Lecture Note 16

Disclaimer: *These notes have not been subjected to the usual scrutiny accorded to formal publications. They may be distributed outside this class only with the permission of the Instructor.*

16.1 Counting the bases of a matroid

Our goal in this lecture is to present a Markov chain Monte Carlo (MCMC) algorithm to approximately count the number of bases of an arbitrary matroid. Specifically, we have the following problem:

Input: A matroid $\mathcal{M} = (E, \mathcal{I})$, where the set E is called the *ground set*, and $\mathcal{I}$ is a collection of subsets of E, called *independent sets* of E. We represent $\mathcal{M}$ by a membership oracle, i.e., a black box that, given any set $J \subseteq E$, specifies whether or not J is an independent set.

Output: The number of bases (i.e., maximal independent sets) in $\mathcal{M}$.

Observation 16.1. *A fully polynomial randomized approximation scheme (fpras) for counting the bases of any matroid immediately implies an fpras for counting the independent sets of any given size k in any matroid. One simply runs the fpras on the truncated matroid $\mathcal{M}' = (E, \mathcal{I} \setminus \{I : I \in \mathcal{I}, |I| > k\})$ that is obtained from the original $\mathcal{M}$ via the elimination of all the independent sets of size greater than k from $\mathcal{I}$.*

16.2 MCMC algorithm setup

To perform approximate counting, we will use the standard reduction from counting to sampling that we saw earlier in the course. We will perform sampling via MCMC, and our goal will be to efficiently sample from a probability distribution over the set of bases that is close in total variation distance to the uniform distribution.

16.2.1 Basis exchange process

Our Markov chain for sampling bases, which is known as the "basis exchange" process, is defined on the state space $\Omega = \{B : B \text{ is a basis of } \mathcal{M}\}$ as follows:

- Let the current state be the basis B. Pick an element $e \in B$ uniformly at random to remove.

- Let $F = \{f \in E : B \setminus \{e\} \cup \{f\} \in \mathcal{I}\}$ denote the set of elements that we can add to $B \setminus \{e\}$ while remaining independent.

- Pick f uniformly at random from F and move to the basis $B \setminus \{e\} \cup \{f\}$.

Observation 16.2. *The basis exchange process is irreducible.*

Proof. Let B, B' be distinct bases. This means they have the same size, $|B| = |B'| = r$, where r is the rank of $\mathcal{M}$. Let the symmetric difference of B and B' be decomposed as $B \oplus B' = \{e_1, \ldots, e_k\} \cup \{e_1', \ldots, e_k'\}$ where $e_i \in B$ and $e_i' \in B'$. Remove e_1 from B. Then, by the augmentation axiom of matroids, there exists an element $e_i' \in B'$ such that $B'' := B \setminus \{e_1\} \cup \{e_i'\}$ is independent (and hence a basis). Now B'', B' are a pair of bases that differ in two fewer elements than B, B', so we can iterate this procedure until we reach B'. $\square$

The basis exchange process is clearly symmetric, so the stationary distribution is uniform. It is also aperiodic (there is always a self-loop as we could choose $f = e$ in the final step of the exchange), and therefore ergodic. Our goal will be to show that the mixing time is $O(r(\log r + \log \log n)) = O(n \log n)$, where $n = |E|$, following the analysis of [CGM19]. The framework below is due to [ALOV18], who used a different analysis based on the spectral gap to obtain a mixing time of $O(r^2 \log n) = O(n^2 \log n)$.

16.2.2   Weight function

While we are interested in the mixing time of our Markov chain over bases, we will actually define Markov chains over independent sets of each size $2 \leq k \leq r$, where r is the rank. We will then show that all of these Markov chains are rapidly mixing by induction on k. This sort of inductive decomposition is typical of log-Sobolev-based mixing time bounds (though in other settings, such as spin systems on lattices $\mathbb{Z}^d$, the induction is often spatial).

