

## Lecture Note 15

*Instructor: Alistair Sinclair*

**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.*

## 15.1 Review of entropy and log-Sobolev constant

In the previous lecture, we saw how we can use entropy and the log-Sobolev constant to upper bound the mixing time of continuous time Markov chains. Since the mixing times of the discrete and continuous time versions of the Markov chain are essentially equivalent (up to constant factors), this also gives bounds on the mixing time of discrete time chains. Recall from the previous lecture:

**Definition 15.1.** *For an irreducible Markov chain with transition matrix  $P$  and stationary distribution  $\pi$ , the (modified) log-Sobolev constant of  $P$  is defined by*

$$\rho := \inf_{\substack{\varphi \geq 0 \\ \text{Ent}_\pi[\varphi] \neq 0}} \frac{\mathcal{E}_P(\varphi, \log \varphi)}{\text{Ent}_\pi[\varphi]}.$$

Here  $\mathcal{E}_P(\varphi, \log \varphi) = \langle \varphi, L\varphi \rangle_\pi$  is the Dirichlet form, and  $\text{Ent}_\pi(\varphi) = \sum_x \pi(x) \varphi(x) \log \left( \frac{\varphi(x)}{\mathbb{E}_\pi[\varphi]} \right)$ . The following result from the previous lecture relates the (continuous time) mixing time to the modified log-Sobolev constant.

**Theorem 15.2.** *For any irreducible  $P$  and any initial state  $x \in \Omega$ , the time to stationarity of the associated continuous time Markov chain (heat kernel) satisfies*

$$\tau_x(\varepsilon) \leq \frac{1}{\rho} (2 \log \varepsilon^{-1} + \log \log \pi(x)^{-1}).$$

From our definition of mixing time, we therefore get

$$\tau_{\text{mix}} = O\left(\frac{1}{\rho} \log \log \pi_{\min}^{-1}\right),$$

where  $\pi_{\min} := \min_{x \in \Omega} \pi(x)$ . In this lecture we will apply this machinery to the benchmark problem of random walk on the hypercube, where it will give a tight mixing time result but with a surprising amount of effort.

## 15.2 Random walk on the hypercube

In this section we consider random walk on the hypercube  $\{0, 1\}^n$  and show that its mixing time is  $O(n \log n)$ , which as we have seen is tight up to constants, and much sharper than the  $O(n^3)$  bound we obtained in Lecture 11 using the Poincaré constant (which is equivalent to spectral analysis since the chain is reversible).

Our approach will be to compute  $\tau_{\text{mix}}$  directly for  $n = 1$ , and then induct on the dimension using the product structure of the space  $\{0, 1\}^n$ ; in doing so we'll follow an argument of Piyush Srivastava. This is typical of analyses of mixing times via log-Sobolev inequalities, where some decomposition of the state space is the key to the computation.

### 15.2.1 Base case: $n = 1$

Perhaps surprisingly, the base case is non-trivial. This is also typical of these analyses. We begin by observing a general relation for asymmetric Dirichlet forms of *reversible* transition matrices  $P$ .

**Fact 15.3.** *Suppose  $P$  is the transition matrix of a reversible Markov chain with finite state space  $\Omega$  and stationary distribution  $\pi$ . Then the (asymmetric) Dirichlet form satisfies*

$$\mathcal{E}_P(\varphi, \psi) = \frac{1}{2} \sum_{x, y \in \Omega} \pi(x) P(x, y) (\varphi(x) - \varphi(y)) (\psi(x) - \psi(y)), \quad \text{for all } \varphi, \psi : \Omega \rightarrow \mathbb{R}. \quad (15.1)$$

*Proof.* Exercise (along similar lines to the proof of equation (10.2) in Lecture 10). □

**Note:** The symmetric expression in (15.1) holds only when  $P$  is reversible, or (for general  $P$ ) when  $\varphi = \psi$  (as we observed in Lecture 10).

Now consider the hypercube with  $n = 1$ , for which  $\Omega = \{0, 1\}$ . The (continuous time) transition matrix here is

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Note that  $P$  is clearly reversible with uniform stationary distribution. (It is not irreducible, but this is not a problem in continuous time; in discrete time we would make it lazy to overcome this problem.)

