Lecture 20 - Average Conductance, Easy Inequality Prover

Friday, October 1

20.1 Average Conductance

In the previous class we proved a theorem relating the congestion quantities $\mathcal{C}_{TV}(x)$, $\mathcal{C}_D(x)$, and $\mathcal{C}_{L^2}(x)$ to mixing times. Originally the set sizes were considered in the context of the *Average Conductance theorem.* This result was originally proven by Lovász and Kannan [1] to bound total variation distance of lazy Markov chains in terms of $\Phi(A)$, improved by Morris and Peres [5] to bound L^2 distance in terms of a congestion quantity, and I have improved it further by extending to the non-lazy case [2] (Morris and Peres give a non-lazy version as well, but ours is more natural and is stronger).

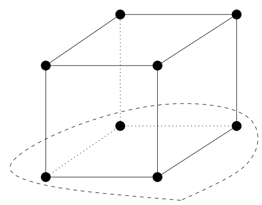
Corollary 20.1 (Average Conductance). *If $\mathcal{M}$ is an ergodic Markov chain then*

$$\tau(\epsilon) \leq L^2(2\epsilon) \leq \int_{\pi_*}^{1/2} \frac{dx}{x(1-x)\tilde{\phi}(x)^2} + \frac{2\log(1/\epsilon)}{\tilde{\phi}^2}$$

where $\tilde{\phi}(x) = \min_{\substack{A\subset\Omega,\\ \pi(A)\leq x}} \tilde{\phi}(A)$ *and* $\tilde{\phi}(A) = \tilde{\Phi}(A)$ *when $\mathcal{M}$ is lazy.*

This follows immediately from Corollary 19.3 because $1 - \mathcal{C}_{L^2}(A) \geq \tilde{\phi}(A)^2/2$.

Example 20.2. Consider the lazy random walk on the boolean cube $\{0,1\}^d$.



Observe that if A is a single point then $\tilde{\Phi}(A) \geq \frac{1}{2d}$, but if A is the lower half space then $\tilde{\Phi}(A) = \frac{1}{d}$. Therefore small sets have very high edge-expansion, while large sets have fairly low edge-expansion.

On the homework you will determine that the log-Sobolev constant is $\rho = \Omega(1/d)$, and so the mixing time is $\tau(\epsilon) = O(d\log d + d\log(1/\epsilon))$. This is in fact correct.

If we consider only **worst-case edge-expansion** $\tilde{\Phi} = 1/d$, with the worst case being when A is a lower half space. It follows that

$$\tau(\epsilon) \leq L^2(2\epsilon) \leq \frac{2}{\tilde{\Phi}^2}\left[\frac{1}{2}\log\frac{1-\pi_*}{\pi_*} + \log(1/2\epsilon)\right] \leq d^3\log d + d^2\log(1/2\epsilon)\,,$$

not particularly good.

Next week we will learn how to study **average-case edge-expansion** and see that $\tilde{\Phi}(x) = \Omega(\log(1/x)/d)$. The Average Conductance theorem implies that

$$\tau(\epsilon) \leq L^2(2\epsilon) = O(d^2 + d^2\log(1/2\epsilon))\,,$$

a fair improvement and only a single order of magnitude from correct.

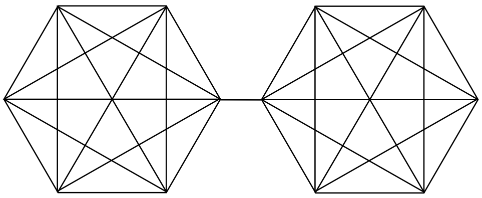
The congestion measure $\mathcal{C}_{L^2}(x)$ considers **average-case edge & vertex-expansion**. It can be shown [3] that $1 - \mathcal{C}_{L^2}(x) = \Omega\left(\frac{\log(1/x)}{d \log d}\right)$, from which the theorem last class implies that

$$\tau(\epsilon) \le L^2(2\epsilon) = O(d \log^2 d + d \log d \log(1/2\epsilon)) .$$

I suspect that in fact $1 - \mathcal{C}_{L^2}(x) = \Omega\left(\frac{\log(1/x)}{d}\right)$, but have not been able to show this.

We see that knowing both vertex and edge expansion can give a substantial improvement over the edge bound alone.

Exercise 1. Compare canonical path, conductance, and average conductance bounds for the mixing time of the lazy max-degree walk on the Barbell $K_n - K_n$ below. What about the regular (non-lazy) max-degree walk?



20.2 Easy Inequality Prover

In the past week we have learned about Evolving sets, shown the connection to mixing times, and found that a non-lazy form of Conductance can be used to study the evolving sets and therefore mixing times. However, evolving sets have one large advantage over conductance, in that they are related to both vertex and edge expansion of a graph. In order to better study the evolving set results we need an easy way to relate measures of congestion to our evolving set congestion bounds.

The following lemma makes it much easier to prove inequalities (see [2, 3, 4]).

Lemma 20.3 (Easy Inequality Prover?). *Given $f : [0,1] \to \mathbb{R}$ concave and $g, g' : [0,1] \to [0,1]$ non-increasing such that $\int_0^1 g(u)\,du = \int_0^1 g'(u)\,du$ and $\forall t \in [0,1] : \int_0^t g(u)\,du \ge \int_0^t g'(u)\,du$, then*

$$\int_0^1 f \circ g(u)\,du \le \int_0^1 f \circ g'(u)\,du .$$

Proof. The concavity of $f(x)$ implies that

$$\forall x \ge y,\, \delta \ge 0 : \, f(x) + f(y) \ge f(x+\delta) + f(y-\delta) . \tag{1}$$

This follows because $y = \lambda\,(y-\delta) + (1-\lambda)\,(x+\delta)$ with $\lambda = 1 - \frac{\delta}{x-y+2\delta} \in [0,1]$ and so by concavity $f(y) \ge \lambda\, f(y-\delta) + (1-\lambda)\, f(x+\delta)$. Likewise, $x = (1-\lambda)\,(y-\delta) + \lambda\,(x+\delta)$ and $f(x) \ge (1-\lambda)\, f(y-\delta) + \lambda\, f(x+\delta)$. Adding these two inequalities gives (1).

