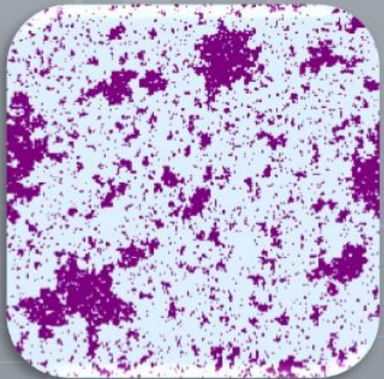
Markov Chain Minicourse

PIMS

Summer school in Probability



lecture 3

Eyal Lubetzky

Microsoft Research

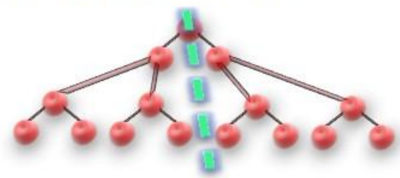
Lower bounds via conductance

- Recall from last lecture:
  - For a chain with transition kernel P and stationary distribution π define:

$$Q(x,y) \triangleq \pi(x)P(x,y) \;\;;\;\; Q(A,B) \triangleq \sum_{x\in A,\, y\in B} Q(x,y).$$

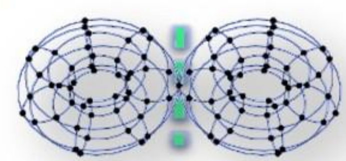
  - The *conductance* (or *bottleneck ratio*) of a set S is

$$\Phi(S) \triangleq \frac{Q(S,S^c)}{\pi(S)}$$



and the *conductance* (*Cheeger constant*) of the chain is

$$\Phi \triangleq \min_{S:\pi(S)\le\frac{1}{2}} \Phi(S).$$

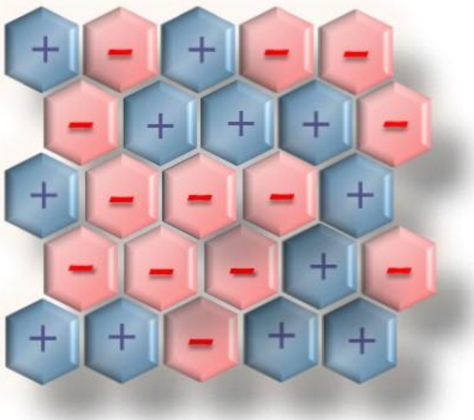


- <u>THEOREM</u>:

Every Markov chain satisfies $t_{\text{mix}}(\frac{1}{4}) \ge \dfrac{1}{4\Phi}$.

Bottlenecks in Glauber for Ising

- Recall the definition of the dynamics:
  - Update sites via *iid* Poisson(1) clocks
  - Each update replaces a spin at $u \in V$ by a new one $\sim \mu$ conditioned on $V \setminus \{u\}$ (heat-bath version).

