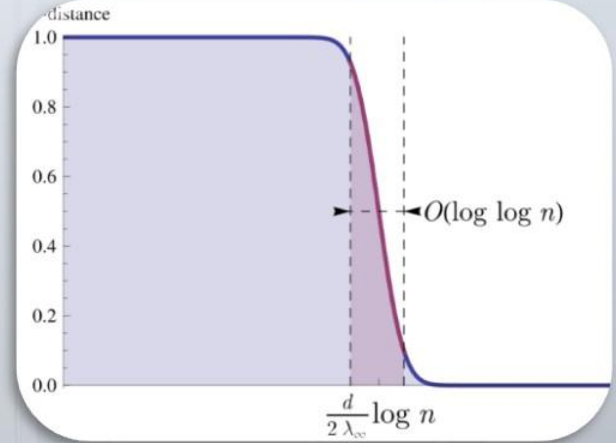


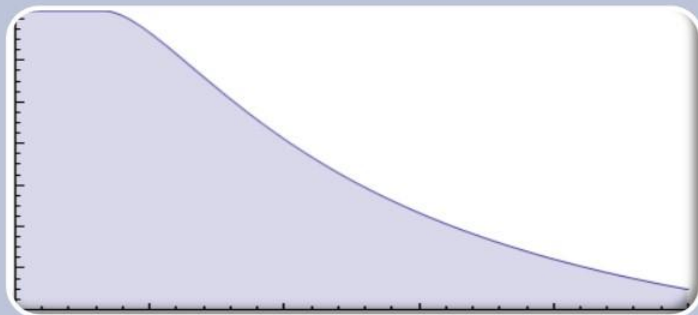
La Pietra 2011
Mini course
lecture 2

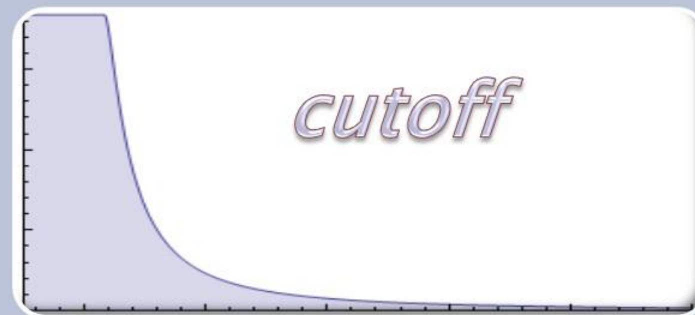
Cutoff for Ising on the lattice

Eyal Lubetzky
Microsoft Research

The Cutoff Phenomenon

- Describes a sharp transition in the convergence of finite ergodic Markov chains to stationarity.





Steady convergence
it takes a while to reach distance ½ from stationarity then a while longer to reach distance ¼, etc.

Abrupt convergence
distance from equilibrium quickly drops from 1 to 0

Cutoff: formal definition

- A family of chains (X_t^n) is said to have ***cutoff*** if:

$$\lim_{n\to\infty} \frac{t_{\mathrm{mix}}(\varepsilon)}{t_{\mathrm{mix}}(1-\varepsilon)} = 1 \quad \forall\ 0 < \varepsilon < 1.$$

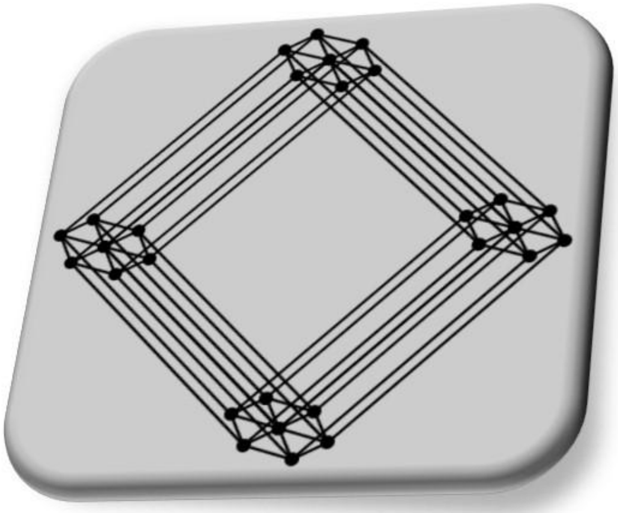
i.e., $t_{\mathrm{mix}}(\alpha) = (1+o(1))\,t_{\mathrm{mix}}(\beta)$ for any $0 < \alpha\ ,\beta < 1$.

