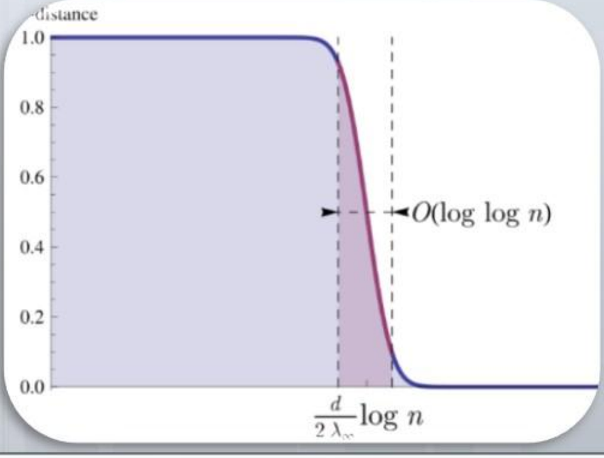



La Pietra 2011
Mini course
*lecture 1*

# Cutoff for Ising on the lattice

Eyal Lubetzky
Microsoft Research

# Course plan

- Lecture 1: crash course on static & stochastic Ising
- Lecture 2: cutoff and two angles on the hypercube
- Lecture 3: reducing $L^1$ to $L^2$ mixing.
- Lecture 4: breaking the dependencies: update supports
- Lecture 5: existence of cutoff and summary.

- Bibliography:

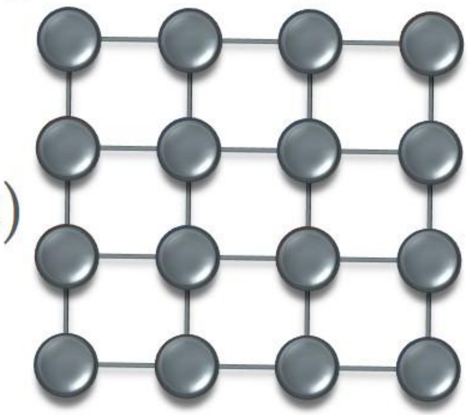
1. F. Martinelli, *Lectures on Glauber dynamics for discrete spin models* Lectures on probability theory and statistics, Saint-Flour, 1997
2. L. Saloff-Coste, *Lectures on finite Markov chains* Lectures on probability theory and statistics, Saint-Flour, 1996
3. D. Levin, Y. Peres & E. Wilmer, *Markov chains and mixing times* American Mathematical Society, 2008

# Definition: the classical Ising model

- Underlying geometry: $\Lambda$ = finite 2D grid.
- Set of possible configurations:

$$\Omega = \{\pm 1\}^{\Lambda}$$

(each *site* receives a plus/minus *spin*)



- Probability of a configuration $\sigma \in \Omega$ given by the *Gibbs distribution*:

$$\mu(\sigma) = \frac{1}{Z(\beta)} \exp\Big(\beta \sum_{x \sim y} \sigma(x)\sigma(y) + h \sum_{x} \sigma(x)\Big)$$

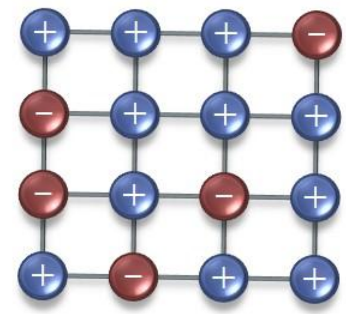
Partition function

Inverse temperature $\beta \geq 0$

External field

# The classical Ising model

- $\mu(\sigma) \propto \exp\Big(\beta \sum_{x \sim y} \sigma(x)\sigma(y)\Big)$ for $\sigma \in \Omega = \{\pm 1\}^\Lambda$
  - Larger $\beta$ favors configurations with aligned spins at neighboring sites.
  - Spin interactions ≈ local, justified by the rapid decay of magnetic force with distance.



- The *magnetization* is the (normalized) sum of spins:

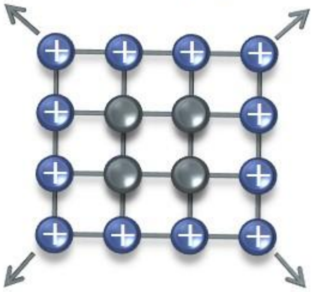
$$M(\sigma) = |\Lambda|^{-1} \sum_{x \in \Lambda} \sigma(x)$$

  - Distinguishes between disorder ($M \approx 0$) and order.

