Applications of Log Concave Polynomials

Arpon Basu

Last updated December 18, 2023

Contents

Notation

Let $n \in \mathbb{N} = \{1, 2, \ldots\}$. Then we refer to the set $\{1, 2, \ldots, n\}$ as $[n]$.

Given any set X, we define $\binom{X}{k} := \{S \subseteq X : |S| = k\}$. We also define $2^X := \{S : S \subseteq X\}$ to be the power-set of X.

For any $S \subseteq [n]$, $\mathbb{1}_S \in \mathbb{R}^n$ is the indicator vector of S, where $(\mathbb{1}_S)_i = 1$ if $i \in S$, and 0 otherwise.

For any $v \in \mathbb{R}^n$ and any $S \subseteq [n]$, we define $v^S := \prod_{i \in S} v_i$.

Unless specified otherwise, all logarithms are assumed to be in base e.

For any two vectors $v, w \in \mathbb{R}^n_{\geq 0}$, we define $v^w := \begin{bmatrix} v_1^{w_1} & \ldots & v_n^{w_n} \end{bmatrix}^\top \in \mathbb{R}^n_{\geq 0}$. We define $0^0 := 1$.

A polynomial $g \in \mathbb{R}[z_1, \ldots, z_n]$ is called d-homogenous if all non-zero monomials in g are of degree d.

A polynomial $g \in \mathbb{R}[z_1, \ldots, z_n]$ is called multilinear if the degree of any variable in g is at most 1. For example, $3xyz - 5y$ is a multilinear polynomial, while $z^2 + 2$ is not.

The differential operator $\frac{\partial}{\partial x}$ is denoted as ∂_x. In case we have indexed variables $x_1, \ldots, x_n$, we abbreviate ∂_{x_i} as ∂_i.

Let $v \in \mathbb{R}^n$. We abbreviate $\sum_{i=1}^n v_i \partial_i$ as ∂_v. We assure the reader that it will be clear from the context if v is a variable or a vector. If f is smooth, then for any $v, w \in \mathbb{R}^n$, $\partial_v \partial_w f = \partial_w \partial_v f$.

Observe that if f is homogenous (resp. multilinear), so is $\partial_v f$ for any $v \in \mathbb{R}^n$.

Consider $\alpha \in \mathbb{Z}^n_{\geq 0}$. Then we define $\partial^\alpha := \prod_{i=1}^n \partial_i^{\alpha_i}$. Also, we define $|\alpha| = \sum_{i=1}^n \alpha_i$.

Let Ω be some non-empty open subset of $\mathbb{R}^n$, and let $h : \Omega \mapsto \mathbb{R}$ be a smooth function. The gradient of f is a $n \times 1$ vector which is denoted as ∇f, where $(\nabla f)_j := \partial_j f$. The *Hessian* of h is a $n \times n$ matrix which is denoted as $\nabla^2 h =: H$, where $H_{ij} := \partial_i \partial_j h$. The Hessian of smooth functions is symmetric.

We will quite often be dealing with logarithms of continuous functions in this survey. Consequently, we'll have an issue with $\log f$ whenever $f = 0$. We remedy this by working with the *extended real line* $\mathbb{R} \cup \{-\infty\}$.

For any $S \subseteq [n]$, the indicator vector $\mathbb{1}_S \in \mathbb{R}^n$ is defined such that $(\mathbb{1}_S)_i = 1$ if $i \in S$, and 0 otherwise.

For any $V \subseteq \mathbb{R}^n$,

$$\operatorname{conv}(V) := \bigcap_{\substack{\mathbb{R}^n \supseteq S \supseteq V \\ S \text{ is convex}}} S$$

denotes the convex hull of V. If V is finite, then $\operatorname{conv}(V)$ is a polytope.

Let $\mu : 2^{[n]} \mapsto [0, 1]$ be a probability distribution on $2^{[n]}$, i.e. $\sum_{S \in 2^{[n]}} \mu(S) = 1$. Let $i \in [n]$ be arbitrary. We define $\mu|_i$ to be the distribution μ conditioned on i, i.e. $\mu|_i$ is a distribution on $2^{[n] \setminus \{i\}}$ such that for any $S \subseteq [n] \setminus \{i\}$, $\mu|_i(S) := \frac{\mu(S \cup \{i\})}{\sum_{S' \subseteq [n] \setminus \{i\}} \mu(S' \cup \{i\})}$. Similarly, $\mu|_{\bar{i}}$ is also a distribution on $2^{[n] \setminus \{i\}}$ such that for any $S \subseteq [n] \setminus \{i\}$, $\mu|_{\bar{i}}(S) := \frac{\mu(S)}{\sum_{S' \subseteq [n] \setminus \{i\}} \mu(S')}$. We say that $\mu|_i$ is "μ conditioned in i", while $\mu|_{\bar{i}}$ is "μ conditioned out i".

Given $\nu \in \mathbb{R}^n_{>0}$, we define the inner product w.r.t ν as $\langle v, w \rangle_\nu := \sum_{i=1}^n \nu_i v_i w_i$.

Acknowledgements

This survey is a personal account of the famous Log-Concave series of papers [AOV18, ALOV19, ALOV18]. Many of the ideas in these papers were also (parallelly) developed by Huh, Adiprasito, Brändén, Katz, Wang and others [AHK18, HW17, HK11, BH22], although we will only focus on the former papers. If the reader wishes for a more expansive introduction to some topics of this survey, then Oveis Gharan's lecture notes [Ove20] are an excellent place to begin. Some more context to the work here can be found in Anari's Simons talk on the same [Ana19], and Gil Kalai's Fields Medal laudatio of June Huh's work [Kal22].

§1. Log Concave Polynomials

We introduce the notion of log-concavity, which forms a bridge between the combinatorial and analytic worlds.

Definition 1.1 (Log-Concavity)**.** A function $g : \mathbb{R}^n_{\geq 0} \mapsto \mathbb{R}_{\geq 0}$ is said to be log-concave if for any $v, w \in \mathbb{R}^n_{\geq 0}, \lambda \in [0, 1]$, $g(\lambda v + (1-\lambda)w) \geq g(v)^\lambda g(w)^{1-\lambda}$. Evidently, the zero function is log-concave.
It is easy to see that if $f : \mathbb{R}^n_{>0} \mapsto \mathbb{R}_{>0}$ is a log-concave function, then $\log f$ is a concave function. Thus if $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$ are such that $a + \lambda b \in \mathbb{R}^n_{>0}$ for every $\lambda \in [0, 1]$, then $\log f(a + \lambda b) \geq \log f(a) + \lambda \log f(b)$ for every $\lambda \in [0, 1]$.
Throughout this survey, we shall only talk of the log-concavity of functions whose domains are a subset of $\mathbb{R}^n_{\geq 0}$.

Remark. Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a continuous function that is concave on some set $U \subseteq \mathbb{R}^n$. Let $\overline{U}$ be the closure of U, and let $x, y \in \overline{U}$ be arbitrary members of $\overline{U}$. Then there exist sequences $\{x_k\}_{k\in\mathbb{N}}, \{y_k\}_{k\in\mathbb{N}}$ in U such that $\lim_{k\to\infty} x_k = x, \lim_{k\to\infty} y_k = y$. Let $\lambda \in [0, 1]$ be some arbitrary number. Then

$$f(\lambda x + (1-\lambda)y) = f\left(\lim_{k\to\infty}(\lambda x_k + (1-\lambda)y_k)\right) \overset{\text{continuity of } f}{=} \lim_{k\to\infty} \underbrace{f(\lambda x_k + (1-\lambda)y_k)}_{\geq \lambda f(x_k) + (1-\lambda)f(y_k)} \geq \lim_{k\to\infty} \lambda f(x_k) + (1-\lambda)f(y_k)$$
$$= \lambda f(x) + (1-\lambda)f(y)$$

Thus f is concave on $\overline{U}$ too.
Consequently, when we want to prove the log-concavity of some function g on $\mathbb{R}^n_{\geq 0}$, we'll just prove the log-concavity of g on $\mathbb{R}^n_{>0}$. This is valid since $\overline{\mathbb{R}^n_{>0}} = \mathbb{R}^n_{\geq 0}$.

We shall mostly be interested in the log-concavity of polynomials with non-negative coefficients. To characterize such polynomials, we describe some *closure properties* which allow us to generate log-concave polynomials from known log-concave polynomials.

Proposition 1 (Closure Properties)**.** Let $p(z_1, \ldots, z_n), q(z_1, \ldots, z_n)$ be log-concave polynomials with non-negative coefficients. Then the following polynomials are log-concave, and have non-negative coefficients:

1. *Affine Transformations*: $p(T(y_1, \ldots, y_m))$, where $T : \mathbb{R}^m \mapsto \mathbb{R}^n : y \mapsto Ay + b$, where $A \in \mathbb{R}^{n \times m}_{\geq 0}, b \in \mathbb{R}^n_{\geq 0}$.
2. *Permutation*: $p(z_{\pi(1)}, \ldots, z_{\pi(n)})$ for any permutation π of $[n]$
3. *External Field*: $cp(c_1 z_1, c_2 z_2, \ldots, c_n z_n)$ for scalars $c, c_1, \ldots, c_n \in \mathbb{R}_{\geq 0}$.
4. *Specialization*: $p(a, z_2, \ldots, z_n) = p(z_1, z_2, \ldots, z_n)\big|_{z_1 = a}$, for $a \in \mathbb{R}_{\geq 0}$.
5. *Product*: $r(z_1, \ldots, z_n) := p(z_1, \ldots, z_n) q(z_1, \ldots, z_n)$

Proof. For the first part, note that $p(T(\lambda v + (1-\lambda)w)) = p(\lambda T(v) + (1-\lambda)T(w)) \geq p(T(v))^\lambda p(T(w))^{1-\lambda}$, where the last inequality follows from the log-concavity of p. Parts 2, 3, 4 follow when one realizes that permutations, scalings and specializations are non-negative linear transformations (for part 3, one further notes that if f is log-concave, then cf is also log-concave for any $c \geq 0$). Finally, for the last part, $r(\lambda v + (1-\lambda)w) = p(\lambda v + (1-\lambda)w) q(\lambda v + (1-\lambda)w) \geq p(v)^\lambda p(w)^{1-\lambda} q(v)^\lambda q(w)^{1-\lambda} = r(v)^\lambda r(w)^{1-\lambda}$. ■

Note that log-concave polynomials are *not* closed under differentiation: $p(z) = \frac{z^4}{4} + z$ is log-concave, yet $q := \partial_z p = z^3 + 1$ is not. Indeed, recall that if f is a smooth concave function, then $\partial_z^2 f \leq 0$ everywhere. Now,

$$\partial_z^2 \log p = \frac{-4(z^3-2)^2}{z^2(z^3+4)^2} \leq 0, \partial_z^2 \log q\big|_{z=1} = \frac{-3z(z^3-2)}{(z^3+1)^2}\Big|_{z=1} = \frac{3}{4} \not\leq 0$$

At this point, we mention a "topological closure" property of log-concave polynomials.

Lemma 1.1. Fix any $d \in \mathbb{N}$. The set of log-concave polynomials of degree $\leq d$ is closed under the topology of pointwise convergence.

Proof. Suppose we have a sequence of polynomials $p_1, p_2, \ldots$ such that $\lim_{k \to \infty} p_k = p$, where the limit of functions is pointwise. Note that since we're dealing with bounded degree polynomials over a field of characteristic 0, pointwise convergence actually means that the coefficients of the polynomials converge. Also, note that the negative semi-definiteness of any matrix is equivalent to a system of polynomial constraints [1]. In particular, the log-concavity of p_k is equivalent to the Hessian of $\log p_k$ being NSD, which further is equivalent to a system of rational function inequalities on the coefficients of p_k [2], i.e. inequalities of the form $f_1(a_k) \geq 0, \ldots, f_r(a_k) \geq 0$, where a_k is the vector of all coefficients of p_k, and $f_1, \ldots, f_r$ are rational functions. Since $f_i(a_k) \geq 0$ for all $i \in [r], k \in \mathbb{N}$, $f_i(a) \geq 0$ [3], where $a = \lim_{k \to \infty} a_k$ is the vector of coefficients of p. Consequently, the p is also log-concave, as desired. ■

[1] The PSDness of a matrix can be enforced by saying that all the principal minors of the matrix must be non-negative. Note that the principal minors of a matrix are polynomials in the entries of the matrix.

[2] Note that the entries of $\nabla^2 \log p_k$ are rational functions in the coefficients of p_k

[3] since rational functions are continuous

Definition 1.2. A smooth function $f : \mathbb{R}^n_{\geq 0} \mapsto \mathbb{R}_{\geq 0}$ is said to be log-concave at $z = a$ if $\nabla^2 \log f|_{z=a} = \left(\nabla^2 \log f\right)(a)$ is a negative semi-definite matrix.

Clearly, a smooth function f is log-concave on $\mathbb{R}^n_{\geq 0}$ if it is log-concave everywhere on $\mathbb{R}^n_{\geq 0}$.

Note that if we want to investigate the log-concavity of f, then we have to check the negative-semi-definiteness of $\nabla^2 \log f$, which is a bit clumsy. Ideally, we would want to characterize the log-concavity of f through $\nabla^2 f$ itself. This is precisely what we shall do now.

Theorem 1.2. Let $f \in \mathbb{R}[z_1, \ldots, z_n]$ be a d-homogenous polynomial with non-negative coefficients, where $d \geq 2$. Let $a \in \mathbb{R}^n_{\geq 0}$ be any point such that $f(a) \neq 0 \iff f(a) > 0$. Define $Q := \nabla^2 f\big|_{z=a}$. Then the following are equivalent:

1. f is log-concave at $z = a$.
2. $z^\top Q z \leq 0$ for every $z \in (Qa)^\perp$.
3. $z^\top Q z \leq 0$ for every $z \in (Qb)^\perp$, where b is *any* vector such that $Qb \neq 0$.
4. $z^\top Q z \leq 0$ for every z in some $(n-1)$-dimensional vector space.
5. $(a^\top Q a)Q - (Qa)(Qa)^\top$ is negative semi-definite.
6. For $d \geq 3$, parts $1 \ldots 5$ and 7 are equivalent to: $\partial_a f = \sum_{i \in [n]} a_i \partial_{z_i} f$ is log-concave at $z = a$.
7. If $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of Q, then $\lambda_1 \geq 0 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

Proof. Applying Lemma A.4 on $\partial_j f$ for every $j \in [n]$ yields $Qa = (d-1) \cdot (\nabla f)(a)$, and then using Lemma A.4 on f yields $a^\top Q a = d(d-1) f(a) \implies a^\top Q a > 0$. Now,

$$\nabla^2 \log f\Big|_{z=a} = \left(\frac{f \cdot \nabla^2 f - \nabla f (\nabla f)^\top}{f^2}\right)\Bigg|_{z=a} = \underbrace{\frac{d(d-1)}{(a^\top Q a)^2}\left(a^\top Q a \cdot Q - \frac{d}{d-1}(Qa)(Qa)^\top\right)}_{=:\mathfrak{M}}$$

$(1 \implies 2)$: Since f is log-concave at $z = a$, then $\nabla^2 \log f\big|_{z=a}$ is negative semi-definite. Now, $z \in (Qa)^\perp \implies z^\top Q a = 0$. Now, if we simplify $z^\top \mathfrak{M} z$ subject to the constraint $z^\top Q a = 0$, we get $z^\top \left(\frac{d(d-1)}{a^\top Q a} Q\right) z$. Thus $z^\top Q z \leq 0$ for every $z \in (Qa)^\perp$, since $d(d-1), a^\top Q a > 0$.

$(2 \implies 4)$: Since $a^\top Q a > 0$, $Qa \neq 0$, and thus $(Qa)^\perp$ is a $(n-1)$-dimensional vector space.

