COUPLING TECHNIQUE

Eric Vigoda

Georgia Tech

EPFL Bernoulli Center, July ’18
1. **Definitions:** Ising Model, Coupling, etc.

2. **Path Coupling:**
   Rapid Mixing on $\mathbb{Z}^2$ for sufficiently high temperature.

3. **Strong Spatial Mixing (SSM) → $O(n \log^2 n)$ 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 $\Omega$ where $N = |\Omega|$.

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

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

Ergodic $\iff$ aperiodic and irreducible

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

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

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

Key Theorem [Doeblin ’38]: For a finite, ergodic Markov chain, there is a unique stationary distribution $\pi$ such that:

for all $x \in \Omega$, $P^t(x, \cdot) = \pi$. 
Consider graph $G = (V, E)$ as $n \times n$ box of $\mathbb{Z}^2$: 

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

For $\sigma \in \Omega$:

- **Monochromatic edges:** $M(\sigma) = |\{(v, w) \in E : \sigma(v) = \sigma(w)\}|$. 
- **Sampling:** Gibbs dist. $\mu(\sigma) = \exp(\beta M(\sigma)) / Z$. 
- **Counting:** Partition function $Z = Z_G = \sum_{\eta \in \Omega} \exp(\beta M(\sigma))$. 

Ferromagnetic (attractive) since $\beta > 0 \rightarrow$ neighbors prefer to align.
Consider graph $G = (V, E)$ as $n \times n$ box of $\mathbb{Z}^2$:

Configurations: $\Omega = \{-1, +1\}^V$. 
Consider graph $G = (V, E)$ as $n \times n$ box of $\mathbb{Z}^2$:

![Graph Image]

**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.
Consider graph $G = (V, E)$ as an $n \times n$ 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: $Z = Z_G = \sum_{\eta \in \Omega} \exp(\beta M(\eta))$.

Ferromagnetic (attractive) since $\beta > 0 \rightarrow$ neighbors prefer to align.
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.: $\mu(\sigma) = \frac{\exp(\beta M(\sigma))}{Z}$

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

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

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

How fast does it converge to $\pi$?
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 $\pi$ is Gibbs distribution $\mu$.

How fast does it converge to $\pi$?

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

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

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

Sub-multiplicative: $T_{\text{mix}}(\epsilon) = \lceil \log_2(\epsilon) \rceil T_{\text{mix}}$
What is a coupling?

For distributions $\mu$ and $\nu$ on $\Omega$, coupling is a joint distribution $\omega$ on $\Omega \times \Omega$ where:

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

Coupling Lemma: For any coupling $\omega$ of $\mu$ and $\nu$:

Choose $(X, Y) \sim \omega$, then $d_{TV}(\mu, \nu) \leq \Pr(X \neq Y)$. 
What is a coupling?

For distributions $\mu$ and $\nu$ on $\Omega$,

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

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

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

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

Sample $(X, Y)$ from $\omega$ in two equivalent ways:

1. Sample $X$ from $\mu$, then apply $\omega$ to get $Y$.
2. Sample $Y$ from $\nu$, then apply $\omega$ to get $X$. 

Coupling Lemma: For any coupling $\omega$ of $\mu$ and $\nu$:

choose $(X, Y) \sim \omega$, then \[d_{\text{TV}}(\mu, \nu) \leq \Pr(X \neq Y)\]
What is a coupling?

For distributions $\mu$ and $\nu$ on $\Omega$, coupling is a joint distribution $\omega$ on $\Omega \times \Omega$ where:

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

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

**Coupling Lemma:** For any coupling $\omega$ of $\mu$ and $\nu$:

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

What is a coupling?
For distributions $\mu$ and $\nu$ on $\Omega$, a coupling is a joint distribution $\omega$ on $\Omega \times \Omega$ where:

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

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

Coupling Lemma: For any coupling $\omega$ of $\mu$ and $\nu$:
choose $(X, Y) \sim \omega$, then $d_{TV}(\mu, \nu) \leq \Pr(X \neq Y)$.