To this end, we now define a weight function $w(I)$ on *all* independent sets as follows:

$$w(I) = \begin{cases} 1 & |I| = r \text{ (basis)}, \\ \sum_{I' \in \mathcal{I}, I' = I + e} w(I') & \text{otherwise.} \end{cases} \tag{16.1}$$

In other words, the weight of a size k independent set I is the sum of the weights of all size $k+1$ independent sets that contain I. See Figure 16.1 for a simple example.
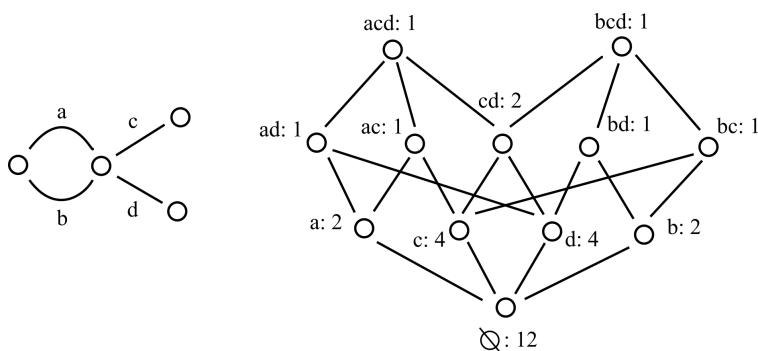

Figure 16.1: The weight function for the graphic matroid of a simple multigraph with edge set $E = \{a, b, c, d\}$ (shown on the left). The diagram on the right shows all independent sets (i.e., forests) organized by size, and their corresponding weights. The rank is $r = 3$.

Note: We can write down an explicit expression for $w(I)$ in terms of the number of bases B that contain I (that is, $B \supseteq I$) by noting that there are $(r - |I|)!$ paths from any such basis set B to I (each path corresponding to an order of removal of elements from B). Therefore, $w(I) = (r - |I|)! \times \#\{B : B \supseteq I\}$.

Let us denote the set of all independent sets of size k by $\mathcal{M}(k)$ (so $\mathcal{M}(r)$ are the bases). The weight function

above induces a probability distribution $\pi_k(I)$ on $\mathcal{M}(k)$ as follows:

$$\pi_k(I) = \frac{w(I)}{Z_k}, \text{ where } Z_k = \sum_{I \in \mathcal{M}(k)} w(I). \tag{16.2}$$

Note that Z_k is a partition function. Recalling that the set of all bases of the matroid $\mathcal{M}$ is denoted by $\Omega = \mathcal{M}(r)$, we have $Z_r = |\Omega|$. Further, it is easy to check that $k! Z_k = Z_0 = w(\emptyset)$. (**Exercise!**)

16.2.3 Contraction

We next introduce the idea of a "contraction", which will be made use of in the proofs of the following sections.

Definition 16.3. *For any independent set I, define the* contraction $\mathcal{M}_I$ *as* $\mathcal{M}_I := (E \setminus I, \mathcal{I}_I)$ *where* $\mathcal{I}_I :=$ $\{J : J \subseteq E \setminus I, \ J \cup I \in \mathcal{I}\}$.

In other words, given an independent set I, we first remove it from the ground set E. Then, to get the independent sets of $E \setminus I$, we consider all independent sets K of E such that $I \subseteq K$, and we say that $J = K \setminus I$ is an independent set of $E \setminus I$. It is a standard fact (**exercise!**) that the contraction $\mathcal{M}_I = (E \setminus I, \mathcal{I}_I)$ is also a matroid.

We next define probability distributions $\pi_{I,k}$ on $\mathcal{M}(|I| + k)$ as follows:

$$\pi_{I,k}(J) = \begin{cases} \frac{k! w(J)}{w(I)} & J \supseteq I, \\ 0 & \text{otherwise,} \end{cases} \tag{16.3}$$

Equivalently, $\pi_{I,k}(J)$ is the probability of J conditioned on J containing I.