$(4 \implies 5)$: Let $\mathcal{L}$ be the $(n-1)$-dimensional vector space over which $z^\top Q z \leq 0$. Consider some arbitrary $b \in \mathbb{R}^n$. Let $P \in \mathbb{R}^{n \times 2}$ be the matrix with columns a and b. Then $P^\top Q P = \begin{bmatrix} a^\top Q a & a^\top Q b \\ b^\top Q a & b^\top Q b \end{bmatrix}$. If $\operatorname{rank}(P^\top Q P) = 1$, then $\det(P^\top Q P) = 0$. Thus assume P has rank 2. Then the column space of P intersects $\mathcal{L}$ non-trivially, i.e. there exists

some $v \in \mathbb{R}^2$ such that $Pv \in \mathcal{L} \setminus \{0\}$, and $v^\top P^\top Q P v \leq 0$. Thus $P^\top Q P$ is not positive-definite. At the same time, the diagonal entry $a^\top Q a$ of $P^\top Q P$ is strictly positive, and thus $P^\top Q P$ can't be negative-definite. Consequently, the eigenvalues of $P^\top Q P$ are of opposite signs, implying that $\det(P^\top Q P) < 0$. Thus, for any $b \in \mathbb{R}^n$, $\det(P^\top Q P) \leq 0$. But $\det(P^\top Q P) = b^\top((a^\top Q a)Q - (Qa)(Qa)^\top)b$. Thus $(a^\top Q a)Q - (Qa)(Qa)^\top$ is negative semi-definite, as desired.
($5 \implies 1$): Ignoring the $\frac{d(d-1)}{(a^\top Q a)^2} > 0$ factor, we must show that $\left(a^\top Q a \cdot Q - \frac{d}{d-1}(Qa)(Qa)^\top\right) \preccurlyeq 0$. Let $b \in \mathbb{R}^n$ be an arbitrary vector. Then

$$b^\top \left(a^\top Q a \cdot Q - \frac{d}{d-1}(Qa)(Qa)^\top\right) b = b^\top \left(a^\top Q a \cdot Q - (Qa)(Qa)^\top\right) b - \frac{(b^\top Q a)^2}{d-1} \leq b^\top \left(a^\top Q a \cdot Q - (Qa)(Qa)^\top\right) b \leq 0$$

where the last inequality follows from the negative semi-definiteness of $\left(a^\top Q a \cdot Q - (Qa)(Qa)^\top\right)$.
($3 \implies 4$) is obvious. For ($4 \implies 3$), note that both (4) and (3) are statements about the matrix Q *only*. In particular, they don't involve f or a. Thus, if we can prove ($4 \implies 3$) for some *particular* f, a for which $(\nabla^2 f)\big|_{z=a} = Q$, we'd be done *for all* f, a for which $(\nabla^2 f)\big|_{z=a} = Q$. Thus, we choose $f(z) := \frac{z^\top Q z}{2}$. Note that $\nabla^2 f$ is identically equal to Q. Thus, for *any* b such that $f(b) = \frac{b^\top Q b}{2} \neq 0 \implies Qb \neq 0$, by (2) ((2) holds since ($4 \iff 2$)), we have that $z^\top Q z \leq 0$ for every $z \in (Qb)^\perp$, as desired.
($4 \iff 6$): Since f is homogenous, $\partial_a f$ is homogenous. Applying Lemma A.4 on $\partial_i \partial_j f$ for all i, j yields that $\left(\nabla^2(\partial_a f)\right)\Big|_{z=a} = (d-2)\left(\nabla^2 f\right)\Big|_{z=a} = (d-2) \cdot Q$.
($7 \implies 4$): Since $(n-1)$ of Q's eigenvalues are non-positive, the quadratic form $z \mapsto z^\top Q z$ is negative semi-definite on the $(n-1)$-dimensional vector space spanned by the eigenvectors of the non-positive eigenvalues.
($\neg 7 \implies \neg 4$): Since Q is real-symmetric, by the spectral theorem, the dimension of the largest subspace (of $\mathbb{R}^n$) over which Q is negative semi-definite is the number of non-positive eigenvalues Q has. Thus, if Q has ≥ 2 (strictly) positive eigenvalues, then Q can't be negative semi-definite over any $(n-1)$-dimensional subspace. If $\lambda_1 < 0$, then Q is negative definite, and thus $P^\top Q P$ is negative semi-definite for any $P \in \mathbb{R}^{n \times 2}$, and thus all diagonal entries of $P^\top Q P$ are non-positive. However, if we choose P as in the proof of ($4 \implies 5$), then the first diagonal entry of $P^\top Q P$ equals $a^\top Q a$, which is strictly positive, thus leading to a contradiction. ∎

Remark. A few remarks are in order:

1. (7) allows us to conclude that $(-1)^n \det(Q) \leq 0$.

We shall now explore variants of the notion of log-concavity, each variant with its applications and uses. We first talk about complete log-concavity, which is a stronger notion of log-concavity.

1.1. Complete Log-Concavity

Definition 1.3 (Complete Log-Concavity). A polynomial $g \in \mathbb{R}[z_1, \ldots, z_n]$ is called *completely log-concave* if for every $k \geq 0$, and every non-negative matrix $V \in \mathbb{R}^{n \times k}_{\geq 0}$, $D_V g$ is a non-negative log-concave function over $\mathbb{R}^n_{>0}$, where

$$D_V g(z) := \left(\prod_{j=1}^{k} \sum_{i=1}^{n} V_{ij} \partial_i\right) g(z_1, \ldots, z_n)$$

Remark. A few remarks are in order:

1. Note that for $k = 0$, the D_V operator is equivalent to the identity operator (since an empty product evaluates to 1). Thus, for $k = 0$, $D_V g = g$, and thus **complete log-concavity implies log-concavity**.

2. Consider $\kappa := (\kappa_1, \ldots, \kappa_n) \in \mathbb{Z}^n_{\geq 0}$. Notice that one can generate the differential operator $\partial_1^{\kappa_1} \cdots \partial_n^{\kappa_n}$ from D_V by choosing a non-negative V appropriately. Also notice that $\partial_1^{\kappa_1} \cdots \partial_n^{\kappa_n} g(z)$ equals the coefficient of $z_1^{\kappa_1} \cdots z_n^{\kappa_n}$ in g, plus monomials containing non-zero powers of $z_1, \ldots, z_n$. Thus, if the coefficient of $z_1^{\kappa_1} \cdots z_n^{\kappa_n}$ in g is negative, then one can derive a contradiction by choosing $z_1, \ldots, z_n$ to be sufficiently small positive numbers which would lead to $\partial_1^{\kappa_1} \cdots \partial_n^{\kappa_n} g(z)$ becoming negative, violating the non-negativity clause in the complete log-concave definition.
   Thus **completely log-concave polynomials have non-negative coefficients**.

3. Let g be a r-homogenous polynomial with non-negative coefficients. Note that if $k \geq r$, then $D_V g$ is a non-negative constant, which is log-concave. If $k = r-1$, then $D_V g = a_1 z_1 + a_2 z_2 + \ldots + a_n z_n$, where $a_1, \ldots, a_n \geq 0$. With some effort, it can be seen that this non-negative linear combination of variables is log-concave too. Thus, for checking the complete log-concavity of r-homogenous polynomials with non-negative coefficients, WLOG one can assume $k \leq r - 2$.

4. Using techniques similar to that in the proof of Lemma 1.1, one can show that **the set of completely log-concave polynomials of bounded degree is also closed**.

5. Using elementary topology, one can show that if $D_V g$ is log-concave for every $V \in \mathbb{R}^{n \times k}_{>0}$, then $D_V g$ is completely log-concave.

As with log-concave polynomials, we prove the closure properties of completely log-concave polynomials.

Proposition 2. Let $g(z_1, z_2, \ldots, z_n)$ be a completely log-concave polynomial. Then the following polynomials are completely log-concave too:

1. *Affine Transformations*: $g(T(y_1, \ldots, y_m))$, where $T : \mathbb{R}^m \mapsto \mathbb{R}^n : y \mapsto Ay + b$, where $A \in \mathbb{R}^{n \times m}_{\geq 0}$, $b \in \mathbb{R}^n_{\geq 0}$.

2. *Permutation*: $g(z_{\pi(1)}, \ldots, z_{\pi(n)})$ for any permutation π of $[n]$

3. *External Field*: $cg(c_1 z_1, c_2 z_2, \ldots, c_n z_n)$ for scalars $c, c_1, \ldots, c_n \in \mathbb{R}_{\geq 0}$.

4. *Specialization*: $g(a, z_2, \ldots, z_n) = g(z_1, z_2, \ldots, z_n)\big|_{z_1 = a}$, for $a \in \mathbb{R}_{\geq 0}$.

5. *Differentiation*: $\partial_v g = \sum_{i \in [n]} v_i \partial_i g$, for $v \in \mathbb{R}^n_{\geq 0}$. We didn't have this for log-concave polynomials.

Proof. For the first part, note that since T is non-negative, and since all coefficients of g are non-negative, $g(T(y))$ has non-negative coefficients. Thus, for any non-negative V, $D_V g$ is a polynomial with non-negative coefficients and is thus non-negative on $\mathbb{R}^n_{>0}$.

Now, note that for any $v \in \mathbb{R}^m$,

$$\partial_v g(T(y)) = \partial_{Av} g(z)\big|_{z=T(y)} \implies \partial_{v_1} \cdots \partial_{v_k} g(T(y)) = \partial_{Av_1} \cdots \partial_{Av_k} g(z)\big|_{z=T(y)} \implies D_V g(T(y)) = D_{AV} g\big|_{z=T(y)}$$

Note that AV is a non-negative matrix, and thus $D_{AV} g$ is a log-concave function (by the assumption on the complete log-concavity of g). By Part 1 of Proposition 1, $(D_{AV} g)(T(y)) = D_{AV} g\big|_{z=T(y)}$ is log-concave, as desired.

As in the proof of Proposition 1, parts 2, 3, and 4 follow from part 1. Finally, part 5 follows directly from the definition of complete log-concavity! ■

Bivariate completely log-concave polynomials have a very powerful property called *ultra log-concavity*. This theorem will become necessary later on when we establish Mason's conjecture (Corollary 2.6).

Theorem 1.3. If $f = \sum_{k=0}^{n} c_k z_1^{n-k} z_2^k \in \mathbb{R}_{\geq 0}[z_1, z_2]$ is completely log-concave, then the sequence $c_0, \ldots, c_n$ is ultra log-concave, i.e. for every $1 \leq k < n$,

$$\left(\frac{c_k}{\binom{n}{k}}\right)^2 \geq \frac{c_{k-1}}{\binom{n}{k-1}} \cdot \frac{c_{k+1}}{\binom{n}{k+1}}$$

Proof. Since f is completely log-concave, the quadratic $q(z_1, z_2) := \partial_{z_1}^{n-k-1} \partial_{z_2}^{k-1} f$ is log-concave on $\mathbb{R}^2_{\geq 0}$. Now,

$$\nabla^2 q = \begin{bmatrix} \partial_{z_1}^2 q & \partial_{z_1}\partial_{z_2} q \\ \partial_{z_2}\partial_{z_1} q & \partial_{z_2}^2 q \end{bmatrix} = n! \begin{bmatrix} \frac{c_{k-1}}{\binom{n}{k-1}} & \frac{c_k}{\binom{n}{k}} \\ \frac{c_k}{\binom{n}{k}} & \frac{c_{k+1}}{\binom{n}{k+1}} \end{bmatrix}$$

By the remark succeeding the proof of Theorem 1.2, $\det(\nabla^2 q) \leq 0$, which yields the desired result. ■

We now investigate an alternate characterization of completely log-concave polynomials, which makes explicit the role of non-zero coefficients in determining the complete log-concavity of a polynomial.

1.1.1. Indecomposability Characterizations of Completely Log-Concave Polynomials

The theorems proved in this sub-subsection are necessary to establish Mason's conjecture (Corollary 2.6). The ultimate aim of these results is to show Theorem 1.6, which relates the complete log-concavity of a polynomial purely to what monomials are non-zero in that polynomial. In a sense, Theorem 1.6 is a "combinatorial characterization" of the analytic property of complete log-concavity.

Lemma 1.4. Let $f, g \in \mathbb{R}[z_1, \ldots, z_n]$ be homogenous with non-negative coefficients satisfying $\partial_b f = \partial_c g \neq 0$ for some $b, c \in \mathbb{R}^n_{\geq 0}$. If f, g are log-concave on $\mathbb{R}^n_{\geq 0}$, then so is $f + g$.

Proof. Note that $\partial_b f, \partial_c g$ are both polynomials. Thus, their equality implies that f, g have the same degree d. We induct on d. For $d = 1$, $f + g$ is a linear polynomial with non-negative coefficients, which can be easily seen to be log-concave from the basic definition. Now, fix any $a \in \mathbb{R}^n_{>0}$, and let $d \geq 2$. Define $Q_1 := \nabla^2 f\big|_{z=a}, Q_2 := \nabla^2 g\big|_{z=a}$. Now, observe that $(Q_1 b)_i = (\partial_i \partial_b f)\big|_{z=a}, (Q_2 c)_i = (\partial_i \partial_c g)\big|_{z=a}$. Since $\partial_b f = \partial_c g$, $Q_1 b = Q_2 c$. Furthermore, since $\partial_b f \neq 0$, $\partial_i \partial_b f \neq 0$ for some i. Since $a \in \mathbb{R}^n_{>0}$ is *strictly* positive, $(\partial_i \partial_b f)(a) \neq 0$, and thus $Q_1 b \neq 0$. By the log-concavity of f, g, both the quadratic forms $z \mapsto z^\top Q_1 z, z \mapsto z^\top Q_2 z$ are negative semi-definite on $(Q_1 b)^\perp = (Q_2 c)^\perp$ by $(1 \implies 2)$ of Theorem 1.2. Consequently, $z \mapsto z^\top (Q_1 + Q_2) z$ is also negative semi-definite on $(Q_1 b)^\perp$, which is a $(n-1)$-dimensional vector space. Thus invoking $(4 \implies 1)$ of Theorem 1.2 we get that *any* homogenous polynomial having Hessian $Q_1 + Q_2$ at $z = a$ is log-concave at $z = a$, and thus $f + g$ is log-concave at $z = a$. Since $a \in \mathbb{R}^n_{>0}$ was arbitrary, $f + g$ is log-concave on $\mathbb{R}^n_{>0}$, and thus $\mathbb{R}^n_{\geq 0}$. ∎

Definition 1.4 (Indecomposable Polynomials)**.** A polynomial $f \in \mathbb{R}[z_1, \ldots, z_n]$ is called indecomposable if f can *not* be written as $f_1 + f_2$, where f_1, f_2 are polynomials on disjoint sets of variables. Equivalently, f is indecomposable if the *indecomposability graph* $G(\{i : \partial_i f \neq 0\}, \{\{i, j\} : \partial_i \partial_j f \neq 0\})$ is connected.

Lemma 1.5. Let $f \in \mathbb{R}[z_1, \ldots, z_n]$ be a d-homogenous indecomposable polynomial with non-negative coefficients, where $d \geq 3$. If $\partial_i f$ is log-concave on $\mathbb{R}^n_{\geq 0}$ for every $i \in [n]$, then $\partial_a f$ is log-concave on $\mathbb{R}^n_{\geq 0}$ for every $a \in \mathbb{R}^n_{\geq 0}$.

Proof. WLOG assume $\partial_i f \neq 0$ for every $i \in [n]$. Since f is indecomposable, we can also relabel $z_1, \ldots, z_n$ such that for every $1 < j \leq n$, there exists $i < j$ such that $\partial_i \partial_j f \neq 0$. Now fix some $a \in \mathbb{R}^n_{>0}$. We want to show that $\sum_{i=1}^n a_i \partial_i f$ is log-concave on $\mathbb{R}^n_{\geq 0}$. We will proceed by inducting on k to show that $\sum_{i=1}^k a_i \partial_i f$ is log-concave on $\mathbb{R}^n_{\geq 0}$ for every $k \in [n]$. Note that this induction on k also proves the statement for $a \in \mathbb{R}^n_{\geq 0}$ [4].
Now, for $k = 1$, the result follows from the assumption. Now, suppose $g := \sum_{i=1}^k a_i \partial_i f$ is log-concave for some $k > 1$. By the induction hypothesis, $h := a_{k+1} \partial_{k+1} f$ is also log-concave. Also, let $b := \begin{bmatrix} a_1 & \ldots & a_k & 0 & \ldots & 0 \end{bmatrix}^\top \in \mathbb{R}^n_{\geq 0}, c := a_{k+1} \mathbb{1}_{k+1} \in \mathbb{R}^n_{\geq 0}$. Then

$$\partial_b h = \partial_c g = \sum_{i=1}^k a_i a_{k+1} \partial_i \partial_{k+1} f$$

By our indecomposability assumption, there is some $i \in [k]$ such that $\partial_i \partial_{k+1} f \neq 0$. Since $a_i a_{k+1} \neq 0$, $\partial_b h \neq 0$. Consequently, we can invoke Lemma 1.4 to obtain that $g + h = \sum_{i=1}^{k+1} a_i \partial_i f$ is log-concave, as desired. ∎

[4] Indeed, if some $a \in \mathbb{R}^n_{\geq 0}$ has k non-zero entries, then we can permute our indices such that the first k entries of a are non-zero, and then the k^{th} step of our induction proves the result

Theorem 1.6. Let $f \in \mathbb{R}[z_1, \ldots, z_n]$ be a d-homogenous polynomial with non-negative coefficients, where $d \geq 2$. If the following conditions hold, then f is completely log-concave:

1. For all $\alpha \in \mathbb{Z}^n_{\geq 0}$ with $|\alpha| \leq d-2$, $\partial^\alpha f = \left(\prod_{i=1}^n \partial_i^{\alpha_i}\right) f$ is indecomposable.