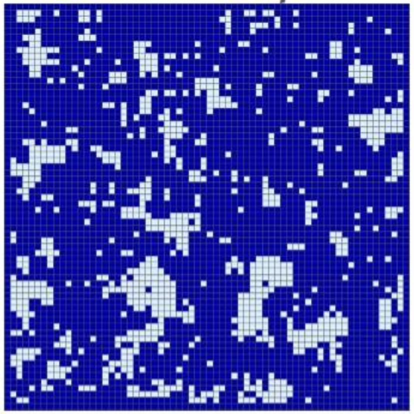
# The Ising phase–transition

- Ferromagnetism in this setting: [recall $M(\sigma) = \frac{1}{|\Lambda|}\sum \sigma(x)$]
  - Condition on the boundary sites all having *plus* spins.
  - Let the system size $|\Lambda|$ tend $\to \infty$

  ($\approx$ a magnetic field with effect $\to 0$).

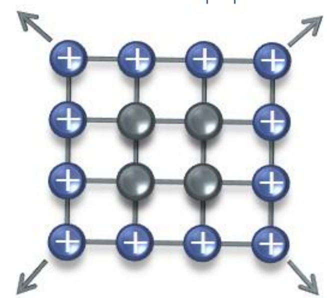


- What is the typical $M(\sigma)$ for large $|\Lambda|$ ?
  Does the effect of *plus* boundary vanish in the limit?



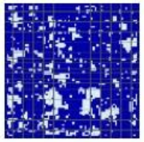
# The Ising phase-transition (ctd.)

- Ferromagnetism in this setting: [recall $M(\sigma) = \frac{1}{|\Lambda|}\sum \sigma(x)$]
  - Condition on the boundary sites all having *plus* spins.
  - Let the system size $|\Lambda|$ tend $\to \infty$



- Expect: phase-transition at some critical $\beta_c$ :

$$\lim_{|\Lambda|\to\infty} \mathbb{E}^{+}[M(\sigma)] = \begin{cases} 0 & \text{if } \beta < \beta_c \\ c_\beta > 0 & \text{if } \beta > \beta_c \end{cases}$$



*all-plus boundary*

*spontaneous magnetization*

# High temperatures

- Def. ([Martinelli & Olivieri ′94]) property $\mathbf{SM}(\Lambda, c, C)$ holds for a set $\Lambda \subset \mathbb{Z}^d$ and $c, C > 0$ iff $\forall\ \Delta \subset \Lambda$ :

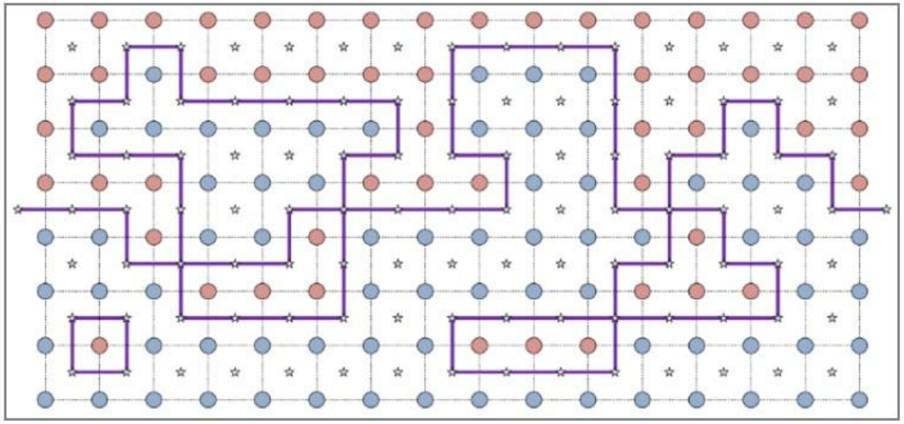
$$\sup_{\substack{\tau \in \{\pm 1\}^\Lambda \\ y \in \partial\Lambda}} \left\| \mu_\Lambda^\tau \Big|_\Delta - \mu_\Lambda^{\tau^y} \Big|_\Delta \right\|_{\mathrm{TV}} \leq C e^{-c\,\mathrm{dist}(y,\Delta)}$$

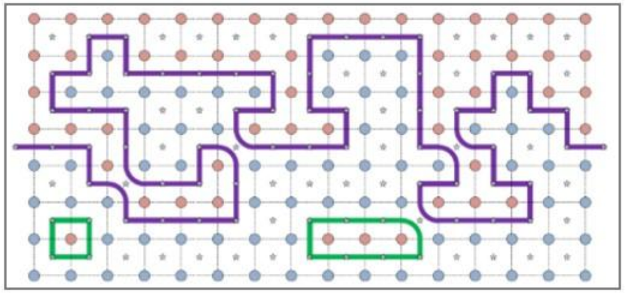
***Strong spatial mixing*** holds iff $\exists\ c, C, L > 0$ so that $\mathbf{SM}(Q, c, C)$ holds for all **cubes** $Q$ of side length $L$.