Proof: \[
\Pr(X \neq Y) \geq 1 - \sum_{z \in \Omega} \min\{\mu(z), \nu(z)\}
\]

\[
= \sum_{z \in \Omega} \mu(z) - \min\{\mu(z), \nu(z)\}
\]

\[
= \sum_{z \in \Omega: \mu(z) \geq \nu(z)} \mu(z) - \min\{\mu(z), \nu(z)\}
\]

\[
= \max_{S \subseteq \Omega} \mu(S) - \nu(S) = d_{TV}(\mu, \nu)
\]
What is a coupling?

For distributions $\mu$ and $\nu$ on $\Omega$, coupling is a joint distribution $\omega$ on $\Omega \times \Omega$ where:

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

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

**Coupling Lemma:** For any coupling $\omega$ of $\mu$ and $\nu$:

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

Always exists optimal coupling:

There exists coupling $\omega$ of $\mu$ and $\nu$ where:

\[ d_{TV}(\mu, \nu) = \Pr(X \neq Y). \]

**Proof:** Constructive.
Consider a Markov chain \((\Omega, P)\).

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 \(a, b, c \in \Omega\),

\[
\Pr(X_{t+1} = c \mid X_t = a, Y_t = b) = P(a, c)
\]

\[
\Pr(X_{t+1} = c \mid X_t = a, Y_t = b) = P(b, c)
\]
Consider a Markov chain \((\Omega, P)\).

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 \(a, b, c \in \Omega\),

\[
\Pr(X_{t+1} = c \mid X_t = a, Y_t = b) = P(a, c)
\]

\[
\Pr(X_{t+1} = c \mid X_t = a, Y_t = b) = P(b, c)
\]

Intuition:

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

Let \(X_0\) be arbitrary, and \(Y_0 \sim \pi\). Once \(X_T = Y_T\) then \(X_T \sim \pi\).
Consider a Markov chain \((\Omega, P)\).

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

\[
X_t \sim P \quad \text{and} \quad Y_t \sim P
\]

More precisely, for all \(a, b, c \in \Omega\),

\[
\Pr(X_{t+1} = c \mid X_t = a, Y_t = b) = P(a, c) \\
\Pr(X_{t+1} = c \mid X_t = a, Y_t = b) = P(b, c)
\]

**Intuition:**

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

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

**Coupling time:**

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

\[
T_{\text{mix}} \leq T_{\text{couple}}
\]
Consider Ising with $\beta = 0$: $w(\sigma) = 1$ for all $\sigma \in \{-1, +1\}^V$.

Note, vertices independent of each other.
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.
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 $s_t$. 
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, $\mathbb{E}[|D_{t+1}| \mid X_t, Y_t] \leq |D_t|(1 - 1/n)$. 
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, $\mathbb{E}[|D_{t+1}| \mid X_t, Y_t] \leq |D_t|(1 - 1/n)$.

$$\Pr(X_T \neq Y_T) \leq \mathbb{E}[|D_t|]$$ since if $X_T \neq Y_T$ then $|D_T| \geq 1$
$$\leq |D_0|(1 - 1/n)^T$$
$$\leq n \exp(-T/n)$$
$$\leq 1/4 \quad \text{for } T \geq n \ln(4n).$$
Consider $\beta > 0$:
Consider $\beta > 0$:

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

- 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. $\propto \exp(\beta d^-)$. 

**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.
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 $d^+ = \# \text{ of } +1 \text{ neighbors of } v_t$, $d^- = \# \text{ of } -1 \text{ neighbors}$. 
- $X_{t+1}(v) = \begin{cases} +1 & \text{if } r_t \leq \frac{\exp(\beta d^+)}{\exp(\beta d^+)+\exp(\beta d^-)} \\ -1 & \text{otherwise} \end{cases}$

**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.
Consider a pair of Ising configurations $X_t$ and $Y_t$:

Identity Coupling: Update same $v^t$, same seed $r^t$ (i.e., same spin as much as possible).

How to analyze???
Consider a pair of Ising configurations $X_t$ and $Y_t$:

Look at $\frac{X_t}{Y_t}$:
Consider a pair of Ising configurations $X_t$ and $Y_t$:

Identity Coupling:
- Update same $v_t$,
- same seed $r_t$ (i.e., same spin as much as possible).
Consider a pair of Ising configurations $X_t$ and $Y_t$:

Identity Coupling:
- Update same $v_t$,
- same seed $r_t$ (i.e., same spin as much as possible).

