CS294 Partition Functions **Fall 2020**

Lecture 14: October 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.*

Note: *This note is shorter than usual as a large part of today's lecture was devoted to completing the Ising model discussion from the previous lecture.*

14.1 Relating mixing times using flows

This brief note describes a flow-based method for bounding the Poincaré constant (or, in the reversible case, the spectral gap) of a Markov chain based on that of another Markov chain, due to Diaconis and Saloff-Coste [DSC93]. Assume we are given two ergodic, lazy Markov chains P and $\tilde{P}$ which share the same stationary distribution π over some state space Ω. Furthermore, assume that we have already proved a lower bound on the Poincaré constant $\tilde{\alpha}$ of $\tilde{P}$ by some other method. By constructing a flow on P that "simulates" the transitions of $\tilde{P}$, we can bound the ratio of the Poincaré constant of P to that of $\tilde{P}$ by a constant determined by the characteristics of the flow.

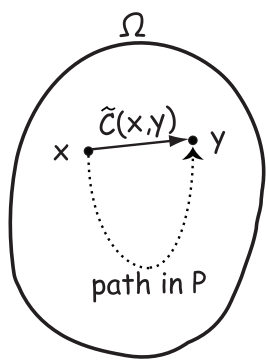

Figure 14.1: A $(P, \tilde{P})$-flow

Definition 14.1. *Let $\mathcal{Q}_{xy}$ denote the set of all simple $x \rightsquigarrow y$ paths in P, and $\mathcal{Q} = \cup_{x,y}\mathcal{Q}_{xy}$. A $(P, \tilde{P})$-flow is a function $f : \mathcal{Q} \to \mathbb{R}^+ \cup \{0\}$ subject to demand constraints*

$$\forall x, y: \quad \sum_{p \in \mathcal{Q}_{xy}} f(p) = \tilde{C}(x, y),$$

where $\tilde{C}(x, y) = \pi(x)\tilde{P}(x, y)$ (the capacity of edge (x, y) in $\tilde{P}$).

Recall that we define the cost of a flow f as

$$\rho(f) = \max_e \frac{f(e)}{C(e)}$$

and the length of a flow as

$$\ell(f) = \max_{p: f(p)>0} |p|.$$

Claim 14.2. *For any $(P, \tilde{P})$ flow f and function $\varphi : \Omega \to \mathbb{R}$, we have*

$$\mathcal{E}_P(\varphi,\varphi) \geq \frac{1}{\rho(f)\ell(f)} \mathcal{E}_{\tilde{P}}(\varphi,\varphi).$$

The proof of this fact follows that of Theorem 11.1 in Lecture 11; it involves expanding out $\mathcal{E}_{\tilde{P}}(\varphi,\varphi)$, applying the Cauchy-Schwarz inequality, and summing over paths.

Exercise: Prove Claim 14.2 via pattern matching on the proof of Theorem 11.1.

Theorem 14.3. *For any $(P, \tilde{P})$-flow f,*

$$\frac{\alpha}{\tilde{\alpha}} \geq \frac{1}{\rho(f)\ell(f)}$$

Proof. Recall from Lecture 10 that

$$\alpha := \inf_{\varphi} \frac{\mathcal{E}_P(\varphi,\varphi)}{\mathrm{Var}_\pi[\varphi]},$$

where the inf is taken over non-constant functions $\varphi : \Omega \to \mathbb{R}$. Thus, we can conclude

$$\begin{aligned} \frac{\alpha}{\tilde{\alpha}} &= \frac{\inf_\varphi \frac{\mathcal{E}_P(\varphi,\varphi)}{\mathrm{Var}_\pi[\varphi]}}{\inf_\varphi \frac{\mathcal{E}_{\tilde{P}}(\varphi,\varphi)}{\mathrm{Var}_\pi[\varphi]}} \\ &\geq \inf_\varphi \frac{\mathcal{E}_P(\varphi,\varphi)}{\mathcal{E}_{\tilde{P}}(\varphi,\varphi)} \\ &\geq \frac{1}{\rho(f)\ell(f)}, \end{aligned}$$

by Claim 14.2. □

Note that the original flow theorem, Theorem 11.1, can be seen as a special case of Theorem 14.3 in which $\tilde{P}$ is the trivial Markov chain with $\tilde{P}(x,y) = \pi(y)$, which mixes in one step and has a Poincaré constant of 1. Theorem 14.3 can be very useful, because other methods of analysis are often rather sensitive to the details of the Markov chain being analyzed; the theorem allows us to deduce bounds on the mixing time of a Markov chain by comparing it to another chain that is simpler to analyze.

