Information Percolation for the Ising model

Eyal Lubetzky
Courant Institute (NYU)
Joint with Allan Sly
Noisy Election Day (on a cycle)

- Setup: (1D noisy voter model with noise $0 < \varepsilon < 1$)
  - $n$ binary voters on a cycle.
  - Every step, a uniformly chosen voter updates its vote:
    - $\text{prob. } 1 - \varepsilon$: copy a random neighbor.
    - $\text{prob. } \varepsilon$: new vote is a fair coin toss.

- How long does it take to reach equilibrium?
  - from all-1? from 01010...? from 001100...?
  - from a typical state? from a random IID state?
Definition: the classical Ising model

- Underlying geometry: $\Lambda = \text{finite 2D grid}$.
- 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:

  $$\mu(\sigma) = \frac{1}{Z(\beta)} \exp\left(\beta \sum_{x \sim y} \sigma(x)\sigma(y)\right)$$

- Partition function
- Inverse temperature $\beta \geq 0$
The classical Ising model

\[ \mu(\sigma) \propto \exp(\beta \sum_{x \sim y} \sigma(x)\sigma(y)) \text{ for } \sigma \in \Omega = \{\pm 1\}^\Lambda \]

- Larger \( \beta \) favors configurations with aligned spins at neighboring sites.
- Spin interactions: local, justified by rapid decay of magnetic force with distance.

- The magnetization is the (normalized) sum of spins:
  \[ M(\sigma) = \frac{1}{|\Lambda|} \sum_{x \in \Lambda} \sigma(x) \]
  - Distinguishes between disorder \( (M \approx 0) \) and order.
- Symmetry: \( \mathbb{E}[M(\sigma)] = 0 \). What if we break the symmetry?
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: \[ M(\sigma) = \frac{1}{|\Lambda|} \sum \sigma(x) \]
  - Condition on the boundary sites all having \textit{plus} spins.
  - Let the system size $|\Lambda|$ tend to $\infty$

- Phase-transition at some critical $\beta_c$:
  \[
  \lim_{|\Lambda| \to \infty} \mathbb{E}^+ [M(\sigma)] = \begin{cases} 
  0 & \text{if } \beta < \beta_c \\
  m_\beta > 0 & \text{if } \beta > \beta_c
  \end{cases}
  \]

  - all-plus boundary
  - spontaneous magnetization
Static vs. stochastic Ising

- Expected behavior for the Ising distribution:
  - $\beta < \beta_c$: $E^+[M(\sigma)] \xrightarrow{|\Lambda| \to \infty} 0$
  - $\beta > \beta_c$: $E^+[M(\sigma)] \xrightarrow{|\Lambda| \to \infty} c_\beta > 0$

- Expected behavior for the mixing time of dynamics:
  - $\beta < \beta_c$: logarithmic
  - $\beta > \beta_c$: exponential
  - $\beta = \beta_c$: power law

Free b.c.
Glauber dynamics for Ising

(a.k.a. the Stochastic Ising model)

- Introduced in 1963 by Roy Glauber. (heat-bath version; famous other flavor: Metropolis)

- Time-dependent statistics of the Ising model


  Cited by 2749

- One of the most commonly used samplers for the Ising distribution $\mu$:
  - Update sites via IID Poisson(1) clocks
  - Each update replaces a spin at $x \in V$ by a new spin $\sim \mu$ given spins at $V \setminus \{x\}$.

