{r! \cdot w(\mathscr{C}(r))}{k!} = \frac{w(\emptyset)}{k!}. \end{aligned}$$

Furthermore, for all $S \in \mathscr{C}$ and $0 \le k \le r_S$,

$$\pi_{S,k}(T) = \frac{k! w_S(T)}{w_S(\emptyset)}, \qquad \forall T \in \mathscr{C}_S(k). \tag{3.9}$$

Notice that the normalising constant $Z_{S,k} = \frac{w(S)}{k!}$.

For $0 \le k \le r$ and a function $f^{(k)} : \mathscr{C}(k) \to \mathbb{R}$, define $f^{(l)} : \mathscr{C}(l) \to \mathbb{R}$ for $0 \le l \le k$ such that

$$f^{(l)} := P_l^{\uparrow} \cdots P_{k-1}^{\uparrow} f^{(k)} = \prod_{i=l}^{k-1} P_i^{\uparrow} f^{(k)}. \tag{3.10}$$

Intuitively, $f^{(l)}$ is the function $f^{(k)}$ "going down" to level l. Similarly, for $T \in \mathscr{C}(l)$ and $0 \le m \le k-l$, we define $f_T^{(m)} := \prod_{i=m}^{k-l-1} P_{T,i}^{\uparrow} f_T^{(k-l)}$, where $f_T^{(m)} : \mathscr{C}_T(m) \to \mathbb{R}$. Because the up operators of a link are equal to a restriction of the up operators on the whole simplicial complex, we can see that $f_T^{(k-l)}(S \setminus T) = f^{(k)}(S)$.

Next we establish some useful properties of $f^{(l)}$.

Lemma 3.5.1. *Let $0 \le k \le r$ and $f^{(k)} : \mathscr{C}(k) \to \mathbb{R}$. Then we have the following:*

1. *for any $0 \le l \le k$, $\mathbb{E}_{\pi_l} f^{(l)} = \mathbb{E}_{\pi_k} f^{(k)}$;*
2. *for any $0 \le l \le k$, $T \in \mathscr{C}(l)$, $f^{(l)}(T) = \mathbb{E}_{\pi_{T,k-l}} f_T^{(k-l)}$.*

Proof. For (1), we have that

$$\begin{aligned} \mathbb{E}_{\pi_l} f^{(l)} &= \pi_l \prod_{i=l}^{k-1} P_i^{\uparrow} f^{(k)} && \text{(by (3.10))} \\ &= \pi_k f^{(k)} && \text{(by (3.8))} \\ &= \mathbb{E}_{\pi_k} f^{(k)}. \end{aligned}$$

For (2), first notice that

$$\begin{aligned} \prod_{i=0}^{k-l-1} P_{T,i}^{\uparrow} &= \pi_{T,0} \prod_{i=0}^{k-l-1} P_{T,i}^{\uparrow} && \text{(because } \pi_{T,0} = (1)\text{)} \\ &= \pi_{T,k-l}, && \text{(by (3.8))} \end{aligned}$$

where $\pi_{T,0}$ is a distribution over one set (the empty set). It follows that

$$\begin{aligned}
\mathbb{E}_{\pi_{T,k-l}} f_T^{(k-l)} = \pi_{T,k-l} f_T^{(k-l)} &= \prod_{i=0}^{k-l-1} P_{T,i}^{\uparrow} f_T^{(k-l)} \\
&= f_T^{(0)} = f_T^{(0)}(\emptyset) && \text{(by (3.10))} \\
&= f^{(l)}(\emptyset \cup T) = f^{(l)}(T). && \square
\end{aligned}$$

Chapter 4

Variance Contraction

Recall that, for $0 \le k \le r$ and a function $f^{(k)} : \mathscr{C}(k) \to \mathbb{R}$, we define $f^{(l)} : \mathscr{C}(l) \to \mathbb{R}$, for $0 \le l \le k$, as

$$f^{(l)} = \prod_{i=l}^{k-1} P_i^{\uparrow} f^{(k)}.$$

The key observation is that this operation contracts the variance (or, as we shall see in the next chapter, the entropy) by a factor which depends on the variance contraction factors of certain local walks. We have that

$$\begin{aligned}
\mathscr{E}_{P_k^{\vee}}\left(f^{(k)}, f^{(k)}\right) &= \left(f^{(k)}\right)^{\mathsf{T}} D_k (I - P_k^{\vee}) f^{(k)} && (4.1)\\
&= \left(f^{(k)}\right)^{\mathsf{T}} D_k f^{(k)} - \left(f^{(k)}\right)^{\mathsf{T}} \left(P_{k-1}^{\uparrow}\right)^{\mathsf{T}} D_{k-1} P_{k-1}^{\uparrow} f^{(k)} && \text{(by (3.6))}\\
&= \mathrm{Var}_{\pi_k}\left(f^{(k)}\right) - \mathrm{Var}_{\pi_{k-1}}\left(f^{(k-1)}\right). && \text{(by Lemma 3.5.1)}
\end{aligned}$$

Thus, a Poincaré inequality for the walk $P_k^{\vee}$ with constant c,

$$\mathscr{E}_{P_k^{\vee}}\left(f^{(k)}, f^{(k)}\right) \ge c\,\mathrm{Var}_{\pi_k}\left(f^{(k)}\right), \qquad \forall f^{(k)} : \mathscr{C}(k) \to \mathbb{R},$$

is equivalent to variance contraction when premultiplying by $P_{k-1}^{\uparrow}$,

$$\mathrm{Var}_{\pi_{k-1}}\left(f^{(k-1)}\right) \le (1-c)\mathrm{Var}_{\pi_k}\left(f^{(k)}\right), \qquad \forall f^{(k)} : \mathscr{C}(k) \to \mathbb{R}.$$

So, since in this setting the notion of variance contraction is equivalent to Poincaré inequalities, we will express the results of this chapter in terms of spectral gap or second largest eigenvalue bounds (recall Equation (2.5)).

Definition 4.0.1. *A γ-local-spectral expander at level k is a weighted simplicial complex $(\mathscr{C}, w)$ for which the local walks G_S, for every $S \in \mathscr{C}(k)$, satisfy $\lambda_2(G_S) \le \gamma < 1$. A $(a_0, \dots, a_{r-2})$-local-spectral expander is a a_k-local-spectral expander at level k, for $0 \le k \le r-2$.*

4.1 Oppenheim's Trickling-Down Theorem

The theorem that we prove in this section allows one to derive a bound for the second eigenvalue of a local walk from bounds of the local walks that are one level above it. This theorem can be used iteratively for the derivation of bounds for all local walks from only those of the highest local walks of the simplicial complex. In the case of the uniform distribution on the top level, the highest local walks are unweighted (and possibly easier to analyze), while the local walks of lower levels are typically weighted in a nontrivial way.

Theorem 4.1.1 (Trickling-Down Theorem)**.** [Opp18]. *Let $(\mathscr{C}, w)$ be a weighted simplicial complex for which the link at S, $(\mathscr{C}_S, w_S)$, is a γ-local-spectral expander at level* 1*, meaning that $\lambda_2(G_{S\cup v}) \leq \gamma < 1$ for every $v \in \mathscr{C}_S(1)$. Then, if $\lambda_2(G_S) < 1$ (connectedness),*

$$\lambda_2(G_S) \leq \frac{\gamma}{1-\gamma}.$$

Proof. It suffices to prove the theorem for $S = \emptyset$, as any link can be viewed as the original complex. Let $D_1 = diag(\pi_1)$ and $D_{v,1} = diag(\pi_{v,1})$, and note the decomposition

$$D_1 = \sum_{v\in\mathscr{C}(1)} \pi_1(v)\overline{D_{v,1}}, \tag{4.2}$$

where the overline denotes the extension by zeros to the appropriate dimension. This can be verified by

$$\begin{aligned} D_1(u,u) &= \sum_{v\in\mathscr{C}(1)} \pi_1(v)\overline{D_{v,1}}(u,u) = \sum_{v\in\mathscr{C}(1):\{u,v\}\in\mathscr{C}(2)} \pi_1(v)\pi_{v,1}(u) \\ &= \sum_{v\in\mathscr{C}(1):\{u,v\}\in\mathscr{C}(2)} \frac{w(v)}{w(\emptyset)}\frac{w(\{u,v\})}{w(v)} = \pi_1(u), \end{aligned}$$

where the last equality follows from the definition of the weight function (Definition 3.2.1). Also, $G_\emptyset = diag(w_1)^{-1}W_\emptyset$, where $W_\emptyset(u,v) := w(\{u,v\})$ (with G_v and W_v analogously defined). Similarly, due to the properties of the weight function, this decomposes as

$$G_\emptyset = diag(w_1)^{-1}W_\emptyset = \frac{1}{w(\emptyset)}D_1^{-1}\sum_{v\in\mathscr{C}(1)}\overline{W_v} = D_1^{-1}\sum_{v\in\mathscr{C}(1)}\frac{\overline{diag(w_{v,1})G_v}}{w(\emptyset)},$$

or, equivalently,

$$G_\emptyset = D_1^{-1}\sum_{v\in\mathscr{C}(1)} \pi_1(v)\overline{D_{v,1}G_v}. \tag{4.3}$$

Moreover, notice that $G_\emptyset(v,u) = \pi_{v,1}(u)$ and for any function $f : \mathscr{C}(1) \to \mathbb{R}$,

$$G_\emptyset f(v) = \sum_{u\in\mathscr{C}(1)} G_\emptyset(v,u)f(u) = \overline{\pi_{v,1}}f. \tag{4.4}$$

Given these, for any function $f : \mathscr{C}(1) \to \mathbb{R}$, we can write

$$\begin{aligned}
\mathscr{E}_{G_\emptyset}(f,f) &= f^{\mathrm{T}}D_1(I-G_\emptyset)f = f^{\mathrm{T}}(D_1 - D_1G_\emptyset)f \\
&= \sum_{v\in\mathscr{C}(1)} \pi_1(v)\left[f^{\mathrm{T}}\overline{D_{v,1}}\left(I-\overline{G_v}\right)f\right] && \text{(by (4.2) and (4.3))} \\
&= \sum_{v\in\mathscr{C}(1)} \pi_1(v)\mathscr{E}_{G_v}(f_v,f_v) && \text{(where } f_v \text{ is } f \text{ restricted to } \mathscr{C}_v(1)) \\
&\geq (1-\gamma)\sum_{v\in\mathscr{C}(1)} \pi_1(v)\mathrm{Var}_{\pi_{v,1}}(f_v) && \text{(by (2.5) and } \lambda_2(G_v)\leq\gamma) \\
&= (1-\gamma)\sum_{v\in\mathscr{C}(1)} \pi_1(v)\left[f^{\mathrm{T}}\overline{D_{v,1}}f - (\overline{\pi_{v,1}}f)^2\right] \\
&= (1-\gamma)\left[\mathrm{Var}_{\pi_1}(f) + (\pi_1 f)^2 - \sum_{v\in\mathscr{C}(1)}\pi_1(v)\left([G_\emptyset f](v)\right)^2\right] \\
& && \text{(by (4.2) and (4.4))} \\
&= (1-\gamma)\left[\mathrm{Var}_{\pi_1}(f) - \mathrm{Var}_{\pi_1}(G_\emptyset f)\right].
\end{aligned}$$

Since the inequality above holds for any f, we can choose $f = v_i$, where v_i is an eigenvector of $G_\emptyset$ corresponding to eigenvalue λ_i. Then,

$$\mathscr{E}_{G_\emptyset}(v_i,v_i) \geq (1-\gamma)\left[\mathrm{Var}_{\pi_1}(v_i) - \mathrm{Var}_{\pi_1}(G_\emptyset v_i)\right],$$

which simplifies into

$$(1-\lambda_i)v_i^{\mathrm{T}}D_1v_i \geq (1-\gamma)(1-\lambda_i^2)v_i^{\mathrm{T}}D_1v_i.$$

Thus, $(1-\lambda_i) \geq (1-\gamma)(1-\lambda_i^2)$. In particular, if $\lambda_2(G_\emptyset) < 1$, then

$$\lambda_2(G_\emptyset) \leq \frac{\gamma}{1-\gamma}. \qquad \square$$

Example 4.1.2. *An example where the Trickling-Down Theorem (4.1.1) behaves poorly.*

Consider the uniform distribution over the proper n-colourings of the graph K_n. We will encode these as the facets of a simplicial complex with ground set $\{i_j \mid 1 \leq i,j \leq n\}$, where each i_j means that vertex i gets colour j. The walk $P_n^\vee$ in this case is not irreducible (it gets stuck in the same state), but the walk $P_{n-1}^\vee$ is (this walk is a type of random transpositions walk). Notice furthermore that the induced distribution π_{n-1}

is uniform and thus we can read out a uniformly distributed n-colouring by extending the $(n-1)$-colouring. Due to the symmetry of the problem, we derive

$$G_{\emptyset} = \left[\frac{1}{n-1}(J_n - I_n)\right]^{\otimes 2},$$

where J denotes the all-ones matrix and $\otimes$ stands for the Kronecker product. The second largest eigenvalue of this matrix is $a_0 = \frac{1}{(n-1)^2}$. Furthermore, by taking into account that links correspond to smaller instances of the same problem, we get, for $0 \leq i \leq n-2$, $a_i = \frac{1}{(n-i-1)^2}$. If we applied the Trickling-Down Theorem (4.1.1) from level $i = n-3$ downwards, we would get a trivial bound after 3 steps ($\frac{1}{4} \to \frac{1}{3} \to \frac{1}{2} \to 1$), and we wouldn't be able to compute a non-trivial bound for $\lambda(P^{\vee}_{n-1})$ by applying the Alev-Lau Local-to-Global Theorem (4.2.2). However, as we can see from the direct calculation of the a_is, the bounds actually improve as we go further down, and we can compute $\lambda(P^{\vee}_{n-1}) \geq \frac{1}{n-1}\prod_{i=0}^{n-3}\left(1 - \frac{1}{(n-i-1)^2}\right) = \frac{1}{n-1}\left(\frac{1}{2}\frac{n}{n-1}\right) \geq \frac{1}{2}\frac{1}{n-1}$.

4.2 Alev-Lau's Local-to-Global Theorem

In this section we present a way to derive bounds on the Poincaré constant of global walks from those of local walks. We will present a proof of Alev-Lau's theorem, which is a strengthening of a theorem by [KO18]. A step towards this is the following lemma.

Lemma 4.2.1. *Let $(\mathscr{C}, w)$ be a weighted simplicial complex that is a $(a_0, \ldots, a_{r-2})$-local-spectral expander. Then, for any $1 \leq k \leq r-1$,*

$$\lambda(P^{\wedge}_k) \geq \frac{k}{k+1}(1 - a_{k-1})\lambda(P^{\vee}_k).$$

Proof. Define $D_{S,1} := diag(\pi_{S,1})$ such that $D_{S,1}\mathbb{1} = \pi^{\mathrm{T}}_{S,1}$, where $\mathbb{1}$ is the all-ones vector. (Recall that we view $\pi_{S,1}$ as a row vector.) Let $f : \mathscr{C}(k) \to \mathbb{R}$, and, for a face $S \in \mathscr{C}(k-1)$, define the restriction $f_S : \mathscr{C}_S(1) \to \mathbb{R}$ with $f_S(v) = f(S \cup v)$. We will use an overline to denote the extension of a matrix by zeros to the appropriate dimension and with the appropriate indexing.

We will also use the decompositions

$$\begin{aligned}
\mathscr{E}_{P_k^\vee}(f,f) &= f^{\mathrm{T}} D_k (I - P_k^\vee) f \\
&= f^{\mathrm{T}} \frac{D_k}{k} \sum_{S\in\mathscr{C}(k-1)} \left(\overline{I_{S,1}} - \overline{P_{S,1}^\vee}\right) f \\
&= \sum_{S\in\mathscr{C}(k-1)} \pi_{k-1}(S) f^{\mathrm{T}} \frac{D_k}{k\pi_{k-1}(S)} \left(\overline{I_{S,1}} - \overline{P_{S,1}^\vee}\right) f \\
&= \sum_{S\in\mathscr{C}(k-1)} \pi_{k-1}(S) f_S^{\mathrm{T}} D_{S,1} \left(I_{S,1} - P_{S,1}^\vee\right) f_S \\
&= \sum_{S\in\mathscr{C}(k-1)} \pi_{k-1}(S) \mathscr{E}_{P_{S,1}^\vee}(f_S, f_S), \qquad (4.5)
\end{aligned}$$

and similarly, for the non-lazy up-down walk $\widetilde{P}_k^\wedge := \frac{k+1}{k} P_k^\wedge - \frac{1}{k}\mathbf{I}$,

$$\mathscr{E}_{\widetilde{P}_k^\wedge}(f,f) = \sum_{S\in\mathscr{C}(k-1)} \pi_{k-1}(S) \mathscr{E}_{\widetilde{P}_{S,1}^\wedge}(f_S, f_S), \qquad (4.6)$$

where $\widetilde{P}_{S,1}^\wedge = G_S$. Then,

$$\begin{aligned}
\mathscr{E}_{P_k^\wedge}(f,f) &= \frac{k}{k+1} \mathscr{E}_{\widetilde{P}_k^\wedge}(f,f) \\
&= \frac{k}{k+1} \sum_{S\in\mathscr{C}(k-1)} \pi_{k-1}(S) \mathscr{E}_{\widetilde{P}_{S,1}^\wedge}(f_S, f_S) && \text{(by (4.6))} \\
&\geq \frac{k}{k+1}(1 - a_{k-1}) \sum_{S\in\mathscr{C}(k-1)} \pi_{k-1}(S) \mathrm{Var}_{\pi_{S,1}}(f_S) && \text{(by (2.5) and } \lambda_2(G_S) \leq a_{k-1}\text{)} \\
&= \frac{k}{k+1}(1 - a_{k-1}) \sum_{S\in\mathscr{C}(k-1)} \pi_{k-1}(S) \left(f_S^{\mathrm{T}} D_{S,1} f_S - (\pi_{S,1} f_S)^2\right) \\
&= \frac{k}{k+1}(1 - a_{k-1}) \sum_{S\in\mathscr{C}(k-1)} \pi_{k-1}(S) \left(\mathscr{E}_{P_{S,1}^\vee}(f_S, f_S) + f_S^{\mathrm{T}} D_{S,1} P_{S,1}^\vee f_S - (\pi_{S,1} f_S)^2\right) \\
&= \frac{k}{k+1}(1 - a_{k-1}) \sum_{S\in\mathscr{C}(k-1)} \pi_{k-1}(S) \mathscr{E}_{P_{S,1}^\vee}(f_S, f_S) && \text{(because } P_{S,1}^\vee = \mathbb{1}\pi_{S,1}\text{)} \\
&= \frac{k}{k+1}(1 - a_{k-1}) \mathscr{E}_{P_k^\vee}(f,f). && \text{(by (4.5))}
\end{aligned}$$

This implies that $\lambda(P_k^\wedge) \geq \frac{k}{k+1}(1 - a_{k-1})\lambda(P_k^\vee)$. □

Theorem 4.2.2 (Alev-Lau Local-to-Global Theorem)**.** [Main Theorem of AL20]. *Let $(\mathscr{C}, w)$ be a weighted simplicial complex that is a $(a_0, \ldots, a_{r-2})$-local-spectral expander. Then, for any $1 \leq k \leq r$,*

$$\lambda(P_k^\vee) = \lambda(P_{k-1}^\wedge) \geq \frac{1}{k} \prod_{i=0}^{k-2} (1 - a_i).$$

Proof. The equality between the spectral gaps is due to Corollary 3.3.2.

The proof is by induction on k. The base case, for $k = 1$, asserts that $\lambda(P_1^\vee) = \lambda(P_0^\wedge) = 1$ (recall that by convention we set the spectral gap equal to 1 if the state space has size 1). This is true because $P_1^\vee = \mathbb{1}\pi_1$, and by Proposition 2.1.1 it has exactly one nonzero eigenvalue which is equal to 1.

