Down-Up Walks and Spectral Independence

1 Overview

- Last time, we introduced Down-Up Walks, Matroids, and Basis Exchange Walks
- Anari-Liu-Oveis-Gharan-Vinzant '19 Theorem: if the generating polynomial of μ is log-concave, then $1 - \lambda_2(P) = \Omega(1/k)$ (μ is uniform probability measure on matroids, or prob measure on $\binom{[n]}{k}$)
- Today we will focus on proving the theorem

2 Log Concavity and Generating Polynomials

Definition 1. The *generating polynomial* is $g_\mu(\lambda) = \sum_{S \in \binom{[n]}{k}} \mu(S) \prod_{u \in S} \lambda_u$ [where $\lambda = (\lambda_1, \cdots, \lambda_n)$]

Definition 2. The *tilt* is the probability measure on $\binom{[n]}{k}$ such that $(\mu \star \lambda)(S) \propto \mu(S) \prod_{u \in S} \lambda_u$

Remark 1. Some notes:

- $\mu = \mu_k$ is a measure on $\binom{[n]}{k}$, which can be interpreted as faces of a simplical complex
- Then $\mu_{k-1} = \mu D_{k \to k-1}$; where $D_{k \to k-1}$ is a Down operator that drops an element uniformly at random
- For this analogy, μ_1 is an induced measure on vertices, and μ_2 is that for edges
- Note that μ_1 is not necessarily uniform; consider the star graph

We can express $(\mu \star \lambda)(S) = \frac{1}{g_\mu(\lambda)} \mu(S) \prod_{u \in S} \lambda_u$, where $g_\mu(\lambda)$ is the normalizing constant for $(\mu \star \lambda)$.

Example 1. Spanning trees on a graph $G = ([n], E)$

- Define $\mu(S) \propto \mathbb{1}(S \text{ is a spanning tree of } G)$, where $S \in \binom{E}{n-1}$
- Then $(\mu \star \lambda)(S) \propto \mathbb{1}(S \text{ is a spanning tree of } G) \prod_{e \in S} \lambda_e$
- Previously, we saw how to sample μ, $(\mu \star \lambda)$ via Kirchoff Matrix-Tree Theorem (see Lecture 2)
- Recall $Z = \det(L_{\lambda,\mu})$ [weighted by λ]

Definition 3. μ (or g_μ) is *log-concave* if $\log g_\mu(\lambda)$ is concave on $\mathbb{R}^n_{>0}$

Remark 2. Why do we care about this definition? When is $g_\mu(\lambda)$ concave?

- This happens exactly when g_μ has degree 1
- Note $g_\mu(z, \cdots, z) = (\text{coeff})(z^k)$ is concave iff $k \leq 1$
- Meanwhile, $\log(z^k) = k \log(z)$ is concave for $z > 0$

Note that since $g_\mu(S) = \sum_S \mu(S) \prod_{i \in S} \lambda_i$, we have $\partial_{\lambda_j}(g_\mu(S)) = \sum_{S \text{ s.t.} j \in S} \mu(S) \prod_{i \in S \setminus \{j\}} \lambda_i$.

Also note that $\partial_{\lambda_j} \log(g_\mu(\lambda)) = \frac{\partial_{\lambda_j} g_\mu(\lambda)}{g_\mu(\lambda)}$.

Computation 1. Next we compute the Hessian of the log generating polynomial at $\lambda = \overrightarrow{1}$

- Clearly $g_\mu(\overrightarrow{1}) = 1$
- Then $\partial_{\lambda_j} \log(g_\mu(\lambda))|_{\lambda=\overrightarrow{1}} = \partial_{\lambda_j} g_\mu(\lambda)|_{\lambda=\overrightarrow{1}} = \sum_{S \text{ s.t.} j \in S} \mu(S) = Pr_\mu[j \in S]$
- For $i \neq j$, $\partial_{\lambda_i} \partial_{\lambda_j} \log(g_\mu(\lambda))_{\lambda=\overrightarrow{1}} = \frac{(g_\mu(\partial_{\lambda_i} \partial_{\lambda_j} g_\mu(\lambda)) - (\partial_{\lambda_i} g_\mu)(\partial_{\lambda_j} g_\mu)}{g_\mu(\lambda)^2}|_{\lambda=\overrightarrow{1}} = (\partial_{\lambda_i} \partial_{\lambda_j} g_\mu) - (\partial_{\lambda_i} g_\mu)(\partial_{\lambda_j} g_\mu)|_{\lambda=\overrightarrow{1}} = Pr_\mu[i \in S, j \in S] - Pr_\mu[i \in S] Pr_\mu[j \in S]$
- For $i = j$, $\partial^2_{\lambda_i} \log(g_\mu(\lambda)) = \partial_{\lambda_i} \frac{\partial_{\lambda_i} g_\mu(\lambda)}{g_\mu(\lambda)}|_{\lambda=\overrightarrow{1}} = -\frac{(\partial_{\lambda_i} g_\mu(\lambda))^2}{g_\mu(\lambda)^2}|_{\lambda=\overrightarrow{1}} = -Pr_\mu[i \in S]^2$

Computation 2. When is this Hessian negative semidefinite?