  - Here $\|\varphi - \nu\|_{\mathrm{TV}} = \sup_{A \subset \Omega} \left[\varphi(A) - \nu(A)\right]$.
  - On $\mathbb{Z}^2$ strong spatial mixing holds for all $\beta < \beta_c$ or whenever $h \neq 0$.
  - Implies $\mathbb{E}^+[M(\sigma)] \to 0$ (*no spontaneous mag*) Exercise

# Low temperatures

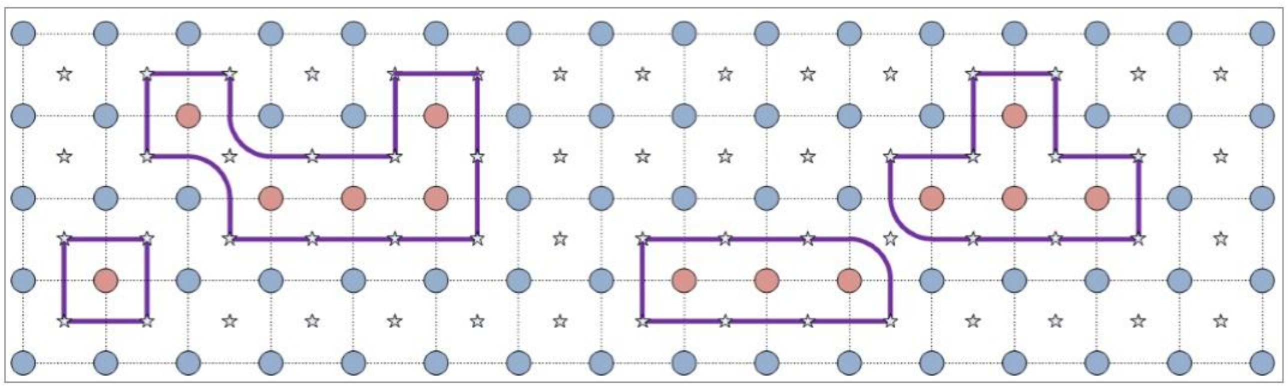
- Ingenius combinatorial argument due to [Peierls '36].
- Key idea: represent Ising configurations as *contours* in the *dual graph*: the edges are dual to disagreeing edges.





# Peierls' phase transition argument

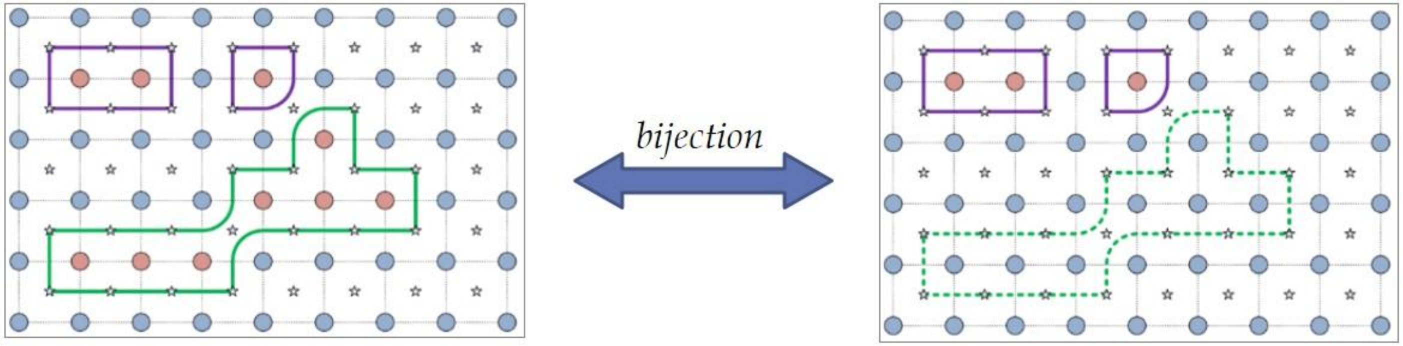
- When all boundary spins are + 's the Peierls contours are all closed [marking "islands" containing of − 's ].



- Goal: show that the fraction of sites inside such − components is bounded away from ½ .

# Peierls' phase transition argument

- Setting: $\Lambda \subset \mathbb{Z}^2$ is an $n \times n$ box with all-plus boundary.
- Fix a contour $C$ of length $\ell$.
- For each $\sigma$ containing $C$ flip all the spins of $C$ *and its interior* to arrive at a unique $\sigma'$ :




- Proof completed by a first moment argument. *Exercise*

# Glauber dynamics / Stochastic Ising

- Glauber dynamics for the Ising model (also known as the *Stochastic Ising model*) introduced in 1963 by Roy J. Glauber (Nobel in Physics 2005).
  - finite ergodic Markov chain on $\Omega = \{\pm 1\}^{\Lambda}$
  - moves between states by flipping a single site.
  - converges to the stationary Ising measure $\mu$.
- Intensively studied over the last 30 years:
  - Natural efficient sampler for the Ising model.
  - Captures its stochastic evolution.