The inequality (1) shows that if a bigger value (x) is increased by some amount, while a smaller value (y) is decreased by the same amount, then the sum $f(x) + f(y)$ decreases. In our setting, the condition that $\forall t \in [0,1] : \int_0^t g(u)\,du \ge \int_0^t g'(u)\,du$ shows that changing from g' to g increased the already large values of $g'(u)$, while the equality $\int_0^1 g(u)\,du = \int_0^1 g'(u)\,du$ assures that this is canceled out by an equal decrease in the already small values. The lemma then follows from (1). $\square$

In our applications we set $g(u) = \pi(A_u)$. The remarkable fact about the lemma is that, given some initial conditions, if there is a distribution of $\pi(A_u)$ which minimizes $\int_0^t \pi(A_u)\,du$ for all $t \in [0,1]$ then this same distribution maximizes $\mathcal{C}_f$ no matter which concave function $f(x)$ is chosen!

Example 20.4. Consider a Markov chain (non-lazy) and a subset $A \subset \Omega$, with $\Psi(A)$ known. Then $\pi(A_u) \in [0,1]$ is non-increasing, $\Psi(A)$ is the area below $\pi(A_u)$ and above $\pi(A)$, and also above $\pi(A_u)$ and below $\pi(A)$.

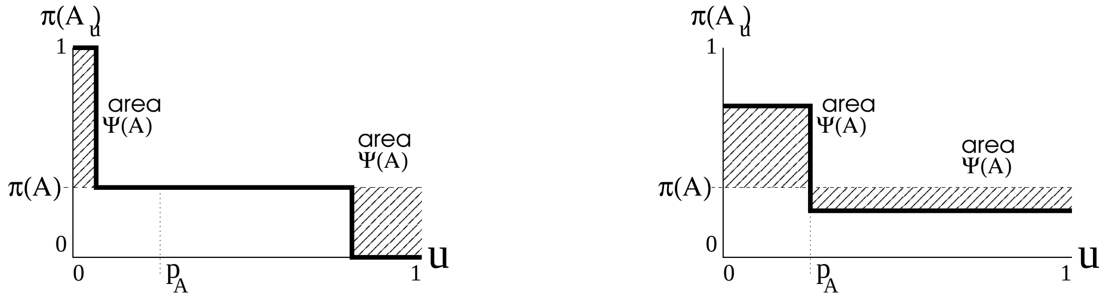


Figure 1: Maximizing $\int_0^t \pi(A_u)\,du$ and minimizing $\int_0^t \pi(A_u)\,du$ with only $\Psi(A)$ fixed.

The choice of $\pi(A_u)$ which maximized $\int_0^t \pi(A_u)\,du$ in the figure is given by

$$\pi(A_u) = \begin{cases} 0 & \text{if } u > 1 - \frac{\Psi(A)}{\pi(A)} \\ \pi(A) & \text{if } u \in \left(\frac{\Psi(A)}{1-\pi(A)},\, 1 - \frac{\Psi(A)}{\pi(A)}\right] \\ 1 & \text{if } u \le \frac{\Psi(A)}{1-\pi(A)} \end{cases}$$

By Lemma 20.3 any choice of $f(x)$ which is concave, with $f(0) = f(1) = 0$ and $\forall x \in (0,1) : f(x) > 0$, will therefore satisfy

$$\begin{aligned} \mathcal{C}_f(A) &= \int_0^1 \frac{f(\pi(A_u))}{f(\pi(A))}\,du \ge \frac{\Psi(A)}{\pi(A)}\,\frac{f(0)}{f(\pi(A))} + \left(1 - \frac{\Psi(A)}{\pi(A)\pi(A^c)}\right)\frac{f(\pi(A))}{f(\pi(A))} + \frac{\Psi(A)}{1-\pi(A)}\,\frac{f(1)}{f(\pi(A))} \\ &= 1 - \frac{\Psi(A)}{\pi(A)\pi(A^c)} = 1 - \tilde{\phi}(A) \end{aligned}$$

This immediately implies all of the upper bounds in Theorem 17.3, and because we explicitly constructed the worst case then this also implies that all of the upper bounds are sharp!

The figure for the lower bound can also be written out explicitly and gives the same result as did our use of Jensen's inequality did. This implies that, at least before simplification was done, all of the lower bounds were sharp as well.

References

[1] L. Lovász and R. Kannan. Faster mixing via average conductance. *Proc. 31st Annual ACM Symposium on Theory of Computing*, pages 282–287, 1999.

[2] R. Montenegro. Evolving sets and bounds on various mixing quantities. *Preprint*, 2004.

[3] R. Montenegro. Evolving sets and cheeger profile bounds on the top and bottom eigenvalues. *Preprint*, 2004.

[4] R. Montenegro. Mixed path and conductance profile bounds for various measures of mixing time. *Preprint*, 2004.

[5] B. Morris and Y. Peres. Evolving sets and mixing. *Proc. 35th Annual ACM Symposium on Theory of Computing*, pages 279–286, 2003.