14.2 Example

We give a very simple example to illustrate the application of Theorem 14.3. Recall the Markov chain on matchings from Lecture 12. This chain has three types of local moves: edge addition, edge deletion and edge exchange. Suppose we're interested in the mixing time of the simpler chain that allows edge additions and deletions only (no exchanges). This is arguably a more natural process, and much simpler to implement; indeed, the exchange moves in the original chain are included to ease the analysis. We can easily bootstrap our mixing time bound from Lecture 12 to this new chain using Theorem 14.3.

We'll denote by $\tilde{P}$ the chain from Lecture 12, and by P the simpler chain. To apply the theorem, we need to specify a flow in P for each transition of $\tilde{P}$. For edge additions and deletions, this is trivial because those transitions are also present in P. For an edge exchange (say, from edge $\{u,v\}$ to $\{v,w\}$), we construct a two-step path in P by first deleting the edge $\{u,v\}$ and then adding $\{v,w\}$.

How much flow is carried by a transition in P? Consider an edge deletion transition $e = (z, z')$, where z' is the matching z with edge $\{u,v\}$ deleted. This carries the same deletion in $\tilde{P}$, whose flow is $\tilde{C}(z,z') = \pi(z)\tilde{P}(z,z') = \frac{\pi(z)}{2\lambda|E|}$, plus all edge exchanges of the form $\{u,v\} \to \{v,w\}$ as w varies. There are at most $\deg(v)$ such exchanges, each of which carries a flow of $\frac{\pi(z)}{2|E|}$. Adding all this up gives $f(e) \le \frac{\pi(z)}{2|E|}(\Delta + \frac{1}{\lambda})$, where Δ is the maximum degree of G. Since $C(e) = \frac{\pi(z)}{2\lambda|E|}$, we get $\frac{f(e)}{C(e)} \le \Delta\lambda + 1$. A similar calculation **[exercise]** holds for edge additions e. Hence we get $\rho(f) \le \Delta\lambda + 1$. Since also $\ell(f) = 2$, we see that the ratio of the two Poincaré constants is bounded by $2(\Delta\lambda + 1)$. Hence, since our bound on the mixing time of the original chain was obtained via its Poincaré constant, we get a bound on the mixing time of the simpler chain that is at most a factor $O(n\lambda)$ larger than that of the original chain (or $O(\lambda)$ in the case of bounded degree graphs).

Note: The above technology can be used even in settings where we have obtained the mixing time $\tilde{\tau}_{\rm mix}$ of the original chain $\tilde{P}$ not via the Poincaré constant but by some other method (such as coupling). In that case we use the fact that $\tilde{\tau}_{\rm mix} \ge \frac{1}{\tilde{\alpha}}$ and compare the Poincaré constants to obtain the bound $\tau_{\rm mix} \le (\frac{\tilde{\alpha}}{\alpha} \log \pi_{\rm min}^{-1})\tilde{\tau}_{\rm mix}$ on the mixing time of the new chain, where the factor $\log \pi_{\rm min}^{-1}$ arises from the fact that we are now using the Poincaré constant. An important example of this type is Vigoda's result, mentioned in an earlier lecture, that Glauber dynamics for colorings has polynomial mixing time when the number of colors $q \le \frac{11}{6}\Delta$ [Vig99]. The proof proceeds via a coupling analysis of a more complicated chain that recolors neighborhoods of radius up to 6 around a vertex; the mixing time bound is then transferred to Glauber dynamics using the comparison method above.

References

[DSC93] P. Diaconis and L. Saloff-Coste. Comparison theorems for reversible Markov chains. *Annals of Applied Probability*, 3:696–730, 1993.

[Vig99] E. Vigoda. Improved bounds for sampling colorings. *Proceedings of the 40th IEEE Symposium on Foundations of Computer Science*, pages 51–59, 1999.