Coupling Technique

Eric Vigoda

Georgia Tech

EPFL Bernoulli Center, July '18

Talk Outline

1. Definitions: Ising Model, Coupling, etc.
2. Path Coupling:
   Rapid Mixing on $\mathbb{Z}^2$ for sufficiently high temperature.
3. Strong Spatial Mixing (SSM) $\rightarrow$ $O(n \log^2 n)$ Mixing:
   On box of $\mathbb{Z}^2$, $O(n \log^2 n)$ mixing for all $\beta < \beta_c(\mathbb{Z}^2)$.
4. Mossel-Sly for General Graphs:
   Any G with max degree d, $O(n \log n)$ mixing for $\beta < \beta_c(\mathbb{T}_d)$.

Markov Chain Basics

Markov Chain:

Finite state space Ω where $N = |\Omega|$.

Transition matrix P is an $N \times N$ stochastic matrix.

$$\textit{Ergodic} = \exists t^*, \forall x, y \in \Omega,\ P^{t^*}(x,y) > 0.$$

$$\text{Ergodic} \iff \text{aperiodic and irreducible}$$

$$\textit{Aperiodic} = \forall x, y \in \Omega, \exists t = t(x,y),\ P^t(x,y) > 0.$$

$$\textit{Irreducible} = \forall z \in \Omega,\ \gcd\{t : P^t(z,z) > 0\} = 1.$$

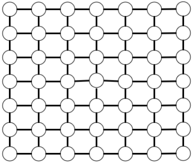
π is a stationary distribution if $\pi P = \pi$.

Key Theorem [Doeblin '38]: For a finite, ergodic Markov chain, there is a unique stationary distribution π such that:

$$\text{for all } x \in \Omega,\ P^t(x,\cdot) = \pi.$$

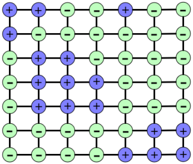
Running Example: Ising Model

Consider graph $G = (V, E)$ as $n \times n$ box of $\mathbb{Z}^2$:



Running Example: Ising Model

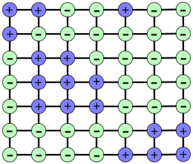
Consider graph $G = (V, E)$ as $n \times n$ box of $\mathbb{Z}^2$:



Configurations: $\Omega = \{-1, +1\}^V$.

Running Example: Ising Model

Consider graph $G = (V, E)$ as $n \times n$ box of $\mathbb{Z}^2$:



Configurations: $\Omega = \{-1, +1\}^V$.

Inverse temperature $\beta > 0$. For $\sigma \in \Omega$:

Monochromatic edges: $M(\sigma) = |\{(v, w) \in E : \sigma(v) = \sigma(w)\}|$

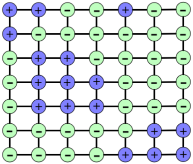
Gibbs dist.: $\mu(\sigma) = \frac{\exp(\beta M(\sigma))}{Z}$

Partition function: $Z = Z_G = \sum_{\eta \in \Omega} \exp(\beta M(\sigma))$.

Ferromagnetic (attractive) since $\beta > 0 \rightarrow$ neighbors prefer to align.

Running Example: Ising Model

Consider graph $G = (V, E)$ as $n \times n$ box of $\mathbb{Z}^2$:



Configurations: $\Omega = \{-1, +1\}^V$.

Inverse temperature $\beta > 0$. For $\sigma \in \Omega$:

Monochromatic edges: $M(\sigma) = |\{(v, w) \in E : \sigma(v) = \sigma(w)\}|$

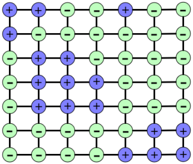
Sampling: Gibbs dist.: $\mu(\sigma) = \frac{\exp(\beta M(\sigma))}{Z}$

Counting: Partition function: $Z = Z_G = \sum_{\eta \in \Omega} \exp(\beta M(\sigma))$.

Ferromagnetic (attractive) since $\beta > 0 \rightarrow$ neighbors prefer to align.

Running Example: Ising Model

Consider graph $G = (V, E)$ as $n \times n$ box of $\mathbb{Z}^2$:



Configurations: $\Omega = \{-1, +1\}^V$. Potts: $\Omega = \{1, 2, \ldots, q\}^V$.

Inverse temperature $\beta > 0$. For $\sigma \in \Omega$:

Monochromatic edges: $M(\sigma) = |\{(v, w) \in E : \sigma(v) = \sigma(w)\}|$

Sampling: Gibbs dist.: $\mu(\sigma) = \dfrac{\exp(\beta M(\sigma))}{Z}$

Counting: Partition function: $Z = Z_G = \sum_{\eta \in \Omega} \exp(\beta M(\sigma))$.

Ferromagnetic (attractive) since $\beta > 0 \rightarrow$ neighbors prefer to align.

Markov Chain for Ising Model

Glauber Dynamics: For $G = (V, E)$, MC (X_t) on $\Omega = \{-1, +1\}^V$.

From $X_t \in \Omega$:

- Choose $v \in V$ uniformly at random.
- For all $w \neq v$, set $X_{t+1}(w) = X_t(w)$.
- Choose $X_t(v)$ from marginal conditional on neighbors spin:

$$\mu(\sigma(v) | \sigma(w) = X_{t+1}(w) \text{ for all } w \in N(v)).$$

Markov Chain for Ising Model

Glauber Dynamics: For $G = (V, E)$, MC (X_t) on $\Omega = \{-1, +1\}^V$.

From $X_t \in \Omega$:

- Choose $v \in V$ uniformly at random.
- For all $w \neq v$, set $X_{t+1}(w) = X_t(w)$.
- Choose $X_t(v)$ from marginal conditional on neighbors spin:

$$\mu(\sigma(v) | \sigma(w) = X_{t+1}(w) \text{ for all } w \in N(v)).$$

Stationary distribution π is Gibbs distribution μ.

Markov Chain for Ising Model

Glauber Dynamics: For $G = (V, E)$, MC (X_t) on $\Omega = \{-1, +1\}^V$.

From $X_t \in \Omega$:

- Choose $v \in V$ uniformly at random.
- For all $w \neq v$, set $X_{t+1}(w) = X_t(w)$.
- Choose $X_t(v)$ from marginal conditional on neighbors spin:

$$\mu(\sigma(v)|\sigma(w) = X_{t+1}(w) \text{ for all } w \in N(v)).$$

Stationary distribution π is Gibbs distribution μ.

How fast does it converge to π?

Markov Chain for Ising Model

Glauber Dynamics: For $G = (V, E)$, MC (X_t) on $\Omega = \{-1, +1\}^V$.

From $X_t \in \Omega$:

- Choose $v \in V$ uniformly at random.
- For all $w \neq v$, set $X_{t+1}(w) = X_t(w)$.
- Choose $X_t(v)$ from marginal conditional on neighbors spin:

$$\mu(\sigma(v)|\sigma(w) = X_{t+1}(w) \text{ for all } w \in N(v)).$$

Stationary distribution π is Gibbs distribution μ.