Now, for the inductive step with $2 \le k \le r$,

$$\begin{aligned}
\lambda(P_k^\vee) &= \lambda(P_{k-1}^\wedge) \\
&\ge \frac{k-1}{k}(1-a_{k-2})\lambda(P_{k-1}^\vee) && \text{(by Lemma 4.2.1)} \\
&\ge \frac{k-1}{k}(1-a_{k-2})\frac{1}{k-1}\prod_{i=0}^{k-3}(1-a_i) && \text{(induction hypothesis)} \\
&= \frac{1}{k}\prod_{i=0}^{k-2}(1-a_i)\,. && \square
\end{aligned}$$

4.3 Related Work and Applications

The combination of the Trickling-Down Theorem (4.1.1) and the Alev-Lau Local-to-Global Theorem (4.2.2) works brilliantly for matroids. With these theorems the rapid mixing of the basis-exchange walk has been proved and the Mihail-Vazirani Conjecture (6.1.1) for the basis-exchange graph has been resolved in the affirmative [ALOV19]. Further than the uniform distribution over the bases, we will also consider strongly log-concave distributions. The relevant definitions and the proof of the mixing time bound appear in Chapter 6.

The framework presented in this chapter however is not limited to independence complexes of matroids and strongly log-concave (SLC) distributions. There are two ways to divert from those: either the underlying complex is still an independence complex but the distribution is not SLC, or the underlying complex is not that of matroid (and thus the distribution is not SLC).

Some other complexes to which this framework can be applied were provided in the paper of [AL20], and these include sampling independent sets of particular small enough size, and some specific instances of the matroid intersection problem, where the matroid intersection complex has faces which are the common independent sets of two matroids.

Spectral Independence. In its generality, the Trickling-Down Theorem (4.1.1) may give rise to tight bounds. However, for a particular system, the bound it yields may deteriorate in a way that does not reflect its true behaviour (see Example 4.1.2).

In order to overcome this limitation in spin systems, the framework of spectral independence has been proposed. This method aims to bound directly the second eigenvalue of each local walk (i.e., more directly than the Trickling-Down Theorem (4.1.1)). It does so by constructing a pairwise influence matrix Ψ whose largest eigenvalue is equal to the second largest eigenvalue of $G_{\emptyset}$. These can then be bounded by any matrix norm of Ψ. This must be done for all links so as to provide results for all local walks, as is required in order to apply the Alev-Lau Local-to-Global Theorem (4.2.2). This method was initially used to prove polynomial-time sampling of independent sets from the hardcore distribution. It was the first result to prove polynomial mixing time for unbounded degree graphs [ALO20].

As it was initially put forth, the method of spectral independence applied to 2-spin systems. For example, for the hardcore model, one creates a simplicial complex with elements $V \times \{0,1\}$ (where the value 0 designates an unoccupied vertex and the value 1 an occupied one), and with maximal faces those that correspond to independent sets. This simplicial complex is not the independence complex of a matroid, and this is why a new approach is required. The spectral independence framework was later generalized to multispin systems, where each vertex may have more than two possible states [FGYZ21, CGSV21]. The down-up chain of the top level of such models corresponds to the widely studied Glauber dynamics, and thus this formulation is a new way to effectively analyze this natural chain.

A general method of proving spectral independence is via stability properties (zero-free regions) of the characteristic polynomial of a distribution. For example, a notion that has been proposed is sector stability, which is a generalization of Hurwitz stability (equivalent to real stability; Real stability $\sim$ Hurwitz stability $\subseteq$ sector stability $\implies$ spectral independence). It has been shown that matchings of a particular size over arbitrary graphs satisfy sector-stability and thus exchange walks over them are rapidly mixing (implementing these walks however is not always tractable, but it is, e.g., for planar graphs) [AASV21]. Spectral independence can also be derived from zero-free regions around a particular point of interest [CLV21b]. Another way of establishing spectral independence is by constructing a contractive path coupling for the Glauber dynamics [BCC$^+$21, Liu21].

Matrix Trickling-Down. A different way of overcoming the inherent limitations of the Trickling-Down Theorem (4.1.1) is to generalize it in a way that takes into account more information about the spectrum of the transition matrices, as opposed to the gross uniform bound of the maximum (worst) of the second largest eigenvalues of the links. Such a generalization was done by [ALO21a] and it was applied to the problem of sampling edge-colourings. Compared to the Trickling-Down Theorem (4.1.1), their theorem is more technical and does not seem easily adaptable to other problems.

Some of the aforementioned results have been strengthened with tighter functional inequalities, namely by establishing entropy contraction instead of variance contraction. This is the subject of the next chapter.

Chapter 5

Entropy Contraction

For entropies, and for $f^{(k)} : \mathscr{C}(k) \to \mathbb{R}_{\geq 0}$, we get an inequality instead of an equality, in contrast to the case of variances (Equation (4.1)):

$$\begin{aligned}
\mathscr{E}_{P_k^\vee}\left(f^{(k)}, \log f^{(k)}\right) &= \left(f^{(k)}\right)^{\mathsf{T}} D_k (I - P_k^\vee) \log f^{(k)} && (5.1)\\
&= \left(f^{(k)}\right)^{\mathsf{T}} D_k \log f^{(k)} - \left(f^{(k)}\right)^{\mathsf{T}} \left(P_{k-1}^\uparrow\right)^{\mathsf{T}} D_{k-1} P_{k-1}^\uparrow \log f^{(k)} \\
& && \text{(by (3.6))}\\
&\geq \mathrm{Ent}_{\pi_k}\left(f^{(k)}\right) - \mathrm{Ent}_{\pi_{k-1}}\left(f^{(k-1)}\right). && \text{(by Jensen's inequality)}
\end{aligned}$$

Thus, entropy contraction when premultiplying by $P_{k-1}^\uparrow$ (or, equivalently, relative entropy contraction for the "half" walk $P_k^\downarrow$) implies a modified log-Sobolev inequality for $P_k^\vee$, but not the other way around, unlike the variance case.

5.1 A Local-to-Global Theorem

Lemma 5.1.1 ([CGM21]). *Let $0 \leq l \leq k \leq r$ and $f^{(k)} : \mathscr{C}(k) \to \mathbb{R}_{\geq 0}$ be a function on $\mathscr{C}(k)$. Then,*

$$\mathrm{Ent}_{\pi_k}\left(f^{(k)}\right) = \sum_{S \in \mathscr{C}(l)} \pi_l(S) \mathrm{Ent}_{\pi_{S,k-l}}\left(f_S^{(k-l)}\right) + \mathrm{Ent}_{\pi_l}\left(f^{(l)}\right),$$

where $f_S^{(k-l)}(T) := f^{(k)}(S \cup T)$ for $T \in \mathscr{C}_S(k-l)$. Moreover, the same decomposition holds for $f^{(k)} : \mathscr{C}(k) \to \mathbb{R}$ with Ent () *replaced by* Var ().

Proof. For any $S \in \mathscr{C}(l)$ and $S \subseteq T \in \mathscr{C}(k)$,

$$\begin{aligned}
\pi_k(T) = \frac{w(T)}{Z_k} &= \frac{Z_{S,k-l}}{Z_k}\frac{w_S(T\setminus S)}{Z_{S,k-l}} \\
&= \frac{k!}{l!(k-l)!}\frac{w(S)}{Z_l}\frac{w_S(T\setminus S)}{Z_{S,k-l}} && (k!Z_k = l!Z_l,\ w(S) = (k-l)!Z_{S,k-l}) \\
&= \binom{k}{l}\pi_l(S)\pi_{S,k-l}(T\setminus S)
\end{aligned}$$

and thus,

$$\pi_k(T) = \sum_{S\in\mathscr{C}(l):S\subseteq T} \pi_l(S)\pi_{S,k-l}(T\setminus S). \tag{5.2}$$

By Lemma 3.5.1, we have that $\mathbb{E}_{\pi_k} f^{(k)} = \mathbb{E}_{\pi_l} f^{(l)}$ and we may assume without loss of generality that $\mathbb{E}_{\pi_k} f^{(k)} = 1$. Then,

$$\begin{aligned}
\mathrm{Ent}_{\pi_k}\left(f^{(k)}\right) &= \sum_{T\in\mathscr{C}(k)} \pi_k(T) f^{(k)}(T)\log f^{(k)}(T) \\
&= \sum_{T\in\mathscr{C}(k)}\sum_{S\in\mathscr{C}(l):S\subseteq T} \pi_l(S)\pi_{S,k-l}(T\setminus S) f^{(k)}(T)\log f^{(k)}(T) && \text{(by (5.2))} \\
&= \sum_{S\in\mathscr{C}(l)} \pi_l(S) \sum_{T\in\mathscr{C}(k):T\supseteq S} \pi_{S,k-l}(T\setminus S) f_S^{(k-l)}(T\setminus S)\log f_S^{(k-l)}(T\setminus S) \\
&= \sum_{S\in\mathscr{C}(l)} \pi_l(S)\mathrm{Ent}_{\pi_{S,k-l}}\left(f_S^{(k-l)}\right) + \sum_{S\in\mathscr{C}(l)} \pi_l(S)\,\mathbb{E}_{\pi_{S,k-l}} f_S^{(k-l)} \log \mathbb{E}_{\pi_{S,k-l}} f_S^{(k-l)} \\
&= \sum_{S\in\mathscr{C}(l)} \pi_l(S)\mathrm{Ent}_{\pi_{S,k-l}}\left(f_S^{(k-l)}\right) + \sum_{S\in\mathscr{C}(l)} \pi_l(S) f^{(l)}(S)\log f^{(l)}(S) \\
& && \text{(by Lemma 3.5.1)} \\
&= \sum_{S\in\mathscr{C}(l)} \pi_l(S)\mathrm{Ent}_{\pi_{S,k-l}}\left(f_S^{(k-l)}\right) + \mathrm{Ent}_{\pi_l}\left(f^{(l)}\right).
\end{aligned}$$

The proof for Var () is analogous and it is an instance of the law of total variance. It suffices to assume without loss of generality that $\mathbb{E}_{\pi_k} f^{(k)} = 0$ and remove all logs. □

Now we are ready to show the local-to-global principle for entropies.

Theorem 5.1.2 (Local-to-Global Entropy Contraction Theorem)**.** [CLV21a, GM20, building upon CGM21]. *Let $(\mathscr{C}, w)$ be a weighted simplicial complex that satisfies the local entropy contraction inequalities*

$$\mathrm{Ent}_{\pi_{S,1}}\left(f_S^{(1)}\right) \le (1-\beta_i)\mathrm{Ent}_{\pi_{S,2}}\left(f_S^{(2)}\right),$$

for any $0 \le i \le r-2$, $S \in \mathscr{C}(i)$, *and* $f_S^{(2)} : \mathscr{C}_S(2) \to \mathbb{R}_{\ge 0}$, *where* $\{\beta_i\}_{i=0}^{r-2}$ *are local entropy contraction parameters in the interval* $(0,1]$.

Then, for any $1 \le k \le r$ *and* $f^{(k)} : \mathscr{C}(k) \to \mathbb{R}_{\ge 0}$, *we get the global entropy contraction inequalities*

$$\mathrm{Ent}_{\pi_{k-1}}\left(f^{(k-1)}\right) \le (1-\gamma_k)\mathrm{Ent}_{\pi_k}\left(f^{(k)}\right),$$

where

$$\gamma_k = \left(\sum_{l=0}^{k-1}\prod_{i=l}^{k-2}\frac{1-\beta_i}{\beta_i}\right)^{-1}.$$

Remark. *The same theorem and proof hold for variance in place of entropy. This is discussed in Section 5.3, where* $\{\beta_k\}$ *is expressed in terms of local-spectral expansion and a comparison with the Alev-Lau Local-to-Global Theorem (4.2.2) is given.*

Proof. We first prove the theorem for an alternative formulation of $\{\gamma_k\}_{k=1}^r$, equivalent to the one stated, which is given by the recursion

$$\gamma_1 = 1, \qquad \gamma_k = \frac{\beta_{k-2}\gamma_{k-1}}{1-\beta_{k-2}+\beta_{k-2}\gamma_{k-1}}, \text{ for } 2 \le k \le r, \tag{5.3}$$

where $\{\gamma_k\}_{k=1}^r$ belong in the interval $(0,1]$. The proof is by induction on k. The base case, for $k=1$, is immediate, given that both sides of the inequality $\mathrm{Ent}_{\pi_0}\left(f^{(0)}\right) \le (1-\gamma_1)\mathrm{Ent}_{\pi_1}\left(f^{(1)}\right)$ are equal to zero and $\gamma_1 \in (0,1]$. For conciseness, we use the notation $\mathrm{Ent}_k = \mathrm{Ent}_{\pi_k}\left(f^{(k)}\right)$ and, more generally, $\mathrm{Ent}_{S,k} = \mathrm{Ent}_{\pi_{S,k}}\left(f_S^{(k)}\right)$. For the induction step with $2 \le k \le r$, we have

$$\begin{aligned}
\mathrm{Ent}_{k-1} &= \sum_{S\in\mathscr{C}(k-2)} \pi_{k-2}(S)\mathrm{Ent}_{S,1} + \mathrm{Ent}_{k-2} && \text{(by Lemma 5.1.1)}\\
&\le (1-\beta_{k-2})\sum_{S\in\mathscr{C}(k-2)} \pi_{k-2}(S)\mathrm{Ent}_{S,2} + \mathrm{Ent}_{k-2} && \text{(local contraction hypothesis)}\\
&= (1-\beta_{k-2})\left(\sum_{S\in\mathscr{C}(k-2)} \pi_{k-2}(S)\mathrm{Ent}_{S,2} + \mathrm{Ent}_{k-2}\right) + \beta_{k-2}\mathrm{Ent}_{k-2} &&\\
&= (1-\beta_{k-2})\mathrm{Ent}_k + \beta_{k-2}\mathrm{Ent}_{k-2} && \text{(by Lemma 5.1.1)}\\
&\le (1-\beta_{k-2})\mathrm{Ent}_k + \beta_{k-2}(1-\gamma_{k-1})\mathrm{Ent}_{k-1}, && \text{(induction hypothesis)}
\end{aligned}$$

which implies that,

$$\begin{aligned}\mathrm{Ent}_{k-1} &\leq \frac{1-\beta_{k-2}}{1-\beta_{k-2}(1-\gamma_{k-1})}\mathrm{Ent}_k \\ &= \left(1-\frac{\beta_{k-2}\gamma_{k-1}}{1-\beta_{k-2}(1-\gamma_{k-1})}\right)\mathrm{Ent}_k \\ &= (1-\gamma_k)\mathrm{Ent}_k.\end{aligned}$$

For the last inequality we used the assumption that $\beta_{k-2} \in (0,1]$ as well as $\gamma_{k-1} \in (0,1]$ (induction hypothesis). Finally, we notice that $\gamma_k \in (0,1]$ because $0 < \frac{\beta_{k-2}\gamma_{k-1}}{1-\beta_{k-2}(1-\gamma_{k-1})} = \gamma_k \leq \frac{\beta_{k-2}\cdot 1}{1-\beta_{k-2}(1-1)} = \beta_{k-2} \leq 1$.

The recursion of (5.3) gives rise to the explicit formula

$$\gamma_k = \left(\sum_{l=0}^{k-1}\prod_{i=l}^{k-2}\frac{1-\beta_i}{\beta_i}\right)^{-1}, \qquad \text{for } 1 \leq k \leq r.$$

In order to show this we use induction again. For $k=1$, we verify that $\gamma_1 = 1$. For $2 \leq k \leq r$, we get

$$\begin{aligned}\frac{1}{\gamma_k} &= \frac{1-\beta_{k-2}+\beta_{k-2}\gamma_{k-1}}{\beta_{k-2}\gamma_{k-1}} = \frac{1-\beta_{k-2}}{\beta_{k-2}}\frac{1}{\gamma_{k-1}}+1 \\ &= \frac{1-\beta_{k-2}}{\beta_{k-2}}\sum_{i=0}^{k-2}\prod_{i=l}^{k-3}\frac{1-\beta_i}{\beta_i}+1 \qquad \text{(induction hypothesis)} \\ &= \sum_{l=0}^{k-2}\prod_{i=l}^{k-2}\frac{1-\beta_i}{\beta_i}+1 = \sum_{l=0}^{k-1}\prod_{i=l}^{k-2}\frac{1-\beta_i}{\beta_i}. \qquad \square\end{aligned}$$

As entropy contraction implies a modified log-Sobolev inequality, we have the following corollary.

Corollary 5.1.3. *Let $(\mathscr{C}, w)$ be a weighted simplicial complex that satisfies the assumptions of the Local-to-Global Entropy Contraction Theorem (5.1.2). Then, for $1 \leq k \leq r$,*

- $\rho(P_k^\vee) \geq \gamma_k$;
- $\rho(P_{k-1}^\wedge) \geq \gamma_k$.

Proof. Both claims are implied by the Local-to-Global Entropy Contraction Theorem (5.1.2). For the down-up walk, combine the global entropy contraction inequalities with Equation (5.1). A detailed proof for both walks is presented in Section 5.2. □

Interestingly, we do not know how to directly relate $\rho(P_k^\vee)$ with $\rho(P_{k-1}^\wedge)$, although it is straightforward to see that both walks have the same spectral gap (Corollary 3.3.2).

5.2 Mixing Time

In this section we prove Corollary 5.1.3, as well as the following mixing time bounds for the global walks.

Corollary 5.2.1. *In the same setting as in the Local-to-Global Entropy Contraction Theorem (5.1.2), we have that, for any* $2 \le k \le r$*,*

- $t_{\text{mix}}(P_k^{\vee}, \varepsilon) \le \frac{1}{\gamma_k}\left(\log\log \pi_{k,\min}^{-1} + \log \frac{1}{2\varepsilon^2}\right)$*;*
- $t_{\text{mix}}(P_{k-1}^{\wedge}, \varepsilon) \le \frac{1}{\gamma_k}\left(\log\log \pi_{k-1,\min}^{-1} + \log \frac{1}{2\varepsilon^2}\right)$*,*

where $\pi_{k,\min} := \min_{S \in \mathscr{C}(k)} \pi(S)$*.*

By recalling (3.10), we observe that the analysis of the "going-down" half—and, similarly, the "going-up" half—of $P_k^{\vee}$ and $P_{k-1}^{\wedge}$ corresponds to premultiplying by $P_{k-1}^{\uparrow}$—or, accordingly, $P_k^{\downarrow}$—to a function f. Hence, an entropy contraction inequality, such as those in the Local-to-Global Entropy Contraction Theorem (5.1.2), implies that the relative entropy contracts by $1-\gamma_k$ in the "going-down" half of the random walks $P_k^{\vee}$ and $P_{k-1}^{\wedge}$. What we show next is that the other half does not increase the relative entropy; a fact which is a special case of the so-called "data processing inequality".

Lemma 5.2.2. *For any* $k \ge 1$ *and* $f : \mathscr{C}(k-1) \to \mathbb{R}_{\ge 0}$*,*

$$\operatorname{Ent}_{\pi_k}\left(P_k^{\downarrow} f\right) \le \operatorname{Ent}_{\pi_{k-1}}(f) . \tag{5.4}$$

Proof. Firstly, we verify that

$$\begin{aligned} \mathbb{E}_{\pi_k} P_k^{\downarrow} f &= \pi_k P_k^{\downarrow} f \\ &= \pi_{k-1} f = \mathbb{E}_{\pi_{k-1}} f. && \text{(by (3.7))} \end{aligned}$$

Thus, we can assume both are 1 without loss of generality. Then,

$$\begin{aligned} \operatorname{Ent}_{\pi_k}\left(P_k^{\downarrow} f\right) &= \pi_k (P_k^{\downarrow} f \odot \log P_k^{\downarrow} f) \\ &\le \pi_k P_k^{\downarrow} (f \odot \log f) && \text{(by Jensen's inequality on } x\log x) \\ &= \pi_{k-1} (f \odot \log f) && \text{(by (3.7))} \\ &= \operatorname{Ent}_{\pi_{k-1}}(f) , \end{aligned}$$

where $\odot$ stands for the Hadamard product. □

With Theorem 5.1.2 and Lemma 5.2.2 in hand, we can show the decay of relative entropy for $P_k^\vee$ and $P_k^\wedge$.

Corollary 5.2.3. *In the same setting as in the Local-to-Global Entropy Contraction Theorem (5.1.2) and for any distribution* τ *on* $\mathscr{C}(k)$,

- *if* $2 \le k \le r$, *then* $D\left(\tau P_k^\vee \parallel \pi_k\right) \le (1-\gamma_k) D(\tau \parallel \pi_k)$;
- *if* $1 \le k \le r-1$, *then* $D\left(\tau P_k^\wedge \parallel \pi_k\right) \le (1-\gamma_{k+1}) D(\tau \parallel \pi_k)$.