R.J. Glauber

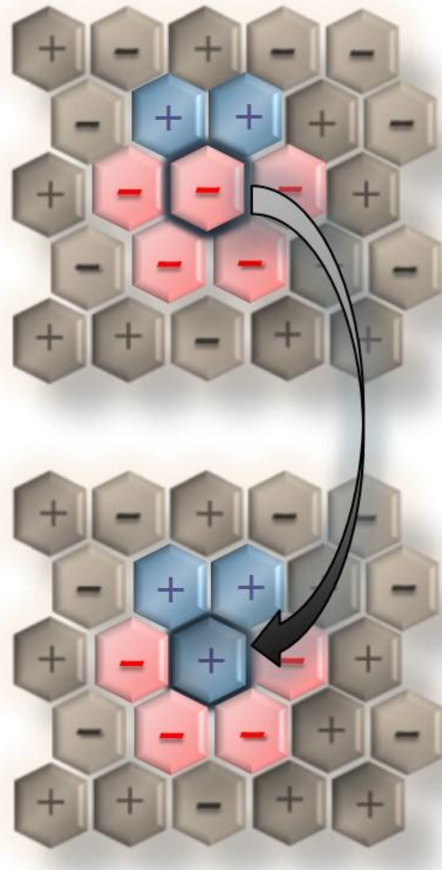
# Glauber dynamics for Ising

- One of the most commonly used MC samplers for the Ising distribution $\mu$ :

Heat-bath version given by the generator

$$\mathcal{L}_\Lambda^\tau(f)(\sigma) = \sum_{x\in\Lambda}[\mu_{\sigma,x}^\tau(f) - f(\sigma)]$$

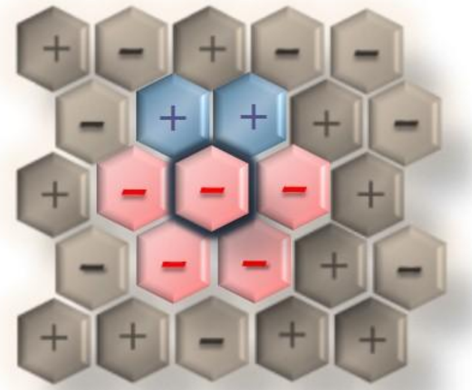
where $\mu_{\sigma,x}^\tau(f) = \mu_\Lambda^\tau(\cdot \mid \sigma_y, y \neq x)$



- Equivalent description:
  - Update sites via *iid* Poisson(1) clocks
  - Each update replaces a spin at $u \in V$ by a new spin $\sim \mu$ conditioned on all remaining spins at $V \setminus \{u\}$.

# Glauber dynamics for Ising

- One of the most commonly used MC samplers for the Ising distribution $\mu$ :
  - Update sites via *iid* Poisson(1) clocks
  - Each update replaces a spin at $u \in V$ by a new spin $\sim \mu$ conditioned on all remaining spins at $V \setminus \{u\}$.



- The above is the *heat-bath* version. Other versions of the dynamics include e.g. Metropolis.
- To sample from the Ising model, start at an arbitrary state (e.g. all-plus) run the chain.
  - How long does it take it to converge to $\mu$ ?

# Notions of convergence to equilibrium

- Spectral gap in the spectrum of the generator:

  $$\text{gap} \;=\; \text{smallest positive eigenvalue of } (-\mathcal{L}_t) \text{ associated with the heat-kernel } H_t\,.$$

- Mixing time : (according to a given metric).
  - Standard choice: $L^1$ (total-variation) mixing time to within $\varepsilon$ is defined as

    $$t_{\text{mix}}(\varepsilon) = \inf\left\{t : \max_{\sigma} \left\| H_t(\sigma,\cdot) - \mu \right\|_{\text{TV}} \le \varepsilon \right\}.$$

    where

    $$\left\| \mu - \nu \right\|_{\text{TV}} = \sup_{A\subset\Omega}\left[\mu\left(A\right) - \nu\left(A\right)\right] = \tfrac{1}{2}\sum_{x\in\Omega}\left|\mu(x) - \nu(x)\right|.$$

# The gap and mixing time

- Mixing time decays exponentially:

$$\max_{\sigma} \left\| H_t(\sigma,\cdot) - \mu \right\|_{\mathrm{TV}} \le e^{-\lfloor t/t_{\mathrm{mix}}(1/2e) \rfloor}$$

- Dirichlet form characterization for the spectral gap:

$$\mathrm{gap} = \inf_{\substack{f \in L^2(\mu_\Lambda^\tau) \\ f \not\equiv 1}} \frac{\mathcal{E}_\Lambda^\tau(f)}{\mathrm{Var}_\Lambda^\tau(f)}$$

where $\mathcal{E}_\Lambda^\tau(f)(\sigma) = \sum_{x \in \Lambda} \mu_\Lambda^\tau\left(\mathrm{Var}_{\sigma,x}^\tau(f)\right)$

- Relating the gap to total-variation mixing:

$$\mathrm{gap}^{-1} \le t_{\mathrm{mix}}(1/2e) \le \mathrm{gap}^{-1} \log(2e/\mu_{\min})$$