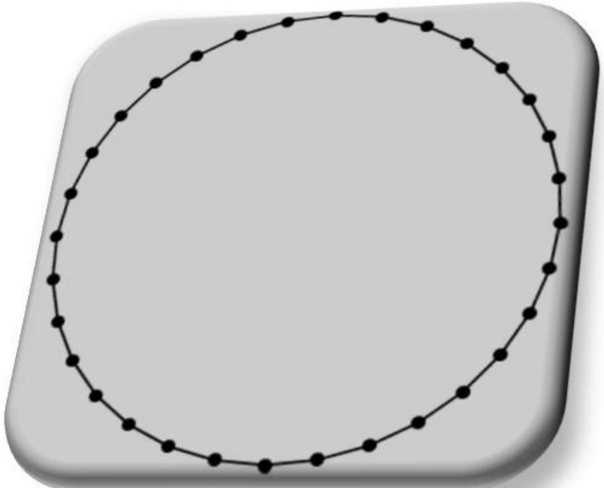
- A sequence (w_n) is called a ***cutoff window*** if

$$w_n = o(t_{\mathrm{mix}}(\tfrac{1}{4}))\ ,$$
$$t_{\mathrm{mix}}(\varepsilon) - t_{\mathrm{mix}}(1-\varepsilon) = O_\varepsilon(w_n) \quad \forall\, 0 < \varepsilon < 1\ .$$

Basic examples

Lazy discrete-time simple random walk



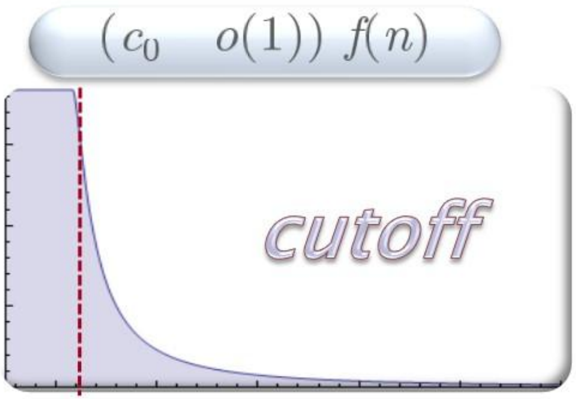


On the hypercube $\{-1,1\}^n$:
☑ Exhibits cutoff at
$\frac{1}{2} \log n + O(1)$
[Aldous '83]

On the n-cycle:
☒ No cutoff.

The importance of cutoff

- Suppose we run Glauber dynamics for the Ising Model satisfying $t_{\mathrm{mix}} \asymp f(n)$ for some $f(n)$.
- Cutoff $\Leftrightarrow$ $\exists$ some $c_0 > 0$ so that:
  - Must run the chain for at least $\sim c_0 \cdot f(n)$ steps to even reach distance $(1-\varepsilon)$ from μ.
  - Running it any longer than that is essentially redundant.



- Proofs usually require (and thus provide) a deep understanding of the chain (its reasons for mixing).
- Many natural chains are *believed* to have cutoff, yet proving cutoff can be extremely challenging.