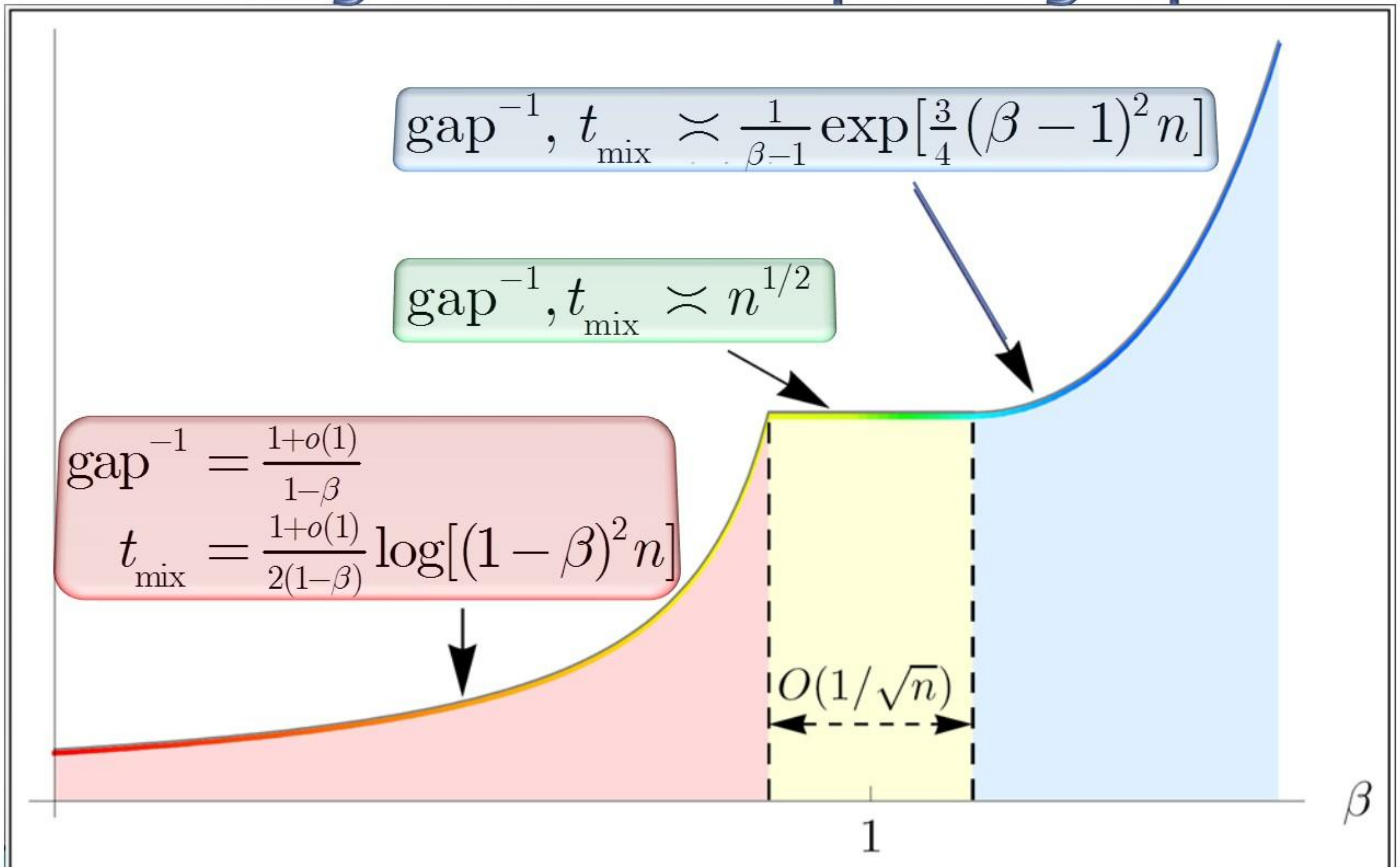


- *How fast does it converge to equilibrium?*
  - Can be exponentially slow in the size of the system: At low temp. (large β) there may be a bottleneck between "plus" and "minus" states (see tutorial).

General (believed) picture for the Glauber dynamics

- Setting: Ising model on the lattice $(\mathbb{Z}/n\mathbb{Z})^d$.
  Belief: For some critical inverse-temperature β_c :
- Low temperature: $(\beta > \beta_c)$
  gap^{-1} and t_{mix} are ***exponential*** in the surface area.
- Critical temperature: $(\beta = \beta_c)$
  gap^{-1} and t_{mix} are ***polynomial*** in the surface area.
- High temperature: $(\beta < \beta_c)$
  - ***Rapid*** mixing: $\text{gap}^{-1} = O(1)$ and $t_{\text{mix}} \asymp \log n$
  - Mixing occurs abruptly, *i.e.* there is ***cutoff***.

Gap/mixing-time evolution for Ising on the complete graph



(Scaling window established in [Ding, L., Peres '09])

Bottleneck in sampling colorings

- A *legal coloring* of an undirected graph $G=(V,E)$ is a mapping $\varphi: V\to\mathbb{N}$ such that $\varphi(u)\neq\varphi(v)$ for all $(u,v)\in E$.
- Problem definition:
  - Input: Undirected graph $G=(V,E)$ and integer q.
  - Goal: Sample a uniform legal coloring via q colors.
- Is there even a single legal coloring?
  - In general this is ***NP-complete*** to determine.
  - Main interest: graphs that are k-colorable for some small k (*e.g.* graphs with maximal degree $\Delta = O(1)$).
- How can we sample a coloring uniformly?