# Coupling

- Well-known tool to bound total-variation distance:

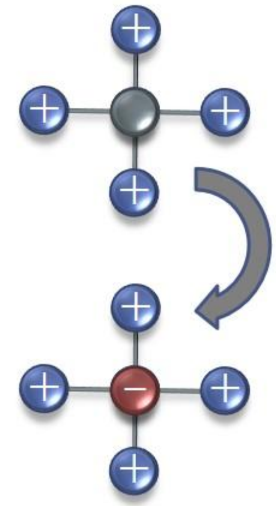
$$\|\mu - \nu\|_{\mathrm{TV}} \le \mathbb{P}(X \ne Y)$$

for $\forall$ coupling $(X,Y)$ with $X \sim \mu$ , $Y \sim \nu$ , and there $\exists$ a **maximal** coupling achieving equality.

- **Monotone coupling** for Glauber dynamics:
  - If $X_t \succcurlyeq Y_t$ then $X_{t+1} \succcurlyeq Y_{t+1}$
  - Consequently:

$$\max_x \|\mathbb{P}_x(X_t \in \cdot) - \mu\|_{\mathrm{TV}} \le \mathbb{P}_{+,-}(X_t \ne Y_t)$$

# Example: fast mixing at high temp



- When all $\Delta$ neighbors of a site are plus, probability of minus is

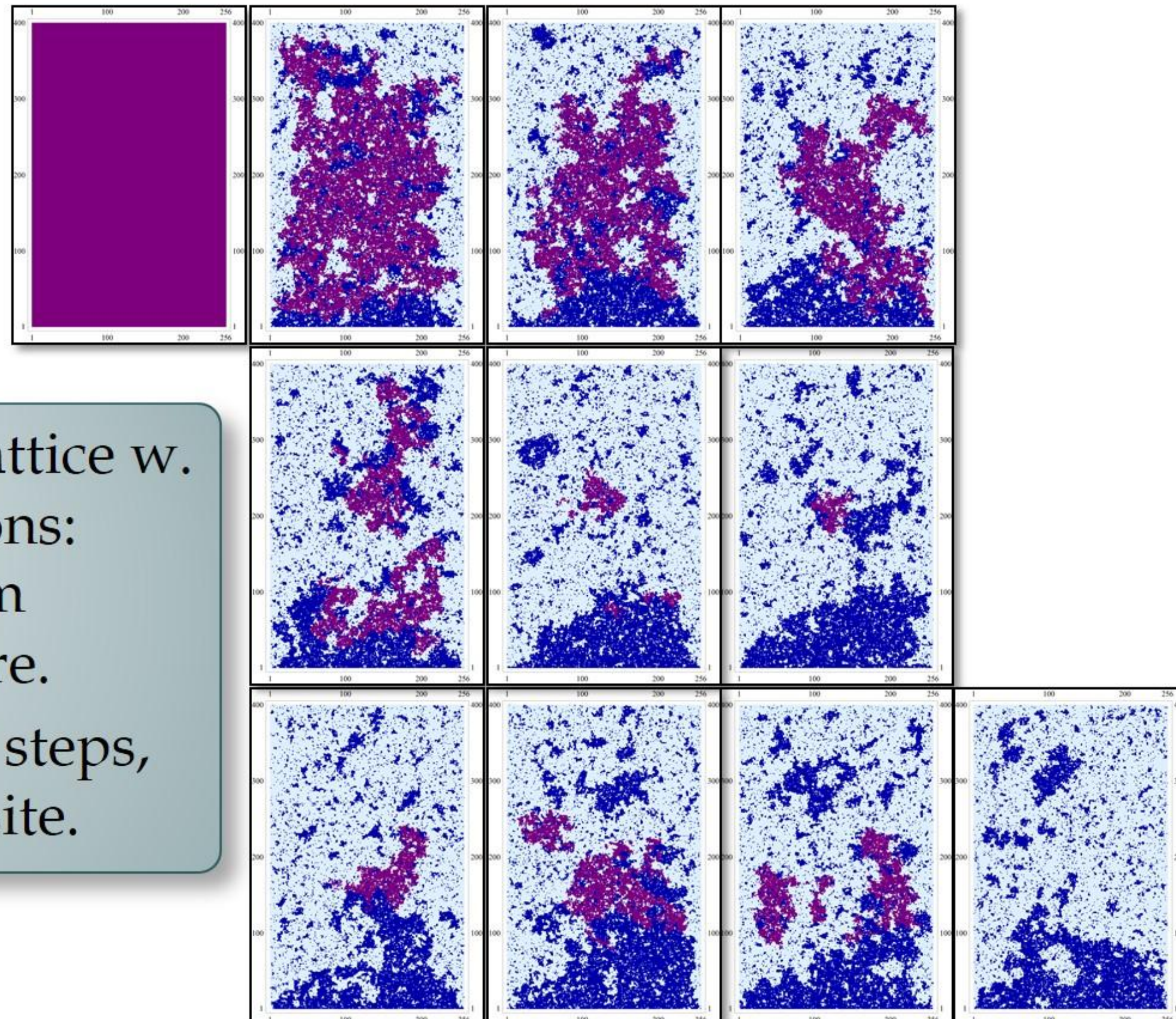
$$\tfrac{1}{2}(1-\tanh(\beta\Delta)) = \frac{e^{-\beta\Delta}}{e^{\beta\Delta}+e^{-\beta\Delta}} = \tfrac{1}{2}-\varepsilon$$