How fast does it converge to π?

$$T_{\mathrm{mix}}(\epsilon) = \max_{X_0 \in \Omega} \min\{t : d_{\mathsf{TV}}(P^t(X_0, \cdot), \pi) \leq \epsilon\}.$$

For dist. μ, ν on Ω, $d_{\mathsf{TV}}(\mu, \nu) = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)| = \max_{S \subset \Omega} \mu(S) - \nu(S)$.

Mixing time: $T_{\mathrm{mix}} = T_{\mathrm{mix}}(1/4)$

Sub-multiplicative: $T_{\mathrm{mix}}(\epsilon) = \lceil \log_2(\epsilon) \rceil T_{\mathrm{mix}}$

What is a coupling?

For distributions μ and ν on Ω,

coupling is a joint distribution ω on $\Omega \times \Omega$ where:

$$\text{for all } x \in \Omega, \quad \sum_{y \in \Omega} \omega(x, y) = \mu(x)$$

$$\text{for all } y \in \Omega, \quad \sum_{x \in \Omega} \omega(x, y) = \nu(y)$$

Coupling Technique

What is a coupling?

For distributions μ and ν on Ω,

coupling is a joint distribution ω on $\Omega \times \Omega$ where:

$$\text{for all } x \in \Omega, \quad \sum_{y \in \Omega} \omega(x, y) = \mu(x)$$

$$\text{for all } y \in \Omega, \quad \sum_{x \in \Omega} \omega(x, y) = \nu(y)$$

Intuition: If $\mu = \nu = \mathit{Uniform}(\Omega)$ then ω defines a fractional perfect matching.

Sample (X, Y) from ω in two equivalent ways:

1. Sample X from μ, then apply ω to get Y.
2. Sample Y from ν, then apply ω to get X.

What is a coupling?

For distributions μ and ν on Ω,

coupling is a joint distribution ω on $\Omega \times \Omega$ where:

$$\text{for all } x \in \Omega, \quad \sum_{y \in \Omega} \omega(x, y) = \mu(x)$$

$$\text{for all } y \in \Omega, \quad \sum_{x \in \Omega} \omega(x, y) = \nu(y)$$

Coupling Lemma: For any coupling ω of μ and ν:

choose $(X, Y) \sim \omega$, then $d_{\mathrm{TV}}(\mu, \nu) \le \mathbf{Pr}(X \ne Y)$.

Coupling Technique

What is a coupling?

For distributions μ and ν on Ω,

coupling is a joint distribution ω on $\Omega \times \Omega$ where:

$$\text{for all } x \in \Omega, \quad \sum_{y \in \Omega} \omega(x, y) = \mu(x)$$

$$\text{for all } y \in \Omega, \quad \sum_{x \in \Omega} \omega(x, y) = \nu(y)$$

Coupling Lemma: For any coupling ω of μ and ν:

choose $(X, Y) \sim \omega$, then $d_{\mathrm{TV}}(\mu, \nu) \leq \mathbf{Pr}\,(X \neq Y)$.

Proof:

$$\begin{aligned}
\mathbf{Pr}\,(X \neq Y) &\geq 1 - \sum_{z \in \Omega} \min\{\mu(z), \nu(z)\} \\
&= \sum_{z \in \Omega} \mu(z) - \min\{\mu(z), \nu(z)\} \\
&= \sum_{\substack{z \in \Omega: \\ \mu(z) \geq \nu(z)}} \mu(z) - \min\{\mu(z), \nu(z)\} \\
&= \max_{S \subset \Omega} \mu(S) - \nu(S) = d_{\mathrm{TV}}(\mu, \nu)
\end{aligned}$$

Coupling Technique

What is a coupling?

For distributions μ and ν on Ω,

coupling is a joint distribution ω on $\Omega \times \Omega$ where:

$$\text{for all } x \in \Omega, \quad \sum_{y \in \Omega} \omega(x, y) = \mu(x)$$

$$\text{for all } y \in \Omega, \quad \sum_{x \in \Omega} \omega(x, y) = \nu(y)$$

Coupling Lemma: For any coupling ω of μ and ν:

choose $(X, Y) \sim \omega$, then $d_{\mathrm{TV}}(\mu, \nu) \leq \mathbf{Pr}(X \neq Y)$.

Always exists optimal coupling:

There exists coupling ω of μ and ν where:

$$d_{\mathrm{TV}}(\mu, \nu) = \mathbf{Pr}(X \neq Y).$$

Proof: Constructive.

Coupling of Markov Chains

Consider a Markov chain (Ω, P).

Coupling is a joint process $\omega = (X_t, Y_t)$ on $\Omega \times \Omega$ where:

$$X_t \sim P \text{ and } Y_t \sim P$$

More precisely, for all $a, b, c \in \Omega$,

$$\begin{aligned} \mathbf{Pr}\left(X_{t+1} = c \mid X_t = a, Y_t = b\right) &= P(a, c) \\ \mathbf{Pr}\left(X_{t+1} = c \mid X_t = a, Y_t = b\right) &= P(b, c) \end{aligned}$$

Coupling of Markov Chains

Consider a Markov chain (Ω, P).

Coupling is a joint process $\omega = (X_t, Y_t)$ on $\Omega \times \Omega$ where:

$$X_t \sim P \text{ and } Y_t \sim P$$

More precisely, for all $a, b, c \in \Omega$,

$$\begin{aligned} \mathbf{Pr}(X_{t+1} = c \mid X_t = a, Y_t = b) &= P(a, c) \\ \mathbf{Pr}(X_{t+1} = c \mid X_t = a, Y_t = b) &= P(b, c) \end{aligned}$$

Intuition:

$(X_t \rightarrow X_{t+1}) \sim P$ and $(Y_t \rightarrow Y_{t+1}) \sim P$ but can correlate by ω.

Let X_0 be arbitrary, and $Y_0 \sim \pi$. Once $X_T = Y_T$ then $X_T \sim \pi$.

Coupling of Markov Chains

Consider a Markov chain (Ω, P).

Coupling is a joint process $\omega = (X_t, Y_t)$ on $\Omega \times \Omega$ where:

$$X_t \sim P \text{ and } Y_t \sim P$$

More precisely, for all $a, b, c \in \Omega$,

$$\begin{aligned} \mathbf{Pr}\left(X_{t+1} = c \mid X_t = a, Y_t = b\right) &= P(a, c) \\ \mathbf{Pr}\left(X_{t+1} = c \mid X_t = a, Y_t = b\right) &= P(b, c) \end{aligned}$$

Intuition:

$(X_t \to X_{t+1}) \sim P$ and $(Y_t \to Y_{t+1}) \sim P$ but can correlate by ω.

Let X_0 be arbitrary, and $Y_0 \sim \pi$. Once $X_T = Y_T$ then $X_T \sim \pi$.

Coupling time:

$$T_{\text{couple}} = \max_{a,b} \min\{t : \mathbf{Pr}\left(X_t \neq Y_t \mid X_0 = a, Y_0 = b\right) \leq 1/4.\}$$

$$T_{\text{mix}} \leq T_{\text{couple}}$$

Coupling Warm-up

Consider Ising with $\beta = 0$: $w(\sigma) = 1$ for all $\sigma \in \{-1,+1\}^V$

Note, vertices independent of each other.