To compute the modified log-Sobolev constant, we consider an arbitrary nonnegative function  $\varphi : \Omega \rightarrow \mathbb{R}$ , specified by  $a := \varphi(0)$  and  $b := \varphi(1)$ , where  $a, b \geq 0$ . Starting from (15.1), and taking  $\psi = \log \varphi$ , it follows that

$$\begin{aligned} \mathcal{E}_P(\varphi, \log \varphi) &= \frac{1}{2} \sum_{x, y \in \Omega} \pi(x) P(x, y) (\varphi(x) - \varphi(y)) \log \frac{\varphi(x)}{\varphi(y)} \\ &= \frac{1}{2} (a - b) \log \frac{a}{b}. \end{aligned} \quad (15.2)$$

Similarly we can compute the entropy as

$$\begin{aligned} \text{Ent}_\pi[\varphi] &= \mathbb{E}_\pi[\varphi \log \varphi] - \mathbb{E}_\pi[\varphi] \log \mathbb{E}_\pi[\varphi] \\ &= \frac{a \log a + b \log b}{2} - \frac{a + b}{2} \log \frac{a + b}{2}. \end{aligned} \quad (15.3)$$

To bound the modified log-Sobolev constant, we seek the smallest  $c$  such that

$$\text{Ent}_\pi[\varphi] - c \mathcal{E}_P(\varphi, \log \varphi) \leq 0,$$

for all nonnegative  $\varphi$  such that  $\text{Ent}_\pi[\varphi] \neq 0$ . This will then ensure that  $\rho \geq \frac{1}{c}$ . Equivalently, by equations (15.2) and (15.3), we seek the smallest  $c$  such that, for all  $a, b \geq 0$ ,

$$h_c(a, b) := \frac{a \log a + b \log b}{2} - \frac{a + b}{2} \log \frac{a + b}{2} - \frac{c}{2} (a - b) \log \frac{a}{b} \leq 0.$$

We now make the following observations about  $h_c(a, b)$ :

1. If  $a = b$ , then  $h_c(a, b) = 0$ , for any  $c$ .
2. The partial derivatives of  $h_c$  satisfy

$$\frac{\partial h_c}{\partial a}(a, b) = \frac{1}{2} \log a - \frac{1}{2} \log \frac{a+b}{2} - \frac{c}{2} \log \frac{a}{b} - \frac{c}{2} \left(1 - \frac{b}{a}\right), \quad \text{and} \quad (15.4a)$$

$$\frac{\partial h_c}{\partial b}(a, b) = \frac{1}{2} \log b - \frac{1}{2} \log \frac{a+b}{2} - \frac{c}{2} \log \frac{b}{a} - \frac{c}{2} \left(1 - \frac{a}{b}\right). \quad (15.4b)$$

In particular, by evaluating the partial derivatives (15.4a) and (15.4b), we see that  $\nabla h_c(a, b) = 0$  when  $a = b$ , for all  $c$ .

3. The matrix of second-order partial derivatives (the Hessian) satisfies

$$\nabla^2 h_c(a, b) = \begin{pmatrix} \frac{b}{2a(a+b)} \left(1 - \frac{c(a+b)^2}{ab}\right) & \frac{1}{2} \left(\frac{c(a+b)^2 - ab}{ab(a+b)}\right) \\ \frac{1}{2} \left(\frac{c(a+b)^2 - ab}{ab(a+b)}\right) & \frac{a}{2b(a+b)} \left(1 - \frac{c(a+b)^2}{ab}\right) \end{pmatrix}$$

Now it is straightforward to see that the diagonal entries (the second-order partial derivatives of  $h_c$  in  $a, b$ ) are nonpositive provided  $c \geq 1/4$ . Additionally, one checks that for all  $a, b$ ,  $\det(\nabla^2 h_{1/4}(a, b)) = 0$ , and therefore the function  $h_{1/4}$  is *concave* in the variables  $a, b$ , since this shows that  $\nabla^2 h_{1/4}$  is a negative semidefinite matrix for all  $a, b$ .

By observation 2,  $h_{1/4}$  has a critical point at any pair  $(a, b)$  with  $a = b$ . By observation 3, this implies that  $h_{1/4}(a, a)$  is a global maximum of the function  $h_{1/4}$ , and by observation 1, this implies that  $h_{1/4}$  is nonpositive on  $\mathbb{R}^2$ . We may therefore conclude that  $\rho \geq 4$  for the base case  $n = 1$ .