- Run monotone coupling: given $k$ disagreements prior to an update:
  - Eliminate one of them with probability $\geq (1-2\varepsilon)\dfrac{k}{n}$
  - Introduce a new one with probability $\leq 2\varepsilon\dfrac{\Delta k}{n}$
  - Contraction for small enough $\varepsilon$ $\Rightarrow\ t_{\mathrm{mix}} = O(\log\ n)$ in cont-time.

$$\Big(-1+2\varepsilon\big(\Delta+1\big)\Big)\frac{k}{n}$$

# Glauber dynamics for critical Ising

- *How fast does the dynamics converge?*



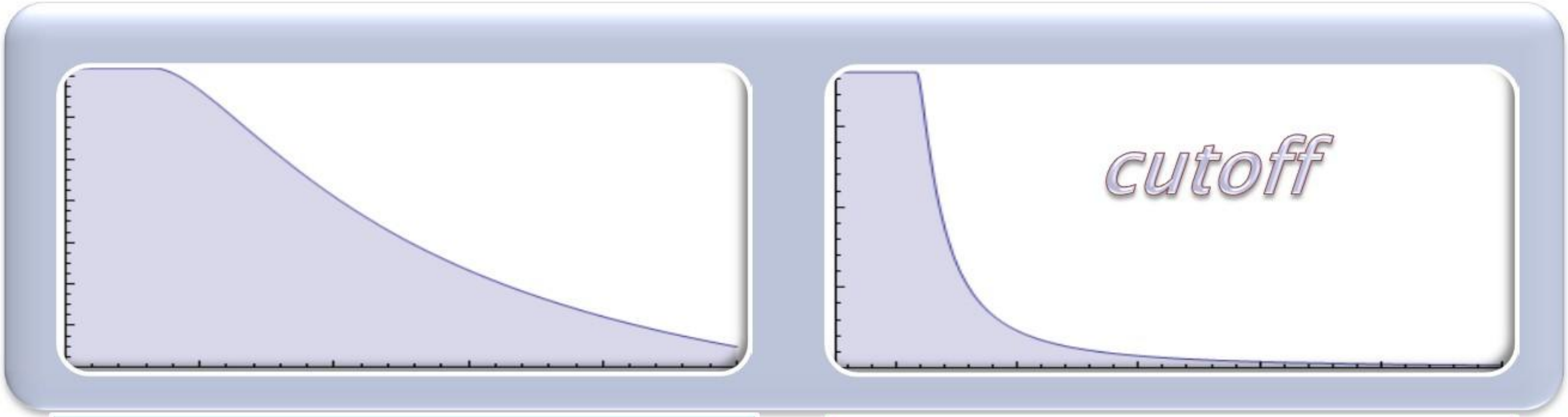
- 256 x 400 square lattice w. boundary conditions:
  - (+) at bottom
  - (−) elsewhere.
- Frame every $\sim 2^{30}$ steps, *i.e.* $\sim 2^{13}$ updates/site.

# 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 $\beta_c$ :
- Low temperature: ($\beta > \beta_c$)
  gap$^{-1}$ and $t_{\text{mix}}$ are ***exponential*** in the surface area.
- Critical temperature: ($\beta = \beta_c$)
  gap$^{-1}$ and $t_{\text{mix}}$ are ***polynomial*** in the surface area.
  - Exponent of gap$^{-1}$ is universal (the *dynamical critical exponent* $z$).
- 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*** .