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

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

If for all \( X_t, Y_t \),

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

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

*Path coupling:* Suffices to consider pairs where \( H(X_t, Y_t) = 1 \).
Let $S \subset \Omega^2$ denote pairs $(X_t, Y_t)$ where $H(X_t, Y_t) = 1$.
Define a coupling $\omega$ for all $(X_t, Y_t) \in S$ where:

$$\mathbb{E}[H(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \leq 1 - C/n.$$
Let $S \subset \Omega^2$ denote pairs $(X_t, Y_t)$ where $H(X_t, Y_t) = 1$. Define a coupling $\omega$ for all $(X_t, Y_t) \in S$ where:

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

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

In graph $(\Omega, S)$, consider a shortest path $A_t$ to $B_t$:

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

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

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

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

1. In graph $(\Omega, S)$, consider a shortest path $A_t$ to $B_t$:

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

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

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

Algorithmic View:

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

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 $\omega$ for all $(X_t, Y_t) \in S$ where:

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

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

In graph $(\Omega, S)$, consider a shortest path $A_t$ to $B_t$:

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

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 $(A_t, B_t)$.

$$\mathbb{E}[H(A_{t+1}, B_{t+1})] \leq \mathbb{E}\left[ \sum_i H(W_{t+1}^{i-1}, W_{t+1}^i) \right] \leq \sum_i \mathbb{E}\left[ H(W_{t+1}^{i-1}, W_{t+1}^i) \right] \leq \sum_i (1 - C/n) \leq H(A_t, B_t)(1 - C/n).$$
Let $S \subseteq \Omega^2$ denote pairs $(X_t, Y_t)$ where $H(X_t, Y_t) = 1$. Define a coupling $\omega$ for all $(X_t, Y_t) \in S$ where:

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

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

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

$$\text{Pr}(A_T \neq B_T) \leq \mathbb{E}[H(A_T, B_T)]$$

$$\leq H(A_0, B_0)(1 - C/n)^T$$

$$\leq n \exp(-C/n)$$

$$\leq 1/4 \quad \text{for } T = O(n \log n).$$

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

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

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

Worst case \(d^+ = d^-\). When \(d = 4\) works for \(\beta < \ln(5/3)\).
Consider a pair \((X_t, Y_t)\) that differ at exactly one vertex \(v^*\):

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

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

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

Goal: All \(\beta < \beta_c := \ln(1 + \sqrt{2})\).
Phase Transition

\[ \Lambda_n = \sqrt{n} \times \sqrt{n} \text{ box of } \mathbb{Z}^2. \]

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

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

To Do:
- What’s the phase transition mean?
- SSM = Strong Spatial Mixing
- Monotonicity of the Ising Model
- SSM \( \rightarrow \) \( T_{\text{mix}} = O(n \log n) \) on \( \Lambda_n \).
  (we’ll do \( O(n \log^2 n) \).)

Let $p^+ + n = \Pr(σ(v^*) = +1 | ∂Λ = +)$ and $p^- - n$ for $−1$ boundary.

$$\lim_{n \to \infty} p^+ + n \approx \lim_{n \to \infty} p^- - n$$

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

$$\lim_{n \to \infty} p_n^+ \overset{?}{=} \lim_{n \to \infty} p_n^-$$
Let $p_n^+ = \Pr(\sigma(v^*) = +1 \mid \partial \Lambda_n = +)$ and $p_n^-$ for $-1$ boundary.

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

Let $\beta_c = \ln(1 + \sqrt{2})$. [Onsager '44]:
- For all $\beta \leq \beta_c$, $p_\infty^+ = p_\infty^-$
- For all $\beta > \beta_c$, $p_\infty^+ \neq p_\infty^-$
Let $p_n^+ = \Pr(\sigma(v^*) = +1 \mid \partial \Lambda_n = +)$ and $p_n^-$ for $-1$ boundary.