Coupling Warm-up

Consider Ising with $\beta = 0$: $w(\sigma) = 1$ for all $\sigma \in \{-1,+1\}^V$

Glauber dynamics: From $X_t \in \{-1,+1\}^V$,

- Choose random vertex v_t and random spin $s_t \in \{-1,+1\}$.
- Set $X_{t+1}(v_t) = s_t$ and, for $w \neq v_t$, $X_{t+1}(w) = X_t(w)$.

Equivalent to (lazy) random walk on n-dimensional hypercube.

Coupling Warm-up

Consider Ising with $\beta = 0$: $w(\sigma) = 1$ for all $\sigma \in \{-1, +1\}^V$

Glauber dynamics: From $X_t \in \{-1, +1\}^V$,

- Choose random vertex v_t and random spin $s_t \in \{-1, +1\}$.
- Set $X_{t+1}(v_t) = s_t$ and, for $w \neq v_t$, $X_{t+1}(w) = X_t(w)$.

Consider arbitrary initial pair (X_0, Y_0).

Coupling at time t: Choose same vertex v_t and spin s_t.

Coupling Warm-up

Consider Ising with $\beta = 0$: $w(\sigma) = 1$ for all $\sigma \in \{-1, +1\}^V$

Glauber dynamics: From $X_t \in \{-1, +1\}^V$,

- Choose random vertex v_t and random spin $s_t \in \{-1, +1\}$.
- Set $X_{t+1}(v_t) = s_t$ and, for $w \neq v_t$, $X_{t+1}(w) = X_t(w)$.

Consider arbitrary initial pair (X_0, Y_0).
Coupling at time t: Choose same vertex v_t and spin s_t.

Disagree vertices: $D_t = \{v : X_t(v) \neq Y_t(v)\}$.

If update disagree $v_t \in D_t$ then: $X_{t+1}(v_t) = Y_{t+1}(v_t) = s_t$.
Hence, $\mathbb{E}\left[|D_{t+1}| \mid X_t, Y_t\right] \leq |D_t|(1 - 1/n)$.

Coupling Warm-up

Consider Ising with $\beta = 0$: $\quad w(\sigma) = 1$ for all $\sigma \in \{-1,+1\}^V$

Glauber dynamics: From $X_t \in \{-1,+1\}^V$,

- Choose random vertex v_t and random spin $s_t \in \{-1,+1\}$.
- Set $X_{t+1}(v_t) = s_t$ and, for $w \neq v_t$, $X_{t+1}(w) = X_t(w)$.

Consider arbitrary initial pair (X_0, Y_0).

Coupling at time t: Choose same vertex v_t and spin s_t.

Disagree vertices: $\quad D_t = \{v : X_t(v) \neq Y_t(v)\}$.

If update disagree $v_t \in D_t$ then: $X_{t+1}(v_t) = Y_{t+1}(v_t) = s_t$.

Hence, $\quad \mathbb{E}\left[|D_{t+1}| \mid X_t, Y_t\right] \leq |D_t|(1 - 1/n)$.

$$\begin{aligned}
\mathbf{Pr}\left(X_T \neq Y_T\right) &\leq \mathbb{E}\left[|D_t|\right] && \text{since if } X_T \neq Y_T \text{ then } |D_T| \geq 1 \\
&\leq |D_0|(1 - 1/n)^T \\
&\leq n\exp(-T/n) \\
&\leq 1/4 \quad \text{for } T \geq n\ln(4n).
\end{aligned}$$

Ising with $\beta > 0$

Consider $\beta > 0$:

ISING WITH $\beta > 0$

Consider $\beta > 0$:
Glauber dynamics: From $X_t \in \{-1, +1\}^V$,

- Choose random vertex v_t .
- For $w \neq v_t$, $X_{t+1}(w) = X_t(w)$.
- Let $d^+ = \#$ of $+1$ neighbors of v_t, $d^- = \#$ of -1 neighbors.
- Set $X_{t+1}(v_t) = +1$ with prob. $\propto \exp(\beta d^+)$
  and $X_{t+1}(v_t) = -1$ with prob. $\propto \exp(\beta d^-)$.

Ising with $\beta > 0$

Consider $\beta > 0$:

Glauber dynamics: From $X_t \in \{-1,+1\}^V$,

- Choose random vertex v_t random $r_t \in [0,1]$.
- For $w \neq v_t$, $X_{t+1}(w) = X_t(w)$.
- Let $d^+ = \#$ of $+1$ neighbors of v_t, $d^- = \#$ of -1 neighbors.
- $X_{t+1}(v) = \begin{cases} +1 & \text{if } r_t \le \frac{\exp(\beta d^+)}{\exp(\beta d^+)+\exp(\beta d^-)} \\ -1 & \text{otherwise} . \end{cases}$

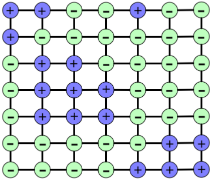
Disagree vertices: $D_t = \{v : X_t(v) \neq Y_t(v)\}$.

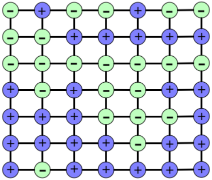
If $\beta = 0$: $D_t \supseteq D_{t+1}$.

If $\beta > 0$: D_t can grow.

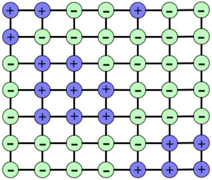
Coupling for Ising Model

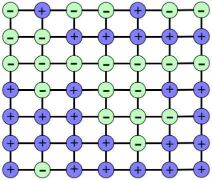
Consider a pair of Ising configurations X_t and Y_t:



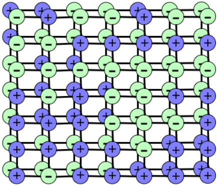


Consider a pair of Ising configurations X_t and Y_t:



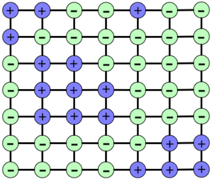


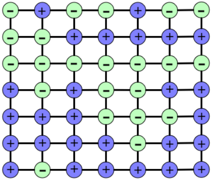
Look at $\frac{X_t}{Y_t}$:



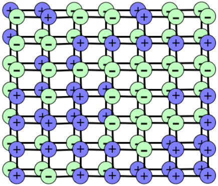
Coupling for Ising Model

Consider a pair of Ising configurations X_t and Y_t:





Look at $\frac{X_t}{Y_t}$:



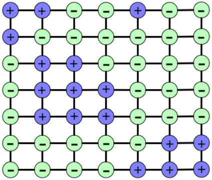
Identity Coupling:

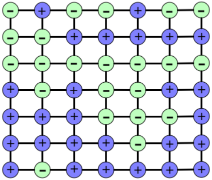
Update same v_t,

same seed r_t (i.e., same spin as much as possible).

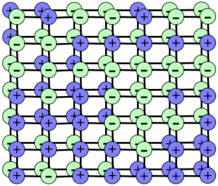
Coupling for Ising Model

Consider a pair of Ising configurations X_t and Y_t:





Look at $\frac{X_t}{Y_t}$:



Identity Coupling:

Update same v_t,

same seed r_t (i.e., same spin as much as possible).

How to analyze???