16.2.4 'Down-up' walk $P_k^\vee$ on $\mathcal{M}(k)$

We will now generalize the basis exchange Markov chain to produce a Markov chain on independent sets of size k, for each k from 2 to r. The Markov chain on $\mathcal{M}(k)$ will have the stationary distribution π_k defined above. This Markov chain will be called the 'down-up' walk over $\mathcal{M}(k)$, and is defined as follows:

- Let the current state be the independent set $I \in \mathcal{M}(k)$. Remove an element $e \in I$ uniformly at random.

- Let $S = \{J \in \mathcal{M}(k) : J = I \setminus \{e\} \cup \{e'\} \text{ for some } e'\}$. Move to $J \in S$ with probability proportional to $w(J)$.

For $k = r$, the down-up walk $P_r^\vee$ is simply the basis exchange process. (**Exercise!**)

The transition probabilities for $I, J \in \mathcal{M}(k)$ can be written down explicitly as [**exercise!**]:

$$P_k^\vee(I, J) = \begin{cases} \frac{w(J)}{k w(I \cap J)} & \text{if } |I \oplus J| = 2; \\ \sum_{e \in I} \frac{w(I)}{k w(I \setminus \{e\})} & \text{if } I = J; \\ 0 & \text{otherwise.} \end{cases} \tag{16.4}$$

Exercise: Verify equation (16.4).

## 16.3   Main theorem: bounding the log-Sobolev constant

Our mixing time bound will follow immediately from the following result:

Theorem 16.4. *For $2 \leq k \leq r$, the modified log-Sobolev constant ρ_k of $P_k^\vee$ satisfies $\rho_k \geq \frac{1}{k}$.*

Corollary 16.5. *The mixing time of the basis exchange walk $P_r^\vee$ is $O(r(\log r + \log \log n)) = O(n \log n)$.*

Proof. We use our general bound on mixing times in terms of the log-Sobolev constant, $\tau_{\text{mix}} = O(\frac{1}{\rho_r} \log \log \pi_{\text{min}}^{-1})$, together with the fact that $\rho_r \geq \frac{1}{r}$ from the above theorem and $\pi_{\text{min}} = \frac{1}{|\Omega|} \geq \frac{1}{\binom{n}{r}} \geq \frac{1}{n^r}$.   □

This mixing time can be seen to be optimal as a function of n with the help of a simple example. Consider the graphic matroid corresponding to the following multigraph with $\frac{n}{2}$ pairs of edges and $\frac{n}{2} + 1$ vertices.
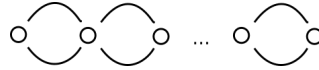
(Equivalently, this is a partition matroid with $\frac{n}{2}$ blocks, each of capacity 1.) The bases correspond to choosing one edge from each pair. The down-up walk is then easily seen to be equivalent to random walk on the hypercube $\{0,1\}^{n/2}$, which as we have seen has mixing time $\Theta(n \log n)$.

Note: Subsequently, Anari *et al.* [ALOV20] sharpened this mixing time bound to $O(r \log r)$, making it independent of n, the size of the ground set.

16.3.1   Setup

It is useful to decompose the transitions of the down-up walk into separate down and up steps. Both of these can be cast in terms of the following matrix.

Definition 16.6. *For $2 \leq k \leq r$, the matrix A_{k-1} with rows indexed by $\mathcal{M}(k-1)$ and columns indexed by $\mathcal{M}(k)$ is defined by*

$$A_{k-1}(I, J) = \begin{cases} 1 & \text{if } J = I \cup \{e\} \text{ for some } e; \\ 0 & \text{otherwise.} \end{cases} \tag{16.5}$$

Equipped with this definition, we can decompose the transition matrix $P_k^\vee$ of the down-up process at level k as $P_k^\vee = P_k^\downarrow P_{k-1}^\uparrow$, where the "down" step is represented by the transition matrix $P_k^\downarrow = \frac{1}{k} A_{k-1}^T$, and the "up" step by the transition matrix $P_{k-1}^\uparrow = A_{k-1}$ with the (I, J) entry set to $\frac{w(J)}{w(I)}$ instead of 1.