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

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
Let \( p_n^+ = \Pr(\sigma(v^*) = +1 \mid \partial \Lambda_n = +) \) and \( p_n^- \) for \(-1\) boundary.

\[
\lim_{n \to \infty} p_n^+ \overset{?}{=} \lim_{n \to \infty} p_n^-
\]

Let \( \beta_c = \ln(1 + \sqrt{2}) \). [Onsager '44]:
- For all \( \beta \leq \beta_c \), \( p_\infty^+ = p_\infty^- \) \( \text{uniqueness} \)
- For all \( \beta > \beta_c \), \( p_\infty^+ \neq p_\infty^- \) \( \text{non-uniqueness} \)

**q-state Potts model**, [Béffara, Duminil-Copin '12]: \( \beta_c(q) = \ln(1 + \sqrt{q}) \).
- Continuous for \( q \leq 4 \) \[Duminil-Copin, Sidoravicius, Tassion '16\]
- Discont. for \( q > 4 \) \[Duminil-Copin, Gagnebin, Harel, Manolescu, Tassion '18\]
Let $p^+_n = \Pr(\sigma(v^*) = +1 \mid \partial \Lambda_n = +)$ and $p^-_n$ for $-1$ boundary.

$$\lim_{n \to \infty} p^+_n \neq \lim_{n \to \infty} p^-_n$$

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

$T_{mix}$ 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]
For a box $\Lambda_n$ and $v \in V$, let $p(v) = Pr(v = +)$.

**Weak Spatial Mixing (WSM):**
Exists $C, \alpha > 0$, for all boxes $\Lambda_n$, all $v \in V$, all boundaries $\sigma, \eta$:

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

For $y \in \partial \Lambda_n$ and boundary $\sigma$, obtain $\sigma^y$ by “flipping” spin at $y$.

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

$$|p^\sigma(v) - p^{\sigma^y}(v)| \leq C \exp(-\alpha \text{dist}(v, y))$$
For a box $\Lambda_n$ and $v \in V$, let $p(v) = \Pr(v = +)$.

**Weak Spatial Mixing (WSM):**
Exist $C, \alpha > 0$, for all boxes $\Lambda_n$, all $v \in V$, all boundaries $\sigma, \eta$:

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

For $y \in \partial \Lambda_n$ and boundary $\sigma$, obtain $\sigma^y$ by “flipping” spin at $y$.

**Strong Spatial Mixing (SSM):**
Exist $C, \alpha > 0$, all $\Lambda_n$, all $v \in V$, all $y \in \partial \Lambda_n$, boundaries $\sigma, \sigma^y$:

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

In 2-dimensions, for all $\beta < \beta_c$: SSM holds.
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 $+1$. 
**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. \( \propto \exp(\beta d^-) \).
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 \(d^+ = \# \text{ of } +1 \text{ neighbors of } v_t\), \(d^- = \# \text{ of } -1 \text{ neighbors}\).
- Set \(X_{t+1}(v) = +1\) if \(r_t \leq \exp(\beta d^+)/\left(\exp(\beta d^+) + \exp(\beta d^-)\right)\).
  Else, set \(X_{t+1}(v_t) = -1\).
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 \(X_0^- = \text{all } -1\) and \(X_0^+ = \text{all } +1\).

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 \(T_{\text{mix}} \leq T_{\text{couple}} \leq T\).
Blanca’s boosting proof [Dyer,Sinclair,V,Weitz ’04]:

Fix the graph $\Lambda_n$ as $\sqrt{n} \times \sqrt{n}$ box of $\mathbb{Z}^2$ – arbitrary b.c.
Let $T^*(n)$ denote an upper bound on $T_{\text{mix}}$, e.g., $T^*(n) \leq 2^n$.
Consider $X_0^- = -1, X_0^+ = +1$.
Goal: Prove $\Pr (X_T^- \neq X_T^+) \leq 1/4$ 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 $v \in V$: $\Pr (X_T^-(v) \neq X_T^+(v)) \leq 1/4n$.
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 $\overline{B}$: $Y_t^- \leq X_t^- \leq X_t^+ \leq Y_t^+$
Blanca’s boosting proof [Dyer,Sinclair,V,Weitz ’04]:

Fix the graph \( \Lambda_n \) as \( \sqrt{n} \times \sqrt{n} \) box of \( \mathbb{Z}^2 \) – arbitrary b.c.. Let \( T^*(n) \) denote an upper bound on \( T_{\text{mix}} \), e.g., \( T^*(n) \leq 2^n \). Consider \( X_0^- = -1, X_0^+ = +1 \). Goal: Prove \( \Pr(X_T^- \neq X_T^+) \leq 1/4 \) 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 \( v \in V \): \( \Pr(X_T^-(v) \neq X_T^+(v)) \leq 1/4n \). 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 \( \bar{B} \): \( Y_t^- \leq X_t^- \leq X_t^+ \leq Y_t^+ \)