Proof. We only prove this corollary for $P_k^\vee$, as the case of $P_k^\wedge$ is similar. We have that $D(\tau \parallel \pi_k) = \mathrm{Ent}_{\pi_k}\left(D_k^{-1}\tau^\mathsf{T}\right)$ where $D_k := \mathrm{diag}(\pi_k)$. Since $P_k^\vee$ is reversible, $D_k^{-1}(P_k^\vee)^\mathsf{T} = P_k^\vee D_k^{-1}$. Therefore,

$$\begin{aligned} D\left(\tau P_k^\vee \parallel \pi_k\right) &= \mathrm{Ent}_{\pi_k}\left(D_k^{-1}(P_k^\vee)^\mathsf{T}\tau^\mathsf{T}\right) = \mathrm{Ent}_{\pi_k}\left(P_k^\vee D_k^{-1}\tau^\mathsf{T}\right) \\ &\le \mathrm{Ent}_{\pi_{k-1}}\left(P_{k-1}^\uparrow D_k^{-1}\tau^\mathsf{T}\right) && \text{(by Lemma 5.2.2)} \\ &\le (1-\gamma_k)\,\mathrm{Ent}_{\pi_k}\left(D_k^{-1}\tau^\mathsf{T}\right) && \text{(by Theorem 5.1.2)} \\ &= (1-\gamma_k)\, D(\tau \parallel \pi_k)\,. \end{aligned}$$ □

It is well-known that the decay of relative entropy implies a modified log-Sobolev inequality.

Proof of Corollary 5.1.3. Given any $f^{(k)} : \mathscr{C}(k) \to \mathbb{R}_{\ge 0}$ such that $\mathbb{E}_{\pi_k} f^{(k)} = 1$, let $\tau = (D_k f^{(k)})^\mathsf{T}$ be the distribution corresponding to $f^{(k)}$. Then,

$$\begin{aligned} & D(\tau \parallel \pi_k) - D\left(\tau P_k^\vee \parallel \pi_k\right) \\ = \; & \sum_{S \in \mathscr{C}(k)} \tau(S) \log\left(\frac{\tau(S)}{\pi_k(S)}\right) - \sum_{S \in \mathscr{C}(k)} \tau P_k^\vee(S) \log\left(\frac{\tau P_k^\vee(S)}{\pi_k(S)}\right) \\ = \; & \sum_{S \in \mathscr{C}(k)} \left[\tau\left(\mathbf{I} - P_k^\vee\right)\right](S) \log\left(\frac{\tau(S)}{\pi_k(S)}\right) - \sum_{S \in \mathscr{C}(k)} \tau P_k^\vee(S) \log\left(\frac{\tau P_k^\vee(S)}{\tau(S)}\right) \\ = \; & \mathscr{E}_{P_k^\vee}\left(f^{(k)}, \log f^{(k)}\right) - D\left(\tau P_k^\vee \parallel \tau\right) \le \mathscr{E}_{P_k^\vee}\left(f^{(k)}, \log f^{(k)}\right). \end{aligned}$$

Thus,

$$\begin{aligned} \mathscr{E}_{P_k^\vee}\left(f^{(k)}, \log f^{(k)}\right) &\ge D(\tau \parallel \pi_k) - D\left(\tau P_k^\vee \parallel \pi_k\right) \\ &\ge \gamma_k D(\tau \parallel \pi_k) = \gamma_k \mathrm{Ent}_{\pi_k}\left(D_k^{-1}\tau^\mathsf{T}\right) && \text{(by Corollary 5.2.3)} \\ &= \gamma_k \mathrm{Ent}_{\pi_k}\left(f^{(k)}\right). \end{aligned}$$

This proves the statement for $P_k^\vee$. The same proof can be used for $P_k^\wedge$ by replacing every occurrence of $P_k^\vee$ with $P_k^\wedge$, and the factor γ_k with γ_{k+1}. □

In fact, the contraction of relative entropy (Corollary 5.2.3) directly implies the mixing time bound of Corollary 5.2.1, as illustrated by the following.

A proof of Corollary 5.2.1. We only prove this for $P_k^{\vee}$; the case of $P_k^{\wedge}$ is similar. Notice that Corollary 5.2.3 implies that

$$\begin{aligned} D\left(\tau_0 \left(P_k^{\vee}\right)^t \,\|\, \pi_k\right) &\le (1-\gamma_k)^t D(\tau_0 \,\|\, \pi_k) \\ &\le e^{-\gamma_k t} D(\tau_0 \,\|\, \pi_k) = e^{-\gamma_k t} \log \pi_k(x_0)^{-1}, \end{aligned}$$

where τ_0 is the initial distribution with $\tau_0(x_0) = 1$ for some $x_0 \in \mathscr{C}(k)$. Then, we use Pinsker's inequality ($2\|\tau - \sigma\|_{\mathrm{TV}}^2 \le D(\tau \,\|\, \sigma)$ for any two distributions τ, σ on the same state space), to show

$$2\left\|\tau_0 \left(P_k^{\vee}\right)^t - \pi_k\right\|_{\mathrm{TV}}^2 \le D\left(\tau_0 \left(P_k^{\vee}\right)^t \,\|\, \pi_k\right).$$

Setting $e^{-\gamma_k t} \log \pi_k(x_0)^{-1} \le 2\varepsilon^2$, we conclude that

$$\left\|\tau_0 \left(P_k^{\vee}\right)^t - \pi_k\right\|_{\mathrm{TV}} \le \varepsilon,$$

whenever

$$t \ge \frac{1}{\gamma_k}\left(\log\log \pi_k(x_0)^{-1} + \log \frac{1}{2\varepsilon^2}\right).$$

This gives us Corollary 5.2.1 for $P_k^{\vee}$. □

5.3 An Alternative Bound for Variance Contraction

In this section we prove an alternative Local-to-Global Variance Contraction Theorem (Theorem 5.3.1) and compare it to the Alev-Lau Local-to-Global Theorem (4.2.2), answering a question raised by [CLV21a].

Theorem 5.3.1. *Let $(\mathscr{C}, w)$ be a weighted simplicial complex that is a $(a_0, \ldots, a_{r-2})$-local-spectral expander. Then, for any $1 \le k \le r$,*

$$\lambda(P_k^{\vee}) = \lambda(P_{k-1}^{\wedge}) \ge \gamma_k,$$

where

$$\gamma_k = \left(\sum_{l=0}^{k-1} \prod_{i=l}^{k-2} \frac{1+a_i}{1-a_i}\right)^{-1}. \tag{5.5}$$

Proof. From the assumption on local-spectral expansion we get that for any $S \in \mathscr{C}(i)$, where $0 \le i \le r-2$,

$$\lambda(P^{\vee}_{S,2}) = \frac{\lambda(G_S)}{2} \ge \frac{1-a_i}{2} =: \beta_i.$$

This implies that for any $f_S^{(2)} : \mathscr{C}_S(2) \to \mathbb{R}$,

$$\mathscr{E}_{P^{\vee}_{S,2}}\left(f_S^{(2)}, f_S^{(2)}\right) \ge \beta_i \mathrm{Var}_{\pi_{S,2}}\left(f_S^{(2)}\right).$$

Remembering (4.1), we obtain

$$\mathrm{Var}_{\pi_{S,1}}\left(f_S^{(1)}\right) \le (1-\beta_i)\mathrm{Var}_{\pi_{S,2}}\left(f_S^{(2)}\right). \tag{5.6}$$

The remainder of the proof is identical to that of Theorem 5.1.2 except that Ent () is replaced by Var (). □

Recall that the Alev-Lau Local-to-Global Theorem (4.2.2) achieves a different lower bound, namely, for $1 \le k \le r$,

$$\gamma_{k,\mathrm{AL}} = \frac{1}{k}\prod_{i=0}^{k-2}(1-a_i).$$

We call a local-spectral profile $(a_0, \ldots, a_{r-2})$ consistent with the Trickling-Down Theorem (4.1.1) if, for $0 \le i \le r-2$,

- $a_i < 1$;
- $a_{i-1} \le \frac{a_i}{1-a_i}$.

Our bound in (5.5) is not better than the bound of the Alev-Lau Local-to-Global Theorem (4.2.2) when $(a_0, \ldots, a_{r-2})$ is consistent with the Trickling-Down Theorem (4.1.1).

Proposition 5.3.2. *Let $(a_0, \ldots, a_{r-2})$ be a local-spectral profile consistent with the Trickling-Down Theorem (4.1.1). Then, for $1 \le k \le r$,*

$$\gamma_{k,\mathrm{AL}} \ge \gamma_k,$$

with equality holding when the Trickling-Down Theorem (4.1.1) is tight.

Proof. By multiplying both numerator and denominator of (5.5) with $\gamma_{k,\mathrm{AL}}$, we get

$$\gamma_k = \frac{\gamma_{k,\mathrm{AL}}}{\frac{1}{k}\sum_{l=0}^{k-1}\prod_{i=0}^{l-1}(1-a_i)\prod_{i=l}^{k-2}(1+a_i)},$$

and thus it suffices to prove that $\sum_{l=0}^{k-1}\prod_{i=0}^{l-1}(1-a_i)\prod_{i=l}^{k-2}(1+a_i) \geq k$. We will do so by induction. For $k=1$ the statement holds. For $2 \leq k \leq r$,

$$\begin{aligned}
&\sum_{l=0}^{k-1}\prod_{i=0}^{l-1}(1-a_i)\prod_{i=l}^{k-2}(1+a_i) \\
&\quad = \sum_{l=0}^{k-2}\prod_{i=0}^{l-1}(1-a_i)\prod_{i=l}^{k-2}(1+a_i) + \prod_{i=0}^{k-2}(1-a_i) \\
&\quad = (1+a_{k-2})\sum_{l=0}^{k-2}\prod_{i=0}^{l-1}(1-a_i)\prod_{i=l}^{k-3}(1+a_i) + \prod_{i=0}^{k-2}(1-a_i) \\
&\quad \geq (1+a_{k-2})(k-1) + \prod_{i=0}^{k-2}(1-a_i) \qquad \text{(induction hypothesis)} \\
&\quad \geq (1+a_{k-2})(k-1) + \prod_{i=0}^{k-2}\frac{1-(k-1-i)a_{k-2}}{1-(k-2-i)a_{k-2}} \\
&\quad = (1+a_{k-2})(k-1) + 1-(k-1)a_{k-2} = k.
\end{aligned}$$

The last inequality is because, for $0 \leq i \leq k-2$, we have $1-a_i > 0$ and $a_i \leq \frac{a_{k-2}}{1-(k-2-i)a_{k-2}}$, as the local-spectral profile is consistent with the Trickling-Down Theorem (4.1.1). Lastly, notice that the equality in the proposition holds when the Trickling-Down Theorem (4.1.1) is tight, as we can perform the same induction with equalities instead of inequalities. In this case, our bound coincides with the bound of the Alev-Lau Local-to-Global Theorem (4.2.2). □

5.4 Concentration of Measure

One application of the modified log-Sobolev inequalities is to show concentration inequalities, via the Herbst argument [see, e.g., BT06, BLM13]. In the discrete setting, concentration inequalities have been obtained by [Goe04, Section 5] and can also be obtained by combining various results by [BG99, Sam05, BHT06]. The following lemma and its proof are a small modification of [HS19, Lemma 5]. For completeness, we include all details.

Lemma 5.4.1. *Let P be the transition matrix of a reversible Markov Chain with stationary distribution π on a finite set Ω, and let $f : \Omega \to \mathbb{R}$ be a function over Ω. Then,*

$$\Pr_{x\sim\pi}(f(x) - \mathbb{E}_\pi f \geq a) \;\leq\; \exp\left(-\frac{\rho(P)a^2}{2v(f)}\right),$$

where $a \geq 0$ *and*

$$v(f) := \max_{x \in \Omega} \left\{ \sum_{y \in \Omega} P(x,y)(f(x)-f(y))^2 \right\}.$$

Proof. For any $x \in \Omega$ and $t \in (0, +\infty)$, let

$$F_t(x) := \exp\left(tf(x) - ct^2\right),$$

where $c := \frac{v(f)}{2\rho(P)}$. We will use the inequality

$$z(e^z+1) \geq 2(e^z-1), \tag{5.7}$$

which holds for $z \geq 0$. To see this, notice that at $z=0$ the equality holds, and for $z>0$ the derivative of the left is larger than that of the right.

If $f(x) \geq f(y)$, we set $z = t(f(x)-f(y))$ in (5.7) and obtain

$$t(f(x)-f(y))(F_t(x)+F_t(y)) \geq 2(F_t(x)-F_t(y)),$$

which in turn implies that

$$(F_t(x)-F_t(y))(f(x)-f(y)) \leq \frac{t}{2}(F_t(x)+F_t(y))(f(x)-f(y))^2. \tag{5.8}$$

Notice that (5.8) also holds even if $f(x) < f(y)$ by swapping x and y. Thus, we have that

$$\begin{aligned}
\mathscr{E}_P(F_t, \log F_t) &= \frac{t}{2} \sum_{x \in \Omega} \sum_{y \in \Omega} \pi(x) P(x,y)(F_t(x)-F_t(y))(f(x)-f(y)) && \text{(by (2.4))} \\
&\leq \frac{t^2}{4} \sum_{x \in \Omega} \sum_{y \in \Omega} \pi(x) P(x,y)(F_t(x)+F_t(y))(f(x)-f(y))^2 && \text{(by (5.8))} \\
&= \frac{t^2}{2} \sum_{x \in \Omega} \pi(x) F_t(x) \sum_{y \in \Omega} P(x,y)(f(x)-f(y))^2 && \text{(by the reversibility of } P) \\
&\leq \frac{t^2}{2} v(f) \, \mathbb{E}_\pi F_t.
\end{aligned}$$

This, together with $\mathscr{E}_P(F_t, \log F_t) \geq \rho(P)\mathrm{Ent}_\pi(F_t)$ (recall the definition of $\rho(P)$), yields

$$\mathrm{Ent}_\pi(F_t) \leq ct^2 \, \mathbb{E}_\pi F_t.$$

By observing that

$$\frac{d}{dt}\left(\frac{\log \mathbb{E}_\pi F_t}{t}\right) = \frac{\mathrm{Ent}_\pi(F_t) - ct^2 \, \mathbb{E}_\pi F_t}{t^2 \, \mathbb{E}_\pi F_t} \leq 0,$$

we deduce that for any $t > 0$,

$$\frac{\log \mathbb{E}_\pi F_t}{t} \leq \lim_{h \to 0^+} \frac{\log \mathbb{E}_\pi F_h}{h} = \mathbb{E}_\pi f,$$

or equivalently,

$$\mathbb{E}_\pi F_t \leq \exp\left(t\, \mathbb{E}_\pi f\right).$$

Finally, by the Markov inequality, for any $a \geq 0$,

$$\begin{aligned} \Pr_{x\sim\pi}(f(x) - \mathbb{E}_\pi f \geq a) = & \Pr_{x\sim\pi}\left(F_t(x) \geq \exp(t\,\mathbb{E}_\pi f - ct^2 + at)\right) \\ \leq & \exp\left(ct^2 - at\right), \end{aligned}$$

where the right hand side is minimized for $t = \frac{a}{2c} = \frac{a\rho(P)}{\nu(f)}$. □

5.5 Related Work and Applications

Similar to the case of variance contraction, the Local-to-Global Entropy Contraction Theorem (5.1.2) lends itself very nicely for matroids and strongly log-concave distributions. After proving the base case for rank 2 matroids, the result follows immediately from an application of the theorem. This is presented in Chapter 6.

Optimal Mixing of Glauber Dynamics. The work of [CLV21a] established some mixing time bounds for the Glauber dynamics over distributions on spin systems with certain properties (most notably, spectral independence η and marginal boundedness b). They gave a bound of the order $O(n \log n)$ for an n-vertex underlying graph, where the maximum degree Δ, η, and b are treated as constants. This result is optimal in settings where η and b are well bounded in terms of the model's parameters, due to Hayes and Sinclair, who proved a $\Omega(n \log n)$ lower bound, with the hidden constant depending on Δ [HS05]. Some applications presented are the hardcore and Ising models in their respective uniqueness regions, colourings on triangle-free graphs, and the monomer-dimer model on all matchings. These results are achieved through the Local-to-Global Entropy Contraction Theorem (5.1.2) along with an extra technical step that deals with issues arising at the highest levels of the simplicial complex (namely proving approximate tensorization from uniform block factorization where the block size is an adequately small linear fraction of n). A drawback of their result is the possibly exponential dependence on Δ, a dependence which was removed altogether by [CFYZ21]; however with a $O(n^2 \log n)$ mixing time bound. Finally, and rather impressively, this series of works culminated in a $O(n \log n)$ bound where the hidden constant only depends on the gap from uniqueness [CFYZ22, CE22].

Entropic Independence. This notion is the entropic analogue of spectral independence, which was discussed in Section 4.3. It was introduced by [AJK$^+$21]. Entropic independence is a generalization of the strongly log-concave (SLC) property. While SLC means that the characteristic polynomial under any conditioning is log-concave at any point of the positive orthant, entropic independence only means that the tangent at the point $\mathbb{1}$ upper bounds a transformed version of the characteristic polynomial (again under any conditioning). One way to establish entropic independence is by proving spectral independence under any external field. Such distributions are also known as fractionally log-concave distributions over the positive orthant. In other words, if one establishes variance contraction results under any external field one gets entropy contraction results. The nice feature of this work is that it does not make any assumptions on marginal boundedness or maximum degree (for graphical models), as is the case in [CLV21a].

Chapter 6

Matroids

A matroid is a combinatorial structure that abstracts the notion of linear independence. We shall define it in terms of its independent sets, although many different equivalent definitions exist. Formally, a matroid $\mathscr{M} = (E, \mathscr{I})$ consists of a finite ground set E and a collection $\mathscr{I}$ of subsets of E (independent sets) that satisfy the following:

- $\emptyset \in \mathscr{I}$;
- if $S \in \mathscr{I}$, $T \subseteq S$, then $T \in \mathscr{I}$;
- if $S, T \in \mathscr{I}$ and $|S| > |T|$, then there exists an element $i \in S \setminus T$ such that $T \cup \{i\} \in \mathscr{I}$.

The first condition guarantees that $\mathscr{I}$ is non-empty, the second implies that $\mathscr{I}$ is downwards closed, and the third is usually called the augmentation axiom. We direct the reader to [Oxl92] for a reference book on matroid theory.

Due to the first two conditions we have that the set family $\mathscr{I}$ is a simplicial complex, which we call the *independence complex*. It is pure, as the augmentation axiom implies that all the maximal independent sets have the same cardinality, namely the rank r of $\mathscr{M}$. The set of bases $\mathscr{B}$ is the collection of maximal independent sets of $(\mathscr{I}, \subseteq)$. Furthermore, we denote by $\mathscr{I}(k)$ the collection of independent sets of size k, where $0 \leq k \leq r$.

Proposition 6.0.1. *Independence complexes are closed under taking links.*

Proof. If $\mathscr{I}$ is the independence complex of a matroid $\mathscr{M}$, then its link at $I \in \mathscr{I}$, namely $\mathscr{I}_I$, is the independence complex of the matroid $\mathscr{M}/I := (E \setminus I, \mathscr{I}_I)$. In matroid terminology this operation is called contraction, and it is well-known that matroids are closed under it. □

Proposition 6.0.2. *The elements in $\mathscr{I}(1)$ can be partitioned into equivalence classes, $\{P_i\}_{i=1}^p$, by the equivalence relation,*

$$i \sim j, \qquad \text{if and only if } \{i,j\} \notin \mathscr{I}(2).$$

Proof. We trivially have that $i \sim i$ and that $i \sim j$ is equivalent to $j \sim i$. It remains to prove that $i \sim j$ and $j \sim k$ implies $i \sim k$. Suppose, towards a contradiction, that $i \not\sim k$. Then, by the augmentation axiom, there must be an element from the independent set $\{i,k\}$ that can augment the independent set $\{j\}$, which yields a contradiction. □

6.1 Basis-Exchange Walk

The basis-exchange walk, denoted by P_{BX}, is defined as follows. In each step, we remove an element from the current basis uniformly at random to get an independent set S. Then, we move to a basis containing S uniformly at random.[1] This chain is irreducible (due to the augmentation axiom) and it converges to the uniform distribution over the bases of the matroid.