Cutoff History

- Random walks on graphs and groups:
  - Discovered:
    - Random transpositions on S_n [Diaconis, Shahshahani ‘81]
    - RW on the hypercube, Riffle-shuffle [Aldous ‘83]
  - Named "Cutoff Phenomenon" in the top-in-at-random shuffle analysis [Diaconis, Aldous ‘86]
  - RWs on finite groups [Saloff-Coste ‘04]
  - RWs on random regular graphs [L., Sly ’10]
- One-dimensional Markov chains:
  - Birth-and-Death chains
    [Diaconis, Saloff-Coste ’06], [Ding, L., Peres ’09]
- No proofs of cutoff except when stationary distribution is completely understood and has many symmetries [*till recently*]

Peres' Product Criterion

- QUESTION [Diaconis '96]: How can we determine whether a given Markov chain exhibits cutoff?
- OBSERVATION [Peres '04]: if a reversible chain has cutoff then $\text{gap} \cdot t_{\text{mix}}(1/4) \to \infty$
  or equivalently: $t_{\text{rel}} = o\big(t_{\text{mix}}(1/4)\big)$.
- PROOF:

> - Key fact: every reversible cont-time MC satisfies
>
> $$t_{\text{mix}}(\varepsilon) \geq \text{gap}^{-1} \log(\tfrac{1}{2\varepsilon}).$$
>
> - Assume that $t_{\text{rel}} \geq \delta\, t_{\text{mix}}(1/4)$ for some $\delta > 0$.
> - It follows that $t_{\text{mix}}(\varepsilon) \geq f(\varepsilon) \cdot t_{\text{mix}}(1/4)$ where $f(\varepsilon) \underset{\varepsilon \to 0}{\longrightarrow} \infty$
> $\Rightarrow$ No (pre) cutoff. ■

Peres' Product Criterion (ctd.)

- The condition $\otimes$ is necessary for cutoff.
  Is it also sufficient, giving a method to determine the existence of cutoff?
- [Aldous '04]: unfortunately *not*: the product-condition $\otimes$ does not imply cutoff (explicit construction).
- Even so, Peres conjectured that for many natural families of chains, *cutoff* occurs iff $\otimes$.
  (*e.g.*, holds for birth-and-death chains [Ding, L., Peres '09]).

$$\otimes \quad \mathrm{gap}\cdot t_{\mathrm{mix}}\left(\tfrac{1}{4}\right) \to \infty \quad \Longleftarrow \quad \overset{?}{\Longrightarrow} \quad \textit{cutoff}$$

- Notable conjectured examples:
  - **Ising** on lattices ; **Potts** model on lattices;
  - **Gas Hard-core** model on lattices; lattice **Colorings** ;
  - **Anti-ferromagnetic** Ising / Potts model, **Spin-glass**,
  - Arbitrary boundary conditions / external field; ...

Cutoff for Ising on lattices

- THEOREM [L., Sly]:

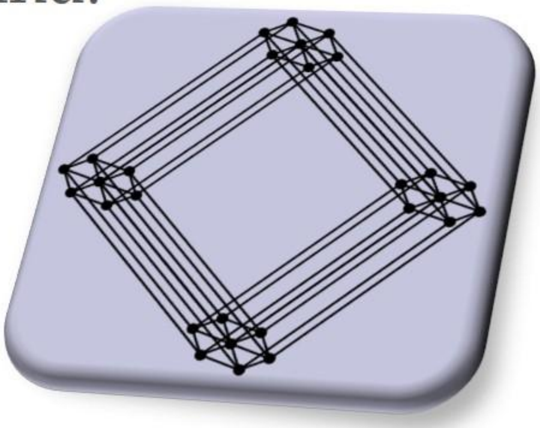
Let $\beta_c = \frac{1}{2}\log(1+\sqrt{2})$ be the critical inverse-temperature for the Ising model on $\mathbb{Z}^2$. Then the continuous-time Glauber dynamics for the Ising model on $(\mathbb{Z}/n\mathbb{Z})^2$ with periodic boundary conditions at $0 \leq \beta < \beta_c$ has cutoff at $(1/\lambda_\infty)\log n$ where λ_∞ is the spectral gap of the dynamics on the infinite volume lattice.