\( T \) updates on \( \Lambda_n \) then whp \( T \times (|B|/n) = T^*(|B|) \log(9n) \) on \( B \).
If \( T_{\text{mix}}(1/4) = T^* \) then \( T_{\text{mix}}(1/9n) = T^* \log(9n) \).
Blanca’s boosting proof [Dyer, Sinclair, V, Weitz '04]:

Fix the graph $\Lambda_n$ as $\sqrt{n} \times \sqrt{n}$ box of $\mathbb{Z}^2$ – arbitrary b.c.. Let $T^*(n)$ denote an upper bound on $T_{\text{mix}}$, e.g., $T^*(n) \leq 2^n$.

Consider $X_0^- = -1, X_0^+ = +1$.

Goal: Prove $\Pr(X_T^- \neq X_T^+) \leq 1/4$ 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 $v \in V$: $\Pr(X_T^-(v) \neq X_T^+(v)) \leq 1/4n$.

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 $\Lambda_n$ then whp $T \times (|B|/n) = T^*(|B|) \log(9n)$ on $B$.

If $T_{\text{mix}}(1/4) = T^*$ then $T_{\text{mix}}(1/9n) = T^* \log(9n)$.

$$\Pr(X_T^-(v) \neq X_T^+(v))$$

$$\leq \Pr(Y_T^-(v) \neq Y_T^+(v))$$

$$\leq \Pr(Y_T^+(v) = +) - \Pr(Y_T^-(v) = +)$$

$$\leq |\Pr(Y_T^+(v) = +) - \mu_B^+(v = +)| + |\mu_B^+(v = +) - \mu_B^-(v = +)|$$

$$+ |\mu_B^-(v = +) - \Pr(Y_T^-(v) = +)|$$

$$\leq 1/9n + \exp(-C' \log n) + 1/9n \quad \text{by induction + SSM}$$

$$\leq 1/4n$$
Blanca’s boosting proof [Dyer, Sinclair, V, Weitz ’04]:

\[ T_{\text{mix}} = O\left(\frac{n}{\log n}\right) \times T^*\left(C\log^2 n\right). \]

\[ 2^n \rightarrow 2^{O\left(\log^2 n\right)}O\left(\frac{n}{\log n}\right) \leq n^{O\left(\log n\right)} \]
Blanca’s boosting proof [Dyer,Sinclair,V,Weitz ’04]:

\[ T_{mix} = O\left( \frac{n}{\log n} \right) \times T^*(C \log^2 n) \].

\[
\begin{align*}
2^n & \rightarrow 2^{O(\log^2 n)} O\left( \frac{n}{\log n} \right) \leq n^{O(\log n)} \\
n^{O(\log n)} & \rightarrow 2^{O((\log \log n)^2)} O\left( \frac{n}{\log n} \right) \leq O(n^{1+\epsilon})
\end{align*}
\]
Blanca’s boosting proof [Dyer, Sinclair, V, Weitz ’04]:

\[ T_{mix} = O\left(\frac{n}{\log n}\right) \times T^*(C \log^2 n) \]

\[
\begin{align*}
2^n & \rightarrow 2^{O(\log^2 n)} O\left(\frac{n}{\log n}\right) \leq n^{O(\log n)} \\
n^{O(\log n)} & \rightarrow 2^{O((\log \log n)^2)} O\left(\frac{n}{\log n}\right) \leq O\left(n^{1+\epsilon}\right) \\
O(n^{1+\epsilon}) & \rightarrow O(n \log^2 n)
\end{align*}
\]
Blanca’s boosting proof [Dyer, Sinclair, V, Weitz '04]:

\[ T_{\text{mix}} = O\left(\frac{n}{\log n}\right) \times T^*(C \log^2 n) \].

\[
\begin{align*}
2^n & \rightarrow 2^{O\left(\log^2 n\right)} O\left(\frac{n}{\log n}\right) \leq n^{O\left(\log n\right)} \\
n^{O\left(\log n\right)} & \rightarrow 2^{O\left((\log \log n)^2\right)} O\left(\frac{n}{\log n}\right) \leq O\left(n^{1+\epsilon}\right) \\
O\left(n^{1+\epsilon}\right) & \rightarrow O\left(n \log^2 n\right)
\end{align*}
\]

Extra Boosting:
Can boost from \( O(n^{1.5-\epsilon}) \), for any \( \epsilon > 0 \), to \( O(n \log n) \).
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 $\beta < \beta_c(\mathbb{T}_d)$,
$$T_{\text{mix}} = O(n \log n).$$

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 $\beta < \beta_c(\mathbb{T}_d)$,

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

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 $\beta_c(T_d)$ denote the critical point for infinite $d$-regular tree $T_d$.