2. For all $\alpha \in \mathbb{Z}^n_{\geq 0}$ with $|\alpha| = d-2$, $\partial^\alpha f$ is log-concave over $\mathbb{R}^n_{\geq 0}$.

Proof. We induct on d. The case $d = 2$ is obvious. Thus assume $d \geq 3$. By the remark following the definition of complete log-concavity, it is enough to show that $D_V g = \partial_{v_1} \cdots \partial_{v_k} g$ is log-concave, for any $V \in \mathbb{R}^{n \times k}_{>0}$, $k \leq d-2$. If $k = 0$, then to show the log-concavity of f at any point a, by Theorem 1.2 it suffices to show that ∂_a is log-concave at $z = a$. Note that the $k = 1$ case requires us to show that $\partial_a g$ is log-concave *everywhere*. Thus the $k = 0$ case can be subsumed into the $k = 1$ case. Thus assume $k > 0$. By our induction hypothesis, $\partial_j f$ is completely log-concave for all $j \in [n]$, and thus $\partial_{v_1} \cdots \partial_{v_{k-1}} \partial_j f$ is log-concave on $\mathbb{R}^n_{\geq 0}$. But $\partial_{v_1} \cdots \partial_{v_{k-1}} \partial_j f = \partial_j \partial_{v_1} \cdots \partial_{v_{k-1}} f$. Now, observe that if f is indecomposable, then $\partial_v f$ is also indecomposable for any $v \in \mathbb{R}^n_{>0}$, and consequently, $\partial_{v_1} \cdots \partial_{v_{k-1}} f$ is also indecomposable, and has degree $= d - k + 1 \geq 3$. We are then done by invoking Lemma 1.5. ■

§2. Matroids, Log-Concavity, and Mason's Conjecture

A combinatorial context where log-concavity arises very naturally is the study of matroids. Matroids [5] are combinatorial entities that were independently introduced by Whitney ([Whi35]) and Nakasawa ([Nak35], [Nak36a], [Nak36b]). One reason why matroids are so ubiquitous throughout combinatorics is that *matroids are exactly the class of combinatorial structures over which a greedy optimization strategy works*, i.e. if $w : [n] \mapsto \mathbb{R}$ is some weight function, then a variant of Kruskal's algorithm for minimum spanning trees gives us the minimum weight base [6].
So let's see what matroids are!

Definition 2.1 (Matroids)**.** A matroid $M = ([n], \mathcal{I})$ is said to be defined over the *ground set* $[n] = \{1, 2, \ldots, n\}$, and is characterized by its non-empty collection of *independent sets* $\mathcal{I} \subseteq 2^{[n]}$ which satisfy the following properties:

1. *Downward Closed*: If $A \in \mathcal{I}$, then $2^A \subseteq \mathcal{I}$, i.e. if A is an independent set, then every subset of A is also an independent set. In particular, since $\mathcal{I}$ is non-empty, it must contain $\varnothing$.
2. *Exchange Property*: If $A, B \in \mathcal{I}$, and $|A| < |B|$, then there exists $i \in B \setminus A$ such that $A \cup \{i\} \in \mathcal{I}$.

The exchange property implies that all maximal independent sets of the matroid have the same size, which is known as the *rank* of that matroid. Any maximal set of a matroid is known as its *basis*. Given a matroid M, $\mathcal{B}_M$ is the set of all bases of M. Note that due to the downward closed property of a matroid, to describe all independent sets it is enough to describe just the bases.
The *generating function* of a matroid is defined as

$$g_M(z_1, \ldots, z_n) := \sum_{B \in \mathcal{B}_M} z^B = \sum_{B \in \mathcal{B}_M} \prod_{i \in B} z_i$$

Clearly g_M is a $\operatorname{rank}(M)$-homogenous polynomial with non-negative coefficients.
For any $S \subseteq [n]$, $\operatorname{rank}(S) := \max_{\substack{I \in \mathcal{I} \\ I \subseteq S}} |I|$. Any set which is not independent is called dependent. A *minimal* [7] dependent set is called a *circuit*. Thus, if $C \subseteq [n]$ is a circuit, then $\operatorname{rank}(C) = |C| - 1$.
An element $i \in [n]$ forms a *loop* if $\{i\}$ is a circuit. Note that if i forms a loop, i doesn't belong to any independent set. We abuse notation slightly to say that $i \in [n]$ is a loop if i forms a loop. Two elements, $i, j \in [n], i \neq j$, such that neither i nor j form loops, are called *parallel* if $\{i, j\}$ form a circuit. A matroid with no loops or parallel elements is called *simple*. If M is simple, then $\{i, j\} \in \mathcal{I}$ for every $i, j \in [n]$.
Let $M = ([n], \mathcal{I})$ be a matroid, and let $S \subseteq [n]$ be the set of elements which are *not* loops. One may note that the parallelism relation on S is an equivalence relation. Thus we can create equivalence classes $S_1, \ldots, S_k$ such that $S = S_1 \cup S_2 \cup \cdots \cup S_k$, and $j, k \in [n]$ are parallel if and only if j, k belong to the same equivalence class S_*. The

[5] Refer to [Oxl11] for a comprehensive introduction to the theory of matroids
[6] where the weight of an independent set I is defined as $w(I) := \sum_{i \in I} w(i)$
[7] under the partial order induced by $\subseteq$

existence of such parallelism equivalence classes is sometimes called the *matroid partition property*.

Let $M = ([n], \mathcal{I})$ be a matroid, and let $I \in \mathcal{I}$. Then the *contraction* M/S is the matroid $M/S := ([n] \setminus S, \{T \subseteq [n] \setminus S : T \cup S \in \mathcal{I}\})$.

Similarly, if $M = ([n], \mathcal{I})$ is a matroid, then for any $k \leq \operatorname{rank}(M)$, the *k-truncation* $M_k := ([n], \{I \in \mathcal{I} : |I| \leq k\})$ is a matroid too. Clearly, $\operatorname{rank}(M_k) = k$, and $\mathcal{B}_{M_k} = \{I \in \mathcal{I} : |I| = k\}$. Thus, if we have an algorithm for calculating (or approximating) $|\mathcal{B}_M|$ for an arbitrary matroid M, then that same algorithm can be used to calculate (or approximate) the number of independent sets of a given size k.

Let $M = ([n], \mathcal{I})$ be a matroid. The *dual matroid* of M, is defined to be $M^* := ([n], \mathcal{I}^*)$, where B^* is a base in $\mathcal{I}^*$ if $[n] \setminus B^*$ is a base in $\mathcal{I}$. Thus $\operatorname{rank}(M^*) = n - \operatorname{rank}(M)$.

Given two matroids $M_1 = (E_1, \mathcal{I}_1), M_2 = (E_2, \mathcal{I}_2)$, the direct sum $M_1 \oplus M_2$ is defined as

$$M_1 \oplus M_2 := (E_1 \sqcup E_2, \{I_1 \sqcup I_2 : I_1 \in \mathcal{I}_1, I_2 \in \mathcal{I}_2\})$$

where $\sqcup$ stands for the disjoint union of sets. Recall the notion of disjoint union: If $A \cap B = \varnothing$, then $A \sqcup B = A \cup B$. If A, B are not disjoint, say for example, $A = \{1,2,3\}, B = \{1,2,4\}$, then $A \sqcup B = \{1_A, 1_B, 2_A, 2_B, 3, 4\}$, i.e. all elements in $A \cap B$ are made "copies" of in $A \sqcup B$.

Two matroids $(E_1, \mathcal{I}_1), (E_2, \mathcal{I}_2)$ are said to be isomorphic if there exists a bijection $\phi : E_1 \to E_2$ such that $I_1 \in \mathcal{I}_1$ if and only if $\phi(I_1) \in \mathcal{I}_2$.

The *Newton polytope* of a matroid is defined as $\mathcal{P}_M := \operatorname{conv}\left(\{\mathbb{1}_B \in \mathbb{R}^n : [n] \supseteq B \in \mathcal{B}_M \subseteq \mathcal{I}\}\right)$, i.e. the Newton polytope is the convex hull of the indicator set of all bases of our matroid. Cunningham [Cun84] showed that $\mathcal{P}_M$ has an "efficient separation oracle" for any matroid M. While we shall not get into what this means, what it implies is that if we have a convex function $f : \mathcal{P}_M \mapsto \mathbb{R}$, then we can, in $\operatorname{poly}(n)$ time, minimize f (and also find the minimizer $p^* \in \mathcal{P}_M$). Similarly, we can, in polynomial time, maximize concave functions over $\mathcal{P}_M$.

2.1. Examples of Matroids

The reason matroids are so useful is because they subsume a wide variety of combinatorial phenomena within themselves. We shall now see a few examples of matroids to get a feel for how powerful this notion is.

1. *Linear Matroids*: Any set of vectors $v_1, \ldots, v_n \in \mathbb{F}^t$, where $\mathbb{F}$ is a field, induce the linear matroid $M = ([n], \mathcal{I})$, where $\mathcal{I} = \{A \subseteq [n] : \{v_i : i \in A\}$ is a linearly independent set$\}$. The rank of this matroid is the rank of the set of vectors, and the bases of this matroid are the sets of indices of vectors in (linear algebraic) bases of $\{v_1, \ldots, v_n\}$.

   If M is isomorphic to a linear matroid induced by vectors in a $\mathbb{F}$-vector space, we say that M is $\mathbb{F}$-representable. There exist matroids (such as the Vámos matroid) which are *not* representable over any field. Also, *in general*, given two fields $\mathbb{F}, \mathbb{H}$, there exist matroids which are $\mathbb{F}$-representable but not $\mathbb{H}$-representable. For example, the *Fano matroid* is $\mathbb{F}_2$-representable but not $\mathbb{R}$-representable.

2. *Graphic Matroids*: Let $G = (V, E)$ be a simple graph. It induces the matroid $M = (E, \mathcal{I})$, where $\mathcal{I} = \{S \subseteq E :$

The edges in S don't create any cycles$\}$. The bases of this matroid are the spanning trees of G.
It can be shown that every graphic matroid is isomorphic to some linear matroid.

3. *Uniform Matroids*: These are the matroids $U_n^r := ([n], \bigcup_{i=0}^{r} \binom{[n]}{r})$, i.e. the set of bases of U_n^r is $\binom{[n]}{r}$. Uniform matroids are linear matroids; indeed, it is easy to see that the linearly independent subsets of n vectors (in general position), in $\mathbb{R}^{r+1}$, induce U_n^r.

4. *'Partition' Matroids*: These are direct sums ('$\oplus$') of uniform matroids. Thus, let $M = \bigoplus_{i=1}^{t} U_{n_i}^{r_i}$. Then the ground set of M is partitioned into t partitions of sizes $n_1, \ldots, n_t$ respectively. A subset of the ground set is independent if it has at most r_i elements from the i^{th} partition, i.e. we can interpret r_i to be the 'capacity' of the i^{th} partition.

5. *Bipartite Matching Matroid (a.k.a Transversal Matroid)*: Let $G = ((A, B), E)$ be a bipartite graph with bipartitions A, B. Consider two partition matroids $M_1 = (E, \mathcal{I}_1), M_2 = (E, \mathcal{I}_2)$ on the ground set E, i.e. the elements of our matroids are the edges of the bipartite graph G.
The partitions of M_1 are given by $S_1, \ldots, S_{|A|}$, where S_i is the set of edges incident on $i \in A$. Set the capacity of each of these sets to be 1. Then the independent sets of M_1 are sets of edges such that every vertex in A is incident on at most one edge from the set.
Similarly, construct M_2, with partitions $T_1, \ldots, T_{|B|}$, where T_j is the set of edges incident on $j \in B$, and set the capacity of each category to 1.
Then note that if I is an independent set in both M_1 and M_2, then I is a matching in G. Furthermore, the bases of $M_1 \cap M_2$ are the maximum matchings of G.
Thus, the **bipartite matching problem is subsumed within the matroid intersection problem**.

6. *Gammoid Matroids*: Let G be a directed graph, and let $S, T \subseteq V(G)$ be sets of vertices, not necessarily disjoint. Define a matroid $\Gamma := (T, \mathcal{I})$, where $I \subseteq T$ is independent if there exists a set of vertex disjoint paths whose starting points belong to S, and endpoints are exactly I.
Thus, $\operatorname{rank}(\Gamma)$ is equal to the *max-flow* between S, T. By the Max Flow-Min Cut theorem, this equals the minimum number of vertices we need to delete to disconnect S, T.

There exist many more examples of matroids that we don't mention here; Nevertheless, we hope that the reader is convinced of the need to study matroid algorithms, for any statement regarding matroids immediately has many combinatorial implications.

2.2. Matroids and Log-Concavity

One of the most remarkable results in matroid theory, which underlies all the breakthroughs in recent years, is the fact that the basis generating polynomial of a matroid is completely log-concave.
This result was shown through advanced mathematical tools (such as combinatorial Hodge theory) by Huh and Wang [HW17], and Adiprasito, Huh and Katz [AHK18].

However, we shall see an alternate, perhaps more elementary, proof of the same given in [ALOV19]. To do that, we shall use the indecomposability criteria (Theorem 1.6) developed in earlier chapters.

Lemma 2.1. Let M be a matroid of rank r, and $q \in (0,1], k \in [n]$ be parameters. Let $\lambda \in \mathbb{R}^n_{>0}$ be arbitrary. Define

$$f_{M,k,q}(x_1, \dots, x_n) := \sum_{S \in \binom{[n]}{k}} q^{-\operatorname{rank}(S)} \prod_{i \in S} \lambda_i x_i$$

$f_{M,k,q}$ is completely log-concave.

Proof. We shall use Theorem 1.6 to prove the complete log-concavity of $f_{M,k,q} =: f$.
We have to first verify that $\partial_{i_1}\partial_{i_2}\dots\partial_{i_\ell} f$ is indecomposable for $\ell \le k-2, i_1, \dots, i_\ell \in [n]$. Note that since f is multilinear, $\partial_i^2 f = 0$ for any $i \in [n]$, and thus WLOG we assume that $i_1, \dots, i_\ell$ are distinct, and let $U = \{i_1, \dots, i_\ell\}$. Note that $\partial_{i_1}\partial_{i_2}\dots\partial_{i_\ell} f = \partial_U f$ has a monomial x^T with non-zero coefficient for **every** $T \subset [n] \setminus U$ with $|T| \le k - |U|$. Thus, since $|U| = \ell < k$, the indecomposability of $\partial_U f$ follows.
We have to now verify the log-concavity of $\partial_U f$ for every $U \subseteq [n], |U| = k-2$. Note that

$$(\partial_U f)(x_1, \dots, x_n) = \lambda^U \sum_{T \in \binom{[n]}{k} : T \supset U} q^{-\operatorname{rank}(T)} \lambda^{T \setminus U} x^{T \setminus U} = \lambda^U \sum_{\{i,j\} \in \binom{[n] \setminus U}{2}} q^{-\operatorname{rank}(U \cup \{i,j\})} \lambda_i \lambda_j x_i x_j$$

Thus (ignoring the constant λ^U factor), the $(i,j)^{\text{th}}$ entry of $\nabla^2 \partial_U f$, where $i \ne j \in [n] \setminus S$ is

$$q^{-\operatorname{rank}(U \cup \{i,j\})} \lambda_i \lambda_j = q^{-\operatorname{rank}(U)} q^{-\operatorname{rank}_{M/U}(\{i,j\})} \lambda_i \lambda_j$$

Thus, if we can show that the matrix $A \in \mathbb{R}^{([n] \setminus S) \times ([n] \setminus S)}$ [8], with $A_{ij} = q^{-\operatorname{rank}_{M/U}(\{i,j\})} \lambda_i \lambda_j$ [9] has at most one positive eigenvalue, then we'd be done by Theorem 1.2 [10]. Now, consider $v \in \mathbb{R}^{[n] \setminus S}$ where $v_i = \lambda_i$ if i is a loop of M/U, and $v_j = q^{-1}\lambda_j$ if j is not a loop of M/U. Then some thought reveals that

$$(vv^\mathsf{T} - A)_{ij} = \begin{cases} (q^{-2} - q^{-1})\lambda_i \lambda_j, & \text{if } i,j \text{ are non-parallel loops in } M/U \\ 0, & \text{otherwise} \end{cases}$$

Thus, splitting the non-loops of M/U into parallelism equivalence classes $B_1, \dots, B_t$, we have $vv^\mathsf{T} - A = (q^{-2} - q^{-1}) \sum_{j=1}^t \boldsymbol{\lambda}_{B_j} \boldsymbol{\lambda}_{B_j}^\mathsf{T}$, where $\boldsymbol{\lambda}$ is the vector with entries such that $\boldsymbol{\lambda}_B(i) = \lambda_i \mathbb{1}_{i \in B}$. Clearly, $vv^\mathsf{T} - A$ is PSD, and thus $A \preceq vv^\mathsf{T}$ and we're done. ∎

[8] Note that $(\nabla^2 \partial_U f)_{ij} = 0$ if i or j is in S. Thus $\nabla^2 \partial_U f$ is just A padded with 0s

[9] once again we ignore the constant $q^{-\operatorname{rank}(U)}$ factor

[10] actually we'd also need to show that the top eigenvalue of A is non-negative: But that follows from the fact that of the top eigenvalue of A were negative, then A is negative definite. Now, if i is a non-loop element in M/U, then $A_{ii} = 0$, which can't be if A is negative definite. If all elements in M/U are loops, then the problem is trivial

One immediately gets as a corollary our desired result:

Theorem 2.2. $g_M(z)$ is completely log-concave for *any* matroid.