A closely related object is the basis-exchange graph, which has as vertices the bases $\mathscr{B}$, and as edges the pairs of bases that differ by one element. This graph can be alternatively defined as the 1-skeleton of the *matroid polytope*,

$$P_{\mathscr{M}} := \mathrm{conv}\{e_B \mid B \in \mathscr{B}\},$$

where $\{e_i \mid i \in E\}$ is the standard basis of $\mathbb{R}^E$ and $e_B = \sum_{i \in B} e_i$. For a nonempty collection of sets $\mathscr{S}$ over E, the polytope $\mathrm{conv}\{e_S \mid S \in \mathscr{S}\}$ is a matroid polytope if and only if every edge is a translation of some $e_i - e_j$ [GGMS87]. The polytopes that result as the convex hull of points with coordinates 0 or 1 (such as the matroid polytope), are called 0-1 polytopes.

Conjecture 6.1.1 (Mihail-Vazirani Conjecture). [MV89, Mih92]. *Let P be a 0-1 polytope, and let $G_P = (V,E)$ be its 1-skeleton. Then, G_P has cutset expansion greater or equal to 1, meaning that,*

$$\min_{S \subseteq V : 0 < |S| \le |V|/2} \frac{|E(S,S^c)|}{|S|} \ge 1.$$

[1] Notice that to implement this step we assume that the matroid is given to us via an independence oracle. Even so, this step may require more than constant time. The chain considered here is sometimes called the modified basis-exchange walk. A common alternative in the literature is to randomly propose an element and then apply a rejection filter.

This conjecture, when specialized to matroid polytopes, is relevant to questions of polynomial mixing of walks over the basis-exchange graph and approximately counting the number of bases.[2] By analyzing the basis-exchange walk and proving a sharp spectral gap bound, [ALOV19] resolved the Mihail-Vazirani Conjecture (6.1.1) for the special class of matroid polytopes.

6.1.1 Proof of Spectral Gap Bound

To begin with, we restrict our attention to the uniform distribution over the bases. However, the results that we prove here hold more generally for strongly log-concave distributions, as we will explain in Section 6.3.

Lemma 6.1.2. *Let $\mathscr{M} = (E, \mathscr{I})$ be a matroid of rank 2. Then, the local walk $G_\emptyset$ of the weighted independence complex $(\mathscr{I}, w)$, where $w(B) = 1$ for $B \in \mathscr{B}$ (uniform distribution over the bases), satisfies $\lambda_2(G_\emptyset) \leq 0$.*

Proof. First notice that, due to the equivalence classes $\{P_i\}_{i=1}^p$ guaranteed by Proposition 6.0.2, the underlying graph of $G_\emptyset$ is a multipartite graph, with degrees given by $w_1 = \{w(v)\}_{v \in \mathscr{I}(1)}$. Thus, we can write

$$G_\emptyset = \operatorname{diag}(w_1)^{-1} \left(\mathbb{1}\mathbb{1}^\intercal - \sum_{i=1}^{p} \mathbb{1}_{P_i} \mathbb{1}_{P_i}^\intercal \right).$$

The matrix $\mathbb{1}\mathbb{1}^\intercal - \sum_{i=1}^{p} \mathbb{1}_{P_i} \mathbb{1}_{P_i}^\intercal$ $(\preceq \mathbb{1}\mathbb{1}^\intercal)$ has at most one positive eigenvalue (by Proposition 2.1.1 on $\mathbb{1}\mathbb{1}^\intercal$ and by Corollary 2.1.4). Premultiplying by the diagonal matrix $\operatorname{diag}(w_1)^{-1}$ does not increase the number of positive eigenvalues (due to Proposition 2.1.5), and so we get that $G_\emptyset$ has at most one positive eigenvalue. $\square$

Theorem 6.1.3. *Let $\mathscr{M} = (E, \mathscr{I})$ be a matroid of rank r and let $(\mathscr{I}, w)$ be the associated weighted independence complex, where $w(B) = 1$ for $B \in \mathscr{B}$ (uniform distribution over the bases). Then, for any $1 \leq k \leq r$,*

$$\lambda(P_k^\vee) = \lambda(P_{k-1}^\wedge) \geq \frac{1}{k}.$$

In particular, for the basis-exchange walk we get $\lambda(P_{\mathrm{BX}}) = \lambda(P_r^\vee) \geq \frac{1}{r}$.

[2]The degree of each vertex in the basis-exchange graph is upper bounded by n^2, where $n = |E|$, and thus the cutset expansion bound can be translated to a bound on the edge expansion $\left(\min_{S \subseteq V: 0 < |E(S)| \leq |E(S^c)|} \frac{|E(S,S^c)|}{|E(S)|}\right)$ with a polynomial loss.

Proof. By Lemma 6.1.2 and the fact that the links of an independence complex are independence complexes (Proposition 6.0.1), we get that $\lambda_2(G_I) \leq 0$, for $I \in \mathscr{I}(r-2)$. Next, by repeated application of the Trickling-Down Theorem (4.1.1), we get that $(\mathscr{I}, w)$ is a $(0,0,\ldots,0)$-local-spectral expander. Finally, the theorem follows by the Alev-Lau Local-to-Global Theorem (4.2.2). □

Example 6.1.4. *A matroid that achieves the spectral gap lower bound of Theorem 6.1.3.*

Consider the graphic matroid that corresponds to the graph which is similar to a path of length r, but with one of the edges replaced by two parallel edges. This matroid has rank r and $r+1$ elements. It has two bases that correspond to the spanning trees of the graph, and the basis-exchange walk transition matrix is $P_{\mathrm{BX}} = (1-\frac{1}{r})I_2 + \frac{1}{2r}J_2$, where J is the all ones matrix. This transition matrix has eigenvalues 1 and $1-\frac{1}{r}$, and thus $\lambda(P_{\mathrm{BX}}) = \frac{1}{r}$. This shows that the spectral gap bound of Theorem 6.1.3 is sharp.

6.2 Strongly Log-Concave Polynomials and Distributions

The characteristic polynomial of a distribution π over $2^{[n]}$ is

$$g_\pi(\mathbf{x}) = \sum_{S\subseteq[n]} \pi(S) \prod_{i\in S} x_i.$$

A polynomial is called *homogeneous* if its non-zero monomial terms have the same degree. It is called *r-homogeneous* if every non-zero monomial term has degree r.

We write ∂_i as shorthand for $\frac{\partial}{\partial x_i}$, and ∂_I for an index set $I = \{i_1,\ldots,i_k\}$ as shorthand for $\partial_{i_1}\ldots\partial_{i_k}$.

Definition 6.2.1 (Strong Log-Concavity). *A polynomial $p \in \mathbb{R}[x_1,\ldots,x_n]$ with non-negative coefficients is* log-concave *at $\boldsymbol{x} \in \mathbb{R}_{\geq 0}$ if the Hessian of its logarithm, $\nabla^2 \log p := \left(\partial_i\partial_j \log p\right)_{ij}$, is negative semi-definite at $\boldsymbol{x}$. We call p* strongly log-concave *if for any index set $I \subseteq [n]$, $\partial_I p$ is log-concave at the all-1 vector $\mathbb{1}$.*

The notion of strong log-concavity was introduced by [Gur09b, Gur09a]. There are also notions of *complete log-concavity* introduced by [AOV18], and *Lorentzian* polynomials introduced by [BH19]. It turns out that for homogeneous polynomials the three notions are equivalent [BH19, Theorem 5.3]. (See also ALOV19.)

The following property of strongly log-concave polynomials is particularly useful [AOV18, BH19].

Proposition 6.2.2. *If p is strongly log-concave, then for any $I \subseteq [n]$, the Hessian matrix $\nabla^2 \partial_I p(\mathbb{1})$ has at most one positive eigenvalue.*

In fact, when p is homogeneous, $\nabla^2 \partial_I p(\mathbb{1})$ having at most one positive eigenvalue is equivalent to $\nabla^2 \log \partial_I p(\mathbb{1})$ being negative semi-definite [AOV18], but we will only need the proposition above.

A distribution π is called homogeneous (or r-homogeneous) if its characteristic polynomial g_π is homogeneous (or, respectively, r-homogeneous). Moreover, a distribution π is called strongly log-concave if g_π is strongly log-concave.

6.3 Variance Contraction for Strongly Log-Concave Distributions

Let $\mathscr{M} = (E, \mathscr{I})$ be a matroid of rank r and let $(\mathscr{I}, w)$ be the associated weighted independence complex, where w_r corresponds to a strongly log-concave distribution. If $|I| \leq r-2$, let W_I be the matrix such that $(W_I)_{uv} = w_I(\{u,v\})$ for any $u, v \in E \setminus I$. Then, notice that

$$\begin{aligned} w_I(\{u,v\}) &= w(I \cup \{u,v\}) \\ &= (r - |I| - 2)! \sum_{B \in \mathscr{B},\, I \cup \{u,v\} \subseteq B} w(B) \\ &= (r - |I| - 2)! Z_r \cdot \partial_u \partial_v \partial_I g_\pi(\mathbb{1}), \qquad \text{(by (6.1))} \end{aligned}$$

because, for any $I \in \mathscr{I}$,

$$\partial_I g_\pi(\mathbb{1}) = \sum_{B \in \mathscr{B}, I \subset B} \pi(B) = \frac{1}{Z_r} \sum_{B \in \mathscr{B}, I \subset B} w(B). \qquad (6.1)$$

In other words, W_I is $\nabla^2 \partial_I g_\pi$ multiplied by the scalar $(r - |I| - 2)! Z_r$. Thus, Proposition 6.2.2 implies the following.

Proposition 6.3.1. *Let π be an r-homogeneous strongly log-concave distribution over $\mathscr{M} = (E, \mathscr{I})$. If $I \in \mathscr{I}$ and $|I| \leq r-2$, then the matrix W_I has at most one positive eigenvalue.*

Proposition 6.3.1 implies the following bound for a quadratic form, which will be useful later.

Lemma 6.3.2. *Let π be an r-homogeneous strongly log-concave distribution over $\mathscr{M} = (E, \mathscr{I})$, and let $I \in \mathscr{I}$ such that $|I| \leq r-2$. Let $f : \mathscr{M}_I(1) \to \mathbb{R}$ be a function such that*

$\mathbb{E}_{\pi_{I,1}} f = 1$. *Then*

$$f^{\mathrm{T}} W_I f \le w(I).$$

Proof. Let $w_I = \{w_I(v)\}_{v \in E \setminus I}$. The constraint $\mathbb{E}_{\pi_{I,1}} f = 1$ implies that $\sum_{v \in E \setminus I} w_I(v) f(v) = w(I)$. Let $D = \mathrm{diag}(w_I)$ and $A = D^{-1/2} W_I D^{-1/2}$. Then, A is a real symmetric matrix. By Proposition 6.3.1, W_I has at most one positive eigenvalue, and thus so does A (by Proposition 2.1.1 and Proposition 2.1.5). By the Spectral Theorem (2.1.2), we may decompose A as

$$A = \sum_{i=1}^{|E \setminus I|} \lambda_i g_i g_i^{\mathrm{T}}, \tag{6.2}$$

where $\{g_i\}$ is an orthonormal basis and $\lambda_i \le 0$ for all $i \ge 2$. Moreover, notice that $\sqrt{w_I}$ is an eigenvector of A with eigenvalue 1. Thus, $\lambda_1 = 1$ and g_1 can be taken as $\sqrt{\pi_{I,1}}$.

The decomposition (6.2) directly implies that

$$W_I = \sum_{i=1}^{|E \setminus I|} \lambda_i h_i h_i^{\mathrm{T}},$$

where $h_i = D^{1/2} g_i$. In particular, $h_1 = \frac{1}{\sqrt{w(I)}} w_I$.

The assumption $\sum_{v \in E \setminus I} w_I(v) f(v) = w(I)$ can be rewritten as $\langle h_1, f \rangle = \sqrt{w(I)}$. Thus,

$$f^{\mathrm{T}} W_I f = \sum_{i=1}^{|E \setminus I|} \lambda_i \langle h_i, f \rangle^2 \le \langle h_1, f \rangle^2 = w(I),$$

where the inequality is due to the fact that $\lambda_1 = 1$ and $\lambda_i \le 0$ for all $i \ge 2$. The lemma follows. □

At the end of this section, let us comment that it is possible to prove the decay of variances (or spectral gap bounds) similar to Theorem 6.1.3. This provides an alternative proof for the spectral gap of $P_{\mathrm{BX},\pi}$ to [KO18, ALOV19], where $P_{\mathrm{BX},\pi}$ is the basis-exchange walk weighted by a strongly log-concave distribution π. Indeed, as is the case in the proof of Theorem 6.1.3, we can apply the Alev-Lau Local-to-Global Theorem (4.2.2), as the base case still holds, yielding a $(0,0,\ldots,0)$-local-spectral expander. However, the proof of the base case needs to be edited as follows.

Lemma 6.3.3. *Let $\mathscr{M} = (E, \mathscr{I})$ be a matroid and let $(\mathscr{I}, w)$ be the associated weighted independence complex, where w_r corresponds to a strongly log-concave distribution. Let $f^{(2)} : \mathscr{I}(2) \to \mathbb{R}$ and $f^{(1)} = P_1^{\uparrow} f^{(2)}$. Then,*

$$\mathrm{Var}_{\pi_1}\left(f^{(1)}\right) \le \frac{1}{2} \mathrm{Var}_{\pi_2}\left(f^{(2)}\right).$$

Proof. We begin by observing that

$$\operatorname{diag}(w_1)\left(2P_1^{\wedge}-I\right)=W_{\emptyset}. \tag{6.3}$$

From this identity and Proposition 6.3.1, we deduce that the symmetric matrix $\operatorname{diag}(w_1)\,(2P_1^{\wedge}-I)$ has at most one positive eigenvalue. Premultiplying by the positive semidefinite matrix $\operatorname{diag}(w_1)^{-1}$, we get that $2P_1^{\wedge}-I$ also has at most one positive eigenvalue (follows from Proposition 2.1.5). Furthermore, the spectra of $2P_1^{\wedge}-I$ and $2P_2^{\vee}-I$ are the same up to some extra -1s (Proposition 2.1.1). So, if $|\mathscr{M}(2)|\geq 2$ (otherwise the lemma holds trivially), $\lambda_2(P_2^{\vee})\leq 1/2$. Then, the spectral gap $\lambda(P_2^{\vee})=1-\lambda_2(P_2^{\vee})\geq 1/2$, which means that

$$\mathscr{E}_{P_2^{\vee}}\left(f^{(2)},f^{(2)}\right)\geq\frac{1}{2}\operatorname{Var}_{\pi_2}\left(f^{(2)}\right).$$

However, this is equivalent to the statement of the lemma, as can be seen by the following equalities:

$$\begin{aligned}
\operatorname{Var}_{\pi_1}\left(f^{(1)}\right) &= \left(f^{(1)}\right)^{\mathsf{T}} D_1 f^{(1)}-\left(\mathbb{E}_{\pi_1}f^{(1)}\right)^2\\
&= \left(f^{(2)}\right)^{\mathsf{T}}\left(P_1^{\uparrow}\right)^{\mathsf{T}} D_1 P_1^{\uparrow} f^{(2)}-\left(\mathbb{E}_{\pi_2}f^{(2)}\right)^2 && \text{(by Lemma 3.5.1)}\\
&= \left(f^{(2)}\right)^{\mathsf{T}} D_2 P_2^{\vee} f^{(2)}-\left(\mathbb{E}_{\pi_2}f^{(2)}\right)^2 && \text{(by (3.6))}\\
&= \operatorname{Var}_{\pi_2}\left(f^{(2)}\right)-\mathscr{E}_{P_2^{\vee}}\left(f^{(2)},f^{(2)}\right). && \square
\end{aligned}$$

6.4 Improved Results Using Entropy Contraction

We begin by establishing the base case, in order to use the Local-to-Global Entropy Contraction Theorem (5.1.2).

Lemma 6.4.1. *Let $\mathscr{M}=(E,\mathscr{I})$ be a matroid and let $(\mathscr{I},w)$ be the associated weighted independence complex, where w_r corresponds to a strongly log-concave distribution. Let $f^{(2)}:\mathscr{I}(2)\to\mathbb{R}_{\geq 0}$ be a non-negative function, and let $f^{(1)}=P_1^{\uparrow}f^{(2)}$. Then,*

$$\operatorname{Ent}_{\pi_1}\left(f^{(1)}\right)\leq\frac{1}{2}\operatorname{Ent}_{\pi_2}\left(f^{(2)}\right).$$

Proof. Without loss of generality we may assume that $\mathbb{E}_{\pi_2}f^{(2)}=1$ and, therefore, by (1) of Lemma 3.5.1, $\mathbb{E}_{\pi_1}f^{(1)}=1$. Note that for $v\in\mathscr{I}(1)$,

$$f^{(1)}(v)=\sum_{S\in\mathscr{I}(2):v\in S}\frac{w(S)}{w(v)}f^{(2)}(S).$$

We will use the following inequality, which is valid for any $a \geq 0$ and $b > 0$,

$$a \log \frac{a}{b} \geq a - b. \tag{6.4}$$

Noticing that $Z_1 = 2Z_2$, we have

$$\begin{aligned}
&\mathrm{Ent}_{\pi_2}\left(f^{(2)}\right) - 2\mathrm{Ent}_{\pi_1}\left(f^{(1)}\right) \\
&= \sum_{S\in\mathscr{I}(2)} \pi_2(S) f^{(2)}(S) \log f^{(2)}(S) - 2\sum_{v\in\mathscr{I}(1)} \pi_1(v) \left(\sum_{S\in\mathscr{I}(2):v\in S} \frac{w(S)}{w(v)} f^{(2)}(S) \right) \log f^{(1)}(v) \\
&= \sum_{S\in\mathscr{I}(2)} \left(\pi_2(S) f^{(2)}(S) \log f^{(2)}(S) - 2\sum_{v\in S} \pi_1(v) \frac{w(S)}{w(v)} f^{(2)}(S) \log f^{(1)}(v) \right) \\
&= \sum_{S\in\mathscr{I}(2)} \left(\frac{w(S)}{Z_2} f^{(2)}(S) \log f^{(2)}(S) - 2\sum_{v\in S} \frac{w(v)}{Z_1} \cdot \frac{w(S)}{w(v)} f^{(2)}(S) \log f^{(1)}(v) \right) \\
&= \sum_{S=\{u,v\}\in\mathscr{I}(2)} \frac{w(S)}{Z_2} f^{(2)}(S) \left(\log f^{(2)}(S) - \log f^{(1)}(v) - \log f^{(1)}(u) \right) \\
&\geq \sum_{S=\{u,v\}\in\mathscr{I}(2)} \frac{w(S)}{Z_2} \left(f^{(2)}(S) - f^{(1)}(v) f^{(1)}(u) \right) \\
&= \sum_{S\in\mathscr{I}(2)} \pi_2(S) f^{(2)}(S) - \sum_{S=\{u,v\}\in\mathscr{I}(2)} \frac{w(S)}{Z_2} \cdot f^{(1)}(v) f^{(1)}(u) \\
&= 1 - \frac{1}{2Z_2} \cdot \left(f^{(1)}\right)^{\mathsf{T}} W_{\emptyset} f^{(1)},
\end{aligned}$$

where the inequality is by (6.4) with $a = f^{(2)}(S)$ and $b = f^{(1)}(u) f^{(1)}(v)$ when $b > 0$, and when $b = 0$ we have $a = 0$ as well. Thus, the lemma follows from Lemma 6.3.2 with $I = \emptyset$ and $w(\emptyset) = Z_1 = 2Z_2$. □

We generalise Lemma 6.4.1 as follows.

Theorem 6.4.2. *Let $\mathscr{M} = (E, \mathscr{I})$ be a matroid of rank r and let $(\mathscr{I}, w)$ be the associated weighted independence complex, where w_r corresponds to a strongly log-concave distribution. Then, for any $1 \leq k \leq r$ and $f^{(k)} : \mathscr{C}(k) \to \mathbb{R}_{\geq 0}$, we get the global entropy contraction inequalities*

$$\mathrm{Ent}_{\pi_{k-1}}\left(f^{(k-1)}\right) \leq \left(1 - \frac{1}{k}\right) \mathrm{Ent}_{\pi_k}\left(f^{(k)}\right),$$

and the modified log-Sobolev constant bounds

- $\rho(P_k^{\vee}) \geq \frac{1}{k}$;
- $\rho(P_{k-1}^{\wedge}) \geq \frac{1}{k}$.