## 15.2.2 Inductive step

Now let's consider the Inductive Step. The strategy here is to decompose high-dimensional hypercubes into products of low-dimensional hypercubes. In particular, the  $n$ -dimensional cube  $\Omega = \{0, 1\}^n$  can be expressed as

$$\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n$$

where  $\Omega_i = \{0, 1\}$  for all  $i$ . Let  $P_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ . Then the transition matrix of the random walk on  $\Omega$  can be written as

$$P = \frac{1}{n} \sum_{i=1}^n \underbrace{I \otimes \dots \otimes I}_{i-1} \otimes P_0 \otimes \underbrace{I \otimes \dots \otimes I}_{n-i}$$

Our goal is to express the modified log-Sobolev constant of  $P$  in terms of that of the constituent matrices  $P_0$ . To do this, we consider the more general setting of a product of two Markov chains, as follows.

**Observation 15.4.** *Let  $(\Omega_1, P_1), (\Omega_2, P_2)$  be two reversible Markov chains with state spaces  $\Omega_1, \Omega_2$ , transition matrices  $P_1, P_2$ , and stationary distributions  $\pi_1, \pi_2$  respectively. Define a new Markov chain  $(\Omega, P)$  with state space  $\Omega = \Omega_1 \otimes \Omega_2$  and transition matrix*

$$P := q_1(P_1 \otimes I) + q_2(I \otimes P_2)$$

where  $q_1 + q_2 = 1$  and  $q_1, q_2 > 0$ . Then  $P$  is reversible w.r.t. the product distribution  $\pi_1 \otimes \pi_2$ .

*Proof.* Easy exercise. □

The following result bounds the modified log-Sobolev constant of the product chain  $P$  in terms of those of  $P_1, P_2$ .

**Theorem 15.5.** *Let  $P = q_1(P_1 \otimes I) + q_2(I \otimes P_2)$  be a reversible product Markov chain as defined in Observation 15.4. If  $\rho, \rho_1, \rho_2$  denote the modified log-Sobolev constants of  $P, P_1, P_2$  respectively, then*

$$\rho \geq \min\{q_1\rho_1, q_2\rho_2\}.$$

The theorem will follow from the following two claims, which decompose the Dirichlet form and entropy, respectively, across  $P_1, P_2$ .

**Claim 15.6.** *For  $P = q_1(P_1 \otimes I) + q_2(I \otimes P_2)$  as defined above, the Dirichlet form  $\mathcal{E}_P(\varphi, \psi)$  can be expressed as:*

$$\mathcal{E}_P(\varphi, \psi) = q_1 \sum_{x_2} \pi_2(x_2) \mathcal{E}_{P_1}(\varphi(\cdot, x_2), \psi(\cdot, x_2)) + q_2 \sum_{x_1} \pi_1(x_1) \mathcal{E}_{P_2}(\varphi(x_1, \cdot), \psi(x_1, \cdot)).$$

**Claim 15.7.** *For  $P = q_1(P_1 \otimes I) + q_2(I \otimes P_2)$  as defined above, the entropy  $\text{Ent}_\pi[\varphi]$  can be upper bounded by:*

$$\text{Ent}_\pi[\varphi] \leq \sum_{x_1} \pi_1(x_1) \text{Ent}_{\pi_2}[\varphi(x_1, \cdot)] + \sum_{x_2} \pi_2(x_2) \text{Ent}_{\pi_1}[\varphi(\cdot, x_2)].$$

Before proving Claims 15.6 and 15.7, we first show that these two claims imply Theorem 15.5.

*Proof of Theorem 15.5.*

$$\begin{aligned} \mathcal{E}_P(\varphi, \log \varphi) &= q_1 \sum_{x_2} \pi_2(x_2) \mathcal{E}_{P_1}(\varphi(\cdot, x_2), \log \varphi(\cdot, x_2)) + q_2 \sum_{x_1} \pi_1(x_1) \mathcal{E}_{P_2}(\varphi(x_1, \cdot), \log \varphi(x_1, \cdot)) \\ &\geq \rho_1 q_1 \sum_{x_2} \pi_2(x_2) \text{Ent}_{\pi_1}[\varphi(\cdot, x_2)] + \rho_2 q_2 \sum_{x_1} \pi_1(x_1) \text{Ent}_{\pi_2}[\varphi(x_1, \cdot)] \\ &\geq \min\{\rho_1 q_1, \rho_2 q_2\} \left[ \sum_{x_2} \pi_2(x_2) \text{Ent}_{\pi_1}[\varphi(\cdot, x_2)] + \rho_2 q_2 \sum_{x_1} \pi_1(x_1) \text{Ent}_{\pi_2}[\varphi(x_1, \cdot)] \right] \\ &\geq \min\{\rho_1 q_1, \rho_2 q_2\} \text{Ent}_\pi[\varphi] \end{aligned}$$

which implies that  $\rho \geq \min\{\rho_1 q_1, \rho_2 q_2\}$ . In the derivation above, the first line comes from Claim 15.6, the second line come from the definitions of  $\rho_1, \rho_2$ , the third line is trivial, and the last inequality comes from Claim 15.7.  $\square$

It remains to prove Claims 15.6 and 15.7.