Proof. Note that $q^r f_{M,r,q}(z_1, \ldots, z_n)$ converges to g_M as $q \to 0$. Since completely log-concave polynomials are closed under pointwise convergence, we're done. ∎

While we shall have the occasion to use Theorem 2.2 soon enough, we now move on towards showing the log-concavity of another polynomial associated with a matroid, which will finally break Mason's conjecture.

2.3. Towards Mason's Conjecture

Mason's conjecture ([Mas72], Corollary 2.6) says that the sequence $\mathcal{I}_M^0, \mathcal{I}_M^1, \ldots, \mathcal{I}_M^r$ is *ultra log-concave* (see Theorem 1.3, Corollary 2.6), where $\mathcal{I}_M^k$ is the number of independent sets of size k of some arbitrary matroid M. As with the previously described results, the resolution of Mason's conjecture for arbitrary matroids is an extremely important achievement, not least because the conjecture had been open for nearly half a century, and breakthroughs had been elusive despite a long line of works ([Sey77, Dow80, Mah85, Zha85, HS89, KN09, HK11, Len13]) by illustrious mathematicians and computer scientists. Mason's conjecture was finally resolved (nearly simultaneously) by Anari, Liu, Oveis Gharan and Vinzant [ALOV18] and Brändén and Huh [BH22]. We shall present the proof in [ALOV18] here.

Let $M = ([n], \mathcal{I})$ be a matroid. Define the *homogenization* of $\mathcal{I}$ to be

$$h_M(y, z_1, \ldots, z_n) := \sum_{I \in \mathcal{I}} y^{n-|I|} \prod_{i \in I} z_i$$

Lemma 2.3. $\partial_y^{n-2} h_M$ is log-concave on $\mathbb{R}^n_{\geq 0}$.

Proof. Note that if $i \in [n]$ is a loop, then i doesn't belong to any independent set, and consequently, z_i is absent from h_M. Thus without loss of generality, assume no element in $[n]$ is a loop. Then $\{i\} \in \mathcal{I}$ for every $i \in [n]$.

Now, observe that

$$\partial_y^{n-2} h_M = (n-2)! \underbrace{\left(\frac{n(n-1)}{2} y^2 + (n-1) \sum_{\{i\} \in \mathcal{I}} y z_i + \sum_{\{i,j\} \in \mathcal{I}} z_i z_j \right)}_{:= q(y, z_1, \ldots, z_n)}$$

Consider $Q := \nabla^2 q$. Note that Q is a $(n+1) \times (n+1)$ matrix. With some effort, one can see that $Q = \begin{bmatrix} n(n-1) & (n-1)\mathbb{1}^\intercal \\ (n-1)\mathbb{1} & B \end{bmatrix}$, where $\mathbb{1} \in \mathbb{R}^n$ is the n-dimensional vector consisting of all 1s, and B is a $n \times n$ matrix where $B_{ij} = \mathbb{1}_{\{i,j\} \in \mathcal{I}}$. Now, fix

$a = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^\intercal \in \mathbb{R}^{n+1}$. If we can show that $(a^\intercal Qa)Q - (Qa)(Qa)^\intercal$ is negative semi-definite, then by $(5 \implies 4)$ in Theorem 1.2 we would have that Q is negative semi-definite on a $(n-1)$-dimensional vector space. Since $\nabla^2 q$ is a constant matrix, by $(4 \implies 1)$ in Theorem 1.2 we would have that q was log-concave at any point in $\mathbb{R}^n_{\geq 0}$, as desired. Now,

$$(a^\intercal Qa)Q - (Qa)(Qa)^\intercal = (n-1)\begin{bmatrix} 0 & 0 \\ 0 & nB - (n-1)\mathbb{1}\mathbb{1}^\intercal \end{bmatrix}$$

Thus it suffices to show that $(nB - (n-1)\mathbb{1}\mathbb{1}^\intercal)$ is negative semi-definite.

Now, recall the *matroid partition property*: We can partition all non-loop elements of a matroid into equivalence classes based on parallelism. Since all elements in $[n]$ are non-loops, we get a partition $[n] = S_1 \cup S_2 \cup \cdots \cup S_k$, where i, j are parallel if and only if they belong to the same equivalence class. Now, note that $B = \mathbb{1}\mathbb{1}^\intercal - \sum_{i=1}^k \mathbb{1}_{S_i}\mathbb{1}_{S_i}^\intercal$, which implies $nB - (n-1)\mathbb{1}\mathbb{1}^\intercal = \mathbb{1}\mathbb{1}^\intercal - n\sum_{i=1}^k \mathbb{1}_{S_i}\mathbb{1}_{S_i}^\intercal$. Now consider arbitrary $x \in \mathbb{R}^n$: Then

$$x^\intercal(nB - (n-1)\mathbb{1}\mathbb{1}^\intercal)x = (\mathbb{1}^\intercal x)^2 - n\sum_{i=1}^k (\mathbb{1}_{S_i}^\intercal x)^2 = \left(\sum_{i=1}^k \mathbb{1}_{S_i}^\intercal x\right)^2 - n\sum_{i=1}^k (\mathbb{1}_{S_i}^\intercal x)^2 \leq k\sum_{i=1}^k (\mathbb{1}_{S_i}^\intercal x)^2 - n\sum_{i=1}^k (\mathbb{1}_{S_i}^\intercal x)^2 \leq 0$$

where the second last inequality follows from the Cauchy-Schwartz inequality, and the last inequality follows from the fact that k, which is the number of partitions of $[n]$, can't exceed n. ■

Theorem 2.4. For any matroid $M = ([n], \mathcal{I})$, the polynomial

$$h_M(y, z_1, \dots, z_n) := \sum_{I \in \mathcal{I}} y^{n-|I|} \prod_{i \in I} z_i$$

is completely log-concave.

Proof. We use Theorem 1.6 to show complete log-concavity. We denote ∂_{z_i} by ∂_i, and for any $\alpha \in \mathbb{Z}^n_{\geq 0}$ we define $\partial^\alpha := \prod_{i=1}^n \partial_i^{\alpha_i}$. We need to show that $\partial_y^k \partial^\alpha h_M$ is indecomposable for $k + |\alpha| \leq n-2$, and log-concave for $k + |\alpha| = n-2$. Note that if $\alpha_i \geq 2$ for some i, then $\partial^\alpha h_M = 0$, and thus assume $\alpha = \mathbb{1}_J$ for some $J \subseteq [n]$. Furthermore, if $J \notin \mathcal{I}$, then also $\partial^{\mathbb{1}_J} h_M = 0$. Thus assume $J \in \mathcal{I}$. Then

$$\partial^{\mathbb{1}_J} h_M = \sum_{I \in \mathcal{I} : J \subseteq I} y^{n-|I|} \prod_{i \in I \setminus J} z_i = h_{M/J}$$

where $M/J = ([n] \setminus J, \{I \setminus J : I \in \mathcal{I}, J \subseteq I\})$ is the contraction matroid.

Thus we have to investigate the indecomposability (and log-concavity) of $\partial_y^k h_{M/J}$ for every $J \in \mathcal{I}$. Now, if some $i \in [n] \setminus J$ is a loop of M/J, then it doesn't appear in h_M and we can ignore it safely. Conversely, if $i \in [n] \setminus J$ is not a loop, then we have the monomial $y^{n-|J|-1} z_i$ in h_M, and thus the monomial $y^{n-|J|-1-k} z_i$ appears in $\partial_y^k h_{M/J}$, and consequently, in the indecomposability graph of $\partial_y^k h_{M/J}$, the node representing y is connected to every non-loop $i \in [n] \setminus J$, and thus $\partial_y^k h_{M/J}$ is indecomposable.

Finally, if $k + |J| = n-2$, invoking Lemma 2.3 on the matroid M/J yields that $\partial_y^{n-|J|-2} h_{M/J} = \partial_y^k \partial^{\mathbb{1}_J} h_M$ is log-concave on $\mathbb{R}^n_{\geq 0}$, as desired. ■

Corollary 2.5. For any matroid $M = ([n], \mathcal{I})$, the polynomial

$$f_M(y,z) := \sum_{k=0}^{\mathrm{rank}(M)} \mathcal{I}_k y^{n-k} z^k$$

is completely log-concave, where $\mathcal{I}_k$ is the number of independent sets of size k.

Proof. Follows by applying Part 1 of Proposition 2 to h_M, with the affine transformation being $T : \mathbb{R}^2_{\geq 0} \mapsto \mathbb{R}^{n+1}_{\geq 0}$, $T(y,z) := (y, z, \ldots, z)$. ∎

Corollary 2.6 (Mason's Conjecture)**.** For any matroid $M = ([n], \mathcal{I})$ of rank r, the sequence $\mathcal{I}_0, \ldots, \mathcal{I}_r$ is ultra log-concave, i.e.

$$\left(\frac{\mathcal{I}_k}{\binom{n}{k}}\right)^2 \geq \frac{\mathcal{I}_{k-1}}{\binom{n}{k-1}} \cdot \frac{\mathcal{I}_{k+1}}{\binom{n}{k+1}}$$

where $\mathcal{I}_r$ is the number of independent sets of size r.

Proof. Apply Theorem 1.3 to Corollary 2.5. ∎

§3. Entropy

We eventually want to be able to design algorithms with our analytic machinery. Now, many algorithms, especially randomized algorithms, boil down to showing that the solution to our problem can be efficiently sampled from some probability distribution. Now, also note that every probability distribution can be associated with its generating function. Thus, we can exploit the power of our machinery by studying probability distributions whose underlying generating functions are log-concave (or some variant thereof). Then we use the mathematical properties of log-concave polynomials to make comments about the probability distributions, and possibly extract algorithmic utility from it.

Before that, we first study some fundamental properties of probability distributions themselves.

Let $\mu : \mathcal{R} \mapsto [0,1]$ be a probability distribution over some set $\mathcal{R}$. We define the *support* of μ as $\operatorname{supp}(\mu) := \{\omega \in \mathcal{R} : \mu(\omega) \neq 0\}$. Now suppose μ is a probability distribution supported over some *finite* set Ω. Then the entropy of μ is defined to be

$$\mathcal{H}(\mu) := \sum_{\omega \in \Omega} \mu(\omega) \log \frac{1}{\mu(\omega)}$$

If X is Bernoulli random variable with parameter p, we use $\mathcal{H}(X)$ and $\mathcal{H}(p)$ interchangeably.

We now state some fundamental facts about the entropy function, such as *subadditivity*, and the fact that *uniform distributions maximize entropy*. Refer to Cover and Thomas [CT06] for proof of these statements.

Proposition 3 (Subadditivity of Entropy)**.** Let X, Y be finitely supported random variables which are not necessarily independent. Let μ be the joint distribution of (X, Y). The marginals μ_X, μ_Y of μ are the distributions of X and Y respectively. Then $\mathcal{H}(\mu) \leq \mathcal{H}(\mu_X) + \mathcal{H}(\mu_Y)$, where equality holds if and only if X and Y are independent.

Proposition 4. Let μ be any finitely supported probability distribution. Then $\mathcal{H}(\mu) \leq \log\left(|\operatorname{supp}(\mu)|\right) = \mathcal{H}(u_{\operatorname{supp}(\mu)})$, where $u_{\operatorname{supp}(\mu)}$ is the uniform distribution over $\operatorname{supp}(\mu)$.

We will be interested in probability distributions over $2^{[n]}$ [11]. Thus, let μ be a distribution over $2^{[n]}$. Then the marginals of μ are defined as $\mu_i := \sum_{S \ni i} \mu(S)$ for every $i \in [n]$. Note that $\sum_{i=1}^n \mu_i = \mathbb{E}_{S \sim \mu}\left[|S|\right]$ may be much greater than 1.

Consider Bernoulli Random Variables $X_1, \ldots, X_n$ with parameters $\mu_1, \ldots, \mu_n$ respectively. It is clear that μ is a particular joint distribution of $X_1, \ldots, X_n$. Thus we can apply Proposition 3 and obtain that

$$\mathcal{H}(\mu) \leq \sum_{i \in [n]} \mathcal{H}(\mu_i) = \sum_{i \in [n]} \mu_i \log \frac{1}{\mu_i} + (1 - \mu_i) \log \frac{1}{1 - \mu_i} \tag{3.1}$$

[11] by identifying $2^{[n]}$ with $\{-1, 1\}^n$, distributions over $2^{[n]}$ are also sometimes termed as distributions over the Boolean hypercube

Equality is achieved if $X_1, \dots, X_n$ are independent, or equivalently, $\mu(S) = \prod_{i \in S} \mu_i \prod_{j \notin S}(1 - \mu_j)$ for every $S \subseteq [n]$.
We now introduce the notion of external fields and Newton Polytopes, which we shall require later on when we are discussing the topology of log-concave distributions.

Definition 3.1 (External Fields)**.** Consider $\lambda := (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n_{>0}$. Then the *external field* distribution $\lambda * \mu$ is a probability distribution on $2^{[n]}$ such that $\Pr_{\lambda * \mu}(S) \propto \lambda^S \mu(S)$, where the proportionality constant is appropriately chosen such that $\sum_{S \subseteq [n]} \Pr_{\lambda * \mu}(S) = 1$.

Definition 3.2 (Newton Polytopes)**.** Let μ be a distribution on $2^{[n]}$. We define the *Newton polytope* of μ to be $\mathcal{P}_\mu := \text{conv}\left(\{\mathbb{1}_S \in \mathbb{R}^n : S \in \text{supp}(\mu)\}\right)$.

We now study the connection between probability distributions and log-concavity, as promised.

3.1. Log-Concave Distributions

Let μ be a distribution on $2^{[n]}$. We define the generating function of μ as

$$g_\mu(z_1, \dots, z_n) := \sum_{S \subseteq [n]} \mu(S) z^S = \sum_{S \subseteq [n]} \mu(S) \prod_{i \in S} z_i$$

Note that $g_\mu(1, \dots, 1) = 1$. Also note that $\mu_i = \Pr_{S \sim \mu}(i \in S) = \partial_{z_i} g_\mu \big|_{z_1 = z_2 = \dots = z_n = 1}$.
We call μ log-concave (resp. completely log-concave) if g_μ is log-concave (resp. completely log-concave).
We now show that for log-concave distributions, there is a corresponding lower bound for Eq. (3.1).

Lemma 3.1. If μ is a log-concave distribution on $2^{[n]}$ with marginals $\mu_1, \dots, \mu_n$, then

$$\mathcal{H}(\mu) \geq \sum_{i \in [n]} \mu_i \log \frac{1}{\mu_i}$$

Proof. Sample $S \sim \mu$, and consider the random variable $X := \mathbb{1}_S$, i.e. $\Pr(X = \mathbb{1}_S) = \mu(S)$. Also define $f(z_1, \dots, z_n) := \log g_\mu\left(\frac{z_1}{\mu_1}, \dots, \frac{z_n}{\mu_n}\right)$. Note that if $\mu_i = 0$ for some i, then g_μ doesn't contain z_i in any non-zero monomial, so f can still be consistently defined. Furthermore, since g_μ is log-concave, so is $g_\mu\left(\frac{z_1}{\mu_1}, \dots, \frac{z_n}{\mu_n}\right)$, by Item 3 of Proposition 1.
Thus $f : \mathbb{R}^n_{\geq 0} \mapsto \mathbb{R} \cup \{-\infty\}$ is a concave function. By Jensen's inequality, we have $f(\mathbb{E}[X]) \geq \mathbb{E}[f(X)]$. Now, note that $\mathbb{E}[X] = \mathbb{E}_{S \sim \mu}[\mathbb{1}_S] = \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix}^\mathsf{T}$, and thus $f(\mathbb{E}[X]) = \log 1 = 0$, implying that $\mathbb{E}[f(X)] \leq 0$.
Now, note that

$$f(\mathbb{1}_S) = \log\left(\sum_{T \subseteq S} \mu(T) \prod_{i \in T} \frac{1}{\mu_i}\right) \geq \log\left(\mu(S) \prod_{i \in S} \frac{1}{\mu_i}\right) = \log \mu(S) + \sum_{i \in S} \log \frac{1}{\mu_i}$$

Thus

$$\mathbb{E}[f(X)] = \sum_{S\subseteq[n]} \mu(S)f(\mathbb{1}_S) \geq \sum_{S\subseteq[n]} \mu(S)\log\mu(S) + \sum_{S\subseteq[n]} \mu(S)\sum_{i\in S}\log\frac{1}{\mu_i} = -\mathcal{H}(\mu) + \sum_{i\in[n]} \underbrace{\left(\sum_{S\ni i}\mu(S)\right)}_{=\mu_i}\cdot\log\frac{1}{\mu_i}$$

Since $\mathbb{E}[f(X)] \leq 0$, we get our desired result. ∎

The above inequality immediately yields an approximation for $\mathcal{H}(\mu)$.