- How long does it take it to converge to $\mu$?
Measuring convergence to equilibrium

- **Mixing time**: (according to a given metric).
  Standard choice: $L^1$ (total-variation) mixing time to within distance $\varepsilon$ is defined as
  \[
  t_{\text{mix}}(\varepsilon) = \inf\left\{ t : \max_{x_0} \left\| p^t(x_0, \cdot) - \mu \right\|_{\text{tv}} \leq \varepsilon \right\}
  \]
  (where $\|\mu - \nu\|_{\text{tv}} = \sup_{A \subset \Omega} [\mu(A) - \nu(A)]$)

- **Dependence on $\varepsilon$**: (cutoff phenomenon [DS81], [A83],[AD86])
  We say there is **cutoff** $\iff t_{\text{mix}}(\varepsilon) \sim t_{\text{mix}}(\varepsilon') \quad \forall$ fixed $\varepsilon, \varepsilon'$
Believed picture for Ising on $\mathbb{Z}_n^d$

- For some critical inverse-temperature $\beta_c$:
  - $\beta < \beta_c$
  - $\beta = \beta_c$
  - $\beta > \beta_c$ (free b.c.)

- $t_{\text{mix}}$:
  - $\sim c_\beta \log n$
  - $\approx n^z$
  - $\approx e^\tau_\beta n^{d-1}$

- Analogous picture verified for:
  - Complete graph [Ding, L., Peres ‘09a, ‘09b], [Levin, Luczak, Peres ‘10]:
    - $\frac{1}{2(1-\beta)} \log n + O(1)$
    - $\approx \sqrt{n}$
    - $\approx \frac{1}{\beta-1} \exp\left[\frac{3}{4}(\beta-1)^2n\right]$
  - Regular tree [Berger, Kenyon, Mossel, Peres ‘05] (high $T$/low $T$)
    - [Ding, L., Peres ‘10] (critical $T$)
  - Potts model on complete graph
    - [Cuff, Ding, L., Louidor, Peres, Sly ‘12]
Glauber dynamics for 2D Ising

Fast mixing at **high** temperatures:
- [Aizenman, Holley ’84]
- [Dobrushin, Shlosman ’87]
- [Holley, Stroock ’87, ’89]
- [Holley ’91]
- [Stroock, Zegarlinski ’92a, ’92b, ’92c]
- [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]

Critical power-law:
- simulations: [Ito ’93], [Wang, Hatano, Suzuki ’95], [Grassberger ’95], …: \( n^{2.17} \), …
- lower bound: [Aizenman, Holley ’84], [Holley ’91]
- upper bound (polynomial mixing): [L., Sly ’12]

\( \beta < \beta_c \)
\[ t_{mix} \approx \log n \]

\( \beta > \beta_c \)
\[ t_{mix} = e^{(\tau_\beta + o(1))n^{d-1}} \]

\( \beta_c \)
\[ n_c^1 \leq t_{mix} \leq n_c^2 \]
Glauber dynamics for 2D Ising

- **$\beta < \beta_c$**
  - $t_{\text{mix}} : O(\log n)$
  - $n^c$ (sim: $n^{2.17...}$)

- **$\beta > \beta_c$**
  - Free b.c.
  - $e^{(\tau_\beta + o(1))n^{d-1}}$

- **High temperature in 2D:**
  - [L., Sly ’13]: **cutoff**
  - For any $\beta < \beta_c = \frac{1}{2} \log(1 + \sqrt{2})$:
    - $t_{\text{mix}}(\varepsilon) = \frac{1}{2} \lambda^{-1} \log n + O(\log \log n)$

- Method caveat: needs **strong spatial mixing**; e.g., breaks on 3D Ising for $\beta$ close to $\beta_c$. 
High temperature unknowns (I)

- High temperature $\leftrightarrow$ Infinite temperature:
  Qualitatively, $\beta < \beta_c$ believed to behave $\approx$ as $\beta = 0$.

\[ t_{\text{mix}}(\epsilon) = c \log n + O(1) \]

- $\beta = 0$: (independent spins) one of the first examples of cutoff:
  [Aldous ’83], [Diaconis Shahshahani ’87]
  [Diaconis, Graham, Morrisson ’90]

  $\Rightarrow$ expect cutoff $\forall \beta < \beta_c$ (conj. [Peres ’04]) & with $O(1)$-window

- Concretely: for 3D Ising (e.g. on a torus) at $\beta = 0.99 \beta_c$:
  does the dynamics exhibit cutoff? if so, where & what is the window?
High temperature unknowns (II)

Warm (random) start vs. cold (ordered) start: random start is better than ordered

- e.g.