Coupling for bounding T_{mix}

For all X_t, Y_t, define a coupling: $(X_t, Y_t) \to (X_{t+1}, Y_{t+1})$.

Look at Hamming distance: $H(X_t, Y_t) = |D_t|$.

If for all X_t, Y_t,

$$\mathbb{E}\left[H(X_{t+1}, Y_{t+1}) \middle| X_t, Y_t \right] \leq (1 - 1/n) H(X_t, Y_t),$$

$$\text{then } T_{\mathrm{mix}} = O(n \log n).$$

Path coupling: Suffices to consider pairs where $H(X_t, Y_t) = 1$.

Path Coupling [Bubley and Dyer '97]

Let $S \subset \Omega^2$ denote pairs (X_t, Y_t) where $H(X_t, Y_t) = 1$.
Define a coupling ω for all $(X_t, Y_t) \in S$ where:

$$\mathbb{E}\left[H(X_{t+1}, Y_{t+1}) \mid X_t, Y_t\right] \leq 1 - C/n.$$

Path Coupling [Bubley and Dyer '97]

Let $S \subset \Omega^2$ denote pairs (X_t, Y_t) where $H(X_t, Y_t) = 1$.
Define a coupling ω for all $(X_t, Y_t) \in S$ where:

$$\mathbb{E}\left[H(X_{t+1}, Y_{t+1}) \mid X_t, Y_t\right] \leq 1 - C/n.$$

For arbitrary $(A_t, B_t) \in \Omega^2$:

In graph (Ω, S), consider a shortest path A_t to B_t:

$$(A_t, W_t^1, W_t^2, \ldots, W_t^{\ell-1}, B_t),\ \ell = H(A_t, B_t).$$

Couplings: $\omega^1 = (W_t^0, W_t^1), \ldots, \omega^\ell = (W_t^{\ell-1}, W_t^\ell)$.

Compose: $\omega = \omega^1 \circ \omega^2 \circ \cdots \circ \omega^\ell$ gives coupling (A_t, B_t).

Path Coupling [Bubley and Dyer '97]

Let $S \subset \Omega^2$ denote pairs (X_t, Y_t) where $H(X_t, Y_t) = 1$.
Define a coupling ω for all $(X_t, Y_t) \in S$ where:

$$\mathbb{E}\left[H(X_{t+1}, Y_{t+1}) \mid X_t, Y_t\right] \leq 1 - C/n.$$

For arbitrary $(A_t, B_t) \in \Omega^2$:

In graph (Ω, S), consider a shortest path A_t to B_t:

$(A_t, W_t^1, W_t^2, \ldots, W_t^{\ell-1}, B_t), \ \ell = H(A_t, B_t)$.

Couplings: $\omega^1 = (W_t^0, W_t^1), \ldots, \omega^\ell = (W_t^{\ell-1}, W_t^\ell)$.

Compose: $\omega = \omega^1 \circ \omega^2 \circ \cdots \circ \omega^\ell$ gives coupling (A_t, B_t).

Algorithmic View:

1. Choose $A_t \to A_{t+1}$ by P,
2. Apply ω^1 to get $W_t^1 \to W_{t+1}^1$,
3. Apply ω^2 to get $W_t^2 \to W_{t+1}^2$, $\ldots$,
4. Get $B_t \to B_{t+1}$.

Path Coupling [Bubley and Dyer '97]

Let $S \subset \Omega^2$ denote pairs (X_t, Y_t) where $H(X_t, Y_t) = 1$.
Define a coupling ω for all $(X_t, Y_t) \in S$ where:

$$\mathbb{E}\left[H(X_{t+1}, Y_{t+1}) \mid X_t, Y_t\right] \le 1 - C/n.$$

For arbitrary $(A_t, B_t) \in \Omega^2$:

In graph (Ω, S), consider a shortest path A_t to B_t:

$$(A_t, W_t^1, W_t^2, \ldots, W_t^{\ell-1}, B_t),\ \ell = H(A_t, B_t).$$

Couplings: $\omega^1 = (W_t^0, W_t^1), \ldots, \omega^\ell = (W_t^{\ell-1}, W_t^\ell)$.
Compose: $\omega = \omega^1 \circ \omega^2 \circ \cdots \circ \omega^\ell$ gives coupling (A_t, B_t).

$$\begin{aligned}
\mathbb{E}\left[H(A_{t+1}, B_{t+1})\right] &\le \mathbb{E}\left[\sum_i H(W_{t+1}^{i-1}, W_{t+1}^i)\right] \\
&\le \sum_i \mathbb{E}\left[H(W_{t+1}^{i-1}, W_{t+1}^i)\right] \\
&\le \sum_i (1 - C/n) \\
&\le H(A_t, B_t)(1 - C/n).
\end{aligned}$$

Path Coupling [Bubley and Dyer '97]

Let $S \subset \Omega^2$ denote pairs (X_t, Y_t) where $H(X_t, Y_t) = 1$.
Define a coupling ω for all $(X_t, Y_t) \in S$ where:

$$\mathbb{E}\left[H(X_{t+1}, Y_{t+1}) \mid X_t, Y_t\right] \leq 1 - C/n.$$

Then for arbitrary $(A_t, B_t) \in \Omega^2$, can construct coupling where:

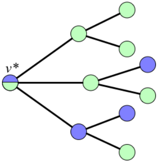
$$\mathbb{E}\left[H(A_{t+1}, B_{t+1})\right] \leq H(A_t, B_t)(1 - C/n).$$

$$\begin{aligned}
\mathbf{Pr}\left(A_T \neq B_T\right) &\leq \mathbb{E}\left[H(A_T, B_T)\right] \\
&\leq H(A_0, B_0)(1 - C/n)^T \\
&\leq n\exp(-C/n) \\
&\leq 1/4 \quad \text{for } T = O(n\log n).
\end{aligned}$$

Hence, $T_{\text{mix}} = O(n \log n)$.

Path Coupling [Bubley and Dyer '97]

Consider a pair (X_t, Y_t) that differ at exactly one vertex v^*:



Update v^* then $H(X_{t+1}, Y_{t+1}) = 0$.
Update $w \in N(v^*)$ then $H(X_{t+1}, Y_{t+1}) = 2$ with probability:

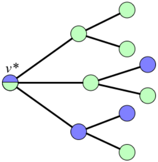
$$\alpha(w) := \frac{\exp(\beta(d^+ + 1))}{\exp(\beta(d^+ + 1)) + \exp(\beta d^-)} - \frac{\exp(\beta d^+)}{\exp(\beta d^+) + \exp(\beta(d^- + 1))}$$

$$\mathbb{E}\left[H(X_{t+1}, Y_{t+1})\right] \le 1 + \frac{1}{n}\left[-1 + \sum_{w \in N(v)} \alpha(w)\right].$$

Worst case $d^+ = d^-$. When $d = 4$ works for $\beta < \ln(5/3)$.