In particular, for the basis-exchange walk we get $\rho(P_{\text{BX}}) = \rho(P_r^\vee) \geq \frac{1}{r}$.

Proof. By Lemma 6.4.1 and the fact that independence complexes (and strongly log-concave distributions) are closed under taking links (Proposition 6.0.1), we get the local entropy contraction inequalities

$$\text{Ent}_{\pi_{S,1}}\left(f_S^{(1)}\right) \leq \frac{1}{2}\text{Ent}_{\pi_{S,2}}\left(f_S^{(2)}\right),$$

for any $0 \leq i \leq r-2$, $S \in \mathscr{C}(i)$, and $f_S^{(2)} : \mathscr{C}_S(2) \to \mathbb{R}_{\geq 0}$. Thus, we can apply the Local-to-Global Entropy Contraction Theorem (5.1.2) with $\beta_i = \frac{1}{2}$, $0 \leq i \leq r-2$. This yields the global entropy contraction inequalities with contraction parameters

$$\gamma_k = \left(\sum_{l=0}^{k-1}\prod_{i=l}^{k-2}\frac{1-1/2}{1/2}\right)^{-1} = \frac{1}{k},$$

for $1 \leq k \leq r$. Finally, the modified log-Sobolev constant bounds follow from Corollary 5.1.3. □

By Corollary 5.2.1, we also have the following result.

Corollary 6.4.3. *In the same setting as Theorem 6.4.2 and for any* $2 \leq k \leq r$*, we have that*

- $t_{\text{mix}}(P_k^\vee, \varepsilon) \leq k\left(\log\log \pi_{k,\min}^{-1} + \log\frac{1}{2\varepsilon^2}\right)$;
- $t_{\text{mix}}(P_{k-1}^\wedge, \varepsilon) \leq k\left(\log\log \pi_{k-1,\min}^{-1} + \log\frac{1}{2\varepsilon^2}\right)$.

In particular, for the basis-exchange walk $P_{\text{BX},\pi}$ *according to a strongly log-concave* π*,*

$$t_{\text{mix}}(P_{\text{BX},\pi}, \varepsilon) \leq r\left(\log\log \pi_{\min}^{-1} + \log\frac{1}{2\varepsilon^2}\right).$$

For the uniform distribution over the bases of a matroid $\mathscr{M}$ of rank r with a ground set of size n, Corollary 6.4.3 implies that the mixing time of the basis-exchange walk is $O(r(\log r + \log\log n))$, which improves upon the $O(r^2 \log n)$ bound of [ALOV19]. Moreover, by arguing that a "warm start" state is reached quickly, the dependence on n can be removed, yielding a $O(r\log r)$ bound [ALO+21b] (however, recall that implementing each step of the chain given oracle access to the independence complex takes $O(n)$ time).

Example 6.4.4. *A matroid for which Corollary 6.4.3 is sharp.*

The mixing time bound in Corollary 6.4.3 is sharp, as there are matroids where the upper bound is achieved [Jer03, Ex. 9.14]. One such example is the graphic matroid defined by a graph which is similar to a path but with two parallel edges connecting every two successive vertices instead of a single edge. Equivalently, this can be viewed as the partition matroid where each block has two elements and each basis is formed by choosing exactly one element from every block. The rank of this matroid is $r = n/2$, and $\pi_{\min} = \frac{1}{2^{n/2}}$. The Markov chain $P_{\mathrm{BX},\pi}$ in this case is just the 1/2-lazy random walk on the $n/2$-dimensional Boolean hypercube, which has mixing time $\Theta(n \log n)$, matching the upper bound in Corollary 6.4.3.

By the standard Herbst argument [see, e.g., Goe04, Sam05, BLM13], Theorem 6.4.2 also implies the following concentration bound.

Corollary 6.4.5. *Let π be an r-homogeneous strongly log-concave distribution with support $\Omega \subset 2^{[n]}$, and $P_{\mathrm{BX},\pi}$ be the corresponding basis-exchange walk. For any function $f : \Omega \to \mathbb{R}$ and $a \geq 0$,*

$$\Pr_{x\sim\pi}(|f(x) - \mathbb{E}_\pi f| \geq a) \;\leq\; 2\exp\left(-\frac{a^2}{2rv(f)}\right),$$

where $v(f)$ is the maximum of one-step variances,

$$v(f) := \max_{x\in\Omega} \left\{ \sum_{y\in\Omega} P_{\mathrm{BX},\pi}(x,y)(f(x) - f(y))^2 \right\}.$$

Corollary 6.4.5 follows from applying Lemma 5.4.1 to both f and $-f$ together with Theorem 6.4.2. For a c-Lipschitz function (under the graph distance in the basis-exchange graph), $v(f) \leq c^2$. For general matroids, an example is the function that counts the number of elements belonging to a specified subset of the ground set, which has Lipschitz constant $c = 1$. More examples were given by [PP14] for graphic matroids, such as functions that count the number of leaves in a spanning tree ($c = 2$), or the number of vertices with odd degrees ($c = 4$).

There have been a number of results concerning concentration inequalities for Lipshitz functions of negatively correlated variables. [PP14] showed concentration for variables satisfying the *stochastic covering property* (**SCP**), which includes strong Rayleigh distributions as special cases. (See also [HS19].) Correcting an earlier proof of [DR98], [GV18] showed concentration for variables with *negative regression* (**NR**), a property even weaker than **SCP**.

Corollary 6.4.5 generalises the concentration bound for Lipschitz functions in strong Rayleigh distributions. However, **SLC** is *not* a negative correlation property. We construct examples in Section 6.5 to show that **SCP** and **SLC** are in fact incomparable. Thus, Corollary 6.4.5 is incomparable to the results of [PP14, HS19, GV18]. It is not clear whether there is a larger class of distributions, generalising both **NR** and **SLC**, which retains this concentration bound.

6.5 Stochastic Covering Property and Strong Log-Concavity

The results obtained by [PP14] and [HS19] only require a property which is weaker than the strong Rayleigh property (**SRP**), namely the *stochastic covering property* (**SCP**). A distribution with the **SRP** is one for which the generating polynomial is nonzero when the imaginary part of each variable is strictly greater than 0. Since strong log-concavity (**SLC**) is also a generalisation of **SRP**, it is natural to wonder about the relationship between **SLC** and **SCP**. In this section we show that **SLC** is incomparable to **SCP**. As a result, Theorem 6.4.2 and Corollary 6.4.5 do not subsume the results of [HS19] and [PP14], respectively. Moreover, Corollary 6.4.5 is also incomparable to the concentration bound of [GV18], whose result requires only *negative regression*, a property weaker than **SCP**.

First, let us define **SCP**. For $S \subseteq [n]$ and $x, y \in \{0,1\}^S$, we say *x covers y*, denoted by $x \rhd y$, if $x = y$ or $x = y + \mathbf{e}_i$ for some i, where $\mathbf{e}_i$ is the unit vector of coordinate i. In other words, x is obtained from y by increasing at most one coordinate. For two distributions μ and ν, we say *μ stochastically covers ν*, if there is a coupling such that $\Pr_{X\sim\mu, Y\sim\nu}(X \rhd Y) = 1$. With slight overload of notation, we also write $\mu \rhd \nu$. A distribution $\mu : \{0,1\}^{[n]} \to \mathbb{R}_{\geq 0}$ satisfies the **SCP** if for any $S \subseteq [n]$ and $x, y \in \{0,1\}^S$ such that $x \rhd y$, $\mu_y \rhd \mu_x$, where $\mu_x : \{0,1\}^{[n]\setminus S} \to \mathbb{R}_{\geq 0}$ is the distribution of μ conditioned on agreeing with x over the index set S.

Furthermore, μ is said to satisfy the *negative cylinder dependence* (**NCD**), if for any $S \subseteq [n]$,

$$\mathbb{E}\prod_{i\in S} X_i \leq \prod_{i\in S} \mathbb{E} X_i,$$

$$\mathbb{E}\prod_{i\in S}(1 - X_i) \leq \prod_{i\in S} \mathbb{E}(1 - X_i),$$

where X_i is the indicator variable of coordinate i. It is known that **SCP** implies **NCD** [PP14]. However, such negative dependence even when $|S| = 2$ is known *not* to hold for the uniform distribution over the bases of some matroids. See [HSW18] for the most comprehensive list of such examples that we are aware of. As the uniform distribution over a matroid's bases is **SLC**, **SLC** does not imply **NCD** or **SCP**.

On the other hand, **SCP** does not imply **SLC** either. We give a concrete example here. Let μ be supported on the bases of the uniform matroid of rank 2 over 4 elements. We choose μ such that

$$\mu(\{1,1,0,0\}) \propto \theta, \qquad \mu(\{1,0,1,0\}) \propto 2, \qquad \mu(\{1,0,0,1\}) \propto 1,$$
$$\mu(\{0,1,1,0\}) \propto 1, \qquad \mu(\{0,1,0,1\}) \propto 1, \qquad \mu(\{0,0,1,1\}) \propto 1.$$

It is straightforward to verify that if $0 \le \theta < 3 - 2\sqrt{2} \approx 0.17157$ or $\theta > 3 + 2\sqrt{2} \approx 5.82843$, then **SLC** fails. However, **SCP** holds as long as $0 \le \theta \le 6$. To see the latter claim, first verify that the distribution conditioned on choosing any $i \in [4]$ stochastically dominates the one conditioned on not choosing i. Then notice that in a homogeneous distribution, such stochastic dominance is the same as stochastic covering.

Here is some insight on how to find an example such as the above. When the generating polynomial g_μ is homogeneous and quadratic, it is **SLC** if and only if it has the **SRP** [BH19], which in turn is equivalent to the following condition as $g_\mu \in \mathbb{R}[x_1, \dots, x_n]$ is multiaffine:

$$\frac{\partial}{x_i} g_\mu(\mathbf{x}) \cdot \frac{\partial}{x_j} g_\mu(\mathbf{x}) \ge g_\mu(\mathbf{x}) \cdot \frac{\partial^2}{\partial x_i \partial x_j} g_\mu(\mathbf{x}), \tag{6.5}$$

for any $i, j \in [n]$ and $\mathbf{x} \in \mathbb{R}^n$. See [Brä07]. If we plug in $\mathbf{x} = \mathbb{1}$, then (6.5) becomes negative dependence for a pair of variables, which is a special case of **NCD** and thus a necessary condition for **SCP**. In our example, we choose μ so that (6.5) holds for $\mathbf{x} = \mathbb{1}$ but not for an arbitrary $\mathbf{x} \in \mathbb{R}^n$. It turns out that our choice is also sufficient for **SCP** in this particular setting.

6.6 Non-Homogeneous Polynomials

As stated in Corollary 6.4.3, strong log-concavity of a homogeneous multiaffine polynomial implies rapid mixing of a certain exchange walk for the corresponding distribution. In this section we will adapt this result to the case of non-homogeneous polynomials.

A common task in sampling problems is to draw from a distribution μ over $\{0,1\}^E$ where E is some ground set. It is equivalent to consider μ as a distribution over all subsets of E. Recall that its generating polynomial is

$$g_\mu(\mathbf{x}) = \sum_{S \subseteq E} \mu(S) \prod_{i \in S} x_i.$$

In general, g_μ is not necessarily homogeneous, and this motivates us to consider the following homogeneous polynomial of degree n, where $|E| = n$,

$$p_\mu(\mathbf{x}, y) = \sum_{S \subseteq E} y^{n-|S|} \mu(S) \prod_{i \in S} x_i.$$

If the degree of every of every variable is at most 1, then we say that the polynomial is *multiaffine*. The polynomial $p_\mu(\mathbf{x}, y)$ is still not the most convenient to work with as it is not multiaffine. We further consider its polarized version,

$$\widehat{p}_\mu(\mathbf{x}, \mathbf{y}) = \sum_{S \subseteq E} \frac{e_{n-|S|}(\mathbf{y})}{\binom{n}{|S|}} \mu(S) \prod_{i \in S} x_i,$$

where $e_k(\mathbf{y}) := \sum_{1 \le i_1 < \cdots < i_k \le n} \prod_{j=1}^{k} y_{i_j}$ is the kth degree elementary symmetric polynomial. In this section we show that polarization preserves log-concavity. Let $Y = \{y_1, \ldots, y_n\}$ and $E^{\text{ext}} = E \cup Y$. The polynomial $\widehat{p}_\mu$ can also be viewed as the generating polynomial of a distribution μ' on $\text{supp}_{\mu'}$ such that

$$\mu'(T) = \frac{\mu(T \cap E)}{\binom{n}{|T \cap Y|}} = \frac{\mu(T \cap E)}{\binom{n}{|T \cap E|}}, \tag{6.6}$$

where $\text{supp}_{\mu'}$ is

$$\text{supp}_{\mu'} := \{T \mid T \subseteq E^{\text{ext}}, |T| = n\}.$$

The projection of μ' on E recovers μ. Namely, for any $S \subseteq E$, $\sum_{T \subseteq E^{\text{ext}} : T \cap E = S} \mu'(T) = \mu(S)$.

Lemma 6.6.1. *Let $p \in \mathbb{R}_{\ge 0}[x_1, \ldots, x_n, y]$ be a d-homogeneous polynomial. If $p(\boldsymbol{x}, y)$ is log-concave at $\mathbb{1}$, then so is*

$$\widehat{p}_\mu(\boldsymbol{x}, \boldsymbol{y}) = \sum_{S \subseteq E} \frac{e_{d-|S|}(\boldsymbol{y})}{\binom{d}{|S|}} \mu(S) \prod_{i \in S} x_i,$$

where $\boldsymbol{y} = \{y_1, \ldots, y_d\}$. (For $d = n$ this applies to the setting discussed previously.)

Proof. Let ∂_i be a shorthand for $\frac{\partial}{\partial x_i}$. An easy calculation yields

$$\partial_i\partial_j(\log p) = \partial_i\left(\frac{\partial_j p}{p}\right) = \frac{p\partial_i\partial_j p - \partial_i p\partial_j p}{p^2}. \tag{6.7}$$

This gives the Hessian $\nabla^2 \log p$ of the logarithm of $p(\mathbf{x})$.

As shown by [ALOV19], a consequence of log-concavity is that, the Hessian of p has at most one positive eigenvalue.

We use (6.7) to compare $\nabla^2 \log p$ with $\nabla^2 \log \widehat{p}$. For all $i, j \in [n]$, let

$$\begin{aligned}
a &= \frac{\partial^2}{\partial y^2}\log p\Big|_{y=1,\mathbf{x}=\mathbb{1}} = \frac{1}{p}\frac{\partial^2 p}{\partial y^2}\Big|_{y=1,\mathbf{x}=\mathbb{1}} - \left(\frac{1}{p}\frac{\partial p}{\partial y}\Big|_{y=1,\mathbf{x}=\mathbb{1}}\right)^2, \\
b_i &= \frac{\partial}{\partial y}\frac{\partial}{\partial x_i}\log p\Big|_{y=1,\mathbf{x}=\mathbb{1}}, \\
C_{ij} &= \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\log p\Big|_{y=1,\mathbf{x}=\mathbb{1}}.
\end{aligned}$$

Then,

$$H := \nabla^2 \log p\Big|_{y=1,\mathbf{x}=\mathbb{1}} = \begin{pmatrix} A & B^{\mathrm{T}} \\ B & C \end{pmatrix},$$

where $A = (a)$, $B = (b_1, \ldots, b_n)^{\mathrm{T}}$, and C is a matrix such that $(C)_{ij} = C_{ij}$. Similarly, we denote

$$\widehat{H} := \nabla^2 \log \widehat{p}\Big|_{\mathbf{y}=\mathbb{1},\mathbf{x}=\mathbb{1}} = \begin{pmatrix} \widehat{A} & \widehat{B}^{\mathrm{T}} \\ \widehat{B} & \widehat{C} \end{pmatrix},$$

where the matrices $\widehat{A}, \widehat{B}, \widehat{C}$ have dimensions $d \times d$, $n \times d$, and $n \times n$ respectively.

A calculation yields that for any $i, j \in [n]$ and any distinct $s, t \in [d]$, $\widehat{p}(\mathbf{x}, \mathbb{1}) = p(\mathbf{x}, 1)$, and

$$\begin{aligned}
\frac{\partial \widehat{p}}{\partial x_i}\Big|_{\mathbf{y}=\mathbb{1},\mathbf{x}=\mathbb{1}} &= \frac{\partial p}{\partial x_i}\Big|_{y=1,\mathbf{x}=\mathbb{1}}, & \frac{\partial \widehat{p}}{\partial y_t}\Big|_{\mathbf{y}=\mathbb{1},\mathbf{x}=\mathbb{1}} &= \frac{1}{d}\frac{\partial p}{\partial y}\Big|_{y=1,\mathbf{x}=\mathbb{1}}, \\
\frac{\partial^2 \widehat{p}}{\partial x_i \partial x_j}\Big|_{\mathbf{y}=\mathbb{1},\mathbf{x}=\mathbb{1}} &= \frac{\partial^2 p}{\partial x_i \partial x_j}\Big|_{y=1,\mathbf{x}=\mathbb{1}}, & \frac{\partial^2 \widehat{p}}{\partial x_i \partial y_s}\Big|_{\mathbf{y}=\mathbb{1},\mathbf{x}=\mathbb{1}} &= \frac{1}{d}\frac{\partial^2 p}{\partial x_i \partial y}\Big|_{y=1,\mathbf{x}=\mathbb{1}}, \\
\frac{\partial^2 \widehat{p}}{\partial y_s^2}\Big|_{\mathbf{y}=\mathbb{1},\mathbf{x}=\mathbb{1}} &= 0, & \frac{\partial^2 \widehat{p}}{\partial y_s \partial y_t}\Big|_{\mathbf{y}=\mathbb{1},\mathbf{x}=\mathbb{1}} &= \frac{1}{d(d-1)}\frac{\partial^2 p}{\partial y^2}\Big|_{y=1,\mathbf{x}=\mathbb{1}}.
\end{aligned}$$

By (6.7), for any $i, j \in [n]$ and any distinct $s, t \in [d]$,

$$\widehat{C}_{i,j} := \frac{\partial^2}{\partial x_i \partial x_j} \log \widehat{p}\Big|_{\mathrm{y}=\mathbb{1},\mathrm{x}=\mathbb{1}} = \frac{\partial^2}{\partial x_i \partial x_j} \log p\Big|_{y=1,\mathrm{x}=\mathbb{1}} = C_{ij}, \tag{6.8}$$

$$\widehat{B}_{i,s} := \frac{\partial^2}{\partial y_s \partial x_i} \log \widehat{p}\Big|_{\mathrm{y}=\mathbb{1},\mathrm{x}=\mathbb{1}} = \frac{1}{d}\frac{\partial^2}{\partial y \partial x_i} \log p\Big|_{y=1,\mathrm{x}=\mathbb{1}} = \frac{b_i}{d}, \tag{6.9}$$

$$\widehat{A}_{s,s} := \frac{\partial^2}{\partial y_s^2} \log \widehat{p}\Big|_{\mathrm{y}=\mathbb{1},\mathrm{x}=\mathbb{1}} = -\frac{1}{d^2}\left(\frac{1}{p}\frac{\partial p}{\partial y}\Big|_{y=1,\mathrm{x}=\mathbb{1}}\right)^2, \tag{6.10}$$

$$\widehat{A}_{s,t} := \frac{\partial^2}{\partial y_s \partial y_t} \log \widehat{p}\Big|_{\mathrm{y}=\mathbb{1},\mathrm{x}=\mathbb{1}} = \frac{1}{d(d-1)}\frac{1}{p}\frac{\partial^2 p}{\partial y^2}\Big|_{y=1,\mathrm{x}=\mathbb{1}} - \frac{1}{d^2}\left(\frac{1}{p}\frac{\partial p}{\partial y}\Big|_{y=1,\mathrm{x}=\mathbb{1}}\right)^2. \tag{6.11}$$

Let $\widehat{\mathbf{z}} = (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^{d+n}$ be a vector, where $\mathbf{u} \in \mathbb{R}^d$ and $\mathbf{v} \in \mathbb{R}^n$. Let $u = \frac{\sum_{i=1}^d u_i}{d}$ and $\mathbf{z} = (u, \mathbf{v}) \in \mathbb{R}^{1+n}$. By our assumption, H is negative semidefinite, and thus $\mathbf{z}^{\mathrm{T}} H \mathbf{z} \le 0$. We claim that $\widehat{\mathbf{z}}^{\mathrm{T}} \widehat{H} \widehat{\mathbf{z}} \le \mathbf{z}^{\mathrm{T}} H \mathbf{z} \le 0$, and therefore $\widehat{H}$ is negatively semidefinite, which implies the lemma.