# The Cutoff Phenomenon

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

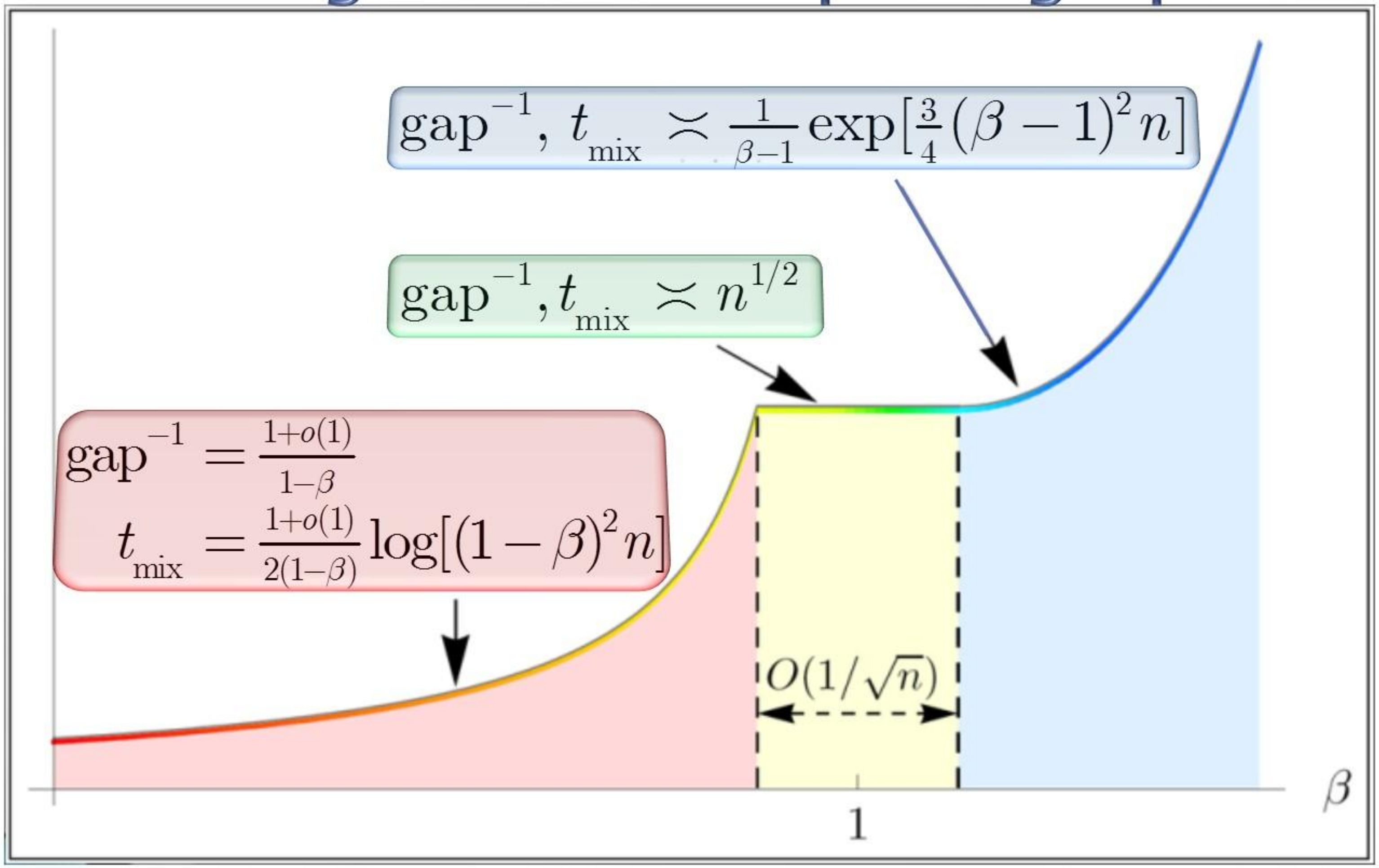



**Steady convergence**
*it takes a while to reach distance ½ from stationarity then a while longer to reach distance ¼, etc.*

**Abrupt convergence**
*distance from equilibrium quickly drops from 1 to 0*

# Example: mixing picture for Ising on the complete graph

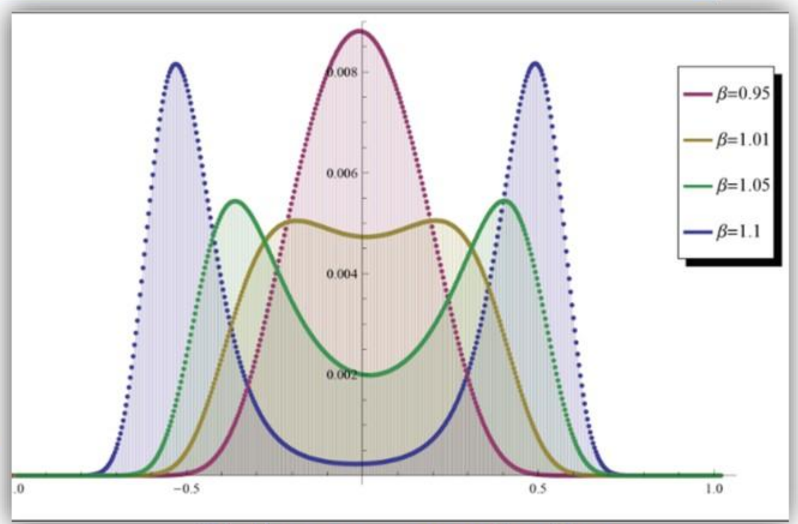
Curie–Weiss model




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

# Critical slowdown

- Intuition: low temperature
  - Exponential mixing due to a bottleneck between the "mostly-plus" and the "mostly-minus" states




Stationary magnetization

- Intuition: high temperature
  - At $\beta = 0$ there is complete independence.
  - For very small $\beta > 0$ a spin is likely to choose the same update given $2$ very different neighborhoods (weak "communication" between sites).
  - States can be coupled quickly, hence rapid mixing.
- Intuition: critical power-law:
  - Doubling the box incurs a constant factor in mixing...