Path Coupling [Bubley and Dyer '97]

Consider a pair (X_t, Y_t) that differ at exactly one vertex v^*:



Update v^* then $H(X_{t+1}, Y_{t+1}) = 0$.
Update $w \in N(v^*)$ then $H(X_{t+1}, Y_{t+1}) = 2$ with probability:

$$\alpha(w) := \frac{\exp(\beta(d^+ + 1))}{\exp(\beta(d^+ + 1)) + \exp(\beta d^-)} - \frac{\exp(\beta d^+)}{\exp(\beta d^+) + \exp(\beta(d^- + 1))}$$

$$\mathbb{E}\left[H(X_{t+1}, Y_{t+1})\right] \leq 1 + \frac{1}{n}\left[-1 + \sum_{w \in N(v)} \alpha(w)\right].$$

Worst case $d^+ = d^-$. When $d = 4$ works for $\beta < \ln(5/3)$.
Goal: All $\beta < \beta_c := \ln(1 + \sqrt{2})$.

Phase Transition

$\Lambda_n = \sqrt{n} \times \sqrt{n}$ box of $\mathbb{Z}^2$.

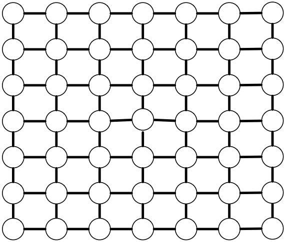
- Path coupling: $T_{\mathrm{mix}} = O(n \log n)$ for all $\beta < \ln(5/3)$.
- Phase transition: $\beta_c = \ln(1 + \sqrt{2})$.

How to prove $T_{\mathrm{mix}} = O(n \log n)$ for all $\beta < \beta_c$?

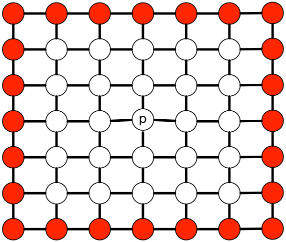
To Do:

- What's the phase transition mean?
- SSM = Strong Spatial Mixing
- Monotonicity of the Ising Model
- SSM $\rightarrow T_{\mathrm{mix}} = O(n \log n)$ on Λ_n.
  (we'll do $O(n \log^2 n)$.)

Ising Model Phase Transition

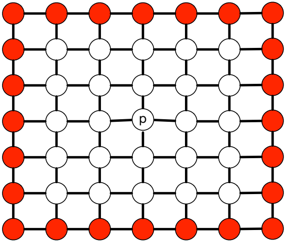


Ising Model Phase Transition



Let $p_n^+ = \mathbf{Pr}\left(\sigma(v^*) = +1 \mid \partial\Lambda_n = +\right)$ and p_n^- for -1 boundary.

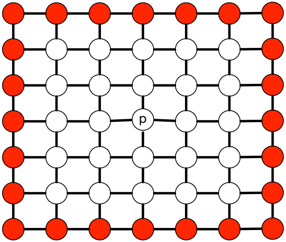
Ising Model Phase Transition



Let $p_n^+ = \mathbf{Pr}\left(\sigma(v^*) = +1 \mid \partial\Lambda_n = +\right)$ and p_n^- for -1 boundary.

$$\lim_{n\to\infty} p_n^+ \stackrel{?}{=} \lim_{n\to\infty} p_n^-$$

Ising Model Phase Transition



Let $p_n^+ = \mathbf{Pr}\left(\sigma(v^*) = +1 \mid \partial\Lambda_n = +\right)$ and p_n^- for -1 boundary.

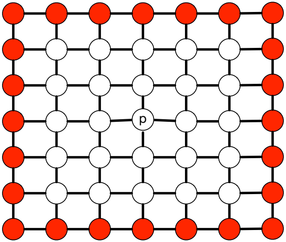
$$\lim_{n\to\infty} p_n^+ \stackrel{?}{=} \lim_{n\to\infty} p_n^-$$

Let $\beta_c = \ln(1+\sqrt{2})$. [Onsager '44]:

For all $\beta \le \beta_c$, $\quad p_\infty^+ = p_\infty^-$

For all $\beta > \beta_c$, $\quad p_\infty^+ \neq p_\infty^-$

Ising Model Phase Transition



Let $p_n^+ = \mathbf{Pr}\left(\sigma(v^*) = +1 \mid \partial\Lambda_n = +\right)$ and p_n^- for -1 boundary.

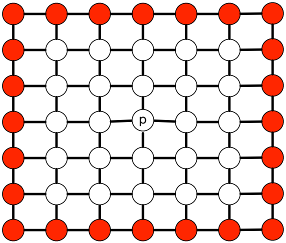
$$\lim_{n\to\infty} p_n^+ \stackrel{?}{=} \lim_{n\to\infty} p_n^-$$

Let $\beta_c = \ln(1+\sqrt{2})$. [Onsager '44]:

For all $\beta \le \beta_c$, $p_\infty^+ = p_\infty^-$ uniqueness

For all $\beta > \beta_c$, $p_\infty^+ \neq p_\infty^-$ non-uniqueness

Ising Model Phase Transition



Let $p_n^+ = \mathbf{Pr}\left(\sigma(v^*) = +1 \mid \partial\Lambda_n = +\right)$ and p_n^- for -1 boundary.

$$\lim_{n\to\infty} p_n^+ \stackrel{?}{=} \lim_{n\to\infty} p_n^-$$

Let $\beta_c = \ln(1+\sqrt{2})$. [Onsager '44]:

For all $\beta \leq \beta_c$, $p_\infty^+ = p_\infty^-$ uniqueness

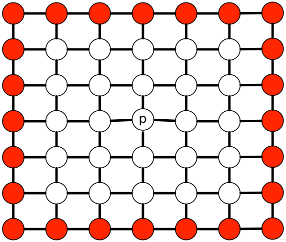
For all $\beta > \beta_c$, $p_\infty^+ \neq p_\infty^-$ non-uniqueness

q-state Potts model, [Beffara, Duminil-Copin '12]: $\beta_c(q) = \ln(1+\sqrt{q})$.

Continuous for $q \leq 4$ [Duminil-Copin, Sidoravicius, Tassion '16]

Discont. for $q > 4$ [Duminil-Copin,Gagnebin,Harel,Manolescu,Tassion '18]

Ising Model Phase Transition



Let $p_n^+ = \mathbf{Pr}\left(\sigma(v^*) = +1 \mid \partial\Lambda_n = +\right)$ and p_n^- for -1 boundary.

$$\lim_{n\to\infty} p_n^+ \stackrel{?}{=} \lim_{n\to\infty} p_n^-$$

Let $\beta_c = \ln(1+\sqrt{2})$. [Onsager '44]:

For all $\beta \le \beta_c$, $p_\infty^+ = p_\infty^-$ uniqueness

For all $\beta > \beta_c$, $p_\infty^+ \ne p_\infty^-$ non-uniqueness

T_{mix} for Glauber dynamics on $\sqrt{n} \times \sqrt{n}$ box:

All $\beta < \beta_c$, all b.c.: $O(n \log n)$ [Martinelli,Olivieri,Schonmann'94]

All $\beta > \beta_c$, wired: $\exp(\Omega(\sqrt{n}))$ [Thomas'89]

$\beta = \beta_c$? See: [Lubetzky,Sly'12][Borgs,Chayes,Tetali'12][Gheissari,Lubetzky'18]

Spatial Mixing

For a box Λ_n and $v \in V$, let $\mathbf{p}(v) = \mathbf{Pr}(v = +)$.