To verify the claim, we have that

$$\begin{aligned}\widehat{\mathbf{z}}^{\mathrm{T}} \widehat{H} \widehat{\mathbf{z}} &= \mathbf{u}^{\mathrm{T}} \widehat{A} \mathbf{u} + 2\mathbf{v}^{\mathrm{T}} \widehat{B} \mathbf{u} + \mathbf{v}^{\mathrm{T}} \widehat{C} \mathbf{v} \\ &= \mathbf{u}^{\mathrm{T}} \widehat{A} \mathbf{u} + 2u\mathbf{v}^{\mathrm{T}} \mathbf{b} + \mathbf{v}^{\mathrm{T}} C \mathbf{v},\end{aligned}$$

where we use (6.8) and (6.9) in the last line, and

$$\mathbf{z}^{\mathrm{T}} H \mathbf{z} = au^2 + 2u\mathbf{v}^{\mathrm{T}} \mathbf{b} + \mathbf{v}^{\mathrm{T}} C \mathbf{v}.$$

We thus only need to verify that $\mathbf{u}^{\mathrm{T}} \widehat{A} \mathbf{u} \le au^2$. Denote by

$$P := \frac{1}{p}\frac{\partial^2 p}{\partial y^2}\Big|_{y=1,\mathrm{x}=\mathbb{1}} \qquad\qquad Q := \left(\frac{1}{p}\frac{\partial p}{\partial y}\Big|_{y=1,\mathrm{x}=\mathbb{1}}\right)^2.$$

Then $a = P - Q$, $\widehat{A}_{s,s} = -\frac{Q}{d^2}$, and $\widehat{A}_{s,t} = \frac{P}{d(d-1)} - \frac{Q}{d^2}$ by (6.10) and (6.11). Thus,

$$\begin{aligned}au^2 - \mathbf{u}^{\mathrm{T}} \widehat{A} \mathbf{u} &= \frac{P-Q}{d^2}\left(\sum_{s=1}^d u_s\right)^2 + \frac{Q}{d^2}\left(\sum_{s=1}^d u_s\right)^2 - \frac{P}{d(d-1)} \sum_{s \ne t \in [d]} u_s u_t \\ &= \frac{P}{d^2}\left(\left(\sum_{s=1}^d u_s\right)^2 - \frac{d}{d-1} \sum_{s \ne t \in [d]} u_s u_t\right) \\ &= \frac{P}{d^2}\left(\frac{d}{d-1}\sum_{s=1}^d u_s^2 - \frac{1}{d-1}\left(\sum_{s=1}^d u_s\right)^2\right) \ge 0,\end{aligned}$$

where in the last line we used the fact that $P \ge 0$ (as $p \in \mathbb{R}_{\ge 0}[x_1, \ldots, x_n, y]$) and the Cauchy-Schwarz inequality. □

In particular, notice that polynomial $\frac{\partial^{k+k'}\widehat{p}}{\partial x_{i_1}\ldots\partial x_{i_k}\partial y_{s_1}\ldots\partial y_{s_{k'}}}(\mathbf{x},\mathbf{y})$ is the polarized form of $\frac{\partial^{k+k'}p}{\partial x_{i_1}\ldots\partial x_{i_k}\partial y^{k'}}(\mathbf{x},y)$ up to a factor of $\frac{(d-k')!}{d!}$. This implies the following corollary.

Corollary 6.6.2. *Let $p \in \mathbb{R}_{\geq 0}[x_1,\ldots,x_n,y]$ be a d-homogeneous polynomial. If $p(\boldsymbol{x},y)$ is strongly log-concave, then $\widehat{p}(\boldsymbol{x},\boldsymbol{y})$ is also strongly log-concave.*

On the other hand, let $q(\mathbf{x}) = \sum_{k=0}^{d} q_{d-k}(\mathbf{x})$ be a polynomial of degree d where $q_{d-k}(\mathbf{x})$ is homogeneous of degree $d-k$. The log-concavity of $q(\mathbf{x})$ does not imply that its homogenisation $p(\mathbf{x},y) = \sum_{k=0}^{d} y^k q_{d-k}(\mathbf{x})$ is also log-concave.

6.6.1 The Exchange Walk

For a homogeneous distribution μ over subsets of E, let $T_t \subset E$ be the current state of the exchange walk. The next step is:

1. choose uniformly at random an element $a \in T_t$ and let $T^- = T_t \setminus \{a\}$;
2. move to $T^- \cup \{b\}$ (where $b \notin T^-$) with probability $\dfrac{\mu'(T^- \cup \{b\})}{\sum_{c \notin T^-} \mu'(T^- \cup \{c\})}$.

Let the transition matrix of this chain be Q_{ex}. Then for $T,T' \in \operatorname{supp}_{\mu'}$ (recall (6.6)),

$$Q_{\text{ex}}(T,T') = \begin{cases} \frac{1}{n} \cdot \frac{\mu'(T')}{\sum_{c \notin T \cap T'} \mu'(T \cap T' \cup \{c\})} & \text{if } |T \cap T'| = n-1; \\ 1 - \sum_{R:|T\cap R|=n-1} \frac{1}{n} \cdot \frac{\mu'(R)}{\sum_{c \notin T \cap R} \mu'(T \cap R \cup \{c\})} & \text{if } T' = T; \\ 0 & \text{otherwise.} \end{cases}$$

When μ is not homogeneous, we will apply the exchange walk on μ', so the chain will move among subsets of E^{ext}. We are particularly interested in the projection chain of Q_{ex} on E, denoted by P_{ex}.

Following [JSTV04], the projection chain P_{ex} is defined by

$$\begin{aligned} P_{\text{ex}}(S,S') &= \frac{1}{\mu(S)} \sum_{T,T':T\cap E=S, T'\cap E=S'} \mu'(T) Q_{\text{ex}}(T,T') \\ &= \frac{1}{\binom{n}{|S|}} \sum_{T,T':T\cap E=S, T'\cap E=S'} Q_{\text{ex}}(T,T'), \end{aligned}$$

for $S,S' \subseteq E$, since $\mu(S) = \sum_{T:T\cap E=S} \mu'(T)$. We now work out the transitions of P_{ex} with respect to μ. Clearly, $P_{\text{ex}}(S,S') > 0$ if and only if $|S \oplus S'| \leq 2$. There are four cases:

- If $|S \oplus S'| = 2$, then it must be that $S' = S \setminus \{a\} \cup \{b\}$ for $a \neq b \in E$. In this case,

$$\begin{aligned} P_{\text{ex}}(S,S') &= \frac{1}{\binom{n}{|S|}} \sum_{T:T\cap E=S} \sum_{T':T'\cap E=S'} Q_{\text{ex}}(T,T') \\ &= \frac{1}{\binom{n}{|S|}} \sum_{T:T\cap E=S} Q_{\text{ex}}(T, T\setminus\{a\}\cup\{b\}) \\ &= \frac{1}{n} \cdot \frac{\mu(S')}{(n-|S|+1)\mu(S\setminus\{a\}) + \sum_{c\in E:c\notin S\setminus\{a\}} \mu(S\setminus\{a\}\cup\{c\})}. \end{aligned}$$

- If $S' = S\setminus\{a\}$ where $a \in S$,

$$\begin{aligned} P_{\text{ex}}(S,S') &= \frac{1}{\binom{n}{|S|}} \sum_{T:T\cap E=S} \sum_{T':T'\cap E=S'} Q_{\text{ex}}(T,T') \\ &= \frac{1}{\binom{n}{|S|}} \sum_{T:T\cap E=S} \sum_{y:y\notin E, y\notin T} Q_{\text{ex}}(T, T\setminus\{a\}\cup\{y\}) \\ &= \frac{1}{n} \cdot \frac{(n-|S'|)\mu(S')}{(n-|S'|)\mu(S') + \sum_{c\in E:c\notin S'} \mu(S'\cup\{c\})}. \end{aligned}$$

- If $S' = S\cup\{a\}$ where $a \notin S$,

$$\begin{aligned} P_{\text{ex}}(S,S') &= \frac{1}{\binom{n}{|S|}} \sum_{T:T\cap E=S} \sum_{T':T'\cap E=S'} Q_{\text{ex}}(T,T') \\ &= \frac{1}{\binom{n}{|S|}} \sum_{T:T\cap E=S} \sum_{y:y\notin E, y\in T} Q_{\text{ex}}(T, T\setminus\{y\}\cup\{a\}) \\ &= \frac{1}{n} \cdot \frac{(n-|S|)\mu(S')}{(n-|S|)\mu(S) + \sum_{c\in E:c\notin S} \mu(S\cup\{c\})}. \end{aligned}$$

- If $S' = S$,

$$P_{\text{ex}}(S,S') = 1 - \sum_{U:1\leq|S\oplus U|\leq 2} P_{\text{ex}}(S,U).$$

It is straightforward to verify the detailed balance condition with respect to μ,

$$\mu(S)P_{\text{ex}}(S,S') = \mu(S')P_{\text{ex}}(S',S).$$

This means that P_{ex} is reversible.

Proposition 6.6.3. *If the homogenized polynomial p_μ, of a distribution μ over subsets of E, is strongly log-concave, then the walks Q_{ex} and P_{ex} have a modified log-Sobolev constant which is at least $1/|E|$.*

Proof. By Lemma 6.6.1 and Theorem 6.4.2, the proposition follows for Q_{ex}. Given also that P_{ex} is a projection of Q_{ex}, it necessarily satisfies its corresponding functional inequalities with the same constants as Q_{ex}. □

6.6.2 The Random-Cluster Model

The Random-Cluster model is a graphical model, defined over a graph $G = (V, E)$, for which we have that, for $S \subseteq E$,

$$\mu(S) \propto p^{|S|}(1-p)^{|E \setminus S|} q^{k(S)},$$

where $q > 0$, $p \in (0, 1)$, and $k(S)$ stands for the number of connected components of the graph (V, S). Although the Random-Cluster model is well-defined for general $q > 0$, we will focus on the case where $0 < q < 1$. In this case, the homogenized generating polynomial of the distribution μ is strongly log-concave [BH18].

By Proposition 6.6.3, we get that, for this model, $\rho(P_{\text{ex}}) \geq 1/|E|$. This in turn implies that the mixing time of P_{ex} is upper bounded by $O(|E| \log |V|)$. A more standard chain for the Random-Cluster model is the so-called *Gibbs Sampling* (which is the subject of the next section) for which we get a $O(|E|^2 \log |V|)$ mixing time bound. As far as we know, there were no established polynomial mixing time bounds for these chains prior to this work. Some relevant results on estimating the partition function of the Random-Cluster model (for $0 < q < 1$) can be found in [ALOV19]. There, the authors obtain a polynomial-time approximation algorithm by estimating the contribution to the partition function of edge sets of each fixed size separately. Compared to their method, our approach is more direct and efficient, as we deal with all edge set sizes together. It is also known that the Glauber dynamics mixes rapidly for $q = 2$ [GJ17], while the region where $q > 2$ is believed to be intractable [GJ12].

6.6.3 Comparison with Gibbs Sampling

Gibbs Sampling, or the "heat-bath" chain, refers to the following Markov chain. Let $S_0 \subseteq E$ be the initial set, and S_t be the set at time t. At time $t+1$,

1. draw uniformly at random an element $x \in E$;
2. Let $S_t' = S_t \setminus x$. Move to S_t' with probability $\frac{\mu(S_t')}{\mu(S_t' \cup \{x\}) + \mu(S_t')}$, and to $S_t' \cup \{x\}$ otherwise.

The $2^n \times 2^n$ transition matrix P_{HB} is given as follows:

$$P_{\text{HB}}(S, S') = \begin{cases} \frac{1}{n} \cdot \frac{\mu(S')}{\mu(S)+\mu(S')} & \text{if } |S \oplus S'| = 1; \\ \sum_{R:|S \oplus R|=1} \frac{1}{n} \cdot \frac{\mu(S)}{\mu(S)+\mu(R)} & \text{if } S' = S; \\ 0 & \text{otherwise.} \end{cases}$$

It is standard to verify that P_{HB} converges to the distribution μ.

We compute upper-bounds for the ratio of $P_{\text{ex}}(S,S')$ and $P_{\text{HB}}^2(S,S')$ for every transition (S,S'). There are several cases:

- If $S' = S \setminus \{a\} \cup \{b\}$ for $a \in S$ and $b \notin S$,

$$\begin{aligned} P_{\text{ex}}(S,S') &= \frac{1}{n} \cdot \frac{\mu(S')}{(n-|S|+1)\mu(S\setminus\{a\}) + \sum_{c \notin S\setminus\{a\}} \mu(S\setminus\{a\}\cup\{c\})} \\ &\leq \frac{1}{n} \cdot \frac{\mu(S')}{\mu(S\setminus\{a\}) + \mu(S')}. \end{aligned}$$

and

$$\begin{aligned} P_{\text{HB}}^2(S,S') &= \left(\frac{1}{n} \cdot \frac{\mu(S\setminus\{a\})}{\mu(S)+\mu(S\setminus\{a\})}\right)\left(\frac{1}{n} \cdot \frac{\mu(S')}{\mu(S\setminus\{a\})+\mu(S')}\right) \\ &\quad + \left(\frac{1}{n} \cdot \frac{\mu(S\cup\{b\})}{\mu(S)+\mu(S\cup\{b\})}\right)\left(\frac{1}{n} \cdot \frac{\mu(S')}{\mu(S\cup\{b\})+\mu(S')}\right) \\ &\geq \frac{1}{n^2}\left(\frac{\mu(S\setminus\{a\})}{\mu(S)+\mu(S\setminus\{a\})}\right)\left(\frac{\mu(S')}{\mu(S\setminus\{a\})+\mu(S')}\right). \end{aligned}$$

Thus,

$$\frac{P_{\text{ex}}(S,S')}{P_{\text{HB}}^2(S,S')} \leq n\left(1 + \frac{\mu(S)}{\mu(S\setminus\{a\})}\right).$$

- If $S' = S \setminus \{a\}$ where $a \in S$,

$$\begin{aligned} P_{\text{ex}}(S,S') &= \frac{1}{n} \cdot \frac{(n-|S'|)\mu(S')}{(n-|S'|)\mu(S') + \sum_{c \notin S'} \mu(S'\cup\{c\})} \\ &\leq \frac{1}{n} \cdot \frac{(n-|S'|)\mu(S')}{\mu(S')+\mu(S)}, \end{aligned}$$

and

$$\begin{aligned} P_{\text{HB}}^2(S,S') &= \frac{1}{n} \cdot \frac{\mu(S')}{\mu(S)+\mu(S')} \cdot \left[\sum_{R:|S\oplus R|=1} \frac{1}{n} \cdot \frac{\mu(S)}{\mu(S)+\mu(R)} \right. \\ &\qquad \left. + \sum_{R:|S'\oplus R|=1} \frac{1}{n} \cdot \frac{\mu(S')}{\mu(S')+\mu(R)}\right] \\ &\geq \frac{1}{n^2} \cdot \frac{\mu(S')}{\mu(S)+\mu(S')} \cdot \left[\frac{\mu(S)}{\mu(S)+\mu(S')} + \frac{\mu(S')}{\mu(S')+\mu(S)}\right] \\ &= \frac{1}{n^2} \cdot \frac{\mu(S')}{\mu(S)+\mu(S')}. \end{aligned}$$

Thus,

$$\frac{P_{\text{ex}}(S,S')}{P_{\text{HB}}^2(S,S')} \leq n(n-|S'|) \leq n^2.$$

- If $S' = S \cup \{a\}$ where $a \notin S$, the computation is similar to the previous case and it yields,

$$\frac{P_{\text{ex}}(S,S')}{P^2_{\text{HB}}(S,S')} \le n(n-|S|) \le n^2.$$

- If $S' = S \cup \{a,b\}$ where $a,b \notin S$, or $S' = S \setminus \{a,b\}$ where $a,b \in S$, then

$$\frac{P_{\text{ex}}(S,S')}{P^2_{\text{HB}}(S,S')} = 0.$$

- If $S' = S$, we don't need to compute a bound as the corresponding terms of the Dirichlet forms are equal to zero.

The above calculations mean that we can compare the Dirichlet forms of P_{ex} and P^2_{HB}. Specifically, we get

$$\mathscr{E}_{P_{\text{ex}}}(f,\log f) \le n \max\left\{n, 1 + \max_S \frac{\mu(S)}{\mu(S\setminus\{a\})}\right\} \mathscr{E}_{P^2_{\text{HB}}}(f,\log f).$$

The Random-Cluster Model. We revisit the above calculations specifically for the Random-Cluster model, which is defined in Section 6.6.2. In this case, we can exploit the following inequalities,

$$l := \frac{p}{1-p} \le \frac{\mu(S)}{\mu(S\setminus\{a\})} \le \frac{1}{q}\frac{p}{1-p} =: u.$$

- If $S' = S \setminus \{a\} \cup \{b\}$ for $a \ne b \in E$,

$$\frac{P_{\text{ex}}(S,S')}{P^2_{\text{HB}}(S,S')} \le n\left(1 + \frac{\mu(S)}{\mu(S\setminus\{a\})}\right) \le n(1+u).$$

- If $S' = S \setminus \{a\}$ where $a \in S$,

$$\begin{aligned} P_{\text{ex}}(S,S') &= \frac{1}{n} \cdot \frac{(n-|S'|)\mu(S')}{(n-|S'|)\mu(S') + \sum_{c\notin S'} \mu(S' \cup \{c\})} \\ &\le \frac{1}{n} \cdot \frac{(n-|S'|)\mu(S')}{(n-|S'|)\mu(S') + (n-|S'|)\cdot l \cdot \mu(S')} \\ &= \frac{1}{n} \cdot \frac{1}{1+l}, \end{aligned}$$

and

$$P^2_{\mathrm{HB}}(S,S') = \frac{1}{n}\cdot\frac{\mu(S')}{\mu(S)+\mu(S')}\cdot\Bigg[\sum_{R:|S\oplus R|=1}\frac{1}{n}\cdot\frac{\mu(S)}{\mu(S)+\mu(R)} + \sum_{R:|S'\oplus R|=1}\frac{1}{n}\cdot\frac{\mu(S')}{\mu(S')+\mu(R)}\Bigg]$$
$$\geq \frac{1}{n^2}\cdot\frac{\mu(S')}{u\cdot\mu(S')+\mu(S')}\cdot\Bigg[|S|\frac{\mu(S)}{\mu(S)+\mu(S)/l}+(n-|S|)\frac{\mu(S)}{\mu(S)+u\cdot\mu(S)} + |S'|\frac{\mu(S')}{\mu(S')+\mu(S')/l}+(n-|S'|)\frac{\mu(S')}{\mu(S')+u\cdot\mu(S')}\Bigg]$$
$$\geq \frac{1}{n}\cdot\frac{2}{1+u}\min\left(\frac{l}{1+l},\frac{1}{1+u}\right).$$

Thus,

$$\frac{P_{\mathrm{ex}}(S,S')}{P^2_{\mathrm{HB}}(S,S')} \leq \frac{1+u}{2}\max\left(\frac{1}{l},\frac{1+u}{1+l}\right).$$

- If $S' = S\cup\{a\}$ where $a\notin S$, the computation is similar to the previous case and it yields,

$$\frac{P_{\mathrm{ex}}(S,S')}{P^2_{\mathrm{HB}}(S,S')} \leq \frac{u}{2l}\max\left(\frac{1+l}{l},1+u\right).$$

The above bounds imply that, for large enough n,

$$\mathcal{E}_{P_{\mathrm{ex}}}(f,\log f) \leq n(1+u)\mathcal{E}_{P^2_{\mathrm{HB}}}(f,\log f),$$

which in turn gives us the following relationship between the modified log-Sobolev constants of the two walks,

$$\rho(P_{\mathrm{ex}}) \leq n(1+u)\rho(P^2_{\mathrm{HB}}).$$

By Proposition 6.6.3, we have that $\rho(P_{\mathrm{ex}}) \geq 1/|E|$ (where $n=|E|$), and thus

$$\rho(P^2_{\mathrm{HB}}) \geq \frac{1}{|E|^2(1+u)},$$

which yields a $O(|E|^2\log|V|)$ mixing time bound for P_{HB}.