- Analogous result holds for *any* dimension $d \geq 1$:
  - Cutoff at $(d/2\lambda_\infty)\log n$.
  - *E.g.*, cutoff at $[2(1-\tanh(2\beta))]^{-1}\log n$ for $d=1$.

Toy example: cutoff at $\beta = 0$

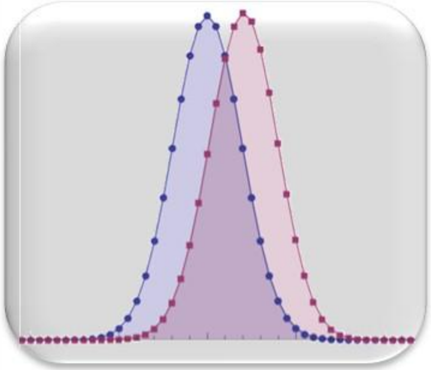
- Glauber dynamics for infinite temperature (β=0) Ising equivalent to cont.-time RW on the hypercube $\{-1,1\}^n$:
  - Stationary distribution is uniform.
  - Spins evolve independently.

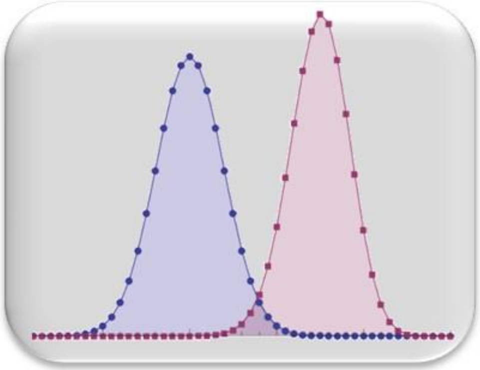
- [Aldous '83]: Cutoff at $\frac{1}{2}\log n + O(1)$.
  - Twice faster than trivial upper bound.
  - Constant window.

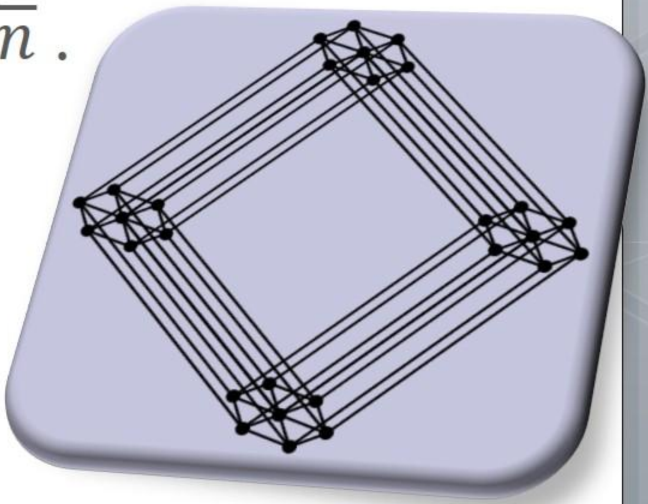


Toy example: cutoff at $\beta = 0$ (ctd.)

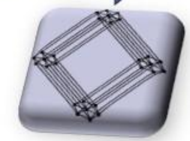
- Magnetization is a birth-and-death chain:
  - By symmetry start at the all-plus state.
  - # of +'s at time t is $\sim \mathrm{Bin}(n, \frac{1}{2}(1+\mathrm{e}^{-t}))$.
  - # of +'s under stationary measure $\sim \mathrm{Bin}(n, \frac{1}{2})$ which has Gaussian fluctuations of $\mathrm{O}(\sqrt{n})$.
  - Mixing occurs when $\frac{1}{2}\, \mathrm{e}^{-t} \asymp \sqrt{n}$.







Toy example: cutoff at $\beta = 0$ (ctd.)