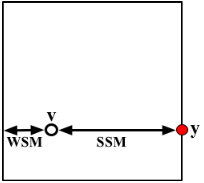
Weak Spatial Mixing (WSM):
Exists $C, \alpha > 0$, for all boxes Λ_n, all $v \in V$, all boundaries σ, η:

$$|\mathbf{p}^{\sigma}(v) - \mathbf{p}^{\eta}(v)| \leq C \exp(-\alpha \mathrm{dist}(v, \partial\Lambda_n))$$

For $y \in \partial\Lambda_n$ and boundary σ, obtain σ^y by "flipping" spin at y.

Strong Spatial Mixing (SSM):
Exists $C, \alpha > 0$, all Λ_n, all $v \in V$, all $y \in \partial\Lambda_n$, boundaries σ, σ^y:

$$|\mathbf{p}^{\sigma}(v) - \mathbf{p}^{\sigma^y}(v)| \leq C \exp(-\alpha \mathrm{dist}(v, y))$$



Spatial Mixing

For a box Λ_n and $v \in V$, let $\mathbf{p}(v) = \mathbf{Pr}\,(v = +)$.

Weak Spatial Mixing (WSM):
Exists $C, \alpha > 0$, for all boxes Λ_n, all $v \in V$, all boundaries σ, η:

$$|\mathbf{p}^{\sigma}(v) - \mathbf{p}^{\eta}(v)| \leq C \exp(-\alpha \mathrm{dist}(v, \partial\Lambda_n))$$

For $y \in \partial\Lambda_n$ and boundary σ, obtain σ^y by "flipping" spin at y.

Strong Spatial Mixing (SSM):
Exists $C, \alpha > 0$, all Λ_n, all $v \in V$, all $y \in \partial\Lambda_n$, boundaries σ, σ^y:

$$|\mathbf{p}^{\sigma}(v) - \mathbf{p}^{\sigma^y}(v)| \leq C \exp(-\alpha \mathrm{dist}(v, y))$$

In 2-dimensions, for all $\beta < \beta_c$: SSM holds.

Monotonicity

Define partial ordering on $\{-1,+1\}^V$:

$$X_t \le Y_t \iff \text{for all } v \in V, X_t(v) \le Y_t(v)$$

Extremal states: $all - 1$ and $all + 1$.

Monotonicity

Define partial ordering on $\{-1,+1\}^V$:

$$X_t \leq Y_t \iff \text{for all } v \in V, X_t(v) \leq Y_t(v)$$

Extremal states: *all* -1 and *all* $+1$.
Grand coupling: All chains choose same v_t, same r_t.

If $X_t \leq Y_t$ then $X_{t+1} \leq Y_{t+1}$.

Glauber dynamics: From $X_t \in \{-1,+1\}^V$,

- Choose random vertex v_t and random $r_t \in [0,1]$.
- For $w \neq v_t$, $X_{t+1}(w) = X_t(w)$.
- Let $d^+ = \#$ of $+1$ neighbors of v_t, $d^- = \#$ of -1 neighbors.
- Set $X_{t+1}(v_t) = +1$ with prob. $\propto \exp(\beta d^+)$
  and $X_{t+1}(v_t) = -1$ with prob. $\propto \exp(\beta d^-)$.

Monotonicity

Define partial ordering on $\{-1,+1\}^V$:

$$X_t \le Y_t \iff \text{for all } v \in V, X_t(v) \le Y_t(v)$$

Extremal states: *all* -1 and *all* $+1$.
Grand coupling: All chains choose same v_t, same r_t.

If $X_t \le Y_t$ then $X_{t+1} \le Y_{t+1}$.

Glauber dynamics: From $X_t \in \{-1,+1\}^V$,

- Choose random vertex v_t and random $r_t \in [0,1]$.
- For $w \neq v_t$, $X_{t+1}(w) = X_t(w)$.
- Let $d^+ = \#$ of $+1$ neighbors of v_t, $d^- = \#$ of -1 neighbors.
- Set $X_{t+1}(v) = +1$ if $r_t \le \exp(\beta d^+)/(\exp(\beta d^+) + \exp(\beta d^-))$.
  Else, set $X_{t+1}(v_t) = -1$.

Monotonicity

Define partial ordering on $\{-1,+1\}^V$:

$$X_t \leq Y_t \iff \text{for all } v \in V, X_t(v) \leq Y_t(v)$$

Extremal states: *all* -1 and *all* $+1$.
Grand coupling: All chains choose same v_t, same r_t.

If $X_t \leq Y_t$ then $X_{t+1} \leq Y_{t+1}$.

Only need to couple extremal states:
Let $X_0^- =$ all -1 and $X_0^+ =$ all $+1$.
Then for all $Y_0 \in \Omega$, for all t:

$$X_t^- \leq Y_t \leq X_t^+$$

Hence, if $X_T^- = X_T^+$ then $T_{\mathrm{mix}} \leq T_{\mathrm{couple}} \leq T$.

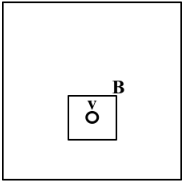
SSM → Rapid Mixing

Blanca's boosting proof [Dyer,Sinclair,V,Weitz '04]:

Fix the graph Λ_n as $\sqrt{n} \times \sqrt{n}$ box of $\mathbb{Z}^2$ – arbitrary b.c..
Let $T^*(n)$ denote an upper bound on T_{mix}, e.g., $T^*(n) \leq 2^n$.
Consider $X_0^- = -1, X_0^+ = +1$.
Goal: Prove $\mathbf{Pr}\left(X_T^- \neq X_T^+\right) \leq 1/4$ for

$$T = \frac{n}{\log^2(n)} T^*(C \log^2 n)) \log 9n = O(n/\log n) \times T^*(C \log^2 n).$$

Suffices to show, for every $v \in V$: $\mathbf{Pr}\left(X_T^-(v) \neq X_T^+(v)\right) \leq 1/4n$.
Fix $v \in V$. Box $B = B_\ell(v), \ell = O(\log n)$ (note, $|B| = O(\log^2 n)$)
Define $Y_0^- = -1$, $Y_0^+ = +1$, but frozen in $\overline{B}$: $Y_t^- \leq X_t^- \leq X_t^+ \leq Y_t^+$



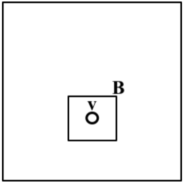
SSM $\to$ Rapid Mixing

Blanca's boosting proof [Dyer,Sinclair,V,Weitz '04]:

Fix the graph Λ_n as $\sqrt{n} \times \sqrt{n}$ box of $\mathbb{Z}^2$ – arbitrary b.c..
Let $T^*(n)$ denote an upper bound on T_{mix}, e.g., $T^*(n) \leq 2^n$.
Consider $X_0^- = -1, X_0^+ = +1$.
Goal: Prove $\mathbf{Pr}\left(X_T^- \neq X_T^+\right) \leq 1/4$ for

$$T = \frac{n}{\log^2(n)} T^*(C\log^2 n)) \log 9n = \ O(n/\log n) \times T^*(C\log^2 n).$$

Suffices to show, for every $v \in V$: $\mathbf{Pr}\left(X_T^-(v) \neq X_T^+(v)\right) \leq 1/4n$.
Fix $v \in V$. Box $B = B_\ell(v), \ell = O(\log n)$ (note, $|B| = O(\log^2 n)$)
Define $Y_0^- = -1$, $Y_0^+ = +1$, but frozen in $\overline{B}$: $Y_t^- \leq X_t^- \leq X_t^+ \leq Y_t^+$



T updates on Λ_n then whp $T \times (|B|/n)$=$T^*(|B|)\log(9n)$ on B.
If $T_{\mathrm{mix}}(1/4) = T^*$ then $T_{\mathrm{mix}}(1/9n) = T^* \log(9n)$.