Lemma 3.2 (Additive Approximation for $\mathcal{H}(\mu)$)**.** A distribution μ is called d-homogenous if g_μ is d-homogenous. If μ is r-homogenous and log-concave, then $\sum_{i\in[n]}\mathcal{H}(\mu_i)$ is an additive r-approximation to $\mathcal{H}(\mu)$, i.e.

$$\sum_{i\in[n]}\mathcal{H}(\mu_i) - r \leq \mathcal{H}(\mu) \leq \sum_{i\in[n]}\mathcal{H}(\mu_i)$$

Proof. $\mathcal{H}(\mu) \leq \sum_{i\in[n]}\mathcal{H}(\mu_i)$ is simply Eq. (3.1).

Now, it is easy to see that if g is r-homogenous, then it only contains monomials of degree r. Consequently, if μ is r-homogenous, then for any $S \in \operatorname{supp}(\mu)$, $|S| = r$. Consequently,

$$\sum_{i\in[n]}(1-\mu_i)\log\frac{1}{1-\mu_i} \leq \sum_{i\in[n]}\mu_i = \mathbb{E}_{S\sim\mu}\left[|S|\right] = r \tag{3.2}$$

where the first inequality follows from the fact that $(1-x)\log\frac{1}{1-x} \leq x$ for any $x \in [0,1]$.

Thus, invoking Lemma 3.1, we get

$$\mathcal{H}(\mu) \geq \sum_{i\in[n]}\mu_i\log\frac{1}{\mu_i} = \sum_{i\in[n]}\mathcal{H}(\mu_i) - \sum_{i\in[n]}(1-\mu_i)\log\frac{1}{1-\mu_i} \geq \sum_{i\in[n]}\mathcal{H}(\mu_i) - r$$

∎

Thus, if we can calculate $\sum_{i\in[n]}\mathcal{H}(\mu_i)$, then we have an additive approximation for $\mathcal{H}(\mu)$. However, additive approximations don't give any *multiplicative* guarantees: Indeed, if $\sum_{i\in[n]}\mathcal{H}(\mu_i) = r+1$, and if $r \gg 1$, then the multiplicative approximation factor for $\mathcal{H}(\mu)$ is r, which is bad.

We seek to remedy this as follows: For any distribution μ on $2^{[n]}$, we define the *dual* of μ, denoted as μ^*, to be $\mu^*(S) := \mu([n]\setminus S)$ for every $S \subseteq [n]$. Note that $\mu_i^* = 1-\mu_i$. Also note that $\mathcal{H}(\mu) = \mathcal{H}(\mu^*)$. Then

Lemma 3.3 (Multiplicative approximation for $\mathcal{H}(\mu)$)**.** Let μ be a distribution on $2^{[n]}$ such that both μ and μ^* are log-concave. Then

$$\frac{1}{2}\sum_{i\in[n]}\mathcal{H}(\mu_i) \leq \mathcal{H}(\mu) \leq \sum_{i\in[n]}\mathcal{H}(\mu_i)$$

Proof. By Lemma 3.1, we have $\mathcal{H}(\mu) \geq \sum_{i\in[n]} \mu_i \log \frac{1}{\mu_i}, \mathcal{H}(\mu^*) \geq \sum_{i\in[n]}(1-\mu_i) \log \frac{1}{1-\mu_i}$ Since $\mathcal{H}(\mu) = \mathcal{H}(\mu^*)$, we have

$$\mathcal{H}(\mu) \geq \frac{1}{2}\left(\sum_{i\in[n]} \mu_i \log \frac{1}{\mu_i} + \sum_{i\in[n]} (1-\mu_i)\log\frac{1}{1-\mu_i}\right) = \frac{1}{2}\sum_{i\in[n]} \mathcal{H}(\mu_i)$$

■

Thus

Theorem 3.4 (Approximation for $\mathcal{H}(\mu)$)**.** Let μ be a distribution on $2^{[n]}$ such that both μ and μ^* are log-concave. Then

$$\max\left(\frac{1}{2}\sum_{i\in[n]}\mathcal{H}(\mu_i), \sum_{i\in[n]}\mathcal{H}(\mu_i) - r\right) \leq \mathcal{H}(\mu) \leq \sum_{i\in[n]}\mathcal{H}(\mu_i)$$

3.1.1. Topology of Log-Concave distributions

Observe that $g_{\lambda*\mu}(z_1,\dots,z_n) \propto g_\mu(\lambda_1\mu_1,\dots,\lambda_n\mu_n)$. Consequently, by Item 3 of Proposition 1 (resp. Proposition 2), if μ is log-concave (resp. completely log-concave), so is $\lambda * \mu$. Similarly, if μ is r-homogenous, so is $\lambda * \mu$.
Now, note that every vector $v \in \mathcal{P}_\mu$ "extrapolates" $\mathrm{supp}(\mu)$ in a sense [12]: We now seek to "extrapolate" μ to distributions $\widetilde{\mu}$ such that the marginals of $\widetilde{\mu}$ are given by some arbitrary vector $v \in \mathcal{P}_\mu$. The following theorem, proven in [AGM+17],[SV14] does exactly that:

Theorem 3.5. Let μ be a probability distribution on $2^{[n]}$. For any $v \in \mathcal{P}_\mu$, and *any* $\varepsilon > 0$, there exist weights $\lambda_1,\dots,\lambda_n \in \mathbb{R}_{>0}$ such that $|v_i - \Pr_{S\sim\lambda*\mu}(i \in S)| \leq \varepsilon$ for every $i \in [n]$. Furthermore, if v lies in the interior of $\mathcal{P}_\mu$, then one may take $\varepsilon = 0$, i.e. one can find λ such that $\Pr_{S\sim\lambda*\mu}(i\in S) = v_i$ for every $i \in [n]$.

This theorem has the following very important corollary.

Corollary 3.6. Let μ be a log-concave distribution on $2^{[n]}$, and let $p \in \mathcal{P}_\mu$. Then there exists a distribution $\widetilde{\mu}$ on $2^{[n]}$ such that $\mathrm{supp}(\widetilde{\mu}) \subseteq \mathrm{supp}(\mu)$, $\widetilde{\mu}_i = p_i$ for every $i \in [n]$. Moreover, $\widetilde{\mu}$ can be obtained as a limit of distributions $\lambda * \mu$ for a sequence of $\lambda \in \mathbb{R}^n_{>0}$.

Proof Sketch. For any $\varepsilon > 0$, by Theorem 3.5, there exist λ_ε such that $|(\lambda_\varepsilon * \mu)_i - p_i| \leq \varepsilon$ for every $i \in [n]$. Thus passing to a convergent subsequence of such λ's, we obtain a distribution $\widetilde{\mu} = \lim_{\varepsilon\to 0}\lambda_\varepsilon * \mu$ such that $\widetilde{\mu}_i = p_i$.

[12] indeed, v is a non-negative linear combination of the vectors in $\mathrm{supp}(\mu)$

Furthermore, since $\mathrm{supp}(\lambda_\varepsilon * \mu) = \mathrm{supp}(\mu)$ for any $\lambda_\varepsilon \in \mathbb{R}^n_{>0}$, we have that $\mathrm{supp}(\widetilde{\mu}) \subseteq \mathrm{supp}(\mu)$. Furthermore, note that $g_{\lambda * \mu}(z_1, \ldots, z_n) \propto g(\lambda_1 z_1, \lambda_2 z_2, \ldots, \lambda_n z_n)$, and thus by Item 3 of Proposition 1, $\lambda * \mu$ are log-concave distributions, since μ is. By Lemma 1.1, $\widetilde{\mu}$ is also log-concave since it is the limit of log-concave distributions (of degree at most n). ■

Remark. [SV14] showed that for the distribution $\widetilde{\mu}$ defined above, we have

$$\log\left(\inf_{z \in \mathbb{R}^n_{>0}} \frac{g_\mu(z)}{z^p}\right) = \sum_{S \in \mathrm{supp}(\widetilde{\mu})} \widetilde{\mu}(S) \log \frac{\mu(S)}{\widetilde{\mu}(S)} \tag{3.3}$$

In particular, if μ is uniform over its support, then the above quantity evaluates to $\mathcal{H}(\widetilde{\mu})$. Since $\mathrm{supp}(\widetilde{\mu}) \subseteq \mathrm{supp}(\mu)$, we can invoke Proposition 4 to obtain $\mathcal{H}(\widetilde{\mu}) \leq \mathcal{H}(\mu) = \log|\mathrm{supp}(\mu)|$, and consequently, $|\mathrm{supp}(\mu)| \geq \sup_{p \in \mathcal{P}_\mu} \inf_{z \in \mathbb{R}^n_{>0}} \frac{g_\mu(z)}{z^p}$.

§4. Matroid Base Counting

Two of the most fundamental problems in algorithmic matroid theory are counting the number of bases of an arbitrary matroid, and counting the *number of common bases* between two matroids. Indeed, being able to estimate the number of bases of an arbitrary matroid would immediately yield corresponding estimates for the number of forests and spanning subgraphs in a given graph.

Now, a consequence of a result proved in [Sno12] tells us that counting the number of bases of an arbitrary matroid is #P-hard. Thus, we turn our attention to approximation algorithms for the same.

We begin by first adapting our machinery of log-concave distributions to the context of matroids, in light of Theorem 2.2.

Lemma 4.1. Let $M = ([n], \mathcal{I})$ be a matroid, and let $p \in \mathcal{P}_M$. Then there is a distribution $\widetilde{\mu}$ such that $\mathrm{supp}(\widetilde{\mu}) \subseteq \mathcal{B}_M$, $\widetilde{\mu}_i = p_i$ for every $i \in [n]$, and $\widetilde{\mu}, \widetilde{\mu}^*$ are both completely log-concave. Furthermore, $\widetilde{\mu}$ (resp. $\widetilde{\mu}^*$) can be obtained as a limit of external fields applied to μ (resp. μ^*), where μ is the uniform distribution on $\mathcal{B}_M$.

Proof. If μ is the uniform distribution on $\mathcal{B}_M$, then $g_\mu(z) \propto g_M(z)$, and thus $g_\mu(z)$ is completely log-concave by Theorem 2.2. Similarly, μ^* is the uniform distribution on $\mathcal{B}_{M^*}$, where M^* is the dual matroid of M, and thus $g_{\mu^*}(z)$ is completely log-concave too. Further note that $((\lambda_1, \ldots, \lambda_n) * \mu)^* = (\lambda_1^{-1}, \ldots, \lambda_n^{-1}) * \mu^*$. Thus, by Item 3 of Proposition 2, both $\lambda * \mu$ and $(\lambda * \mu)^*$ are completely log-concave.

Let $\widetilde{\mu}$ be the distribution as defined in the proof of Corollary 3.6. Since the set of completely log-concave polynomials is also closed, $\widetilde{\mu}$ is completely log-concave, and so is $\widetilde{\mu}^*$, since if $\widetilde{\mu} = \lim_{\varepsilon \to 0} \lambda_\varepsilon * \mu$, then $\widetilde{\mu}^* = \lim_{\varepsilon \to 0} (\lambda_\varepsilon * \mu)^*$. ■

Theorem 4.2 (Anari-Oveis Gharan-Vinzant's deterministic Matroid Base Counting Algorithm)**.** Let $M = ([n], \mathcal{I})$ be *any* matroid of rank r, and let $\mathcal{O}$ be an independence oracle for M, i.e. given any $S \subseteq [n]$, it tells us, in $\mathcal{O}(1)$ time, if $S \in \mathcal{I}$. Then there is a *deterministic* $\mathrm{poly}(n)$-time algorithm which outputs $\beta \in \mathbb{R}$ such that

$$\max\left(2^{-\mathcal{O}(r)}\beta, \sqrt{\beta}\right) \leq |\mathcal{B}_M| \leq \beta$$

Thus $|\mathcal{B}_M|$ can be approximated within a factor of $2^{\mathcal{O}(r)}$.

Proof. Consider

$$p^{\mathrm{opt}} = \left(p_1^{\mathrm{opt}}, p_2^{\mathrm{opt}}, \ldots, p_n^{\mathrm{opt}}\right) = \underset{p=(p_1,\ldots,p_n)\in\mathcal{P}_M}{\mathrm{argmax}} \sum_{i\in[n]} \mathcal{H}(p_i)$$

As discussed above, p^{opt} can be found in $\mathrm{poly}(n)$ time. Also, let $\tau = \sum_{i\in[n]} \mathcal{H}(p_i^{\mathrm{opt}})$.

Now, let μ be the uniform distribution over $\mathcal{B}_M$. Note that $\mathcal{H}(\mu) = \log|\mathcal{B}_M|$. Now, since $(\mu_1, \ldots, \mu_n) \in \mathcal{P}_M$, $\tau \geq$

$\sum_{i\in[n]} \mathcal{H}(\mu_i) \overset{\text{Eq. (3.1)}}{\geq} \mathcal{H}(\mu)$. Also, since $p^{\text{opt}} \in \mathcal{P}_M = \mathcal{P}_\mu$, by Lemma 4.1, there exists a distribution $\widetilde{\mu}$ such that both $\widetilde{\mu}, \widetilde{\mu}^*$ are completely log-concave, $\text{supp}(\widetilde{\mu}) \subseteq \text{supp}(\mu)$, and $\widetilde{\mu}_i = p_i^{\text{opt}}$. Thus by Theorem 3.4,

$$\mathcal{H}(\widetilde{\mu}) \geq \max\left(\frac{1}{2}\sum_{i\in[n]} \mathcal{H}(\widetilde{\mu}_i), \sum_{i\in[n]} \mathcal{H}(\widetilde{\mu}_i) - r\right)$$

But $\sum_{i\in[n]} \mathcal{H}(\widetilde{\mu}_i) = \sum_{i\in[n]} \mathcal{H}(p_i) = \tau$, and thus $\mathcal{H}(\widetilde{\mu}) \geq \max\left(\frac{\tau}{2}, \tau - r\right)$. Now, since $\text{supp}(\widetilde{\mu}) \subseteq \text{supp}(\mu)$, by Proposition 4, $\mathcal{H}(\mu) \geq \mathcal{H}(\widetilde{\mu})$. Thus $\tau \geq \mathcal{H}(\mu) \geq \mathcal{H}(\widetilde{\mu}) \geq \max\left(\frac{\tau}{2}, \tau - r\right)$. The statement of the theorem follows when one notices that $\mathcal{H}(\mu) = \log|\mathcal{B}_M|$, and sets $\beta = e^\tau$. Furthermore, as promised in the theorem, τ can be calculated, in deterministic polynomial time, through standard convex program solving algorithms such as the *ellipsoid method*. ■

Remark. Azar, Brode, and Frieze ([ABF94]) showed that any deterministic polynomial time algorithm having only independence oracle access to some arbitrary matroid M with rank r can approximate $|\mathcal{B}_M|$ only up to a factor of $2^{\Omega\left(\frac{r}{\log^2 n}\right)}$, provided $r \gg \log n$. Thus the above algorithm is almost optimal.

Corollary 4.3. Let M be an arbitrary matroid, and assume we have an independence oracle for it. Then for any $k \in \mathbb{N}$, we have a deterministic polynomial time algorithm to calculate a number β such that

$$\max\left(2^{-\mathcal{O}(k)}\beta, \sqrt{\beta}\right) \leq |\mathcal{I}_M^k| \leq \beta$$

where $\mathcal{I}_M^k := \{I \in \mathcal{I} : |I| = k\}$.

Proof. Follows from the fact that the k-truncation of a matroid is also a matroid. ■

Before the next algorithm, we state (without proof [13]), an analytic statement about log-concave polynomials.

Lemma 4.4. Let $g \in \mathbb{R}[y_1, \ldots, y_n, z_1, \ldots, z_n]$ be a completely log-concave, multilinear polynomial. Let $p \in [0,1]^n$. Then

$$\left(\prod_{i=1}^{n}(\partial_{y_i} + \partial_{z_i})\right) g(y,z)\Big|_{y=z=0} \geq \left(\frac{p}{e^2}\right)^p \inf_{y,z\in\mathbb{R}^n_{>0}} \frac{g(y,z)}{y^p z^{1-p}}$$

where $y = (y_1, \ldots, y_n), z = (z_1, \ldots, z_n)$ and $1 - p = (1 - p_1, \ldots, 1 - p_n)$.