# Mixing for Ising on the 2D lattice

- Fast mixing at high temperatures:
  - [Aizenman, Holley ’84]
  - [Dobrushin, Shlosman ’87]
  - [Holley, Stroock ’87, ’89]
  - [Holley ’91]
  - [Stroock, Zegarlinski ’92a, ’92b, ’92c]
  - [Zegarlinski ’90, ’92]
  - [Lu, Yau ’93]
  - [Martinelli, Olivieri ’94a, ‘94b]
  - [Martinelli, Olivieri, Schonmann ’94]
- Slow mixing at low temperatures:
  - [Schonmann ’87]
  - [Chayes, Chayes, Schonmann ’87]
  - [Martinelli ’94]
  - [Cesi, Guadagni, Martinelli, Schonmann ’96].

# The gap and log-Sobolev const

- Recall: $\mathrm{gap} = \inf_{\substack{f \in L^2(\mu_\Lambda^\tau) \\ f \not\equiv 1}} \frac{\mathcal{E}_\Lambda^\tau(f)}{\mathrm{Var}_\Lambda^\tau(f)}$ where

$$\mathcal{E}_\Lambda^\tau(f)(\sigma) = \sum_{x\in\Lambda} \mu_\Lambda^\tau\big(\mathrm{Var}_{\sigma,x}^\tau(f)\big).$$

- The ***log-Sobolev constant*** is given by

$$\alpha_{\mathrm{s}} = \inf_{\substack{f \in L^2(\mu_\Lambda^\tau) \\ f \not\equiv 1}} \frac{\mathcal{E}_\Lambda^\tau(f)}{\mathrm{Ent}_\Lambda^\tau(f)}$$

where $\mathrm{Ent}_\Lambda^\tau(f) = \mathbb{E}_{\mu_\Lambda^\tau}\left[f^2(\sigma)\log\left(f^2(\sigma) \,/\, \mathbb{E}_{\mu_\Lambda^\tau} f^2(\sigma)\right)\right].$

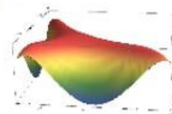
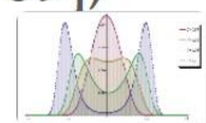
- Relating the gap to $L^2$ mixing:

$$\left\|\mathbb{P}_x(X_s \in \cdot) - \nu\right\|_{L^2(\nu)} \le \exp\left[1 - \mathrm{gap}\left(s - \frac{1}{4\alpha_{\mathrm{s}}}\log_+ \log\frac{1}{\nu(x)}\right)\right]$$

# Mixing on the square lattice

- High temperature regime: if there is strong spatial mixing (on $\mathbb{Z}^2$ this covers $\forall\, \beta < \beta_c$) :
  - $O(1)$ inverse spectral-gap constant.
  - $O(1)$ inverse log-Sobolev constant and as a result $O(\log\ n)$ total-variation mixing.
  - ? Dynamics conjectured to exhibit *cutoff* [Peres'04].

# Cutoff for the Glauber dynamics

- Till recently: *only* spin-systems where cutoff was verified are Ising and Potts models on the *complete graph* [Levin, Luczak, Peres '10], [Ding, L., Peres '09], [Cuff, Ding, L., Louidor, Peres, Sly]  
- Conjectured to believe at high temperatures for:
  - ? Ising on the lattice, e.g. with periodic or free boundary.
  - ? Potts model on the lattice.
  - ? Gas Hard-core model on lattices.
  - ? Colorings of lattices.
  - ? Arbitrary boundary conditions / external field.
  - ? Anti-ferromagnetic Ising/Potts models, Spin-glass, Other lattices / amenable transitive graphs,…

Unknown even in 1 dimension (Q. of Peres)…

# Cutoff for Ising on lattices

- <u>THEOREM</u> [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 $\lambda_\infty$ 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 $d=1$ .

# Cutoff for Ising on the lattice

- Main result hinges on an $L^1$-$L^2$ reduction, enabling the application of log-Sobolev inequalities.
- Generic method gives further results on many other models conjectured to have cutoff:
  - ✔ Ising on the lattice, e.g. with periodic or free boundary.
  - ✔ Potts model on the lattice.
  - ✔ Gas Hard-core model on lattices.
  - ✔ Colorings of lattices.
  - ✔ Arbitrary boundary conditions / external field.
  - ✔ Anti-ferromagnetic Ising/Potts models, Spin-glass, Other lattices / amenable transitive graphs,…

# Key tool: breaking dependencies…