Concretely: for 3D Ising at $\beta = 0.01$:

$?\quad$what is $t_{\text{mix}}^U(\epsilon) = \inf\left\{ t : \left\| \frac{1}{|\Omega|} \sum_{x_0} p^t(x_0, \cdot) - \mu \right\|_{\text{tv}} \leq \epsilon \right\}$?

how does it compare with $t_{\text{mix}}(\epsilon)$?
High temperature unknowns (III)

- **Universality of cutoff:**
  on any locally finite geometry there should be cutoff if the temperature is high enough (function of max-degree)

  - $\exists c_0 > 0$: The Ising model on any graph $G$ on $n$ vertices with maximal degree $d$ at $\beta < c_0/d$ has $t_{\text{mix}} = O(\log n)$
    - [Dobrushin ’71], [Holley ’72], [Dobrushin-Shlosman ’85], [Aizenman-Holley ’87]

  - $\Rightarrow$ expect cutoff $\forall \beta < \kappa/d$, and with $O(1)$-window.

- Concretely: for Ising on a binary tree at $\beta = 0.01$:
  - does the dynamics exhibit cutoff? 
  - if so, where & what is the window?
Recipe for stochastic Ising analysis

Traditional approach to sharp mixing results
1. Establish spatial properties of static Ising measure
2. Use to drive a multi-scale analysis of dynamics.

Example: best-known results on 2D Ising (torus \( \mathbb{Z}_n^2 \)):
- [L., Sly ‘13]: cutoff at \( \beta < \beta_c = \frac{1}{2} \log(1 + \sqrt{2}) \)
  - used log-Sobolev ineq. & strong spatial mixing.
- [L., Sly ‘12]: power-law at \( \beta_c \)
  - used SLE behavior of critical interfaces.
- [L., Martinelli, Sly, Toninelli ‘13]: at \( \beta > \beta_c \)
  - quasi-polynomial mixing under all-plus b.c.
  - uses interface convergence to Brownian bridges
New framework for the analysis

- Traditional approach to sharp mixing results
  1. Establish spatial properties of static Ising measure
  2. Use to drive a multi-scale analysis of dynamics.

- New approach: study these simultaneously examining information percolation clusters in the space-time slab:
  - track update lineage back in time.
  - update either (a) branches out, or (b) terminates ("oblivious")
  - analyze RED/GREEN/BLUE clusters...

Eyal Lubetzky, Courant Institute
Results: cutoff up to $\beta_c$ in 3D Ising

- Confirm Peres’s conj. on $\mathbb{Z}_n^d$ for any $d$, with $O(1)$-window.
- **Theorem:** ([L.-Sly ’14+])

  $\forall d \geq 1$ and $\beta < \beta_c$ there is **cutoff** with an $O(1)$-window at

  $t_m = \inf \left\{ t : \mathbb{E}_+ \left[ M(\sigma_t) \right] \leq \sqrt{n^d} \right\}$

- Examples:
  - $d = 1$: $t_m = \frac{1}{2(1-\tanh(2\beta))} \log n$.
  - $\beta = 0$: $t_m = \frac{1}{2} \log n$ (matching [Aldous ’83])

[recall $M(\sigma) = \frac{1}{|\Lambda|} \sum \sigma(x)$]
Results: initial states

- Warm start is twice faster:
  - All-plus starting state is worst (up to an additive $O(1)$)
    [but twice faster than naïve monotone coupling bound].
  - Uniform initial state $\approx$ twice faster than all-plus.
  - Almost $\forall$ deterministic initial state $\approx$ as bad as all-plus.