Theorem 4.5 (Anari-Oveis Gharan-Vinzant's deterministic Matroid Base Intersection Counting Algorithm)**.** Let M, N be two matroids on the same ground set $[n]$ such that $\text{rank}(M) = \text{rank}(N) = r$. Also assume that we have an

[13] the proof is just "analytic bashing", and isn't particularly insightful, which is why it was skipped

independence oracle for both M, N. Then there is a *deterministic* $\text{poly}(n)$-time algorithm which outputs $\beta \in \mathbb{R}$ such that

$$2^{-\mathcal{O}(r)}\beta \leq |\mathcal{B}_M \cap \mathcal{B}_N| \leq \beta$$

Thus $|\mathcal{B}_M \cap \mathcal{B}_N|$ can be approximated within a factor of $2^{\mathcal{O}(r)}$.

Proof. Note that $\mathcal{P}_M \cap \mathcal{P}_N$ is a convex polytope, and it is as easy to optimize concave functions over $\mathcal{P}_M \cap \mathcal{P}_N$ as it is over $\mathcal{P}_M$. Thus, we can, in polynomial time, calculate $\tau = \max_{p \in \mathcal{P}_M \cap \mathcal{P}_N} \sum_{i \in [n]} \mathcal{H}(p_i)$. As usual, let μ be the uniform distribution on $\mathcal{B}_M \cap \mathcal{B}_N$. Then $(\mu_1, \ldots, \mu_n) \in \mathcal{P}_M \cap \mathcal{P}_N$, and consequently, $\tau \geq \sum_{i \in [n]} \mathcal{H}(\mu_i) \overset{\text{Eq. (3.1)}}{\geq} \mathcal{H}(\mu) = \log |\mathcal{B}_M \cap \mathcal{B}_N|$.

Now, let $g_M(y)$ be the generating polynomial of M, and let $g_{N^*}(z)$ be the generating polynomial of N^*, the dual matroid of N. Then $g_M(y)g_{N^*}(z)$ is the generating polynomial of $M \oplus N^*$. Further, note that

$$|\mathcal{B}_M \cap \mathcal{B}_N| = \sum_{S \subseteq [n]} \underbrace{\left(\prod_{i \in S} \partial_{y_i}\right) g_M(y)}_{=\mathbb{1}_{S \in \mathcal{I}_M}} \underbrace{\left(\prod_{j \in [n] \setminus S} \partial_{z_j}\right) g_{N^*}(z)}_{=\mathbb{1}_{[n] \setminus S \in \mathcal{I}_{N^*}}} = \left(\prod_{i=1}^{n} (\partial_{y_i} + \partial_{z_i})\right) g_M(y) g_{N^*}(z)$$

where the first equality follows since $\mathbb{1}_{S \in \mathcal{I}_M} \cdot \mathbb{1}_{[n] \setminus S \in \mathcal{I}_{N^*}} = \mathbb{1}_{S \in \mathcal{B}_M \cap \mathcal{B}_N}$.

Thus, $\left(\prod_{i=1}^n (\partial_{y_i} + \partial_{z_i})\right) g_M(y) g_{N^*}(z)$ is a constant function, and consequently, $\left(\prod_{i=1}^n (\partial_{y_i} + \partial_{z_i})\right) g_M(y) g_{N^*}(z)\Big|_{y=z=0} = \left(\prod_{i=1}^n (\partial_{y_i} + \partial_{z_i})\right) g_M(y) g_{N^*}(z)$. Now, applying Lemma 4.4, yields

$$\left(\prod_{i=1}^{n} (\partial_{y_i} + \partial_{z_i})\right) g_M(y) g_{N^*}(z)\Bigg|_{y=z=0} \geq \left(\frac{p}{e^2}\right)^p \cdot \inf_{y,z \in \mathbb{R}^n_{>0}} \frac{g_M(y) g_{N^*}(z)}{y^p z^{1-p}} = \left(\frac{p}{e^2}\right)^p \cdot \inf_{y \in \mathbb{R}^n_{>0}} \frac{g_M(y)}{y^p} \cdot \inf_{z \in \mathbb{R}^n_{>0}} \frac{g_{N^*}(z)}{z^{1-p}}$$

$$\implies \log |\mathcal{B}_M \cap \mathcal{B}_N| \geq \log \left(\frac{p}{e^2}\right)^p + \log \left(\inf_{y \in \mathbb{R}^n_{>0}} \frac{g_M(y)}{y^p}\right) + \log \left(\inf_{z \in \mathbb{R}^n_{>0}} \frac{g_{N^*}(z)}{z^{1-p}}\right) \tag{4.1}$$

Now, let $p \in \mathcal{P}_M \cap \mathcal{P}_N$ be some arbitrary vector. By Lemma 4.1, there exist distributions ν, ω such that $\nu, \nu^*, \omega, \omega^*$ are completely log-concave distributions and for every $i \in [n]$, $\nu_i = p_i, \omega_i = 1 - p_i$. By Eq. (3.3) and the remark accompanying it, we know that $\log \left(\inf_{y \in \mathbb{R}^n_{>0}} \frac{g_M(y)}{y^p}\right) = \mathcal{H}(\nu)$. Furthermore, since $p \in \mathcal{P}_M \cap \mathcal{P}_N$, $p \in \mathcal{P}_N$, which further implies that $1 - p \in \mathcal{P}_{N^*}$, which then implies that $\log \left(\inf_{z \in \mathbb{R}^n_{>0}} \frac{g_{N^*}(z)}{z^{1-p}}\right) = \mathcal{H}(\omega^*) = \mathcal{H}(\omega)$. Now, note that the marginals of both ν, ω^* are p: Then by Lemma 3.1, we have $\min \left(\mathcal{H}(\nu), \mathcal{H}(\omega^*)\right) \geq \sum_{i \in [n]} p_i \log \frac{1}{p_i}$. Thus, simplifying Eq. (4.1),

$$\log |\mathcal{B}_M \cap \mathcal{B}_N| \geq \log \left(\frac{p}{e^2}\right)^p + \log \left(\inf_{y \in \mathbb{R}^n_{>0}} \frac{g_M(y)}{y^p}\right) + \log \left(\inf_{z \in \mathbb{R}^n_{>0}} \frac{g_{N^*}(z)}{z^{1-p}}\right) = \sum_{i \in [n]} p_i \log p_i - 2 \sum_{i \in [n]} p_i + \mathcal{H}(\nu) + \mathcal{H}(\omega^*)$$

$$\geq \sum_{i \in [n]} p_i \log p_i - 2 \sum_{i \in [n]} p_i + 2 \sum_{i \in [n]} p_i \log \frac{1}{p_i} = \sum_{i \in [n]} p_i \log \frac{1}{p_i} - 2 \sum_{i \in [n]} p_i = \sum_{i \in [n]} \mathcal{H}(p_i) - \sum_{i \in [n]} (1 - p_i) \log \frac{1}{1 - p_i} - 2 \sum_{i \in [n]} p_i$$

Now, $\sum_{i \in [n]} (1 - p_i) \log \frac{1}{1 - p_i} \leq r$ by Eq. (3.2). Also, $\ell : \mathbb{R}^n \mapsto \mathbb{R} : \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^\top \mapsto \sum_{i \in [n]} x_i$ is a continuous convex function, and thus ℓ is maximized over some boundary point of $\mathcal{P}_M \cap \mathcal{P}_N$ since $\mathcal{P}_M \cap \mathcal{P}_N$ is convex and compact. But note that ℓ is equal to r on the boundary of $\mathcal{P}_M \cap \mathcal{P}_N$. Thus $\sum_{i\in[n]} p_i \le r$, and thus $\log|\mathcal{B}_M \cap \mathcal{B}_N| \ge \sum_{i\in[n]} \mathcal{H}(p_i) - 3r$ for every $p \in \mathcal{P}_M \cap \mathcal{P}_N$, and consequently $\log|\mathcal{B}_M \cap \mathcal{B}_N| \ge \max_{p\in\mathcal{P}_M\cap\mathcal{P}_N} \sum_{i\in[n]} \mathcal{H}(p_i) - 3r = \tau - 3r$. Thus $\tau \ge \log|\mathcal{B}_M \cap \mathcal{B}_N| \ge \tau - 3r$, and the theorem follows by setting $\beta = e^\tau$. ■

Remark. There are a few easy extensions of the result above:

1. The matroid common base problem is "self-reducible": Informally, this means that the problem of counting the number of bases which include $i_1, \ldots, i_k \in [n]$ and exclude $j_1, \ldots, j_m \in [n]$, reduces to the problem of counting the number of bases between matroids, obtained by contracting $i_1, \ldots, i_k \in [n]$ and deleting $j_1, \ldots, j_m \in [n]$. Then from a result of Sinclair and Jerrum ([SJ89]), we get that there exists a randomized algorithm, which, given two parameters $\varepsilon, \delta > 0$, outputs a β such that $\Pr((1-\varepsilon)\beta \le |\mathcal{B}_M \cap \mathcal{B}_N| \le \beta) \ge 1 - \delta$, i.e. there exists a randomized algorithm to approximate $|\mathcal{B}_M \cap \mathcal{B}_N|$ with arbitrary precision.
   Furthermore, the runtime of this algorithm is $2^{\mathcal{O}(r)}\,\mathrm{poly}(n, \frac{1}{\varepsilon}, \log\frac{1}{\delta})$. Consequently, if $r = \mathcal{O}(\log n)$, then this algorithm is a **F**ully **P**olynomial-time **R**andomized **A**pproximation **S**cheme (**FPRAS**).

2. Since the framework of generating polynomials can be easily transported to a weighted setting, we have a deterministic polynomial time algorithm, which, given any $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n_{\ge 0}$, outputs β such that

$$2^{-\mathcal{O}(r)}\beta \le \sum_{B\in\mathcal{B}_M\cap\mathcal{B}_N} \prod_{i\in B} \lambda_i \le \beta$$

§5. Simplicial Complexes, an FPRAS for Matroid Base Counting, and the Mihail-Vazirani Conjecture

We first define a generalization of matroids.

Definition 5.1 (Simplicial Complexes)**.** A simplicial complex $X = ([n], \Sigma)$ is said to be defined over the *ground set* $[n] = \{1, 2, \ldots, n\}$, and is characterized by its non-empty collection of *faces* $\Sigma \subseteq 2^{[n]}$ which satisfy the following property:

1. *Downward Closed*: $\sigma \in \Sigma \implies 2^{\sigma} \subseteq \Sigma$, i.e. if σ is a face, then every subset of σ is also a face. In particular, since Σ is non-empty, it must contain $\varnothing$.

Let $X = (U, \Sigma)$ be an arbitrary simplicial complex:

1. The dimension of X is defined to be the size of its largest face.
2. For any integer k, $X(k) := \{\sigma \in \Sigma : |\sigma| = k\}$.
3. For a face $\tau \in \Sigma$, the *link* of τ is defined to be the simplicial complex $X_\tau := (U \setminus \tau, \{\sigma \setminus \tau : \sigma \in \Sigma, \sigma \supseteq \tau\})$. Note that $X_\varnothing = X$.

We say X is *pure* if all maximal faces in X have the same size. A pure d-dimensional simplicial complex $X = (U, \Sigma)$ is said to be d-partite if there exists a partition $U_1, \ldots, U_d$ of U such that for every maximal face σ of X, $|\sigma \cap U_i| = 1$ for every $i \in [d]$.
We often consider *weighted simplicial complexes*: Let $X = (U, \Sigma)$ be a pure d-dimensional simplicial complex. Consider a weight function $w : X(d) \mapsto \mathbb{R}_{>0}$. Then:

1. For any $\tau \in \Sigma$, we extend w to τ as $w(\tau) := \sum_{\sigma \in X(d) : \sigma \supseteq \tau} w(\sigma)$ [14].
2. For any $\tau \in \Sigma$, given any maximal face σ' of X_τ, we endow it with weight $w_\tau(\sigma') := w(\tau \cup \sigma')$. We can then extend w_τ to X_τ as above.

The *1-skeleton* of the link X_τ is defined to be a weighted graph as follows: Every $i \in U \setminus \tau$ such that $\{i\}$ is a face of X_τ, is a vertex of our graph. Two vertices $i, j, i \neq j$ are connected if $\{i, j\}$ is a face of X_τ. The edge $\{i, j\}$ has weight $w_\tau(\{i, j\})$.

[14] note that the weight function is strictly positive. Thus the weight of every face in the complex is non-zero

5.1. Walks on Simplicial Complexes

As usual, let X be a pure d-dimensional simplicial complex.
We define a random walk, known as the *lower k-walk*, on $X(k+1)$ [15] as follows: Suppose our walk is currently at $\sigma \in X(k+1)$.

1. Remove a uniformly random element $i \in \sigma$ from σ.

2. Add a $j \notin \sigma \setminus \{i\}$ to $\sigma \setminus \{i\}$ with probability proportional to $w\left((\sigma \setminus \{i\}) \cup \{j\}\right)$.

The transition probabilities for this walk can be written as (σ, σ' are arbitrary elements of $X(k+1)$):

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

We can also define a counterpart of the above random walk, known as the *upper k-walk* [16], which is defined over $X(k)$ as follows. Suppose we have some $\sigma \in X(k)$. Then:

1. Consider $\mathcal{T} := \{\tau \in X(k+1) : \tau \supset \sigma\}$, i.e. the set of all $(k+1)$-dimensional faces which contain σ. Sample τ from $\mathcal{T}$ with probability proportional to $w(\tau)$, i.e. τ is sampled with probability $\frac{w(\tau)}{\sum_{\eta \in \mathcal{T}} w(\eta)}$.

2. Delete one of the $k+1$ elements of τ uniformly at random.

Similar to the above calculation, the transition probabilities are (σ, σ' are arbitrary elements of $X(k)$):

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

First off, note that both the random walks defined above are *reversible*, i.e. for every $\sigma, \sigma' \in X(k)$, we have $w(\sigma)P^{\wedge}_k(\sigma, \sigma') = w(\sigma')P^{\wedge}_k(\sigma', \sigma)$ and $w(\sigma)P^{\vee}_k(\sigma, \sigma') = w(\sigma')P^{\vee}_k(\sigma', \sigma)$. Furthermore, we also get that $P^{\wedge}_k, P^{\vee}_k$ have the same stationary distribution [17], i.e. the probability of $\tau \in X(k)$ is proportional to $w(\tau)$.
We now show that both the aforementioned walks have the same spectra.

Lemma 5.1. For any $k \in [d-1]$, $P^{\wedge}_k$ and $P^{\vee}_{k+1}$ have the same (with multiplicity) non-zero eigenvalues.

[15] where $k \in [d-1]$
[16] again, $k \in [d-1]$
[17] if $\mathcal{M}$ is a finite reversible Markov chain, and π satisfies the detailed balance conditions for $\mathcal{M}$, then π is a stationary distribution for $\mathcal{M}$. If $\mathcal{M}$ is irreducible, then π is the unique stationary distribution. Note that WLOG we'll be assuming the underlying graphs in $P^{\wedge}_k, P^{\vee}_k$ are connected

Proof. Construct a bipartite graph $G_k = (X(k) \sqcup X(k+1), E)$. We connect $\tau \in X(k)$ with $\sigma \in X(k+1)$ if $\tau \subset \sigma$. Then the transition matrix for the simple random walk on G_k is $P_k = \begin{bmatrix} 0 & P_k^{\downarrow} \\ P_k^{\uparrow} & 0 \end{bmatrix}$ where $P_k^{\downarrow} \in \mathbb{R}^{X(k+1)\times X(k)}$, $P_k^{\uparrow} \in \mathbb{R}^{X(k)\times X(k+1)}$ are stochastic matrices (i.e. their rows sum to 1). Furthermore, observe that $P_k^{\wedge} = P_k^{\uparrow}P_k^{\downarrow}, P_{k+1}^{\vee} = P_k^{\downarrow}P_k^{\uparrow}$: Indeed, in $P_k^{\wedge}$ we take a suitably weighted sample of a higher-dimensional face before coming back down, while in $P_{k+1}^{\vee}$ we do the reverse.