Chapter 7

The Broken Circuit Complex

Historically, there has been a persistent effort towards the characterization of the coefficients of the chromatic polynomial of a graph, a central object of algebraic graph theory. An important realization towards this direction was made by [Whi32], who proved that the absolute value of each of its coefficients corresponds to the count of certain edge sets without "broken circuits". Later on, [Wil73] noticed that these edge sets form a downwards closed family which can be viewed as a simplicial complex, the *broken circuit complex*; a notion that was generalized to matroids by [Bry77]. The number of faces of each level of this complex corresponds to one coefficient of the chromatic polynomial. Given that the possible face vectors of all simplicial complexes are fully characterized, this provided a partial characterization of the possible coefficients of the chromatic polynomial. This was later refined by a full characterization of Cohen-Macaulay complexes, a family which includes the broken circuit complexes [Sta07]. Moreover, it was shown that the number of faces of the broken circuit complex of a graph equals the number of its acyclic orientations [Sta73], the evaluation of which is a #P-hard problem of unknown approximability [GJ08]. Around the 1970s, some conjectures were expressed regarding the unimodality and log-concavity of the coefficients of the chromatic polynomial of a graph and the characteristic polynomial of a matroid, which proved true by the work of [AHK18], further refining our knowledge of these polynomials.

Around the time of the resolution of these conjectures another series of works took place, which concerned the properties of random walks over simplicial complexes. This provided a highly effective framework of proving rapid or optimal mixing of various exchange walks, such as the basis-exchange walk of matroids [ALOV19] and Glauber dynamics [CLV21a]. In particular, the basis-exchange walk of matroids is a special case of the top dimensional exchange walk of the broken circuit complex [Bry77]. Such mixing results, when established, can allow for the approximation of the size of the state space of a Markov chain, for example the number of bases of a matroid or – in the case of the broken circuit complex – a coefficient of the chromatic polynomial.

With this background in mind, we are interested in studying the mixing properties of the exchange walks over the broken circuit complex. As a starting point, we only consider the lowest dimensional interesting case, which is derived from matroids of rank 3.

Let $\mathscr{M} = (E, \mathscr{B})$ be a matroid, where $\mathscr{B}$ is the set of bases. Alternatively, $\mathscr{M}$ can be defined by its circuits, denoted by $\mathscr{C}$, or its independent sets, denoted by $\mathscr{I}$. The following circuit definition [Oxl92, Theorem 1.1.4] is particularly useful.

Definition 7.0.1. *A matroid $\mathscr{M}$ over a ground set E is defined by a set $\mathscr{C} \subseteq 2^E$ (circuits) that satisfies the following axioms:*

(C1) $\emptyset \notin \mathscr{C}$;

(C2) if C_1 and C_2 are members of $\mathscr{C}$ and $C_1 \subseteq C_2$, then $C_1 = C_2$;

(C3) if C_1 and C_2 are distinct members of $\mathscr{C}$ and $e \in C_1 \cap C_2$, then there is a member C_3 of $\mathscr{C}$ such that $C_3 \subseteq (C_1 \cup C_2) \setminus e$.

Item (C3) is called the *circuit elimination axiom*.

The ground set E can be considered to be the set $[n] := \{0, 1, \ldots, n\}$, which has the ordering $0 < 1 < \cdots < n$. A *broken circuit* is a set $C \subseteq E$ for which there exists $i \in E$ such that $\{i\} \cup C \in \mathscr{C}$ and i is smaller than all elements in C. The set of *broken circuit bases*, denoted by $\mathscr{BC}$, is the subset of bases $\mathscr{B}$ such that $B \in \mathscr{BC}$ does not contain any broken circuit. For a matroid of rank r, it is known that the basis-exchange walk over all bases mixes in $O(r \log r)$ steps [CGM21, ALO$^+$21b]. Here we want to show that the basis-exchange walk is still rapid mixing even when restricted to $\mathscr{BC}$.

The basis-exchange walk P_{BX} is the following. Suppose the current basis is B. Then,

- uniformly at random select an element $b \in B$ and remove it to get $B' = B \setminus b$; and
- uniformly at random select an element b' among those for which $B' \cup \{b'\} \in \mathscr{B}$.

To restrict to $\mathscr{BC}$, we replace $\mathscr{B}$ in the second step by $\mathscr{BC}$. Denote the restricted walk by P_{BC}. The basis-exchange walk P_{BX} converges to the uniform distribution over $\mathscr{B}$, and P_{BC} converges to the uniform distribution over $\mathscr{BC}$. The key to analyzing P_{BX} is to view it as the down-up walk of the matroid's independence complex at the top level r. We will take the same viewpoint for P_{BC}, as independent sets with no broken circuits are also downwards closed.

Let $\mathscr{BCI}$ denote the set of independent sets that do not contain a broken circuit. From now on we assume that $\mathscr{M}$ is loopless, in order for $\mathscr{BCI}$ to be nonempty. Let the *broken circuit complex* be the pure simplicial complex defined by $\mathscr{BCI}$. For a detailed introduction to the broken circuit complex we refer the reader to [Bjö92]. The broken circuit complex is *shellable* [see, e.g., Theorem 7.4.3 of Bjö92]. This means that there exists a linear ordering of its facets such that the complex can be constructed by adding its facets (along with their faces) one by one so that each facet – when added – intersects the complex constructed up to that point in a pure complex which is of one dimension lower. Furthermore, the links and skeletons of a shellable complex are also shellable [see, e.g., 10.12 and 10.14 of BW97]. These facts imply that all exchange walks over the broken circuit complex are irreducible (for example, one may notice by induction that there always exists a sequence of exchanges connecting the first facet in the linear ordering to every other facet).

The link at 0 of the broken circuit complex is the same as its subcomplex over $[n]\setminus\{0\}$. That is because for any $I \in \mathscr{BCI}$ that doesn't contain 0, we necessarily have that $I\cup\{0\} \in \mathscr{BCI}$ (otherwise the set I would contain a broken circuit), and similarly for any $I \in \mathscr{BCI}$ that does contain 0, we trivially have that $I\setminus\{0\} \in \mathscr{BCI}$. Thus, we may restrict our attention to the link at 0 of the broken circuit complex, which is called the *reduced broken circuit complex* (RBC complex).

In the case of a matroid $\mathscr{M}$, the link at an independent set I of its independence complex corresponds to the independence complex of $\mathscr{M}_I$, which is the matroid $\mathscr{M}$ with I contracted. However, there is no analogous result for the broken circuit or RBC complex.

7.1 A Decomposition of the Exchange Graph

Definition 7.1.1. *A graph $G=(V,E)$ admits a* star decomposition *if there exist $A_1,\dots,A_k \subseteq V$ such that*

1. $|A_i| \geq 2$;

2. *for any $u,v \in V$, $\{u,v\}$ is a subset of* exactly one A_i *for some $1 \leq i \leq k$; and*

3. *each* $G[A_i]$ *is a star.*

The *flats* of a matroid are subsets of the ground set for which the addition of any element would increase the cardinality of a largest independent set contained therein.

The *simplification* of a matroid is the matroid that results from removing loops and retaining only one representative from each of the equivalence classes formed by parallel elements (see Proposition 6.0.2). For the purposes of analyzing the broken circuit complex, we assume that the simplification always retains the smallest element of each class.

Lemma 7.1.2. *The 1-skeleton of the broken circuit complex of a matroid* $\mathscr{M}$*, a graph denoted by* $G_{BC} = (V_{BC}, E_{BC})$*, admits a star decomposition* $\{F_i\}_{i=1}^k$*, where* $\{F_i\}_{i=1}^k$ *is the set of flats of rank* 2 *of the simplification of* $\mathscr{M}$*.*

Proof. We shall assume that the matroid $\mathscr{M}$ is simple without loss of generality. This is because the broken circuit complex ignores any parallel elements and only retains the smallest of each equivalence class.

Firstly, condition (1) of Definition 7.1.1 is trivially satisfied because the flats are of rank 2. For a simple matroid, the flats of rank ≤ 1 have size equal to their rank. This implies that the intersection of two flats F_i and F_j of rank 2 (which is itself a flat of rank less than 2) satisfies $|F_i \cap F_j| \leq 1$. Also, every pair of elements appears in (at least) one flat of rank 2. This establishes condition (2).

In G_{BC}, $\{u,v\}$ is a non-edge if and only if there exists a circuit $\{t,u,v\}$ where $t < u,v$. Thus, the smallest element, 0, is connected to all other elements.

The smallest element u of F_i is necessarily connected in G_{BC} to every other element v in F_i. Otherwise, there would exist a $t < u,v$, $t \notin F_i$, such that $\{t,u,v\} \in \mathscr{C}$ which contradicts with the fact that F_i is closed $(t \in cl(\{u,v\}) \subseteq cl(F_i) = F_i \not\ni t)$. Moreover, three elements from F_i form a circuit as F_i is rank 2 and there is no circuit of rank 2. This implies that any other pair of elements in F_i is disconnected as it is a broken circuit (it forms a circuit with the smallest element of F_i). This establishes condition (3) of Definition 7.1.1. □

We remark that a more general decomposition result holds for the exchange graph of $P_{r-2}^{\wedge}$ (over the broken circuit complex) for *paving* matroids of arbitrary rank. A matroid is paving if each of its circuits has size r or $r+1$. By deleting element 0, the exchange graph of $P_{r-2}^{\wedge}$ corresponds to the walk $P_{r-2}^{\wedge}$ over the reduced broken circuit complex (whose spectral gap can be related to the spectral gap of top level down-up walk of the reduced broken circuit complex). Let $\mathscr{M} = (E, \mathscr{C})$ be a paving matroid of rank r. Then, its flats of rank $r-1$ (hyperplanes) form an $(r-1)$-partition. Let $F_1, \ldots, F_k$ be an enumeration of the hyperplanes. For any F_i, any size r subset of F_i is a circuit. Conversely, if a set of size r is not a subset of any F_i, then it is a basis. Moreover, the broken circuits of $\mathscr{M}$ are either size $r-1$ subsets of some F_i which do not contain $\min F_i$, or bases that do not contain 0. The exchange graph of $P^{\wedge}_{r-2}$, similar to Lemma 7.1.2, has a decomposition given by the flats $\{F_i\}_{i=1}^{k}$. The vertices of the exchange graph of $P^{\wedge}_{r-2}$ are the $(r-2)$-subsets of E. The $(r-2)$-subsets of each F_i form a "generalized star", where two distinct $(r-2)$-subsets are connected by an edge if they differ by exactly one element and if at least one of them contains $\min F_i$. Furthermore, this describes all edges, as every $(r-1)$-subset appears in exactly one of the hyperplanes. If the $(r-1)$-subset contains $\min F_i$ then it gives rise to $\frac{(r-1)(r-2)}{2}$ edges, otherwise it is a broken circuit and it produces no edges.

7.2 Expansion for Matroids of Rank 3

In this section we prove constant expansion for graphs that admit a star decomposition (Definition 7.1.1). This implies expansion of the 1-skeleton of the broken circuit complex. In order to also prove expansion for the 1-skeleton of the reduced broken circuit complex, we will employ a slightly more delicate result which can handle the removal of the special vertex 0, a vertex which is connected to all other vertices. Finally, this can be related to the RBC-basis-exchange walk P_{RBC} of a matroid of rank 3 and provide a lower bound on its spectral gap.

For a graph $G = (V, E)$ and $S \subseteq V$, let $e(S) := |E(S)|$ and $e(S, S^c) := |E(S, S^c)|$.

Lemma 7.2.1. *Let $G = (V, E)$ be a graph with a star decomposition $\{A_i\}_{i \in [s]}$ and $S \subseteq V$. Then,*

$$e(S, S^c) \geq \frac{3\sqrt{17} - 5}{8} \min\{e(S), e(S^c)\},$$

where $\frac{3\sqrt{17}-5}{8} \approx 0.921.$

Proof. Let $p := \frac{|S^c|^2}{|S|^2 + |S^c|^2}$ and consider

$$\Delta := \left(3\sqrt{p(1-p)} + 1 - \frac{3}{4}\min\{p, 1-p\}\right) e(S, S^c) - 2pe(S) - 2(1-p)e(S^c).$$

For each $i \in [s]$, let $a_i := |A_i \cap S|$ and $b_i := |A_i \cap S^c|$. Note that if the center of A_i is in S, then its contribution to Δ is

$$\left(3\sqrt{p(1-p)}+1-\frac{3}{4}\min\{p,1-p\}\right)b_i-2p(a_i-1),$$

whereas if the center of A_i is in S^c, then its contribution to Δ is

$$\left(3\sqrt{p(1-p)}+1-\frac{3}{4}\min\{p,1-p\}\right)a_i-2(1-p)(b_i-1).$$

Thus,

$$\begin{aligned}\Delta \geq \sum_{i\in[s]}\min\{&\left(3\sqrt{p(1-p)}+1-\frac{3}{4}\min\{p,1-p\}\right)b_i-2p(a_i-1),\\ &\left(3\sqrt{p(1-p)}+1-\frac{3}{4}\min\{p,1-p\}\right)a_i-2(1-p)(b_i-1)\}. \qquad (7.1)\end{aligned}$$

Now, notice that

$$\begin{aligned}&\left(3\sqrt{p(1-p)}+1-\frac{3}{4}\min\{p,1-p\}\right)b_i-2p(a_i-1)\\ &\qquad\qquad -2\left(\sqrt{p(1-p)}a_ib_i-p\binom{a_i}{2}-(1-p)\binom{b_i}{2}\right)\\ &=\left(\sqrt{p}a_i-\sqrt{1-p}b_i-\sqrt{p}\right)\left(\sqrt{p}a_i-\sqrt{1-p}b_i-2\sqrt{p}\right)+pb_i-\frac{3}{4}\min\{p,1-p\}b_i\\ &\geq\left(\sqrt{p}a_i-\sqrt{1-p}b_i-\sqrt{p}\right)\left(\sqrt{p}a_i-\sqrt{1-p}b_i-2\sqrt{p}\right)+\frac{pb_i}{4}\geq 0,\end{aligned}$$

where the last inequality is due to the following: if $b_i = 0$, the last expression simplifies into $p(a_i-1)(a_i-2) \geq 0$ as a_i is an integer; if $b_i \geq 1$, it has the form $x(x-\sqrt{p})+pb_i/4 \geq -p/4+pb_i/4 \geq 0$ where x is $\sqrt{p}a_i-\sqrt{1-p}b_i-\sqrt{p}$.

Similarly,

$$\begin{aligned}&\left(3\sqrt{p(1-p)}+1-\frac{3}{4}\min\{p,1-p\}\right)a_i-2(1-p)(b_i-1)\\ &\qquad\qquad -2\left(\sqrt{p(1-p)}a_ib_i-p\binom{a_i}{2}-(1-p)\binom{b_i}{2}\right)\\ &\geq\left(\sqrt{1-p}b_i-\sqrt{p}a_i-\sqrt{1-p}\right)\left(\sqrt{1-p}b_i-\sqrt{p}a_i-2\sqrt{1-p}\right)+\frac{(1-p)a_i}{4}\geq 0.\end{aligned}$$

By (7.1) and the two inequalities above,

$$\Delta \geq \sum_{i\in[s]} 2\left(\sqrt{p(1-p)}a_ib_i-p\binom{a_i}{2}-(1-p)\binom{b_i}{2}\right).$$

Then, by double counting pairs in S, S^c, or $S \times S^c$ and property (2) in Definition 7.1.1, we have

$$\binom{|S|}{2} = \sum_{i\in[s]} \binom{a_i}{2}, \tag{7.2}$$

$$|S|\,|S^c| = \sum_{i\in[s]} a_i b_i. \tag{7.3}$$

$$\binom{|S^c|}{2} = \sum_{i\in[s]} \binom{b_i}{2}, \tag{7.4}$$

Thus, by (7.2), (7.3), (7.4),

$$\Delta \geq 2\sqrt{p(1-p)}\,|S|\,|S^c| - 2p\binom{|S|}{2} - 2(1-p)\binom{|S^c|}{2} = \frac{|S|\,|S^c|\,(|S|+|S^c|)}{|S|^2+|S^c|^2} \geq 0. \tag{7.5}$$

On the other hand,

$$\begin{aligned}\Delta &= \left(3\sqrt{p(1-p)}+1-\frac{3}{4}\min\{p,1-p\}\right)e(S,S^c) - 2pe(S) - 2(1-p)e(S^c)\\ &= \frac{|S|^2+|S^c|^2+3|S||S^c|-\frac{3}{4}\min\{|S|^2,|S^c|^2\}}{|S|^2+|S^c|^2}e(S,S^c)\\ &\qquad -2\frac{|S^c|^2}{|S|^2+|S^c|^2}e(S) - 2\frac{|S|^2}{|S|^2+|S^c|^2}e(S^c),\end{aligned}$$

which implies that (taking also (7.5) into account)

$$e(S,S^c) \geq 2\frac{|S^c|^2 e(S) + |S|^2 e(S^c)}{|S|^2+|S^c|^2+3|S||S^c|-\frac{3}{4}\min\{|S|^2,|S^c|^2\}},$$

and

$$\begin{aligned}\frac{e(S,S^c)}{\min\{e(S),e(S^c)\}} &\geq 2\frac{|S^c|^2+|S|^2}{|S|^2+|S^c|^2+3|S||S^c|-\frac{3}{4}\min\{|S|^2,|S^c|^2\}}\\ &= \frac{2}{1+3\frac{|S||S^c|-\frac{1}{4}\min\{|S|^2,|S^c|^2\}}{|S|^2+|S^c|^2}} \geq \frac{3\sqrt{17}-5}{8},\end{aligned}$$

as $\frac{xy-\frac{1}{4}\min\{x^2,y^2\}}{x^2+y^2} \leq \frac{1}{8}(\sqrt{17}-1)$ for $x,y \geq 0$. The lemma follows. □

Lemma 7.2.2. *Let $G = (V \cup \{0\}, E)$ be a graph with a star decomposition $\{A_i\}_{i\in S}$ and a special vertex 0, which is the center of all stars that it is in. Let $U \subseteq V$, $U^c = V \setminus U$, and $p := \frac{|U^c|^2}{|U|^2+|U^c|^2}$. Then,*

$$e(U,U^c) \geq \frac{3\sqrt{17}-5}{8}\min\{e(U),e(U^c)\}.$$

Proof. For each $i \in S$, let $a_i := |A_i \cap U|$ and $b_i := |A_i \cap U^c|$. Let $S_0 \subset S$ be the set of indices of all stars involving 0. Then we have that

$$\sum_{i \in S_0} a_i = |U| \qquad \sum_{i \in S_0} b_i = |U^c|. \tag{7.6}$$

Similar to Lemma 7.2.1, consider

$$\Delta := \left(3\sqrt{p(1-p)} + 1 - \frac{3}{4}\min\{p, 1-p\}\right) e(U, U^c) - 2pe(U) - 2(1-p)e(U^c).$$

Note that, if $i \in S_0$, then because the center of A_i is 0, A_i does not contribute anything to Δ. On the other hand, for any $i \in S - S_0$, if the center of A_i is in U, then its contribution to Δ is

$$\left(3\sqrt{p(1-p)} + 1 - \frac{3}{4}\min\{p, 1-p\}\right) b_i - 2p(a_i - 1),$$

whereas if the center of A_i is in U^c, then its contribution to Δ is

$$\left(3\sqrt{p(1-p)} + 1 - \frac{3}{4}\min\{p, 1-p\}\right) a_i - 2(1-p)(b_i - 1).$$

Thus,

$$\begin{aligned}\Delta \geq \sum_{i \in S - S_0} \min\{ &\left(3\sqrt{p(1-p)} + 1 - \frac{3}{4}\min\{p, 1-p\}\right) b_i - 2p(a_i - 1), \\ &\left(3\sqrt{p(1-p)} + 1 - \frac{3}{4}\min\{p, 1-p\}\right) a_i - 2(1-p)(b_i - 1)\}.\end{aligned} \tag{7.7}$$

Similar to the proof of Lemma 7.2.1, both expressions inside the min are greater than $2\left(\sqrt{p(1-p)}a_ib_i - p\binom{a_i}{2} - (1-p)\binom{b_i}{2}\right)$. Thus, by (7.7),

$$\Delta \geq \sum_{i \in S - S_0} 2\left(\sqrt{p(1-p)}a_ib_i - p\binom{a_i}{2} - (1-p)\binom{b_i}{2}\right).$$