Note that these matrices satisfy the following "detailed balance" condition:

Claim 16.7. *For all I, J, $\pi_k(I) P_k^\downarrow(I, J) = \pi_{k-1}(J) P_{k-1}^\uparrow(J, I)$ (**Exercise!**)*

Given this setup, we can then naturally extend a function $f^{(k)} : \mathcal{M}(k) \to \mathbb{R}_{\geq 0}$ over independent sets of size k to functions $f^{(i)}$ over independent sets of size $i < k$ as follows:

Definition 16.8. *For any function $f^{(k)} : \mathcal{M}(k) \to \mathbb{R}_{\geq 0}$, define $f^{(i)} : \mathcal{M}(i) \to \mathbb{R}_{\geq 0}$ by:*

$$f^{(i)} := \prod_{j=i}^{k-1} P_j^\uparrow f^{(k)} = P_i^\uparrow P_{i+1}^\uparrow \ldots P_{k-1}^\uparrow f^{(k)}.$$

In other words, to apply $f^{(i)}$ to an independent set $I \in \mathcal{M}(i)$, we take $k - i$ random "up" steps to reach $J \in \mathcal{M}(k)$ and then apply the original function $f^{(k)}$ to J. Another way to interpret $f^{(i)}$ is as an expectation over the conditional measure $\pi_{J,k-i}$, namely: $f^{(i)}(J) = E_{\pi_{J,k-i}} f^{(k)}$ **[exercise!]**.

The following simple properties of these functions will be useful; the proofs are left as an **exercise!**

Claim 16.9. *1. If $f^{(k)}$ is normalized so that $E_{\pi_k} f^{(k)} = 1$, then $E_{\pi_i} f^{(i)} = 1$ for all $i < k$.*

2. $\pi_{k-1}(P^{\uparrow}_{k-1} f^{(k)}) = (\pi_k f^{(k)}) P^{\downarrow}_k$.

16.3.2   Main lemma

We now state a key lemma that will lead fairly directly to a proof of our main goal, Theorem 16.4. This lemma essentially says that, under one "down" step of the process, entropy is contracted by a factor of at least $1 - \frac{1}{k}$.

Lemma 16.10. *For any $k \geq 2$ and $f^{(k)} : \mathcal{M}(k) \to \mathbb{R}_{\geq 0}$ with $E_{\pi_k} f^{(k)} = 1$,*

$$\operatorname{Ent}_{\pi_{k-1}}(f^{(k-1)}) \leq \left(1 - \frac{1}{k}\right) \operatorname{Ent}_{\pi_k}(f^{(k)}). \tag{16.6}$$

We'll see how to prove this lemma by induction on k in the next subsection. First, we show how to use the lemma to prove Theorem 16.4.

Proof of Theorem 16.4. Recall that, to prove the theorem, we need to show that

$$\rho_k = \inf_{\substack{f^{(k)} \geq 0, \\ \operatorname{Ent}_{\pi_k} f^{(k)} \neq 0}} \frac{\mathcal{E}_{P^{\vee}_k}(f^{(k)}, \log f^{(k)})}{\operatorname{Ent}_{\pi_k} f^{(k)}} \geq \frac{1}{k}. \tag{16.7}$$

Let $f^{(k)} : \mathcal{M}(k) \to \mathbb{R}_{\geq 0}$ be any function, and assume without loss of generality that $E_{\pi_k} f^{(k)} = 1$ (since rescaling $f^{(k)}$ does not affect the ratio in (16.7)). Our first step will be to apply the definitions of $\operatorname{Ent}_{\pi_k}$ and $f^{(k-1)}$ to show that

$$\operatorname{Ent}_{\pi_{k-1}}(f^{(k-1)}) \geq \operatorname{Ent}_{\pi_k}(f^{(k)}) - \mathcal{E}_{P^{\vee}_k}(f^{(k)}, \log f^{(k)}), \tag{16.8}$$

which relates the entropies at the two adjacent levels via the Dirichlet form. From (16.8), we can directly apply Lemma 16.10 to bound the log-Sobolev constant as follows:

$$\mathcal{E}_{P^{\vee}_k}\left(f^{(k)}, \log f^{(k)}\right) \geq \operatorname{Ent}_{\pi_k}(f^{(k)}) - \operatorname{Ent}_{\pi_{k-1}}(f^{(k-1)})$$

$$\geq \frac{1}{k} \operatorname{Ent}_{\pi_k}(f^{(k)}).$$

Thus, it suffices to prove (16.8). We'll do this by applying the definitions and using Jensen's inequality:

$$
\begin{aligned}
\operatorname{Ent}_{\pi_{k-1}}(f^{(k-1)}) &= \sum_{I \in \mathcal{M}(k-1)} \pi_{k-1}(I) f^{(k-1)}(I) \log f^{(k-1)}(I) \\
&= \sum_{I \in \mathcal{M}(k-1)} \pi_{k-1}(I) [P_{k-1}^{\uparrow} f^{(k)}](I) \log f^{(k-1)}(I) \\
&= \sum_{\substack{I \in \mathcal{M}(k-1) \\ J \in \mathcal{M}(k)}} \pi_{k-1}(I) P_{k-1}^{\uparrow}(I, J) f^{(k)}(J) \log \left(\sum_{L \in \mathcal{M}(k)} P_{k-1}^{\uparrow}(I, L) f^{(k)}(L) \right) \\
&= \sum_{\substack{I \in \mathcal{M}(k-1) \\ J \in \mathcal{M}(k)}} \pi_k(J) f^{(k)}(J) P_k^{\downarrow}(J, I) \log \left(\sum_{L \in \mathcal{M}(k)} P_{k-1}^{\uparrow}(I, L) f^{(k)}(L) \right) \qquad \text{by Claim 16.7} \\
&\geq \sum_{\substack{I \in \mathcal{M}(k-1) \\ J, L \in \mathcal{M}(k)}} \pi_k(J) f^{(k)}(J) P_k^{\downarrow}(J, I) P_{k-1}^{\uparrow}(I, L) \log f^{(k)}(L) \qquad \text{by Jensen's inequality} \\
&= \sum_{J, L \in \mathcal{M}(k)} \pi_k(J) f^{(k)}(J) P_k^{\vee}(J, L) \log f^{(k)}(L) \\
&= \operatorname{Ent}_{\pi_k}(f^{(k)}) - \mathcal{E}_{P_k^{\vee}} \left(f^{(k)}, \log f^{(k)} \right),
\end{aligned}
$$

where in the last line we used the definition $\mathcal{E}_{P_k^{\vee}}(f^{(k)}, \log f^{(k)}) = \langle f^{(k)}, L \log f^{(k)} \rangle_{\pi_k} = \operatorname{Ent}_{\pi_k}(f^{(k)}) - \langle f^{(k)}, P \log f^{(k)} \rangle_{\pi_k}$. This completes the proof of (16.7) and hence of the theorem. $\qquad\square$

16.3.3 Inductive proof of the main lemma

Now we turn to proving Lemma 16.10 via induction on k. We will show the inductive step and leave the base case $k = 2$ to the end of the proof. (As we saw in the hypercube example in the last lecture, in log-Sobolev analyses the base case is often the more challenging part, and will be again here.)

Proof of inductive step in Lemma 16.10. We wish to show that the inequality

$$
\operatorname{Ent}_{\pi_{k-1}}(f^{(k-1)}) \leq \left(1 - \frac{1}{k} \right) \operatorname{Ent}_{\pi_k}(f^{(k)}), \tag{16.9}
$$

implies the analogous inequality with $k - 1$ replaced by k, i.e.,

$$
\operatorname{Ent}_{\pi_k}(f^{(k)}) \leq \left(1 - \frac{1}{k+1} \right) \operatorname{Ent}_{\pi_{k+1}}(f^{(k+1)}). \tag{16.10}
$$