Sampling recipe for legal colorings

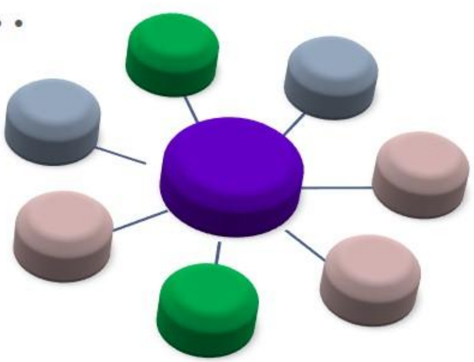
- Glauber dynamics for colorings :
  - Markov chain on $\Omega =$ legal colorings ($\Omega \subseteq [q]^V$) .
  - Start at an arbitrary legal coloring.
  - Transition rule:
    - Choose a uniform vertex $v \in V$.
    - Replace its color by a uniformly chosen color out of all legal ones (*i.e.* not occupied by neighbors).
- Reversible with respect to the uniform distribution π since the transition kernel is symmetric.
- How long does it take the chain to converge to π ?
  - (*We will later see that* $t_{\mathrm{mix}} = O(|V| \log |V|)$ *when* $q > 2\Delta$)

Slow mixing with large degrees

- PROPOSITION:

The Glauber dynamics for colorings of the n-vertex star via $q \geq 3$ colors has $t_{\mathrm{mix}} \geq \frac{1}{16} n\, \mathrm{e}^{n/(q-1)}$.

- Few colors here analogous to low temperature Ising...
- In this example we can easily color the graph using **2** colors yet sampling a **100**-coloring uniformly via Glauber is exponentially slow in n...
- Where is the bottleneck?
  - Let S be all colorings assigning the color **1** to middle vertex...



Slow coloring of the star (ctd.)

- Def.: $S = \{\sigma \in \Omega : \sigma(v_0) = 1\}$. ($|S| = q^{n-1}$)
- For all $\sigma \in S$, $\sigma' \in S^c$ we have $Q(\sigma, \sigma') = 0$ unless:
  - $\sigma(v_0) = 1$ and $\sigma'(v_0) \neq 1$,
  - $\sigma(u) = \sigma(u)$ for every leaf u , and
  - $\sigma(u) \notin \{1, \sigma'(v_0)\}$ for every leaf u .
- Since there are $(q-1)(q-2)^{n-1}$ such pairs, each satisfying $Q(\sigma, \sigma') \leq 1/(|\Omega| n)$, we get

$$Q(S, S^c) \leq \frac{1}{|\Omega| n}(q-1)(q-2)^{n-1},$$

and so

$$\frac{Q(S,S^c)}{\pi(S)} \leq \frac{(q-1)(q-2)^{n-1}}{n(q-1)^{n-1}} \leq \frac{(q-1)^2}{n(q-2)} e^{-n/(q-1)}. \quad \blacksquare$$

Path coupling (⇒upper bound for coloring)

- Def.: a *premetric* on Ω is a connected undirected graph $H=(\Omega,E)$ with positive edge weights $w{:}E\to\mathbb{R}^+$ so that
  - If $e=(x,y)\in E$ then $w(e)\le w(\Gamma)$ $\forall$ path Γ between x,y .
- Let d_H denote the metric extending the premetric H.
- THEOREM: [Bubley,Dyer '97]

Let $H=(\Omega,E_H)$ be a premetric for Ω and suppose that for some $\rho>0$ and $\forall\ x,y\in\ E_H$ there $\exists$ a coupling such that

$$\mathbb{E}\left[d_H(X_1,Y_1)\mid X_0=x,Y_0=y\right]\le(1-\rho)d_H(x,y).$$

Then there $\exists$ such a coupling for $\forall x,y\in\Omega$.