*Proof of Claim 15.6.* Starting from (15.1) we may write

$$\begin{aligned} \mathcal{E}_P(\varphi, \psi) &= \frac{1}{2} \sum_{x, y \in \Omega} \pi(x) P(x, y) (\varphi(x) - \varphi(y)) (\psi(x) - \psi(y)) \\ &= \frac{q_1}{2} \sum_{x_1, x_2, y_1} \pi_1(x_1) \pi_2(x_2) P_1(x_1, y_1) (\varphi(x_1, x_2) - \varphi(y_1, x_2)) (\psi(x_1, x_2) - \psi(y_1, x_2)) \\ &\quad + \frac{q_2}{2} \sum_{x_1, x_2, y_2} \pi_1(x_1) \pi_2(x_2) P_2(x_2, y_2) (\varphi(x_1, x_2) - \varphi(x_1, y_2)) (\psi(x_1, x_2) - \psi(x_1, y_2)) \\ &= q_1 \sum_{x_2} \pi_2(x_2) \mathcal{E}_{P_1}(\varphi(\cdot, x_2), \psi(\cdot, x_2)) + q_2 \sum_{x_1} \pi_1(x_1) \mathcal{E}_{P_2}(\varphi(x_1, \cdot), \psi(x_1, \cdot)), \end{aligned}$$

where in the second equality we decompose  $x = (x_1, x_2)$  and  $y = (y_1, y_2)$  with  $x_i, y_i \in \Omega_i$ .  $\square$

*Proof of Claim 15.7.* Using the definition and the decomposition  $x = (x_1, x_2)$  for  $x \in \Omega, x_1 \in \Omega_1, x_2 \in \Omega_2$ , we have:

$$\begin{aligned} \text{Ent}_\pi[\varphi] &= \sum_x \pi(x) \varphi(x) \log \left( \frac{\varphi(x)}{\mathbb{E}_\pi[\varphi]} \right) \\ &= \sum_{x_1, x_2} \pi_1(x_1) \pi_2(x_2) \varphi(x_1, x_2) \log \frac{\varphi(x_1, x_2)}{\sum_{x_1, x_2} \pi_1(x_1) \pi_2(x_2) \varphi(x_1, x_2)} \\ &= \underbrace{\sum_{x_1, x_2} \pi_1(x_1) \pi_2(x_2) \varphi(x_1, x_2) \log \frac{\varphi(x_1, x_2)}{\sum_{x_2} \pi_2(x_2) \varphi(x_1, x_2)}}_{\text{first sum}} + \underbrace{\sum_{x_1, x_2} \pi_1(x_1) \pi_2(x_2) \varphi(x_1, x_2) \log \frac{\sum_{x_2} \pi_2(x_2) \varphi(x_1, x_2)}{\sum_{x_1, x_2} \pi_1(x_1) \pi_2(x_2) \varphi(x_1, x_2)}}_{\text{second sum}} \end{aligned}$$

Using the definition of  $\text{Ent}_\pi[\varphi]$ , the first sum can be simplified as:

$$\sum_{x_1, x_2} \pi_1(x_1) \pi_2(x_2) \varphi(x_1, x_2) \log \frac{\varphi(x_1, x_2)}{\sum_{x_2} \pi_2(x_2) \varphi(x_1, x_2)} = \sum_{x_1} \pi_1(x_1) \text{Ent}_{\pi_2}(\varphi(x_1, \cdot))$$