The key to our induction will be the following decomposition of π_k:

$$
\pi_k = \sum_{e \in E} \pi_1(e) \pi_{e, k-1}, \tag{16.11}
$$

where we recall that $\pi_{e, k-1}$ is the distribution over $\mathcal{M}(k)$ conditioned on the set containing e. [**Exercise:** Check this!] This allows us to use the following standard factorization of the entropy:

$$
\operatorname{Ent}_{\pi_k} f^{(k)} = \operatorname{Ent}_{\pi_1} f^{(1)} + \sum_{e \in E} \pi_1(e) \operatorname{Ent}_{\pi_{e, k-1}} f^{(k)}. \tag{16.12}
$$

[**Exercise:** Check this factorization, by using the decomposition (16.11); you will also need the fact that $E_{\pi_{e,k-1}}(f^{(k)}) = f^{(1)}(e)$, which can be verified by induction on k.]

Now by the induction hypothesis applied to the functions $f^{(k)}, f^{(k-1)}, \ldots, f^{(2)}$, we have

$$\text{Ent}_{\pi_k} f^{(k)} \geq \frac{k}{k-1} \frac{k-1}{k-2} \cdots \frac{2}{1} \; \text{Ent}_{\pi_1} f^{(1)} = k\text{Ent}_{\pi_1} f^{(1)}. \tag{16.13}$$

Applying Equations (16.12) and (16.13) together gives us

$$\sum_{e \in E} \pi_1(e) \text{Ent}_{\pi_{e,k-1}} f^{(k)} = \text{Ent}_{\pi_k} f^{(k)} - \text{Ent}_{\pi_1} f^{(1)} \geq (k-1)\text{Ent}_{\pi_1} f^{(1)}. \tag{16.14}$$

Now, noting that the conditional distribution $\pi_{e,k}$ corresponds to a distribution on the smaller contracted matroid $\mathcal{M}_e$, we can use induction again to conclude:

$$\text{Ent}_{\pi_{k+1}} f^{(k+1)} = \text{Ent}_{\pi_1} f^{(1)} + \sum_{e \in E} \pi_1(e) \text{Ent}_{\pi_{e,k}} f^{(k+1)}$$

$$\text{(induction hypothesis on } \mathcal{M}_e) \geq \text{Ent}_{\pi_1} f^{(1)} + \sum_{e \in E} \pi_1(e) \frac{k}{k-1} \text{Ent}_{\pi_{e,k-1}} f^{(k)}$$

$$= \text{Ent}_{\pi_1} f^{(1)} + \frac{1}{k(k-1)} \sum_{e \in E} \pi_1(e) \text{Ent}_{\pi_{e,k-1}} f^{(k)} + \frac{k+1}{k} \sum_{e \in E} \pi_1(e) \text{Ent}_{\pi_{e,k-1}} f^{(k)}$$

$$\text{(by (16.14))} \geq \text{Ent}_{\pi_1} f^{(1)} + \frac{k-1}{k(k-1)} \text{Ent}_{\pi_1} f^{(1)} + \frac{k+1}{k} \sum_{e \in E} \pi_1(e) \text{Ent}_{\pi_{e,k-1}} f^{(k)}$$

$$= \frac{k+1}{k} \left(\text{Ent}_{\pi_1} f^{(1)} + \sum_{e \in E} \pi_1(e) \text{Ent}_{\pi_{e,k-1}} f^{(k)} \right)$$

$$= \frac{k+1}{k} \text{Ent}_{\pi_k} f^{(k)},$$

as desired. $\qquad\square$

To complete the proof of Lemma 16.10, it remains to verify the base case $k = 2$, which we do in the next section.

16.3.4 Proof of base case of Lemma 16.10

Rewriting the inequality in the theorem and substituting $k = 2$, our goal is to prove

$$\text{Ent}_{\pi_2} f^{(2)} - 2\text{Ent}_{\pi_1} f^{(1)} \geq 0. \tag{16.15}$$

Recall that we are assuming the normalization $E_{\pi_2} f^{(2)} = 1$, which as we've seen implies $E_{\pi_1} f^{(1)} = 1$. (These normalizations are w.l.o.g. as they don't affect the ratio in the log-Sobolev constant, as we observed last time.)