Note, $\beta_c(T_d) = 2(d - 1) \tanh(\beta)$.

[Mossel-Sly '13]: For all $G$ of max. degree $d$, all $\beta < \beta_c(T_d)$,

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

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
For a graph $G = (V, E)$, for $v \in V$, let $p(v) = \Pr(v = +)$.
For radius $R$, let $B = B_R(v)$ be the ball of radius $R$ around $v$.
For $y \in \partial B$ and boundary $\sigma$, obtain $\sigma^y$ 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 $p(v) = \Pr(v = +)$. For radius $R$, let $B = B_R(v)$ be the ball of radius $R$ around $v$. For $y \in \partial B$ and boundary $\sigma$, obtain $\sigma^y$ 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 $\exp(-\alpha) < 1/d$. 
Once again, $X_0^+ = +1$, $X_0^- = -1$.

Goal: for every $v \in V$, $\Pr(X_T^+(v) \neq X_T^-(v)) \leq 1/4n$.

Ball $B = B_R(v)$ where $R$ is radius from ASSM.

Let $T^*$ denote, $\max v$, worst boundary $\sigma$, mixing time on $B^\sigma_R(v)$.

Let $T' = T^* \times \log(8d|B|)$.

Induction, for all $s \geq 0$,

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

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

Fix $v$ and time $s$. Let’s prove:

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

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 $\Pr \left( Y_{s+T'}^+(v) \neq Y_{s+T'}^-(v) \right)$. 

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 $\sigma^- = X^-_s(A) = Y^-_s(A)$.

$$
\Pr \left( Y^+_{s+T'}(v) \neq Y^-_{s+T'}(v) \mid \sigma^+, \sigma^- \right)
\leq |\Pr \left( Y^+_{s+T'}(v) = + \mid Y^+_s(A) = \sigma^+ \right) - \mu^+_{\sigma^+}(v = +)|
+ |\mu^+_{\sigma^+}(v = +) - \mu^-_{\sigma^-}(v = +)|$
+ $|\Pr \left( Y^+_{s+T'}(v) = + \mid Y^-_s(A) = \sigma^- \right) - \mu^-_{\sigma^-}(v = +)|$

Now bound each term on RHS.
Recall, $T' = T^* \times \log(8|A|)$.

Outer terms:

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

Middle term:

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

Combining we get:

$$\Pr \left( Y_{s+T'}^+ (v) \neq Y_{s+T'}^- (v) \mid \sigma^+, \sigma^- \right)$$

$$\leq \frac{1}{4|A|} + \sum_{w \in \partial B} 1(\sigma^+(w) \neq \sigma^-(w)) a_w.$$
Unconditioning...

Combining we got:

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

If \( Y_s^+(A) = Y_s^-(A) \) then \( Y_t^+(A) = Y_t^-(A) \) for \( t \geq s \).

\[
\Pr \left( Y_{s+T'}^+ (v) \neq Y_{s+T'}^- (v) \right) \\
\leq \sum_{\sigma^+, \sigma^-} \Pr \left( Y_{s+T'}^+ (v) \neq Y_{s+T'}^- (v) \mid \sigma^+, \sigma^- \right) \Pr \left( \sigma^+, \sigma^- \right) \\
\leq \frac{\Pr \left( Y_s^+ (A) \neq Y_s^- (A) \right)}{4|A|} + \sum_{w \in \partial B} \Pr \left( Y_s^+ (w) \neq Y_s^- (w) \right) a_w \\
\leq \frac{1}{2} \max_v \Pr \left( Y_s^+ (v) \neq Y_s^- (v) \right)
\]
Next lecture (Eric): Estimating the permanent

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

Thank you