To give an upper bound on the second sum, we use the Log-Sum Inequality which states that:

$$\sum_i a_i \log \frac{\sum a_i}{\sum b_i} \leq \sum a_i \log \left( \frac{a_i}{b_i} \right)$$

for all non-negative  $a_i, b_i$  with  $b_i > 0$ . Applying this to the second sum, we have:

$$\begin{aligned} \text{Second sum} &= \sum_{x_1, x_2} \pi_1(x_1) \pi_2(x_2) \varphi(x_1, x_2) \log \frac{\sum_{x_2} \pi_2(x_2) \varphi(x_1, x_2)}{\sum_{x_1, x_2} \pi_1(x_1) \pi_2(x_2) \varphi(x_1, x_2)} \\ &\leq \sum_{x_1} \pi_1(x_1) \sum_{x_2} \pi_2(x_2) \varphi(x_1, x_2) \log \frac{\pi_2(x_2) \varphi(x_1, x_2)}{\pi_2(x_2) \sum_{x_1} \pi_1(x_1) \varphi(x_1, x_2)} \\ &= \sum_{x_2} \pi_2(x_2) \sum_{x_1} \pi_1(x_1) \varphi(x_1, x_2) \log \frac{\varphi(x_1, x_2)}{\sum_{x_1} \pi_1(x_1) \varphi(x_1, x_2)} \\ &= \sum_{x_2} \pi_2(x_2) \text{Ent}_{\pi_1}(\varphi(\cdot, x_2)). \end{aligned}$$

This concludes the proof.  $\square$

Now we are ready to apply Theorem 15.5 to random walk on the hypercube. For the 2-dimensional cube  $\{0, 1\}^2$ , we have the product decomposition  $P_1 = P_2 = P_0$  (where  $P_0$  is the matrix of the 1-dimensional cube as before) and  $q_1 = q_2 = \frac{1}{2}$ . From earlier we know that  $\rho_1 = \rho_2 \geq 4$ , so Theorem 15.5 implies that  $\rho \geq \min\{\rho_1 q_1, \rho_2 q_2\} = 2$ .

We can use the same idea to prove by induction on  $n$  that the log-Sobolev constant for random walk on  $\{0, 1\}^n$  satisfies  $\rho \geq \frac{4}{n}$ . To do this, we decompose  $\Omega = \{0, 1\}^n$  as  $\Omega_1 \times \Omega_2$ , with  $\Omega_1 = \{0, 1\}$  and  $\Omega_2 = \{0, 1\}^{n-1}$ , and  $q_1 = \frac{1}{n}, q_2 = \frac{n-1}{n}$ . We then have  $\rho_1 \geq 4$ , and we may assume by induction that  $\rho_2 \geq \frac{4}{n-1}$ . Now by Theorem 15.5, we get

$$\rho \geq \min \left\{ \frac{1}{n} \times 4, \frac{n-1}{n} \times \frac{4}{n-1} \right\} = \frac{4}{n},$$

completing the proof by induction.

Finally, by Theorem 15.2 we get

$$\tau_{\text{mix}} \leq \underbrace{\frac{1}{\rho}}_{O(n)} \underbrace{\log \log \pi_{\min}^{-1}}_{O(\log n)} = O(n \log n),$$

as claimed.

## 15.3 Matroid bases

The hypercube example in the previous section illustrates that the modified log-Sobolev inequality can sometimes be used to find a tight bound on the mixing time of certain Markov chains. In this section, we will see how the same technique can be applied to the much more involved problem of counting the bases (or the independent sets) of a matroid. The first fpras for this problem was obtained recently by Anari *et al.* [ALOV18] by analyzing the spectrum of a natural Markov chain (i.e., via the Poincaré constant). We will instead look at the recent paper of Cryan, Guo and Mousa [CGM19], who analyze the same Markov chain using the modified log-Sobolev constant, and get a sharper result.

We begin by recalling some basic facts about matroids.

**Definition 15.8.** A matroid  $\mathcal{M} = (E, \mathcal{I})$  consists of a ground set  $E$  and a non-empty set of subsets  $\mathcal{I} \subseteq 2^E$ , called independent sets, with the following properties:

1. Downward closure, i.e.,  $I \in \mathcal{I}, J \subseteq I \implies J \in \mathcal{I}$
2. Augmentation property, i.e., if  $I, J \in \mathcal{I}$  and  $|I| > |J|$ , then there exists  $e \in I \setminus J$  such that  $J \cup \{e\} \in \mathcal{I}$

Note in particular that the emptyset  $\emptyset$  is always an independent set. A subset of the ground set  $E$  that is not independent is called *dependent*.

**Definition 15.9.** A basis of a matroid  $\mathcal{M}$  is a maximal independent set, i.e., an independent set that becomes dependent upon adding any element of  $E$ .