SSM $\to$ Rapid Mixing

Blanca's boosting proof [Dyer,Sinclair,V,Weitz '04]:

Fix the graph Λ_n as $\sqrt{n} \times \sqrt{n}$ box of $\mathbb{Z}^2$ – arbitrary b.c..
Let $T^*(n)$ denote an upper bound on $T_{\rm mix}$, e.g., $T^*(n) \leq 2^n$.
Consider $X_0^- = -1, X_0^+ = +1$.
Goal: Prove $\mathbf{Pr}\left(X_T^- \neq X_T^+\right) \leq 1/4$ for

$$T = \frac{n}{\log^2(n)} T^*(C \log^2 n)) \log 9n = O(n/\log n) \times T^*(C \log^2 n).$$

Suffices to show, for every $v \in V$: $\mathbf{Pr}\left(X_T^-(v) \neq X_T^+(v)\right) \leq 1/4n$.
Fix $v \in V$. Box $B = B_\ell(v), \ell = O(\log n)$ (note, $|B| = O(\log^2 n)$)
Define $Y_0^- = -1$, $Y_0^+ = +1$, but frozen in $\overline{B}$: $Y_t^- \leq X_t^- \leq X_t^+ \leq Y_t^+$
T updates on Λ_n then whp $T \times (|B|/n) = T^*(|B|) \log(9n)$ on B.
If $T_{\rm mix}(1/4) = T^*$ then $T_{\rm mix}(1/9n) = T^* \log(9n)$.

$$\begin{aligned}
&\mathbf{Pr}\left(X_T^-(v) \neq X_T^+(v)\right) \\
&\quad \leq \mathbf{Pr}\left(Y_T^-(v) \neq Y_T^+(v)\right) \\
&\quad \leq \mathbf{Pr}\left(Y_T^+(v) = +\right) - \mathbf{Pr}\left(Y_T^-(v) = +\right) \\
&\quad \leq \left|\mathbf{Pr}\left(Y_T^+(v) = +\right) - \mu_B^+(v = +)\right| + \left|\mu_B^+(v = +) - \mu_B^-(v = +)\right| \\
&\qquad + \left|\mu_B^-(v = +) - \mathbf{Pr}\left(Y_T^-(v) = +\right)\right| \\
&\quad \leq 1/9n + \exp(-C' \log n) + 1/9n \quad \text{by induction + SSM} \\
&\quad \leq 1/4n
\end{aligned}$$

Boosting

Blanca's boosting proof [Dyer,Sinclair,V,Weitz '04]:

$T_{\rm mix} = O(n/\log n) \times T^*(C \log^2 n)).$

$$2^n \quad \rightarrow \quad 2^{O(\log^2 n)} O(n/\log n) \leq n^{O(\log n)}$$

Blanca's boosting proof [Dyer,Sinclair,V,Weitz '04]:

$$T_{\mathrm{mix}} = O(n/\log n) \times T^*(C \log^2 n)).$$

$$\begin{aligned} 2^n &\rightarrow 2^{O(\log^2 n)} O(n/\log n) \le n^{O(\log n)} \\ n^{O(\log n)} &\rightarrow 2^{O((\log\log n)^2)} O(n/\log n) \le O(n^{1+\epsilon}) \end{aligned}$$

Boosting

Blanca's boosting proof [Dyer,Sinclair,V,Weitz '04]:

$T_{\mathrm{mix}} = O(n/\log n) \times T^*(C \log^2 n))$.

$$\begin{aligned} 2^n &\rightarrow 2^{O(\log^2 n)} O(n/\log n) \leq n^{O(\log n)} \\ n^{O(\log n)} &\rightarrow 2^{O((\log\log n)^2)} O(n/\log n) \leq O(n^{1+\epsilon}) \\ O(n^{1+\epsilon}) &\rightarrow O(n \log^2 n) \end{aligned}$$

Blanca's boosting proof [Dyer,Sinclair,V,Weitz '04]:

$T_{\mathrm{mix}} = O(n/\log n) \times T^*(C\log^2 n)$.

$$
\begin{aligned}
2^n &\rightarrow 2^{O(\log^2 n)} O(n/\log n) \le n^{O(\log n)} \\
n^{O(\log n)} &\rightarrow 2^{O((\log\log n)^2)} O(n/\log n) \le O(n^{1+\epsilon}) \\
O(n^{1+\epsilon}) &\rightarrow O(n\log^2 n)
\end{aligned}
$$

Extra Boosting:

Can boost from $O(n^{1.5-\epsilon})$, for any $\epsilon > 0$, to $O(n\log n)$.

General graphs

Let $\beta_c(\mathbb{T}_d)$ denote the critical point for infinite d-regular tree $\mathbb{T}_d$.

Note, $\beta_c(\mathbb{T}_d) = 2(d-1)\tanh(\beta)$.

[Mossel-Sly '13]: For all G of max. degree d, all $\beta < \beta_c(\mathbb{T}_d)$,

$$T_{\mathrm{mix}} = O(n \log n).$$

Previous approach for grid:

Box/ball B of radius $\Omega(\log n)$ around v.

But for arbitrary G, $|B|$ can be $|G|$.

Can only do constant radius R for B.

Hence, need to do multistages, can't couple in one round.

Let $\beta_c(\mathbb{T}_d)$ denote the critical point for infinite d-regular tree $\mathbb{T}_d$.

Note, $\beta_c(\mathbb{T}_d) = 2(d-1)\tanh(\beta)$.

[Mossel-Sly '13]: For all G of max. degree d, all $\beta < \beta_c(\mathbb{T}_d)$,

$$T_{\mathrm{mix}} = O(n \log n).$$

Weitz's approach: running time $n^{1+C(\beta,d)}$

Previous approach for grid:

Box/ball B of radius $\Omega(\log n)$ around v.
But for arbitrary G, $|B|$ can be $|G|$.
Can only do constant radius R for B.
Hence, need to do multistages, can't couple in one round.

General graphs

Let $\beta_c(\mathbb{T}_d)$ denote the critical point for infinite d-regular tree $\mathbb{T}_d$.

Note, $\beta_c(\mathbb{T}_d) = 2(d-1)\tanh(\beta)$.

[Mossel-Sly '13]: For all G of max. degree d, all $\beta < \beta_c(\mathbb{T}_d)$,

$$T_{\mathrm{mix}} = O(n \log n).$$

Weitz's approach: running time $n^{1+C(\beta,d)}$

Previous approach for grid:

Box/ball B of radius $\Omega(\log n)$ around v.
But for arbitrary G, $|B|$ can be $|G|$.
Can only do constant radius R for B.
Hence, need to do multistages, can't couple in one round.