- Example: the 1D Ising model ($\mathbb{Z}_n$):
  **Theorem:** ([L.-Sly ’14+])

$$\text{Fix } \beta > 0 \text{ and } 0 < \varepsilon < 1 \text{; set } t_m = \frac{1}{2(1 - \tanh(2\beta))} \log n.$$  

1. (Annealed) $t_{\text{mix}}^{(U)}(\varepsilon) \sim \frac{1}{2} t_m$
2. (Quenched) $t_{\text{mix}}^{(x_0)}(\varepsilon) \sim t_{\text{mix}}^{(+)}(\varepsilon) \sim t_m$ for almost $\forall x_0$
Results: universality of cutoff

- **Paradigm:** cutoff for *any locally finite geometry* at high enough temperature (including expanders, trees, ...)

- **Theorem:** ([L.-Sly ’14+])
  \[ \exists \kappa > 0 \text{ so that, if } G \text{ is any } n\text{-vertex graph with degrees } \leq d \text{ and } \beta < \kappa/d, \text{ then } \exists \text{ cutoff with an } O(1)\text{-window at} \]
  \[ t_m = \inf \left\{ t : \sum_x \mathbb{E}_+ \left[ M(\sigma_t(x))^2 \right] \leq 1 \right\} . \]

- Moreover:
  \[ t_{mix}^{(U)} \leq \left( \frac{1}{2} + \epsilon_\beta \right) t_m \text{ yet } t_{mix}^{(x_0)} \geq \left( 1 - \epsilon_\beta \right) t_m \text{ a.e. } x_0. \]
The new framework (revisited)

- Information percolation clusters in the space-time slab:
  - track update lineage back in time.
  - update either (a) branches out, or (b) terminates ("oblivious")
Information percolation clusters

**Blue:**
dies out quickly in space & time.

**Red:**
top spins are affected by initial state.

**Green:**
o/w.

- Rough idea: condition on Green, let the effect of Red clusters vanish among Blue (show $\mathbb{E} \left[ 2^{R \cap R'} | G \right] \to 1$).
Example: the framework in 1D

- In 1D: $\theta = \mathbb{P}(\text{oblivious update}) = 1 - \tanh 2\beta$
- Update history: continuous-time RW killed at rate $\theta$.
- $\mathbb{P}(\text{surviving to time } t_m) \approx 1/\sqrt{n}$.
- Cutoff at $t_m = \frac{1}{2\theta} \log n$
- Effect of the initial state on the final state is in terms of the bias of the cont.-time RW...

the 3 cluster classes (R/G/B) in $\mathbb{Z}_{256}$
Example: random initial state

- Handling a uniform (IID) starting configuration:
  - Compare the dynamics directly with Ising measure: develop history to time $-\infty$ (coupling from the past).
  - Redefine **RED** clusters (coalesce before time 0).
Losing red clusters in a blue sea

**Lemma**: ([Miller, Peres ’12])

Let $\mu$ be a measure on $\sigma \in \Omega = \{\pm 1\}^n$ as follows:

1. draw a random variable $R \subseteq [n]$ via a law $\tilde{\mu}$;
2. let $\sigma_R \sim \text{law } \phi_R$ and $\sigma_{[n]\setminus R} \sim \text{IID Bernoulli} \begin{pmatrix} +1 & 1/2 \\ -1 & 1/2 \end{pmatrix}$

$$\Rightarrow \|\mu - \nu\|_{L^2(\nu)}^2 \leq \mathbb{E}\left[2|R \cap R'|\right] - 1$$

- the set $R$ embodies the nontrivial part of $\mu$
- has a negligible effect provided the exponential moment can be controlled...
Open problems

- High temperature regime for other spin-systems (Potts / Independent sets / Legal colorings / Spin glass,...):
  - asymptotic mixing on the lattice up to $\beta_c$
  - cutoff on a transitive expander
  - asymptotic mixing from random starting states (e.g., compare ordered/disordered start in Potts)

- 3D Ising:
  - no cutoff at criticality
  - power-law behavior at criticality
  - sub-exponential upper bound at low temperatures under all-plus b.c.
Thank you