It is easy to check from the definitions [**exercise!**] that all bases of  $\mathcal{M}$  have the same cardinality, which we call the *rank* of  $\mathcal{M}$ .

### 15.3.1 Examples of matroids

- **Graphic matroids.** Let  $G = (V, E)$  be a connected graph. The *graphic matroid* associated with  $G$  is defined as  $\mathcal{M} = (E, \mathcal{I})$ , where:
  - $E$  is the set of edges of  $G$
  - $\mathcal{I}$  is the set of *forests* (acyclic subgraphs) of  $G$

The bases of this matroid are maximal forests, i.e., *spanning trees* of  $G$ .

**Exercise:** Check that the graphic matroid defined above satisfies the matroid properties.

- **Partition matroids.** Let  $B_1, B_2, \dots, B_k$  be disjoint sets. Let  $d_1, \dots, d_k$  be capacities such that  $0 \leq d_i \leq |B_i|$ . The associated partition matroid is defined as  $\mathcal{M} = (E, \mathcal{I})$ , where:

- $E = \bigcup_{i=1}^k B_i$
- $\mathcal{I} = \{I \subseteq E : |I \cap B_i| \leq d_i \forall i\}$

I.e., independent sets are collections of elements from  $E$  that contain no more than  $d_i$  elements of each block  $B_i$ . The bases are the sets  $I \subseteq E$  such that  $|I \cap B_i| = d_i \forall i$ .

- **Transversal matroids.** Let  $G = (U, V, E)$  be a bipartite graph. We can define a matroid  $\mathcal{M} = (E, \mathcal{I})$  as follows:

- $E = U$
- $\mathcal{I}$  is the set of subsets of  $U$  that are the endpoints of some matching of  $G$

Note that the bases here are subsets of  $U$  that are covered by *maximum* (not just *maximal*) matchings, and that any independent set may correspond to multiple matchings.

- **Linear (or representable) matroids.** Let  $A$  be an  $m \times n$  matrix over some field  $\mathbb{F}$ , where  $m \geq n$ . The associated linear matroid  $\mathcal{M} = (E, \mathcal{I})$  is defined by:

- $E$  is the set of columns of  $A$
- $\mathcal{I}$  are the linearly independent subsets of columns of  $A$

Here bases are maximal linearly independent columns, and the rank is the usual algebraic rank of  $A$ .

Most “natural” matroids are in fact representable over some field, and many (such as graphic and partition matroids) are representable over any field, including  $GF(2)$ . (By contrast, transversal matroids are representable but only over very large fields.) It is also possible to construct exotic matroids that are not representable at all.

## 15.4 Counting the bases of a matroid

As mentioned earlier, our goal is to count the number of bases of a given matroid, or more generally the number of independent sets of a given size. These problems have many applications in various fields, including (for graphic matroids) the network reliability problem we discussed in an earlier lecture. Both of these problems are known to be #P-complete in general; in fact, counting independent sets (of all sizes) is #P-complete even for graphic matroids [JVV90] (though counting bases of graphic matroids corresponds to counting spanning trees, which we’ve already seen in an earlier lecture can be done in polynomial time); and counting bases is #P-complete even for transversal matroids [CPV95].

Specifically, we will consider the following problem:

**Input:** A matroid  $\mathcal{M} = (E, \mathcal{I})$ , represented 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 of  $\mathcal{M}$ .

In the next lecture, we will see an fpras for this problem for an arbitrary matroid  $\mathcal{M}$ , running in time  $\text{poly}(n, \varepsilon^{-1})$  where  $n = |E|$ . We note that this actually implies an fpras for counting independent sets of *any* size  $k$ , since the *truncation* of  $\mathcal{M}$  to independent sets of size at most  $k$  is also a matroid, whose bases are precisely the independent sets of  $\mathcal{M}$  of size  $k$ . (Note, however, that this implication would not hold if we only had an algorithm for bases of a particular class of matroids, such as graphic matroids, since this class may not be closed under truncation.)

## 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.
- [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.
- [CPV95] C.J. Colbourn, J. Scott Provan, and D. Vertigan. The complexity of computing the Tutte polynomial on transversal matroids. *Combinatorica*, 15:1–10, 1995.
- [JWV90] F. Jaeger, D. Vertigan, and D. Welsh. On the computational complexity of the Jones and Tutte polynomials. *Mathematical Proceedings of the Cambridge Philosophical Society*, 108:35–53, 1990.