- Note that $Cov(X) = \mathbb{E}[XX^T] - \mathbb{E}[X]\mathbb{E}[X^T]$, so $Cov(\mathbb{1}_S) \succcurlyeq 0$
- Also $\partial_{\lambda_i} \partial_{\lambda_j} \log(g_\mu(S)) = Cov(\mathbb{1}_S)_{i,j}$
- Hence $\nabla^2 \log(g_\mu(\lambda))|_{\lambda=\overrightarrow{1}} = Cov(\mathbb{1}_S) - diag(Pr_\mu[i \in S, i \in S] - Pr_\mu[i \in S]^2) + diag(-Pr_\mu[i \in S]^2) \preccurlyeq 0$ if and only if $(0 \preccurlyeq)$ $Cov(\mathbb{1}_S) \preccurlyeq diag(Pr_\mu[i \in S]) = diag(\mathbb{E}[\mathbb{1}_S])$

3 Spectral Independence

Definition 4. μ is *C-spectrally independent* if $Cov_{S \sim \mu}(\mathbb{1}_S) \preccurlyeq C \cdot diag(\mathbb{E}[\mathbb{1}_S])$

Hence from the previous computations, μ log-concave implies 1-spectral independence, since this necessity condition comes from evaluating the Hessian at $\lambda = \overrightarrow{1}$.

Theorem 1. [AJKPV '24] Suppose P is the down-up walk and $1 - \lambda_2(P) = \frac{1}{Ck}$, then μ is C-SI.
This theorem shows that spectral independence is necessary for a large spectral gap, although we won't be using it for proving the direction we want to prove.

C-SI is necessary for rapid mixing, next we show 1-SI is sufficient for rapid mixing.

Definition 5. (Induced measure at link S) Let $S \subseteq [n]$, $|S| < k$. Define $\mu_S(T) \propto \mu(S \cup T)$ for $T \cap S = \emptyset$, $|T| + |S| = k$. Note $\mu_S(T) \leftrightarrow Pr_{R \sim \mu}(\cdot | S \subseteq R)$.

Fact 1. g_μ log-concave $\Rightarrow$ g_{μ_S} log-concave $\Rightarrow$ μ_S is 1-SI
We can prove this with the observation $g_{\mu_S}(\lambda_{\sim S}) \propto \lim_{r \to \infty} \frac{g_\mu(r \mathbb{1}_S, \lambda_{\sim S})}{r^{|S|}}$, then take the log, etc.

Fact 2. $(\mu \star \lambda)$ is 1-SI for all λ
We don't use this fact in our proof, since we only need to evaluate at $\lambda = \overrightarrow{1}$

Note $g_{\mu_S}(\lambda_{\sim S}) = \sum_T \mu(S \cup T) \prod_{i \in T} \lambda_i = \sum_{S \subseteq R} \mu(R) \prod_{i \in R \setminus S} \lambda_i \prod_{i \in S} 1$

Theorem 2. Suppose μ is 1-SI at all links, then $1 - \lambda_2(P) \geq 1/k$ [here $P = D_{k \to k-1} U_{k-1 \to k}$]
In the lecture notes, the RHS was written as $\prod_{i=0}^{k-2} (1 - \frac{1}{k-i})$, which is the same as $1/k$ by telescoping

For a brief sketch of the proof, first note that the Poincaré inequality for P is equivalent to $Var(f) \leq \frac{1}{1-\lambda_2}\mathbb{E}_{S_{k-1}\sim\mu_{k-1}}Var(f(S)|S_{k-1})$. The proof is basically the same as that in Glauber Dynamics.

Recall 1. The Law of Total Variance says $Var(f) = \mathbb{E}[Var(f(S)|\Delta_1] + Var(\mathbb{E}[f(S)|\Delta_1])$, where $S = \{\Delta_1, \cdots, \Delta_k\}$

Lemma 1. C-SI is equivalent to $Var(\mathbb{E}[f(S)|\Delta_1]) \leq \frac{C}{k}Var(f)$ [proof next class]

Assuming this lemma, we can use the Law of Total Variance to bound $(1-\frac{1}{k})Var(f) \leq \mathbb{E}[Var(f(S)|\Delta_1]$, then 1-SI at all links lets us show $(1-\frac{1}{k-1})Var(f(S)|\Delta_1) \leq \mathbb{E}[Var(f(S)|\Delta_1, \Delta_2]$, then combining these give us $(1-\frac{1}{k})(1-\frac{1}{k-1})Var(f) \leq \mathbb{E}[Var(f(S)|\Delta_1, \Delta_2]$, etc. and use induction to prove the theorem.