- Symmetry $\Rightarrow$ Start at the all-plus state.
- Symmetry $\Rightarrow$ Mixing of *magnetization* $S_t = \sum_{i=1}^{n} X_t(i)$ [a birth & death chain] determines entire mixing:

$$\left\| \mathbb{P}_+(X_t \in \cdot) - \pi \right\|_{\mathrm{TV}} = \left\| \mathbb{P}_+(S_t \in \cdot) - \pi_S \right\|_{\mathrm{TV}} .$$

  - To bound the coupling-time of this 1d chain it thus suffices to couple it from its extreme ends $+$, $-$.
- Magnetizations contract to within $\sqrt{n}$ from each other:

$$\mathbb{E}_+[S_t] = ne^{-t} \ , \ \mathbb{E}_-[S_t] = -ne^{-t} .$$

  - At time $t = \frac{1}{2}\log n$ the expected distance between the chains is $O(\sqrt{n})$.
- Afterwards: distance is a biased RW drifting towards 0. Comparing to SRW $\Rightarrow$ $O(n)$ extra discrete steps to hit 0.

$\blacksquare$

General result on product chains

- PROPOSITION:

Let $X_t = (X_t^1, \dots, X_t^n)$ be a product chain where each X_t^i is ergodic with stationary measures π_i and $\pi = \prod_i \pi_i$. Let

$$\mathfrak{M}_t = \sum_{i=1}^{n} \left\| \mathbb{P}(X_t^i \in \cdot) - \pi_i \right\|_{L^2(\pi_i)}^2 .$$

For $\forall\, \delta > 0$ there $\exists\, \varepsilon > 0$ so that if for some $t > 0$

$$\max_i \left\| \mathbb{P}(X_t^i \in \cdot) - \pi_i \right\|_{L^\infty(\pi_i)} < \varepsilon$$

then

$$\left| \left\| \mathbb{P}(X_t \in \cdot) - \pi \right\|_{\mathrm{TV}} - \left(2\Phi\left(\tfrac{\sqrt{\mathfrak{M}_t}}{2} \right) - 1 \right) \right| < \delta .$$

Example: the hypercube (once more)

- COROLLARY: Aldous' hypercube result:

Let X_t be the lazy random walk on the hypercube $\{\pm 1\}^n$.
Then X_t exhibits cutoff at $\frac{1}{2}\log n$ and furthermore,
if $t = \frac{1}{2}\log n + c$ for some $c \in \mathbb{R}$ then

$$\left\| \mathbb{P}(X_t \in \cdot) - \pi \right\|_{\mathrm{TV}} = 2\Phi\left(\tfrac{1}{2}e^{-c}\right) - 1 + o(1).$$

- PROOF:

In the notation of the proposition, the (X_t^i)'s are i.i.d. with stationary measure uniform on $\{\pm 1\}$, thus

$$\left\| \mathbb{P}(X_t^i \in \cdot) - \pi_i \right\|_{L^2(\pi_i)}^2 = e^{-2t} \text{ and } \mathfrak{M}_t = ne^{-2t}. \qquad \blacksquare$$

Pf of product chain proposition

- Let $\nu = \mathbb{P}(X_t \in \cdot)$ and $\nu_i = \mathbb{P}(X_t^i \in \cdot)$ for $i = 1, \dots, n$.
- Let $U_1, \dots, U_n$ be independent r.v's $\sim \pi_1, \dots, \pi_n$ resp. and
$$Y_i = Y_i(t) = \nu_i(U_i)/\pi_i(U_i)$$
- By def. $\mathbb{E}Y_i = 1$ and:
$$\mathrm{Var}(Y_i) = \sum_x \left|\frac{\nu_i(x)}{\pi_i(x)} - 1\right|^2 \pi_i(x) = \left\|\nu_i - \pi_i\right\|_{L^2(\pi_i)}^2$$
and (✶) implies that $\|Y_i - 1\|_\infty < \varepsilon$ for all i .
- $\Rightarrow \mathbb{E}|Y_i - 1|^3 \le \|Y_i - 1\|_\infty \mathrm{Var}(Y_i) < \varepsilon\,\mathrm{Var}(Y_i)$.
- Set $Z_i = \log Y_i$ and look at a Taylor expansion:

$$\mathbb{E}Z_i = \mathbb{E}(Y_i - 1) - \tfrac{1}{2}\mathbb{E}(Y_i - 1)^2 + O\left(\mathbb{E}\left|Y_i - 1\right|^3\right) \quad \left(-\tfrac{1}{2} + O(\varepsilon)\right)\mathrm{Var}\,Y_i$$

$$\mathbb{E}Z_i^2 = \mathbb{E}(Y_i - 1)^2 + O\left(\mathbb{E}\left|Y_i - 1\right|^3\right) \quad (1 + O(\varepsilon))\mathrm{Var}\,Y_i$$

Pf of product chain (ctd.)

- Z_i's independent with $\|Z_i\|_\infty = O(\varepsilon)$ and
  $\mathbb{E} Z_i = -\frac{1+O(\varepsilon)}{2} \operatorname{Var} Y_i$, $\operatorname{Var} Z_i = (1+O(\varepsilon)) \operatorname{Var} Y_i$.
- By Berry-Esseen:
  $$\sup_x \left| \mathbb{P}\left(\frac{\sum_{i=1}^n (Z_i - \mathbb{E} Z_i)}{\sqrt{\sum_{i=1}^n \operatorname{Var}(Z_i)}} < x \right) - \Phi(x) \right| < \frac{O(\varepsilon)}{\sqrt{\mathfrak{M}_t}} .$$
  and so $\sum_{i=1}^n Z_i \to \mathcal{N}(-\frac{1}{2}\mathfrak{M}_t, \mathfrak{M}_t)$ in dist as $\varepsilon \to 0$.
- Relating this to $\|\nu - \pi\|_{\mathrm{TV}}$:
  $$\|\nu - \pi\|_{\mathrm{TV}} = \sum_{x_1,\ldots,x_n} \left| \prod_i \frac{\nu_i(x_i)}{\pi_i(x_i)} - 1 \right|^- \prod_i \pi_i(x_i)$$
  $$= \mathbb{E} \left| \prod_i Y_i - 1 \right|^- = \mathbb{E} \left| e^{\sum_i Z_i} - 1 \right|^- . \qquad \blacksquare$$

Products of i.i.d.'s

- COROLLARY:

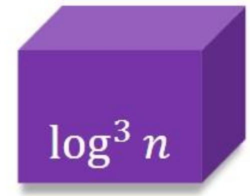
Let X_t be a product chain made of n i.i.d. copies of a finite ergodic chain Y_t with spectral-gap and log-Sobolev const gap and α_s resp. and stationary measure φ . If

$$\log \varphi_{\min}^{-1} \le n^{o(\alpha_s/\mathrm{gap})}$$

then X_t exhibits cutoff at $\frac{1}{2}\,\mathrm{gap}^{-1}\log n$ with window of order $O(\alpha_s^{-1}\log_+\log\varphi_{\min}^{-1})$.

Intuition: cutoff on the lattice

- Break up $\mathbb{Z}_n^d$ to cubes of side-length $\log^3 n$.
  Dynamics on such a cube:
  - $\alpha_s^{-1} = O(1)$
  - $\log \varphi_{\min}^{-1}(\sigma) = O(\log^{3d} n) = n^{o(1)}$



- Take non-adjacent cubes $Q_1, \dots, Q_m$ $(m \asymp (n/\log^3 n)^d)$ and ***suppose as if*** the projection on those would predict mixing for the entire system:
  - Distance between cubes turn them $\approx$ independent.
  - Expect cutoff at $\frac{1}{2\text{gap}} \log m = \frac{1}{2\text{gap}} \log n + O(\log\log n)$ with window $O(\log\log n)$.

Random support of update seq.