Key ideas:

- Stronger Spatial Mixing condition
- Induction on disagreement probability for a vertex
- Censoring

Spatial Mixing

For a graph $G = (V, E)$, for $v \in V$, let $\mathbf{p}(v) = \mathbf{Pr}(v = +)$.
For radius R, let $B = B_R(v)$ be the ball of radius R around v.
For $y \in \partial B$ and boundary σ, obtain σ^y by "flipping" spin at y.

Strong Spatial Mixing (SSM):
Exists $C, \alpha > 0$,

$$a_y \leq C(\exp(-\alpha R).$$

Aggregate Strong Spatial Mixing (ASSM):
Exists R,

$$\sum_{y \in \partial B} a_y \leq \frac{1}{4}.$$

Spatial Mixing

For a graph $G = (V, E)$, for $v \in V$, let $\mathbf{p}(v) = \mathbf{Pr}(v = +)$.
For radius R, let $B = B_R(v)$ be the ball of radius R around v.
For $y \in \partial B$ and boundary σ, obtain σ^y by "flipping" spin at y.

Strong Spatial Mixing (SSM):
Exists $C, \alpha > 0$,

$$a_y \leq C(\exp(-\alpha R).$$

Aggregate Strong Spatial Mixing (ASSM):
Exists R,

$$\sum_{y \in \partial B} a_y \leq \frac{1}{4}.$$

Roughly: ASSM means $\exp(-\alpha) < 1/d$

.

Mossel-Sly Proof Approach

Once again, $X_0^+ = +1, X_0^- = -1$.
Goal: for every $v \in V$, $\mathbf{Pr}\left(X_T^+(v) \neq X_T^-(v)\right) \leq 1/4n$.

Ball $B = B_R(v)$ where R is radius from ASSM.
Let T^* denote, max v, worst boundary σ, mixing time on $B_R^\sigma(v)$.
Let $T' = T^* \times \log(8d|B|)$.
Induction, for all $s \geq 0$,

$$\max_v \mathbf{Pr}\left(X_{s+T'}^+(v) \neq X_{s+T'}^-(v)\right) \leq \frac{1}{2} \max_v \mathbf{Pr}\left(X_s^+(v) \neq X_s^-(v)\right)$$

Then we get $T_{\mathrm{mix}} = T' \log(n/\epsilon)$.

Mossel-Sly Inductive Proof

Fix v and time s. Let's prove:

$$\max_v \mathbf{Pr}\left(X^+_{s+T'}(v) \neq X^-_{s+T'}(v)\right) \leq \frac{1}{2} \max_v \mathbf{Pr}\left(X^+_s(v) \neq X^-_s(v)\right)$$

New chains: $Y^+_0 = +1$, $Y^-_0 = -1$.
For times $t \leq s$, same: $X^+_t = Y^+_t$ and $X^-_t = Y^-_t$.
For times $t > s$, Y^+_t and Y^-_t ignore moves outside B.
Note, $X^+_s(\overline{B}) = Y^+_s(\overline{B})$ is arbitrary.
Can go $X^+_t(\overline{B}) \geq Y^+_t(\overline{B})$ since we didn't fix extremal (all $+$) on $\overline{B}$.

Instead use:
[Peres-Winkler '13] Censoring: "Extra moves don't hurt":

$$Y^+_t \geq X^+_t \geq X^-_t \geq Y^+_t$$

Hence, suffices to bound $\mathbf{Pr}\left(Y^+_{s+T'}(v) \neq Y^-_{s+T'}(v)\right)$.

Ball $B = B_R(v)$ where R is from ASSM.
Outer boundary $\partial B = S_{R+1}(v)$.
All $A = B_{R+1}(v) = B \cup \partial B$.
Let $\sigma^+ = X_s^+(A) = Y_s^+(A)$ and $\sigma^- = X_s^-(A) = Y_s^-(A)$.

$$\begin{aligned}\mathbf{Pr}\left(Y^+_{s+T'}(v) \neq Y^-_{s+T'}(v) \mid \sigma^+, \sigma^-\right) &\leq \left|\mathbf{Pr}\left(Y^+_{s+T'}(v) = + \mid Y_s^+(A) = \sigma^+\right) - \mu_B^{\sigma^+}(v=+)\right| \\ &+\left|\mu_B^{\sigma^+}(v=+) - \mu_B^{\sigma^-}(v=+)\right| \\ &+\left|\mathbf{Pr}\left(Y^+_{s+T'}(v) = + \mid Y_s^-(A) = \sigma^-\right) - \mu_B^{\sigma^-}(v=+)\right|\end{aligned}$$

Now bound each term on RHS.

Induction

Recall, $T' = T^* \times \log(8|A|)$.
Outer terms:

$$\left|\mathbf{Pr}\left(Y^+_{s+T'}(v) = + \mid \sigma^+\right) - \mu_B^{\sigma^+}(v = +)\right| \leq 1/8|A|.$$

Middle term:

$$\left|\mu_B^{\sigma^+}(v = +) - \mu_B^{\sigma^-}(v = +)\right| \leq \sum_{w \in \partial B} \mathbf{1}\big(\sigma^+(w) \neq \sigma^-(w)\big) a_w.$$

Combining we get:

$$\mathbf{Pr}\left(Y^+_{s+T'}(v) \neq Y^-_{s+T'}(v) \mid \sigma^+, \sigma^-\right) \leq \frac{1}{4|A|} + \sum_{w \in \partial B} \mathbf{1}\big(\sigma^+(w) \neq \sigma^-(w)\big) a_w.$$

Unconditioning...

Combining we got:

$$\mathbf{Pr}\left(Y^+_{s+T'}(v) \neq Y^-_{s+T'}(v) \mid \sigma^+, \sigma^-\right) \leq \frac{1}{4|A|} + \sum_{w \in \partial B} \mathbf{1}\left(\sigma^+(w) \neq \sigma^-(w)\right) a_w.$$

If $Y^+_s(A) = Y^-_s(A)$ then $Y^+_t(A) = Y^-_t(A)$ for $t \geq s$.

$$\begin{aligned}
&\mathbf{Pr}\left(Y^+_{s+T'}(v) \neq Y^-_{s+T'}(v)\right) \\
&\leq \sum_{\sigma^+,\sigma^-} \mathbf{Pr}\left(Y^+_{s+T'}(v) \neq Y^-_{s+T'}(v) \mid \sigma^+, \sigma^-\right) \mathbf{Pr}\left(\sigma^+, \sigma^-\right) \\
&\leq \frac{\mathbf{Pr}\left(Y^+_s(A) \neq Y^-_s(A)\right)}{4|A|} + \sum_{w \in \partial B} \mathbf{Pr}\left(Y^+_s(w) \neq Y^-_s(w)\right) a_w \\
&\leq \frac{1}{2} \max_v \mathbf{Pr}\left(Y^+_s(v) \neq Y^-_s(v)\right)
\end{aligned}$$

Next lecture (Eric): Estimating the permanent

4th lecture (Daniel): Gadgets for inapproximability of counting
$=$ random bipartite d-regular graphs

Thank you