Recalling from last lecture the definition $f^{(k-1)} = P^\uparrow_{k-1} f(k)$, we first express $f^{(1)}$ in terms of $f^{(2)}$:

$$f^{(1)}(e) = (P^\uparrow_1 f(2))(e) = \sum_{s:e \in s} \frac{w(s)}{w(e)} f^{(2)}(s).$$

(Throughout this proof, we will use e, e' to denote elements of the ground set, and s to denote an independent set of size 2.) This leads to the following expression for the entropy:

$$\mathrm{Ent}_{\pi_1} f^{(1)} = \sum_e \pi_1(e) \left(\sum_{s:e \in s} \frac{w(s)}{w(e)} f^{(2)}(s) \right) \log f^{(1)}(e)$$

$$= \sum_s \sum_{e \in s} \pi_1(e) \frac{w(s)}{w(e)} f^{(2)}(s) \log f^{(1)}(e)$$

$$= \sum_{s=(e,e')} \frac{w(s)}{Z_1} f^{(2)}(s) \left(\log f^{(1)}(e) + \log f^{(1)}(e') \right),$$

where in the last line we recall that $\pi_1(e) = \frac{w(e)}{Z_1}$. Similarly, we can use the definition of entropy to write

$$\mathrm{Ent}_{\pi_2} f^{(2)} = \sum_s \pi_2(s) f^{(2)}(s) \log f^{(2)}(s)$$

$$= \sum_s \frac{w(s)}{Z_2} f^{(2)}(s) \log f^{(2)}(s).$$

As a result

$$\mathrm{Ent}_{\pi_2}(f^{(2)}) - 2\mathrm{Ent}_{\pi_1}(f^{(1)}) = \sum_{s=(e,e')} \frac{w(s)}{Z_2} f^{(2)}(s) \left[\log f^{(2)}(s) - \log \left(f^{(1)}(e) f^{(1)}(e') \right) \right]$$

$$\geq \sum_{s=(e,e')} \pi_2(s) f^{(2)}(s) - \sum_{s=(e,e')} \frac{w(s)}{Z_2} f^{(1)}(e) f^{(1)}(e')$$

$$= 1 - \frac{1}{2Z_2} f^{(1)^T} W f^{(1)}.$$

where in the first line we used $Z_1 = 2Z_2$ (because every element in $\mathcal{M}(2)$ contributes its weight to exactly two elements in $\mathcal{M}(1)$), and in the second line we used the elementary inequality

$$a \log \frac{a}{b} \geq a - b, \quad \text{for all } a \geq 0, b > 0.$$

The matrix W above is symmetric with entries $W(e, e') = w(\{e, e'\})$. To finish the base case, comparing with (16.15) it suffices to show that

$$f^{(1)^\top} W f^{(1)} \leq 2Z_2 = Z_1. \tag{16.16}$$

The proof of (16.16) involves an excursion into the geometry of polynomials and log-concave distributions. There is a large recent literature on this topic, but we will just cover what is necessary to prove the inequality.

Let $\pi : 2^{|E|} \to \mathbb{R}_{\geq 0}$ be a discrete probability distribution. The *generating polynomial* of π is defined as

$$g_\pi(x_1, \ldots, x_{|E|}) = \sum_{S \subseteq E} \pi(S) \prod_{i \in S} x_i.$$

Definition 16.11. *The polynomial g is*

- ***log-concave*** *at x if $\log g$ is concave at x, i.e., $\nabla^2 \log g(x)$ is negative semi-definite.*

- ***strongly log-concave*** *(SLC) if $\nabla^2 \log \partial_I g(1)$ is negative semidefinite for any $I \subseteq \mathcal{E}$.*
  Here, ∂_I represents the partial derivatives taken with respect to all x_i with $i \in I$.