Then, by double counting pairs of vertices (not just the edges) in U, U^c, or $U \times U^c$,and by property (2) in Definition 7.1.1, we have

$$\begin{aligned}\binom{|U|}{2} &= \sum_{i \in S - S_0} \binom{a_i}{2} + \sum_{i \in S_0} \binom{a_i}{2}, \\ \binom{|U^c|}{2} &= \sum_{i \in S - S_0} \binom{b_i}{2} + \sum_{i \in S_0} \binom{b_i}{2}, \\ |U||U^c| &= \sum_{i \in S - S_0} a_ib_i + \sum_{i \in S_0} a_ib_i.\end{aligned} \tag{7.8}$$

Thus,

$$\begin{aligned}
\frac{\Delta}{2} &\geq \sqrt{p(1-p)}\,|U|\,|U^c| - p\binom{|U|}{2} - (1-p)\binom{|U^c|}{2} && \text{(by (7.8))} \\
&\quad - \left(\sqrt{p(1-p)}\sum_{i\in S_0} a_i b_i - p\sum_{i\in S_0}\binom{a_i}{2} - (1-p)\sum_{i\in S_0}\binom{b_i}{2}\right) \\
&= \frac{p}{2}|U| + \frac{1-p}{2}|U^c| + \frac{1}{2}\sum_{i\in S_0}\left(\sqrt{p}a_i - \sqrt{1-p}b_i\right)^2 - p\sum_{i\in S_0}\frac{a_i}{2} - (1-p)\sum_{i\in S_0}\frac{b_i}{2} \\
&\geq \frac{p}{2}|U| + \frac{1-p}{2}|U^c| - \frac{p}{2}|U| - \frac{1-p}{2}|U^c| = 0. && \text{(by (7.6))}
\end{aligned}$$

Having proven that $\Delta \geq 0$, the rest of the proof is similar to that of Lemma 7.2.1. □

Corollary 7.2.3. *Let $G_{RBC} = (V_{RBC}, E_{RBC})$ denote the 1-skeleton of the reduced broken circuit complex. Let $S \subset V_{RBC}$ with $|E_{RBC}[S]| \leq |E_{RBC}|/2$, and let ∂S be the set of edges between S and $V \setminus S$. Then,*

$$|\partial S| \geq \frac{3\sqrt{17}-5}{8}\,|E_{RBC}[S]|\,.$$

Proof. The graph G_{RBC} is the same as the graph G_{BC} (which admits a star decomposition by Lemma 7.1.2) but with the vertex 0 missing. The result then follows immediately from Lemma 7.2.2. □

Corollary 7.2.4. *For a matroid of rank 3, the spectral gap of the down-up walk over the facets of the reduced broken circuit complex, P_{RBC}, is at least $\frac{1}{2} - \frac{1}{4}\sqrt{78\sqrt{17}-318} \approx 0.0255$.*

Proof. Let P be the simple random walk over G_{RBC}, with $P(x,y) = 1/\deg(x)$ if $(x,y) \in E_{RBC}$ and $P(x,y) = 0$ otherwise. The stationary distribution of this walk is $\pi(x) = \deg(x)/(2|E_{RBC}|)$. Then, the conductance of P is

$$\begin{aligned}
h(P) &:= \min_{S:\pi(S)\leq 1/2} \frac{\sum_{x\in S, y\in S^c}\pi(x)P(x,y)}{\pi(S)} \\
&= \min_{S:\pi(S)\leq 1/2} \frac{|\partial S|}{|\partial S| + 2|E_{RBC}[S]|} \\
&\geq \frac{13-3\sqrt{17}}{2} \approx 0.315. && \text{(by Corollary 7.2.3)}
\end{aligned}$$

Note that for the RBC complex of rank 2 with uniform weights over its facets, $P_1^{\wedge} = (P+I)/2$ and

$$\begin{aligned}
\lambda(P_{\mathrm{RBC}}) = \lambda(P_1^{\wedge}) &= \frac{1}{2}\lambda(P) && \text{(by Corollary 3.3.2)} \\
&\geq \frac{1}{2}\left(1 - \sqrt{1-h(P)^2}\right) && \text{(by Theorem 2.3 in [Chu97])} \\
&\geq \frac{1}{2} - \frac{1}{4}\sqrt{78\sqrt{17}-318}.
\end{aligned}$$

□

We conjecture that the optimal expansion constant for Lemma 7.2.2 is 1, while the optimal lower bound for the spectral gap in Corollary 7.2.4 is 1/4 (or 1/2 for the non-lazy walk). These bounds, if true, would be tight. A family of RBC exchange graphs that achieves these is of the form $(K_n \cup K_n) + E_{n-1}$, for $n \geq 2$, where E_{n-1} is the graph on $n-1$ vertices with no edges. For example, for $n = 2$, we get the “bow tie” graph. These graphs are indeed the exchange graph of an RBC complex, as can be seen by the following construction. Let E_{n-1} have the $n-1$ smallest elements (after 0). The two non-adjacent cliques have n^2 non-edges between them. Each element of $[n-1]$ (including 0) removes n of these edges by forming a circuit with them (n edge-disjoint matchings of size n between the two K_n's). Lastly, the elements of $[n-1]$ are all disconnected by 0 (one large star). For odd n, if one picks the elements of one K_n along with half of the elements of E_{n-1} to be the set S, then we indeed have that $|\partial S| = |E_{RBC}[S]|$.

Interestingly, the spectral gap of the simple walks over these graphs seems to be among the worst possible for a 1-dimensional RBC complex, with a value of 1/2 (this value is 1 for the independence complex of a matroid in the worst case). This has been computationally verified for all matroids of rank 3 and up to 9 elements. Further computation becomes very time-consuming because all possible orderings of the ground set must be checked.

Unlike the base case of matroid independence complexes, which corresponds to rank 2 matroids and to complete multipartite graphs as the exchange graphs, the case of broken circuit complexes seems much more challenging to analyze. A further discouraging fact is that even if these bounds that we are aiming for are proved, and although they would be the optimal possible, they are still too weak for the Trickling-Down Theorem (4.1.1) and for the local-to-global theorems that we have at our disposal. This means that we don't have a method to go from 1-dimensional reduced broken circuit complexes to a result for higher dimensional ones (unlike the case of matroid independence complexes). Nevertheless, proving the optimal constants for this problem in its low-dimensional case may give rise to new ideas, and thus the resolution of that would be a good step forwards.

Chapter 8

Open Problems

Entropy Contraction Trickling-Down and other Local-to-Global Theorems. Is there a way to lower bound the entropy contraction parameter in terms of the entropy contraction parameters of its links? An immediate attempt towards this, inspired by the proof of the Trickling-Down Theorem (4.1.1), does not seem to work. Also, can we prove other interesting Local-to-Global theorems? For example, based on the proof of Lemma 4.2.1, and towards a Local-to-Global log-Sobolev constant, one can show that

$$\begin{aligned}
\mathscr{E}_{P_k^\wedge}\left(\sqrt{f^{(k)}}, \sqrt{f^{(k)}}\right) &= \frac{k}{k+1} \sum_{S \in \mathscr{C}(k-1)} \pi_{k-1}(S) \mathscr{E}_{G_S}\left(\sqrt{f_S}, \sqrt{f_S}\right) \\
&\geq \frac{k}{k+1} \alpha_{k-1} \sum_{S \in \mathscr{C}(k-1)} \pi_{k-1}(S) \mathrm{Ent}_{\pi_{S,1}}\left(f_S^{(1)}\right) \\
&\qquad\qquad \text{(assuming } \alpha(G_S) \geq \alpha_{k-1}\text{)} \\
&= \frac{k}{k+1} \alpha_{k-1} \left(\mathrm{Ent}_{\pi_k}\left(f^{(k)}\right) - \mathrm{Ent}_{\pi_{k-1}}\left(f^{(k-1)}\right)\right) \\
&\qquad\qquad \text{(by Lemma 5.1.1)} \\
&\geq \frac{k}{k+1} \alpha_{k-1} \mathscr{E}_{P_k^\vee}\left(\sqrt{f^{(k)}}, \sqrt{f^{(k)}}\right). \\
&\qquad\qquad \text{(Proof of Lemma 5.9 in [CLV21a])}
\end{aligned}$$

Thus, $\alpha(P_k^\wedge) \geq \frac{k}{k+1}\alpha_{k-1}\alpha(P_k^\vee)$. Is there a strong continuation of this argument, by relating $\alpha(P_k^\wedge)$ and $\alpha(P_{k-1}^\wedge)$?

Matrix Trickling-Down Theorem and Matrix Local-to-Global. Is there a more canonical version of the Matrix Trickling-Down Theorem [ALO21a]? i.e., one that provides a general and versatile method for proving second eigenvalue bounds such as the Trickling-Down Theorem (4.1.1). Given also that the local-to-global theorems, such as the Alev-Lau Local-to-Global Theorem (4.2.2), use the worst case second eigenvalue bound for the links of each level, does there exist a Matrix Local-to-Global Theorem that works in a more averaging fashion on the influence of the links?

Perfect Sampling of Matroid Bases. Is there an efficient perfect sampling procedure for uniform matroid bases or, more generally, for strongly log-concave distributions? Could coupling from the past and ideas like the ones in [Hub98] achieve this? Can we extend the Aldous-Broder algorithm [Ald90, Bro89] or similar algorithms to regular matroids?

Broken Circuit Complex. Can the result on the expansion of the reduced broken circuit complex of a rank 3 matroid (Lemma 7.2.2) be improved to constant 1? Can good expansion or mixing time results (similar to the ones that we have for matroid independence complexes) be proved for the higher-dimensional exchange walks? A step towards this would be to investigate the case of paving matroids, as there is a decomposition into induced subgraphs that one may exploit, similar to the star decomposition of the low-dimensional case (Lemma 7.1.2). A good handle of these cases may provide the insight to tackle the general case of arbitrary matroids. This is of interest as it would, for example, give us an efficient way to approximately count the number of acyclic orientations of a graph.

Greedoids. In light of the log-concave inequalities proved by [CP21], and given that the work of [AHK18] preceded the results we presented for matroids (e.g., functional inequalities, mixing times), are there analogues of these results for antimatroids or, more generally, for other classes of greedoids? One such example is the vertex search greedoid.

0-1 polytopes. Is the Mihail-Vazirani Conjecture (6.1.1) true for all 0-1 polytopes?

Bibliography

[AASV21] Yeganeh Alimohammadi, Nima Anari, Kirankumar Shiragur, and Thuy-Duong Vuong. Fractionally log-concave and sector-stable polynomials: counting planar matchings and more. In *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 433–446. ACM, 2021.

[AHK18] Karim Adiprasito, June Huh, and Eric Katz. Hodge theory for combinatorial geometries. *Ann. of Math. (2)*, 188(2):381–452, 2018.

[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. *CoRR*, abs/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. ACM, 2020.

[Ald90] David J. Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discret. Math.*, 3(4):450–465, 1990.

[ALO20] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 1319–1330. IEEE, 2020.

[ALO21a] Dorna Abdolazimi, Kuikui Liu, and Shayan Oveis Gharan. A matrix trickle-down theorem on simplicial complexes and applications to sampling colorings. In *62nd IEEE Annual Symposium on Foundations of Com-*

puter Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022, pages 161–172. IEEE, 2021.

[ALO+21b] 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 *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 408–420. ACM, 2021.

[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.

[AOV18] Nima Anari, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids. In *FOCS*, pages 35–46, 2018. Available at abs/1807.00929.

[BCC+21] Antonio Blanca, Pietro Caputo, Zongchen Chen, Daniel Parisi, Daniel Stefankovic, and Eric Vigoda. On mixing of Markov chains: Coupling, spectral independence, and entropy factorization. *CoRR*, abs/2103.07459, 2021.

[BG99] Sergey G. Bobkov and Friedrich Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.

[BH18] Petter Brändén and June Huh. Hodge-Riemann relations for Potts model partition functions. *arXiv preprint arXiv:1811.01696*, 2018.

[BH19] Petter Brändén and June Huh. Lorentzian polynomials. *arXiv*, abs/1902.03719, 2019.

[BHT06] Sergey G. Bobkov, Christian Houdré, and Prasad Tetali. The subgaussian constant and concentration inequalities. *Israel J. Math.*, 156:255–283, 2006.

[Bjö92] Anders Björner. The homology and shellability of matroids and geometric lattices. *Matroid applications*, 40:226–283, 1992.

[BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

[Brä07] Petter Brändén. Polynomials with the half-plane property and matroid theory. *Adv. Math.*, 216(1):302–320, 2007.

[Bro89] Andrei Z. Broder. Generating random spanning trees. In *30th Annual Symposium on Foundations of Computer Science, Research Triangle Park, North Carolina, USA, 30 October - 1 November 1989*, pages 442–447. IEEE Computer Society, 1989.

[Bry77] Tom Brylawski. The broken-circuit complex. *Trans. Amer. Math. Soc.*, 234(2):417–433, 1977.

[BT06] Sergey G. Bobkov and Prasad Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.*, 19(2):289–336, 2006.

[BW97] Anders Björner and Michelle L. Wachs. Shellable nonpure complexes and posets. II. *Transactions of the American Mathematical Society*, pages 3945–3975, 1997.

[CE22] Yuansi Chen and Ronen Eldan. Localization schemes: A framework for proving mixing bounds for Markov chains. *CoRR*, abs/2203.04163, 2022.

[CFYZ21] Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Rapid mixing of Glauber dynamics via spectral independence for all degrees. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 137–148. IEEE, 2021.

[CFYZ22] Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Optimal mixing for two-state anti-ferromagnetic spin systems. *CoRR*, abs/2203.07771, 2022.

[CGM21] Mary Cryan, Heng Guo, and Giorgos Mousa. Modified log-Sobolev inequalities for strongly log-concave distributions. *Ann. Probab.*, 49(1):506–525, 2021.

[CGSV21] Zongchen Chen, Andreas Galanis, Daniel Stefankovic, and Eric Vigoda. Rapid mixing for colorings via spectral independence. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, pages 1548–1557. SIAM, 2021.

[Chu97] Fan R. K. Chung. *Spectral graph theory*. American Mathematical Society, 1997.

[CLV21a] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of Glauber dynamics: entropy factorization via high-dimensional expansion. In *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 1537–1550. ACM, 2021.

[CLV21b] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Spectral independence via stability and applications to holant-type problems. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 149–160. IEEE, 2021.

[CP21] Swee Hong Chan and Igor Pak. Log-concave poset inequalities. *CoRR*, abs/2110.10740, 2021.

[DK17] Irit Dinur and Tali Kaufman. High dimensional expanders imply agreement expanders. In *FOCS*, pages 974–985. IEEE Computer Society, 2017.

[DR98] Devdatt Dubhashi and Desh Ranjan. Balls and bins: a study in negative dependence. *Random Structures Algorithms*, 13(2):99–124, 1998.

[DSC96] Persi Diaconis and Laurent Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 1996.

[FGYZ21] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Rapid mixing from spectral independence beyond the Boolean domain. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, pages 1558–1577. SIAM, 2021.

[GGMS87] Israel M. Gelfand, R. Mark Goresky, Robert D. MacPherson, and Vera V. Serganova. Combinatorial geometries, convex polyhedra, and Schubert cells. *Advances in Mathematics*, 63(3):301–316, 1987.

[GJ08] Leslie Ann Goldberg and Mark Jerrum. Inapproximability of the Tutte polynomial. *Information and Computation*, 206(7):908–929, 2008.

[GJ12] Leslie Ann Goldberg and Mark Jerrum. Approximating the partition function of the ferromagnetic potts model. *J. ACM*, 59(5):25:1–25:31, 2012.

[GJ17] Heng Guo and Mark Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, January 16-19*, pages 1818–1827. SIAM, 2017.

[GM20] Heng Guo and Giorgos Mousa. Local-to-global contraction in simplicial complexes. *CoRR*, abs/2012.14317, 2020.

[Goe04] Sharad Goel. Modified logarithmic Sobolev inequalities for some models of random walk. *Stochastic Process. Appl.*, 114(1):51–79, 2004.

[Gur09a] Leonid Gurvits. On multivariate Newton-like inequalities. In *Advances in combinatorial mathematics*, pages 61–78. Springer, Berlin, 2009.

[Gur09b] Leonid Gurvits. A polynomial-time algorithm to approximate the mixed volume within a simply exponential factor. *Discrete Comput. Geom.*, 41(4):533–555, 2009.

[GV18] Kevin Garbe and Jan Vondrák. Concentration of Lipschitz functions of negatively dependent variables. *arXiv*, abs/1804.10084, 2018.

[HJ12] Roger A. Horn and Charles R. Johnson. *Matrix Analysis, 2nd Ed.* Cambridge University Press, 2012.

[HS05] Thomas P. Hayes and Alistair Sinclair. A general lower bound for mixing of single-site dynamics on graphs. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2005), 23-25 October 2005, Pittsburgh, PA, USA, Proceedings*, pages 511–520. IEEE Computer Society, 2005.

[HS19] Jonathan Hermon and Justin Salez. Modified log-Sobolev inequalities for strong-Rayleigh measures. *arXiv*, abs/1902.02775, 2019.

[HSW18] June Huh, Benjamin Schröter, and Botong Wang. Correlation bounds for fields and matroids. *arXiv*, abs/1806.02675, 2018.

[Hub98] Mark Huber. Exact sampling and approximate counting techniques. In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 31–40. ACM, 1998.

[Jer03] Mark Jerrum. *Counting, sampling and integrating: algorithms and complexity*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003. Chapter 9 (unpublished) is available at `http://www.maths.qmul.ac.uk/~mj/ETHbook/chapter9.pdf`.

[JSTV04] Mark Jerrum, Jung-Bae Son, Prasad Tetali, and Eric Vigoda. Elementary bounds on Poincaré and log-Sobolev constants for decomposable Markov chains. *Ann. Appl. Probab.*, 14(4):1741–1765, 2004.

[KM17] Tali Kaufman and David Mass. High dimensional random walks and colorful expansion. In *8th Innovations in Theoretical Computer Science Conference, ITCS 2017, January 9-11, 2017, Berkeley, CA, USA*, volume 67 of *LIPIcs*, pages 4:1–4:27. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[KO18] Tali Kaufman and Izhar Oppenheim. High order random walks: Beyond spectral gap. In *APPROX-RANDOM*, volume 116 of *LIPIcs*, pages 47:1–47:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

[Liu21] Kuikui Liu. From coupling to spectral independence and blackbox comparison with the down-up walk. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2021, August 16-18, 2021, University of Washington, Seattle, Washington, USA (Virtual Conference)*, volume 207 of *LIPIcs*, pages 32:1–32:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[LP17] David A. Levin and Yuval Peres. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017. Second edition.

[Mih92] Milena Mihail. On the expansion of combinatorial polytopes. In *MFCS*, volume 629 of *Lecture Notes in Computer Science*, pages 37–49, 1992.

[MV89] Milena Mihail and Umesh Vazirani. On the magnification of 0-1 polytopes. Technical Report TR-03-89, Harvard University, 1989.

[Opp18] Izhar Oppenheim. Local spectral expansion approach to high dimensional expanders part I: Descent of spectral gaps. *Discret. Comput. Geom.*, 59(2):293–330, 2018.

[Oxl92] James G. Oxley. *Matroid theory*. Oxford University Press, 1992.

[PP14] Robin Pemantle and Yuval Peres. Concentration of Lipschitz functionals of determinantal and other strong Rayleigh measures. *Combin. Probab. Comput.*, 23(1):140–160, 2014.

[Sam05] Marcus D. Sammer. *Aspects of mass transportation in discrete concentration inequalities*. 2005. Thesis (Ph.D.)–Georgia Institute of Technology.

[Sta73] Richard P Stanley. Acyclic orientations of graphs. *Discrete Mathematics*, 5(2):171–178, 1973.

[Sta07] Richard P Stanley. *Combinatorics and commutative algebra*, volume 41. Springer Science & Business Media, 2007.

[Whi32] Hassler Whitney. A logical expansion in mathematics. *Bulletin of the American Mathematical Society*, 38(8):572–579, 1932.

[Wil73] Herbert Wilf. Which polynomials are chromatic? In *Proc. Colloq. Combinatorial Theory, Rome*, volume 352, 1973.