Path coupling bounds mixing

- COROLLARY:

Let $H=(\Omega,E_H)$ be a premetric for Ω with integer weights. Suppose that for some $\rho>0$ and $\forall\, x,y \in E_H$ there exists a coupling such that

$$\mathbb{E}\left[d_H(X_1,Y_1) \mid X_0 = x, Y_0 = y\right] \leq (1-\rho) d_H(x,y).$$

Then the mixing time of (X_t) satisfies

$$t_{\text{mix}}(\varepsilon) \leq \frac{1}{\rho}\left[\log(\text{diam}(\Omega)) + \log(\tfrac{1}{\varepsilon})\right],$$

where $\text{diam}(\Omega) \triangleq \max\{d_H(x,y) : x,y \in \Omega\}$.

Path coupling (ctd.)

- PROOF:

Let $x,y \in \Omega$ (not necessarily adjacent in H), and let

$$\Gamma = (x = u_0, u_1, \ldots, u_k = y)$$

be a shortest path between x,y in H.

Couple X_1, Y_1 started at x,y by composing couplings:

- Base: couple X, Y started at (x,u_1) satisfying ✶.
- Extend a coupling of (X,Y) *from* (x,u_i) to a coupling of (X,Z) *from* (x,u_{i+1}) via a coupling of (Y,Z) *from* (u_i,u_{i+1}) [generate (X_1,Y_1) then generate (Y_1,Z_1) conditioned on Y_1].
- This satisfies ✶ since:

$$\mathbb{E}_{x,u_{i+1}}\left[d_H(X_1,Z_1)\right] \le \mathbb{E}_{x,u_i}\left[d_H(X_1,Y_1)\right] + \mathbb{E}_{u_i,u_{i+1}}\left[d_H(Y_1,Z_1)\right]$$
$$\le (1-\rho)\Big(d_H(x,u_i) + d_H(u_i,u_{i+1})\Big) = (1-\rho)d_H(x,u_{i+1}). \quad \blacksquare$$

Example: Sampling legal coloring

- THEOREM: ([Jerrum '95], [Salas, Sokal '97])

Let G be a graph on n vertices with maximum degree Δ. If $q > 2\Delta$ then the Glauber dynamics for legal colorings of G via q colors has $t_{\text{mix}}(\varepsilon) \leq \frac{q-\Delta}{q-2\Delta} n[\log(n) + \log(\frac{1}{\varepsilon})]$.

- PROOF:

Premetric: connect $x, y \in [q]^n$ (possibly illegal) in H iff they differ in a single coordinate (extends to Hamming distance).

The statement of the theorem will follow from providing a path coupling satisfying the contraction ✶ where:

$$\rho = \frac{q - 2\Delta}{(q-\Delta)n}.$$

Sampling legal colorings (ctd.)

- A contracting coupling on H:
  Take two states x,y that differ only at vertex v.
  - Update the vertex v itself : coalesce
  - Update some u not adjacent to v : identity.
  - Update u adjacent to v : available color lists are $\mathcal{C}_x \triangleq \mathcal{C} \setminus x(v)$ and $\mathcal{C}_y \triangleq \mathcal{C} \setminus y(v)$ for some $\mathcal{C} \subseteq [q]$.
    - If $|\mathcal{C}_x| = |\mathcal{C}_y| \in \mathcal{C}$: couple $\mathcal{C}_x$, $\mathcal{C}_y$ via swapping $x(v)$,$y(v)$ and the identity-coupling elsewhere.
    - Else: w.l.o.g. $|\mathcal{C}_x|=|\mathcal{C}_y|-1$. Let $y'(u) \in \mathcal{C}_y$ uniformly.
      - If $y'(u) \neq x(v)$ then reuse it for $x'(u)$.
      - Else: Let $x'(u) \in \mathcal{C}_x$ uniformly.

Sampling legal colorings (ctd.)

- Accounting:
  - Eliminating a disagreement ⇔ Updating v. $1/n$
  - New disagreement ⇔ Updating $u \sim v$ and selecting the color $x(v)$ for $y'(u)$. $\leq \frac{\Delta}{n} \cdot \frac{1}{q-\Delta}$
- Altogether:

$$\mathbb{E}_{x,y}\left[d_H(X_1,Y_1)\right] \leq 1 - \frac{1}{n}\left(1 - \frac{\Delta}{q-\Delta}\right) = 1 - \frac{q-2\Delta}{(q-\Delta)n}. \quad \blacksquare$$