We have the following beautiful theorem due to Anari et.al [ALOV18] and Brändén-Huh [**?**]:

Theorem 16.12. *The generating polynomial of the uniform distribution over the bases of any matroid is SLC.*

Note: A converse-like statement is also true: the support of any SLC distribution forms the bases of some matroid (though we will not need this here). Note also that all the results presented here actually hold in the more general setting of an arbitrary (homogeneous) SLC distribution; the proof is essentially identical, with the more general SLC distribution replacing the uniform weights on the bases in $\mathcal{M}(r)$ and all other weights inherited from them.

The generating polynomial for the uniform distribution over bases is

$$g_\pi = \frac{1}{|\mathcal{B}|} \sum_{B \in \mathcal{B}} \prod_{i \in B} x_i.$$

The derivatives $\partial_I g_\pi(1)$ also have a nice interpretation as the probability that a uniform random basis contains I:

$$\partial_I g_\pi(1) = \frac{1}{|\mathcal{B}|} |\{B : B \supseteq I\}| = \frac{w(I)}{|\mathcal{B}| \cdot (r - |I|)!}, \tag{16.17}$$

where we used the explicit expression for $w(I)$ stated earlier.

We need another result, which is a consequence of SLC [ALOV18].

Lemma 16.13. *If g is SLC, then $\nabla^2 \partial_I g(1)$ has at most one positive eigenvalue.*

This result gives useful information about the eigenvalues of the matrix W, because from (16.17) W can be written as

$$W = (r-2)! \cdot |\mathcal{B}| \cdot \nabla^2 g_\pi(1).$$

which is just a scaling of the Hessian of $g_\pi(1)$ at 1. Thus Lemma 16.13 tells us that W has at most one positive eigenvalue.

Let $f : \mathcal{M}(1) \to \mathbb{R}_{\geq 0}$ be a function with $\mathbb{E}_{\pi_1} f = \sum_e \pi_1(e) f(e) = 1$, or equivalently, $\sum_e w(e) f(e) = Z_1$. Let $D = \mathrm{diag}(w(e))$ and $A = D^{-1/2} W D^{-1/2}$, so $A(e, e') = \frac{1}{\sqrt{w(e)w(e')}} W(e, e')$ is real symmetric and admits the spectral decomposition

$$A = \sum_i \lambda_i g_i g_i^\top.$$

Here, the g_i are an orthonormal bases of eigenvectors of A, and the λ_i are the corresponding eigenvalues. **Exercise:** Show that A also has at most one positive eigenvalue, and that $\sqrt{w(e)}$ is an eigenvector with eigenvalue 1. Thus, we can take $g_1 = \sqrt{\pi_1}$, $\lambda_1 = 1$, and $\lambda_i \leq 0$ for all $i \geq 2$.

From the decomposition of A, we can recover the following decomposition of W:

$$W = \sum_i \lambda_i h_i h_i^\top,$$

where $h_i = D^{1/2} g_i$. In particular $h_1 = D^{1/2} g_1 = \sqrt{\pi_1} \sqrt{w} = \frac{w}{\sqrt{Z_1}}$. From this, we obtain

$$\langle h_1, f \rangle = \sum_e \frac{w(e)}{\sqrt{Z_1}} f(e) = \sqrt{Z_1},$$

because $\sum_e w(e)f(e) = Z_1$ by assumption. Finally,

$$f^\top W f = \sum_i \lambda_i \langle h_i, f \rangle^2 \le \langle h_1, f \rangle^2 = Z_1,$$

because $\lambda_i \le 0$ for $i \ge 2$ and $\lambda_1 = 1$. This completes the proof of (16.15) and hence the proof of Lemma 16.10.

References

[ALOV18] N. Anari, K. Liu, S. Oveis Gharan, and C. Vinzant. Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids. *Proceedings of the 59th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 35–46, 2018.

[ALOV20] N. Anari, K. Liu, S. Oveis Gharan, and C. Vinzant. Log-concave polynomials IV: Approximate exchange, tight mixing times, and faster sampling of spanning trees. *arXiv, abs/2004.07220*, 2020.

[CGM19] M. Cryan, H. Guo, and G. Mousa. Modified log-Sobolev inequalities for strongly log-concave distributions. *Proceedings of the 60th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 1358–1370, 2019.