Information Percolation for the Ising model

Eyal Lubetzky
Microsoft Research
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: \[ M(\sigma) = \frac{1}{|\Lambda|} \sum \sigma(x) \]
  - 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:**
  - $\beta < \beta_c$: $\mathbb{E}^+ [M(\sigma)]_{|\Lambda| \to \infty} \to 0$
  - $\beta > \beta_c$: $\mathbb{E}^+ [M(\sigma)]_{|\Lambda| \to \infty} \to c_\beta > 0$

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

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

$$
t_{\text{mix}}(\varepsilon) = \inf \left\{ t : \max_{x_0} \| p^t(x_0, \cdot) - \mu \|_{\text{tv}} \leq \varepsilon \right\}
$$

(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** $\iff t_{\text{mix}}(\varepsilon) \sim t_{\text{mix}}(\varepsilon')$ for all fixed $\varepsilon, \varepsilon'$.
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]
Glauber dynamics for 2D Ising

\[ t_{\text{mix}} : \begin{cases} 
O(\log n) & \beta < \beta_c \\
\sim n^{2.17} & n^c \\
e^{(\tau \beta + o(1))n^{d-1}} & \beta > \beta_c 
\end{cases} \]

High temperature in 2D:

- [L., Sly ‘13]: cutoff
  for any \( \beta < \beta_c = \frac{1}{2} \log(1 + \sqrt{2}) \):

\[ t_{\text{mix}}(\varepsilon) = \frac{1}{2} \lambda_\infty^{-1} \log n + O(\log \log n) \]

- 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 $\beta = 0$.

- $\beta = 0$: (independent spins) one of the first examples of cutoff:
  \[ t_{\text{mix}}(\varepsilon) = c \log n + O(1) \]
  [Aldous ’83], [Diaconis Shahshahani ‘87]
  [Diaconis, Graham, Morrow ‘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 $\beta = 0.99 \beta_c$:
  - $\exists$ 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 $\beta = 0.01$:

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

  - how does it compare with $t_{\text{mix}}(\varepsilon)$?
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-Holey '87]

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

- Concretely: for Ising on a binary tree at $\beta = 0.01$:
  - $\Rightarrow$ 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...
Results: cutoff up to $\beta_c$ in 3D Ising

- Confirm Peres’s conj. on $\mathbb{Z}^d_n$ for any $d$, with $O(1)$-window.
- **THEOREM:** ([L.-Sly ’14+])
  \[
  \forall d \geq 1 \text{ and } \beta < \beta_c \text{ there is cutoff with an } O(1)\text{-window at } t_m = \inf \left\{ t : \mathbb{E}_+[M(\sigma_t)] \leq \sqrt{n^d} \right\}
  \]

  **cutoff window:** $O(\log(1/\epsilon))$

- 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+])

  Fix $\beta > 0$ and $0 < \varepsilon < 1$; 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+])

  $\exists \kappa > 0$ so that, if $G$ is any $n$-vertex graph with degrees $\leq d$ and $\beta < \kappa/d$, then $\exists$ cutoff with an $O(1)$-window at

  $$t_m = \inf\left\{t : \sum_x \mathbb{E}_+ \left[M(\sigma_t(x))^2\right] \leq 1 \right\}.$$  

- Moreover:

  $t_{\text{mix}}^{(U)} \leq \left(\frac{1}{2} + \varepsilon_\beta\right)t_m$ yet $t_{\text{mix}}^{(x_0)} \geq (1 - \varepsilon_\beta)t_m \text{ a.e. } x_0.$
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")

$\mathbb{Z}_2^{200}$ cluster (top/side view)
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 $\mathbb{E} [2^{R_n R'} | G] \to 1$).
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 $\sigma_R \sim$ some law $\phi_R$ and $\sigma_{R^c} \sim$ IID Bernoulli $\frac{1}{2}$

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

- the set $R$ embodies the nontrivial part of $\mu$
- it has a negligible effect on 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