By elementary linear algebra, for any two matrices A, B such that AB and BA are defined, the non-zero spectra of AB and BA are identical, and consequently $P_k^{\wedge}$ and $P_{k+1}^{\vee}$ have the same non-zero eigenvalues (with the same multiplicities). ■

Now, let us take a closer look at $P_1^{\wedge}$: $P_1^{\wedge}$ is a walk on $X(1)$, i.e. the "vertices" of X. Furthermore, for any vertex $v \in X(1)$, $P_1^{\wedge}(v,v) = \frac{1}{2}$, and for any other vertex $w \in X(1)$, the probability of transitioning to w is proportional to $w(\{v,w\})$. In other words, $P_1^{\wedge}$ is the lazy random walk on the 1-skeleton of $X_{\varnothing} = X$ [18]. For purposes of our study, we'll also have to define the non-lazy random walk on X:

$$\widetilde{P}_1^{\wedge} := 2(P_1^{\wedge} - I/2)$$

Similarly, let $P_{\tau,1}^{\wedge}$ be the upper random walk on the 1-skeleton of X_τ, and let $\widetilde{P}_{\tau,1}^{\wedge} := 2(P_{\tau,1}^{\wedge} - I/2)$ be the non-lazy version of $P_{\tau,1}^{\wedge}$.

At this point, we also define the very useful notion of *local spectral expansion*, which was introduced by [KO20]:

Definition 5.2 (Local Spectral Expanders)**.** Given a pure d-dimensional weighted simplicial complex (X, w), we call X a λ-local-spectral expander if $\lambda_2(\widetilde{P}_{\tau,1}^{\wedge}) \le \lambda$ for every $\tau \in X(k)$, for every $0 \le k \le d-2$.

Remark. Note that $\lambda_2(\widetilde{P}_{\tau,1}^{\wedge})$ refers to the second largest eigenvalue of $\widetilde{P}_{\tau,1}^{\wedge}$.

We now connect the property of being a local-spectral-expander to the spectral properties of $P_k^{\wedge}$.

Theorem 5.2. Let (X, w) be a pure d-dimensional weighted simplicial complex which is also a 0-local-expander. Fix some k, where $0 \le k < d$. Then, for all $-1 \le i \le k$, $P_k^{\wedge}$ has at most $|X(i)| \le \binom{n}{i}$ eigenvalues of value $> 1 - \frac{i+1}{k+1}$.

Remark. We set $X(-1) := \varnothing, \binom{n}{-1} := 0$. Note that putting $i = 0$ in the above theorem yields that $P_k^{\wedge}$ has at most 1 eigenvalue greater than $\frac{k}{k+1}$. Indeed, since $P_k^{\wedge}$ has 1 as an eigenvalue on the account of being stochastic, we get that the second largest eigenvalue of $P_k^{\wedge}$ is at most $\frac{k}{k+1}$.

[18] Since $P_1^{\wedge}$, and in general, $P_k^{\wedge}$ are lazy random walks on weighted graphs, they have real eigenvalues by standard Markov chain theory

Before proving this theorem, we first prove an auxiliary lemma. For the lemma, we define a new inner product on $\mathbb{R}^{X(k)}$, in which we simply reweight every $X(k)$ by the appropriate weight function, i.e.

$$\langle \phi, \psi \rangle_* := \sum_{\tau \in X(k)} w(\tau)\phi(\tau)\psi(\tau)$$

Lemma 5.3. Let (X, w) be as in Theorem 5.2. Then $P_k^\wedge \preceq_* \frac{k}{k+1} P_k^\vee + \frac{1}{k+1} I$ for every $0 \le k < d$, where $\preceq_*$ is defined w.r.t the inner product $\langle \cdot, \cdot \rangle_*$.

Proof. Set $M := P_k^\wedge - \frac{k}{k+1} P_k^\vee - \frac{1}{k+1} I$. Fix an arbitrary $\eta \in X(k-1)$. Construct the matrix M_η as follows:

$$M_\eta(\tau, \sigma) := \begin{cases} M(\tau, \sigma), & \text{if } \tau \neq \sigma, \eta = \tau \cap \sigma \\ -\frac{1}{k+1} \cdot \frac{w(\tau)}{w(\eta)}, & \text{if } \tau = \sigma, \eta \subset \tau \\ 0, & \text{otherwise} \end{cases}$$

Some calculation reveals that $M = \sum_{\eta \in X(k-1)} M_\eta$. Thus, we'll show that $M_\eta \preceq_* 0$, and be done.
Now, note that if $\tau \neq \sigma$, and $\tau \cap \sigma = \eta \in X(k-1)$, then

$$M_\eta(\tau, \sigma) = M(\tau, \sigma) = \frac{1}{k+1}\left(\frac{w(\tau \cup \sigma)}{w(\tau)} - \frac{w(\sigma)}{w(\tau \cap \sigma)}\right) = \frac{1}{(k+1)w(\eta)} \cdot w(\tau)^{-1} \cdot \big(w(\eta)w(\tau \cup \sigma) - w(\tau)w(\sigma)\big)$$

Also, for $\tau \in X(k)$ with $\tau \supset \eta$,

$$M_\eta(\tau, \tau) = \frac{-w(\tau)}{(k+1)w(\eta)} = \frac{1}{(k+1)w(\eta)} \cdot w(\tau)^{-1} \cdot \big(0 - w(\tau) \cdot w(\tau)\big)$$

Given the above expressions, it's not too hard to see that:

$$M_\eta = \frac{1}{(k+1)w(\eta)} \cdot \operatorname{diag}(w_\eta)^{-1}\left(w(\eta) \cdot A_\eta - w_\eta w_\eta^\intercal\right)$$

where w_η is a $|X(k)|$-dimensional vector whose non-zero entries are given by $w(\tau)$ for $\tau \supset \eta$, and A_η is a $|X(k)| \times |X(k)|$ matrix whose non-zero entries are given by $w(\tau \cup \sigma)$ for $\tau, \sigma \in X(k)$ satisfying $\tau \cap \sigma = \eta$.
Now, clearly $M_\eta \preceq_* 0$ if and only if $\operatorname{diag}(w_k)M_\eta \preceq 0$, where w_k is a $|X(k)|$-dimensional vector indexed by $w(\sigma)$ for every $\sigma \in X(k)$. But note that $\operatorname{diag}(w_k)M_\eta = \operatorname{diag}(w_\eta)M_\eta$, and thus it suffices to show that $A_\eta \preceq \frac{w_\eta w_\eta^\intercal}{w(\eta)}$.
Now, note that A_η is the weighted adjacency matrix of the 1-skeleton of X_η. In that light, it is not difficult to see that $\widetilde{P}_{\eta,1}^\wedge = \frac{1}{k+1} \operatorname{diag}(w_\eta)^{-1} A_\eta$, since $\widetilde{P}_{\eta,1}^\wedge$ is the random walk matrix on the same graph. Since (X, w) is a 0-local-spectral expander, $\widetilde{P}_{\eta,1}^\wedge$ has atmost one positive eigenvalue, and consequently by Lemma A.2, A_η has atmost 1 positive eigenvalue. A simple application of Lemma A.3 then shows that $A_\eta \preceq \frac{w_\eta w_\eta^\intercal}{w(\eta)}$. ■

Proof of Theorem 5.2. We induct on k. When $k = 1$, $P_1^\wedge = \frac{\widetilde{P}_1^\wedge + I}{2}$. Since (X, w) is a 0-local-spectral expander, $\widetilde{P}_1^\wedge$ has exactly one positive eigenvalue, which is 1. Thus $P_1^\wedge$ has eigenvalue 1 with multiplicity 1, and all other eigenvalues of $P_1^\wedge$ are ≤ 0. In particular, we have $|X(1)| - 1$ many eigenvalues $\le 0 < \frac{1}{2}$, and thus the base case is proved.

Suppose the claim holds for some $k < d-1$. By Lemma 5.1, $P_{k+1}^{\vee}$ and $P_k^{\wedge}$ have the same non-zero eigenvalues, and thus, by the induction hypothesis, $P_k^{\vee}$ has at most $|X(i)|$ eigenvalues $> 1 - \frac{i+1}{k+1}$ for $-1 \le i \le k$. Since by Lemma 5.3, $P_{k+1}^{\wedge} \preceq \frac{k+1}{k+2} P_{k+1}^{\vee} + \frac{1}{k+2} I$, $P_{k+1}^{\wedge}$ has at most $|X(i)|$ eigenvalues $> \frac{k+1}{k+2} \cdot \left(1 - \frac{i+1}{k+1}\right) + \frac{1}{k+2} = 1 - \frac{i+1}{k+2}$. For $i = k+1$, $P_{k+1}^{\wedge}$ trivially has $\le |X(k+1)|$ eigenvalues > 0, since $P_{k+1}^{\wedge}$ is a $|X(k+1)| \times |X(k+1)|$ matrix. ∎

5.2. Polynomials, Distributions and Simplicial Complexes

To apply our log-concave machinery to the simplicial complex setting, we need to set up the basic connections first, which is what we now do.

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be a multilinear d-homogenous polynomial, i.e. $p = \sum_{S \in \binom{[n]}{d}} c_S x^S = \sum_{S \in \binom{[n]}{d}} c_S \prod_{i \in S} x_i$. Define a d-dimensional (pure) simplicial complex X^p, where $X^p(d) := \left\{ S \in \binom{[n]}{d} : c_S \neq 0 \right\}$, and then "complete" the simplicial complex by taking all subsets of faces in $X^p(d)$. Similarly, assign $w(S) := c_S$ for $S \in X^p(d)$, and then extend w to all faces of X^p in the usual way, i.e. $w(\tau) = \sum_{\sigma \in X^p(d): \sigma \supseteq \tau} w(\sigma)$, for any face τ. Then:

Proposition 5. *For any $0 \le k \le d$, and any $\sigma \in X^p(k)$, $w(\sigma) = (d-k)! p_\sigma(\mathbb{1})$, where $p_\sigma := \left(\prod_{i \in \sigma} \partial_i \right) p$.*

Proof. We prove this by induction on $d-k$ (or "reverse induction on k"). If $k = d$, $p_\sigma(\mathbb{1}) = c_\sigma$, and we're done. So suppose the statement holds for all $\tau \in X^p(k+1)$, and consider some $\sigma \in X^p(k)$. Then

$$w(\sigma) = \sum_{\eta \in X^p(d): \eta \supset \sigma} w(\eta) = \sum_{\tau \in X^p(k+1): \tau \supset \sigma} w(\tau) = \sum_{i \in X_\sigma^p(1)} w(\sigma \cup i) = (d-k-1)! \sum_{i \in X_\sigma^p(1)} p_{\sigma \cup i}(\mathbb{1})$$

$$= (d-k-1)! \sum_{i \in X_\sigma^p(1)} (\partial_i p_\sigma)(\mathbb{1}) = (d-k-1)! \left(\left(\sum_{i \in X_\sigma^p(1)} \partial_i \right) p_\sigma \right)(\mathbb{1}) \overset{\text{Lemma A.4}}{=} (d-k)! p_\sigma(\mathbb{1})$$

∎

We now connect log-concavity to the spectral properties of simplicial complexes.

Lemma 5.4. *Let p be a multilinear completely log-concave polynomial. Then X^p is a 0-local spectral expander.*

Proof. Define:

$$\widetilde{\nabla}^2 p_\tau := \frac{1}{d-k-1} \operatorname{diag}\left((\nabla p)(\mathbb{1}) \right)^{-1} \left(\nabla^2 p_\tau \right)(\mathbb{1})$$

We claim that $\widetilde{\nabla}^2 p_\tau = \widetilde{P}_{\tau,1}^{\wedge}$: Indeed, note that:

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

On the other hand,

$$(\widetilde{\nabla}^2 p_\tau)(i,j) = \frac{(\partial_i \partial_j p_\tau)(\mathbb{1})}{(d-k-1)\cdot(\partial_i p_\tau)(\mathbb{1})}$$

But the above expressions are equal by Proposition 5.

Now, the non-zero entries of $\operatorname{diag}\left((\nabla p)(\mathbb{1})\right)$ are the coefficients of p, which are non-negative [19], and thus $\operatorname{diag}\left((\nabla p)(\mathbb{1})\right)$ is PSD. Furthermore, since p is completely log-concave, $\left(\nabla^2 p\right)(\mathbb{1})$ has at most one positive eigenvalue by Theorem 1.2. Thus by Lemma A.2, $\widetilde{\nabla}^2 p_\tau$ has at most one positive eigenvalue, and thus $\widetilde{P}^{\wedge}_{\tau,1}$ has at most one positive eigenvalue, implying that $\lambda_2(\widetilde{P}^{\wedge}_{\tau,1}) \leq 0$. Since τ was arbitrary, we get that X^p is a 0-local spectral expander. ∎

5.3. An FPRAS for Matroid Base Counting

We can now finally construct an FPRAS for Matroid Base Counting. But before that, we quickly recall a very standard result from the theory of Markov chains:

Theorem 5.5. Let $\mathcal{M} := (\Omega, P, \pi)$ be an irreducible and reversible Markov chain with stationary distribution π, and let $\tau \in \Omega$ and $\varepsilon > 0$ be arbitrary. Then

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

where $\lambda^*(P) := \max\{|\lambda_2|, |\lambda_n|\}$ is the second eigenvalue of $\mathcal{M}$, which has eigenvalues $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq -1$, and $t_\tau(\varepsilon) := \min\{t \in \mathbb{Z}_{\geq 0} : \|P^t(\tau, \cdot) - \pi\|_1 \leq \varepsilon\}$.

Let $\mu : 2^{[n]} \mapsto \mathbb{R}_{\geq 0}$ be a d-homogenous completely log-concave distribution. We turn it into a weighted pure d-dimensional simplicial complex X^μ in the usual way, with $X^\mu(d) := \operatorname{supp}(\mu)$, and $w(\sigma) := \mu(\sigma)$ for any $\sigma \in X^\mu(d)$, and we extend $X^\mu(d)$ and w to construct our weighted simplicial complex.

Lemma 5.6. Let $\mu : 2^{[n]} \mapsto \mathbb{R}_{\geq 0}$ be a d-homogenous completely log-concave distribution, and consider X^μ as defined above. Consider the lower random walk $P_d^{\vee}$ on $X^\mu(d) = \operatorname{supp}(\mu)$, which we start from τ. Then

$$t_\tau(\varepsilon) \leq d \log\left(\frac{1}{\varepsilon \mu(\tau)}\right)$$

Proof. By Theorem 5.5 it is enough to show that $\lambda^*(P_d^{\vee}) \leq 1 - 1/d$. Since μ is completely log-concave, X^p is a 0-local-spectra-expander by Lemma 5.4. Thus,

$$\lambda^*(P_d^{\vee}) \overset{\text{Lemma 5.1}}{=} \lambda^*(P_{d-1}^{\wedge}) \overset{\text{Theorem 5.2}}{\leq} 1 - \frac{1}{(d-1)+1} = 1 - \frac{1}{d}$$

[19] recall that a completely log-concave polynomial must have non-negative coefficients

as desired. ■

At this point, we are done: Indeed, for any matroid $M = ([n], \mathcal{I})$ with set of bases $\mathcal{B}_M$, construct a graph $\mathcal{G}_M$ over $\mathcal{B}_M$, where two bases B, B' are connected if $|B \triangle B'| = |(B \setminus B') \cup (B' \setminus B)| = 2$. By the basis exchange property of matroids, $\mathcal{G}_M$ is connected.

Now, let μ be the uniform distribution over $\mathcal{B}_M$. By Theorem 2.2, μ is completely log-concave. Thus, by Lemma 5.6, the lower random walk on the maximal faces of X^μ mixes fast. But the lower random walk on the maximal faces of X^μ is precisely a walk on $\mathcal{B}_M$, and furthermore, this walk converges on the uniform distribution over $\mathcal{B}_M$! This walk is also known as the *basis exchange walk*, and we have the following theorem about it:

Theorem 5.7 (FPRAS for Matroid Base Counting)**.** For any matroid $M = ([n], \mathcal{I})$ of rank r, any basis B of M, and any $\varepsilon \in (0, 1)$, the mixing time of the basis exchange walk, starting at B is

$$t_B(\varepsilon) \leq r \log\left(\frac{n^r}{\varepsilon}\right) \leq r^2 \log\left(\frac{n}{\varepsilon}\right) \leq n^2 \log\left(\frac{n}{\varepsilon}\right)$$

Thus the basis exchange walk converges to the uniform distribution over $\mathcal{B}_M$ in $\text{poly}(n, \log \frac{1}{\varepsilon})$ time. Equivalently, we can sample (with ε-ℓ_1 error) from the uniform distribution over matroid bases in $\text{poly}(n, \log \frac{1}{\varepsilon})$ time. Furthermore, for any $\epsilon, \delta \in (0, 1)$, we can produce an (randomized) estimate 'β' of $|\mathcal{B}_M|$, in $\text{poly}(n, r, \frac{1}{\epsilon}, \log \frac{1}{\delta})$ time, such that $\Pr((1-\epsilon)\beta \leq |\mathcal{B}_M| \leq (1+\epsilon)\beta) \geq 1 - \delta$. In other words, we have an FPRAS for calculating $|\mathcal{B}_M|$.

Proof. We can conclude from the above discussion and the fact that $\frac{1}{\mu(B)}$ = Number of bases of $M \leq \binom{n}{r} \leq n^r$. The randomized algorithm calculating β is simply the Sinclair-Jerrum theorem ([SJ89]) [20], which says that if we can sample (in polynomial time) the uniform distribution over a set to some given error, then we can also estimate the size of that set very accurately with high probability. ■

The above result is pathbreaking, for it immediately resolves a host of open questions, which we shall see now.

5.4. Applications of Theorem 5.7

5.4.1. The Mihail-Vazirani Conjecture

Given a simple (unweighted) undirected graph $G(V, E)$, we define the *expansion* [21] of a set $S \subseteq V$ to be

$$h(S) := \frac{E(S, \overline{S})}{|S|}$$

The expansion of the whole graph is defined to be $h(G) := \min_{|S| \leq |\overline{S}|} h(S)$. Consider the basis exchange graph $\mathcal{G}_M$ defined in the previous section, where two bases B, B' are connected if $|B \triangle B'| = 2$. In the 1990s, Mihail and Vazirani conjectured that $h(\mathcal{G}_M) \geq 1$ for *every* matroid. We will prove that now.

[20] which we also saw when we demonstrated an FPRAS for the common base problem (see Theorem 4.5 and the remarks following it).
[21] note that expansion is different from conductance. However, for regular graphs, expansion is proportional to conductance.

Before that, we recall the very famous Cheeger's inequality from spectral graph theory: Let $H(V, E, w)$ be a weighted graph, with the weight of an edge e being $w(e)$. For any $v \in V$, define $w(v) := \sum_{v \in e} w(e)$ [22]. We call a weighted graph d-regular if $w(v) = d$ for every $v \in V$. Then:

Theorem 5.8 (Weighted Cheeger's inequality). For any d-regular weighted graph $H(V, E, w)$,

$$\frac{d - \lambda_2}{2} \leq \Phi(H) \leq \sqrt{2(d - \lambda_2)}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of A_H, where $A_H(i, j) := w(\{i, j\})$ if $\{i, j\}$ is an edge, and 0 otherwise.

Recall that $\Phi(H)$ is the conductance of H, where

$$\Phi(S) := \frac{w(E(S, \overline{S}))}{\text{vol}(S)} = \frac{\sum_{e \in E(S, \overline{S})} w(e)}{\sum_{v \in S} w(v)}, \Phi(H) := \min_{\text{vol}(S) \leq \text{vol}(\overline{S})} \Phi(S)$$

Theorem 5.9 (Mihail-Vazirani Theorem). For any matroid M, the expansion of its basis exchange graph is at least 1.

Proof. Define the usual simplicial complex on M, and let the basis exchange graph be denoted as $\mathcal{G}_M$. Also, let $\text{rank}(M) = r$.

Now, for any $\tau_1 \neq \tau_2, \tau_3 \neq \tau_4$, if $P_r^{\vee}(\tau_1, \tau_2)$ and $P_r^{\vee}(\tau_3, \tau_4)$ are both non-zero, then they must be equal, since the simplicial complex assigned equal weight to all its bases. Furthermore, for any τ, τ', $P_r^{\vee}(\tau, \tau) = P_r^{\vee}(\tau', \tau')$. Thus, if we write $\xi := P_r^{\vee}(\tau, \tau)$, then $P_r^{\vee}$ is a ξ-lazy random walk on $\mathcal{G}_M$, i.e. with probability ξ, $P_r^{\vee}$ stays on the same vertex, and with $\frac{1-\xi}{\ell}$ probability $P_r^{\vee}$ goes to a neighbor of the current vertex, where ℓ is the degree of any vertex in $\mathcal{G}_M$ [23].

Then by Theorem 5.8,

$$\Phi(\mathcal{G}_M) \geq \frac{1 - \lambda_2(P_r^{\vee})}{2} \geq \frac{1 - (1 - 1/r)}{2} = \frac{1}{2r}$$

where we recall $\lambda_2(P_r^{\vee}) \leq 1 - 1/r$ from the proof of Theorem 5.7.

On the other hand, fix some $S \subset \mathcal{B}_M$ with $|S| \leq |\mathcal{B}_M|/2$. Then

$$\frac{1}{2r} \leq \Phi(\mathcal{G}_M) \leq \Phi(S) = \frac{\sum_{\tau \in S, \tau' \notin S} P_r^{\vee}(\tau, \tau')}{|S|} \overset{\text{Proposition 6}}{\leq} \frac{\sum_{\tau \in S, \tau' \notin S} \frac{1}{2r}}{|S|} = \frac{\frac{1}{2r}|E(S, \overline{S})|}{|S|} = \frac{h(S)}{2r}$$

Thus $h(S) \geq 1$ for every $|S| \leq |\mathcal{B}_M|/2$, and we're done. ■

Proposition 6. $P_r^{\vee}(\tau, \tau') \leq \frac{1}{2r}$ for any $\tau, \tau' \in \mathcal{B}_M, \tau \neq \tau'$.

Proof. If $P_r^{\vee}(\tau, \tau') \neq 0$, then $P_r^{\vee}(\tau, \tau') = \frac{w(\tau')}{rw(\tau \cap \tau')}$. But $w(\tau \cap \tau') \geq w(\tau) + w(\tau') = 2w(\tau')$, and we're done. ■

[22] we basically treat the graph as a 2-dimensional simplicial complex, and extend the weights from edges to vertices accordingly

[23] using basis exchange properties, one may show that the basis exchange graph is regular

5.4.2. The Random Cluster Model

We'll now see another application of Theorem 5.7 in statistical physics.

Given a matroid $M = ([n], \mathcal{I})$ of rank r and parameters p, q, we define the *partition function* [24], as the following polynomial:

$$Z_M(p, q) := \sum_{S \subseteq [n]} q^{r+1-\mathrm{rank}(S)} p^{|S|}$$

When M is the graphic matroid, $r + 1 - \mathrm{rank}(S)$ calculates the number of connected components of S. Before Theorem 5.7 was proved, one could only compute/approximate $Z_M(p, 2)$ (see [JS93], [GJ17]). Using Theorem 5.7, we can now approximate $Z_M(p, q)$ for all $p \geq 0, q \in (0, 1]$.

We can further define the *Tutte polynomial* of a matroid as

$$T_M(x, y) := \sum_{S \subseteq [n]} (x-1)^{r-\mathrm{rank}(S)} (y-1)^{|S|-\mathrm{rank}(S)}$$

Clearly,

$$T_M(x, y) = \frac{1}{(x-1)(y-1)^{r+1}} Z_M(y-1, (x-1)(y-1))$$

Thus we also have an FPRAS for estimating the Tutte polynomial in the region $y \geq 1, (x-1)(y-1) \in [0, 1]$.

Thus without further ado, let's see how the FPRAS for approximating $Z_M(p, q)$ comes about.

Theorem 5.10. There is a FPRAS for estimating $Z_M(p, q)$.

Proof. $Z_M(p, q) = q^{r+1} \sum_{k=0}^{n} p^k f_{M,k,q}(\mathbb{1})$, where $f_{M,k,q}$ is as defined in Lemma 2.1. By Lemma 2.1, since $f_{M,k,q}$ is completely log-concave, using ideas similar to the proof of Lemma 5.6, we get our desired result. ■

[24] It was introduced by Fortuin and Kasteleyn. See this book by Grimmett ([Gri06]) for further details

References

[ABF94] Yossi Azar, Andrei Z. Broder, and Alan M. Frieze. On the problem of approximating the number of bases of a matroid. *Inf. Process. Lett.*, 50(1):9–11, 1994.

[AGM+17] Arash Asadpour, Michel Goemans, Aleksander Mądry, Shayan oveis gharan, and Amin Saberi. An $O(\log n/\log\log n)$-approximation algorithm for the asymmetric traveling salesman problem. *Operations Research*, 65, 06 2017. doi:10.1287/opre.2017.1603.

[AHK18] Kareem Adiprasito, June Huh, and Eric Katz. Hodge theory for combinatorial geometries. *Annals of Mathematics*, 188(2):381 – 452, 2018. doi:10.4007/annals.2018.188.2.1.

[ALOV18] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials iii: Mason's ultra-log-concavity conjecture for independent sets of matroids. *ArXiv*, abs/1811.01600, 2018. URL: https://api.semanticscholar.org/CorpusID:53214942.

[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. New York, NY, USA, 2019. Association for Computing Machinery. doi:10.1145/3313276.3316385.

[Ana19] Nima Anari. Completely log-concave polynomials and distributions, 2019. URL: https://simons.berkeley.edu/sites/default/files/docs/11174/presentation1.pdf.

[AOV18] Nima Anari, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 35–46, Los Alamitos, CA, USA, oct 2018. IEEE Computer Society. URL: https://doi.ieeecomputersociety.org/10.1109/FOCS.2018.00013, doi:10.1109/FOCS.2018.00013.

[BH22] Petter Brändén and June Huh. Lorentzian polynomials, 2022. arXiv:1902.03719.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, USA, 2006.

[Cun84] William H Cunningham. Testing membership in matroid polyhedra. *Journal of Combinatorial Theory, Series B*, 36(2):161–188, 1984. URL: https://www.sciencedirect.com/science/article/pii/0095895684900236, doi:10.1016/0095-8956(84)90023-6.

[Dow80] T.A. Dowling. On the independent set numbers of a finite matroid. In M. Deza and I.G. Rosenberg, editors, *Combinatorics 79 Part I*, volume 8 of *Annals of Discrete Mathematics*, pages 21–28. Elsevier, 1980. URL: https://www.sciencedirect.com/science/article/pii/S0167506008708422, doi:10.1016/S0167-5060(08)70842-2.

[GJ17] Heng Guo and Mark Jerrum. *Random cluster dynamics for the Ising model is rapidly mixing*, pages 1818–1827. 2017. URL: https://epubs.siam.org/doi/abs/10.1137/1.9781611974782.118, arXiv:https://epubs.siam.org/doi/pdf/10.1137/1.9781611974782.118, doi:10.1137/1.9781611974782.118.

[Gri06] G.R. Grimmett. *The Random-Cluster Model*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2006.

[HK11] June Huh and Eric Katz. Log-concavity of characteristic polynomials and the bergman fan of matroids. *Mathematische Annalen*, 354:1103–1116, 2011.

[HS89] Yahya Ould Hamidoune and Isabelle Salaün. On the independence numbers of a matroid. *Journal of Combinatorial Theory, Series B*, 47(2):146–152, 1989. URL: https://www.sciencedirect.com/science/article/pii/0095895689900154, doi:10.1016/0095-8956(89)90015-4.

[HW17] June Huh and Botong Wang. Enumeration of points, lines, planes, etc. *Acta Mathematica*, 218(2):297 – 317, 2017. doi:10.4310/ACTA.2017.v218.n2.a2.

[JS93] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the ising model. *SIAM J. Comput.*, 22(5):1087–1116, oct 1993. doi:10.1137/0222066.

[Kal22] Gil Kalai. The work of june huh. Proc. Int. Cong. Math. 2022, Vol. 1, 2022.

[KN09] Jeff D. Kahn and M. Neiman. A strong log-concavity property for measures on boolean algebras. *J. Comb. Theory, Ser. A*, 118:1749–1760, 2009. URL: https://api.semanticscholar.org/CorpusID:15696589.

[KO20] Tali Kaufman and Izhar Oppenheim. High order random walks: Beyond spectral gap. *Combinatorica*, 40(2):245–281, apr 2020. doi:10.1007/s00493-019-3847-0.

[Len13] Matthias Lenz. The f-vector of a representable-matroid complex is log-concave. *Advances in Applied Mathematics*, 51(5):543–545, 2013. URL: https://www.sciencedirect.com/science/article/pii/S0196885813000705, doi:10.1016/j.aam.2013.07.001.

[Mah85] Carolyn Mahoney. On the unimodality of the independent set numbers of a class of matroids. *Journal of Combinatorial Theory, Series B*, 39(1):77–85, 1985. URL: https://www.sciencedirect.com/science/article/pii/0095895685900383, doi:10.1016/0095-8956(85)90038-3.

[Mas72] John Mason. Matroids: unimodal conjectures and motzkin's theorem. *Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst.) Inst. Math. Appl., Southend-on-Sea*, 1972.

[Nak35] Takeo Nakasawa. Zur axiomatik der linearen abhängigkeit. i. *Sci. Rep. Tokyo Bunrika Daigaku Sect.*, A(2), 1935.

[Nak36a] Takeo Nakasawa. Zur axiomatik der linearen abhängigkeit. ii. *Sci. Rep. Tokyo Bunrika Daigaku Sect.*, A(3), 1936.

[Nak36b] Takeo Nakasawa. Zur axiomatik der linearen abhängigkeit. iii schluss. *Sci. Rep. Tokyo Bunrika Daigaku Sect.*, A(3), 1936.

[Ove20] Shayan Oveis Gharan. Cse 599: Polynomial paradigm in algorithm design, 2020. URL: https://homes.cs.washington.edu/~shayan/courses/polynomials/.

[Oxl11] James Oxley. *Matroid Theory*. Oxford University Press, 02 2011. doi:10.1093/acprof:oso/9780198566946.001.0001.

[Sey77] P.D. Seymour. The matroids with the max-flow min-cut property. *Journal of Combinatorial Theory, Series B*, 23(2):189–222, 1977. URL: https://www.sciencedirect.com/science/article/pii/0095895677900314, doi:10.1016/0095-8956(77)90031-4.

[Sil00] John R. Silvester. Determinants of block matrices. *The Mathematical Gazette*, 84(501):460–467, 2000. URL: http://www.jstor.org/stable/3620776.

[SJ89] Alistair Sinclair and Mark Jerrum. Approximate counting, uniform generation and rapidly mixing markov chains. *Information and Computation*, 82(1):93–133, 1989. URL: https://www.sciencedirect.com/science/article/pii/0890540189900679, doi:10.1016/0890-5401(89)90067-9.

[Sno12] Michael Snook. Counting bases of representable matroids. *The Electronic Journal of Combinatorics*, 19(4), 2012.

[SV14] Mohit Singh and Nisheeth K. Vishnoi. Entropy, optimization and counting. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 50–59, New York, NY, USA, 2014. Association for Computing Machinery. doi:10.1145/2591796.2591803.

[Whi35] Hassler Whitney. On the abstract properties of linear dependence. *American Journal of Mathematics*, 57(3), 1935. URL: http://www.jstor.org/stable/2371182?origin=JSTOR-pdf.

[Zha85] Cui Kui Zhao. A conjecture on matroids. *Neimenggu Daxue Xuebao*, 16(3), 1985.

§A. Appendix

Lemma A.1. Let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, where A, B, C, D are square matrices (over a field) of the same order such that $CD = DC$. Then $\det(M) = \det(AD - BC)$.

Proof. Refer to [Sil00], Theorem 3. ■

Lemma A.2. Let A be a symmetric matrix with at most 1 positive eigenvalue. Then for any strictly PSD matrix $P \succ 0$, PA has at most 1 positive eigenvalue.

Proof. Write $P = B^{\mathsf{T}}B$ for some matrix B. Now, elementary linear algebra tells us that if X, Y are two matrices such that XY, YX are defined, then XY and YX have the same non-zero spectra.
Thus, $PA = B^{\mathsf{T}}BA$ has the same non-zero eigenvalues as BAB^{T}. Since B is invertible, BAB^{T} preserves the signs of the eigenvalues of A, and we're done. ■

Lemma A.3. Let A be a symmetric matrix with at most 1 positive eigenvalue. Suppose all entries of A are non-negative. Let w be a vector such that $w(i) := \sum_j A_{ij}$. Then $\frac{ww^{\mathsf{T}}}{\sum_i w(i)} - A$ is PSD.

Proof. Write $W = \mathrm{diag}(w)$. Clearly, $B := W^{-1/2}AW^{-1/2}$ has atmost 1 positive eigenvalue. Observe that $B\sqrt{w} = \sqrt{w}$, where $\sqrt{w}$ is the entry-wise square-root of w. Thus $\sqrt{w}$ is the only eigenvector of B corresponding to a positive eigenvalue, and thus $\sqrt{w}$ corresponds to the largest eigenvalue of B, implying that

$$B \preceq \frac{\sqrt{w}\sqrt{w}^{\mathsf{T}}}{\|\sqrt{w}\|^2} = \frac{\sqrt{w}\sqrt{w}^{\mathsf{T}}}{\sum_i w(i)}$$

Multiplying both sides of the inequality, left and right by $W^{1/2}$ proves the desired statement. ■

Lemma A.4 (Euler's Identity)**.** Let $g \in \mathbb{R}[z_1, \ldots, z_n]$ be a d-homogenous polynomial. Then $\sum_{i=1}^n z_i \partial_i g = d \cdot g$. Consequently, for any $a \in \mathbb{R}^n$, $\partial_a g\big|_{z=a} = \left(\sum_{i=1}^n a_i \partial_i g\right)(a) = d \cdot g(a)$.