**Review**

- **Designing Chains:** \( \mathbb{M} \in \mathbb{R}_{>0}^{2} \), \( N \in \mathbb{R}_{>0}^{2 \times 2} \)

\[ \Pi(x, y) = \mu(x) N(x, y), \quad N^0(y, x) = \Pi(x, y) \]

**Time-reversible Chain:** \( NN^0 \) with Stationary \( \mu \)

**Examples:** Glauber dynamics, block dynamics, delete/add or add/delete in STJ, hit and run, restricted Gaussian, Langevin dynamics

- **Transport/Wasserstein:**
  \[ W(\nu, \nu') = \min \left\{ \mathbb{E} d(X, Y) \mid X \sim \nu, \ X' \sim \nu' \right\} \]

- **Contraction of \( W \Rightarrow \) mixing time bound**

**Strategy:** Couple two \( X_0 \rightarrow X_1 \), for all \( X_0' \rightarrow X_1' \) of (random)

\[ \mathbb{E} [d(X_1, X_1')] < (1 - c) d(X_0, X_0') \]

---

- **Coloring:** If \( q \geq 4A + 1 \) we get

\[ W(v, v') \leq (1 - \frac{q - 4A}{q^2}) W(v, v') \]

- **Path Coupling:** If \( d \) is a shortest-path metric, enough to look at adjacent starts.

\[ x = x_0 \sim x_1 \sim x_2 \sim \ldots \sim x_0 = x' \]

\[ W(1, P, 1', P) \leq \sum W(x_i, x_i') = (1 - c) d(x, x') \]

---

**Plan for Today**

- Coloring with \( q \geq 2A + 1 \)
- Dobrushin's Conditions
  * Hardcore model
  * Ising model
- Intro to Spectral Analysis
- Fourier analysis
Path Coupling for Colorings \[\text{[Jerrum]}\]

Take adjacent colorings \(X_0, X'_0\) even improper
\(X_0(w) = a, \quad X'_0(w) = b \quad (a \neq b)\)

Coupling:
- Pick same vertex \(v\).
- If \(v = w\), or \(v \not= w\), pick same \(c\).
- If \(v \not= w\):
  \(C: 1, 2, \ldots, a, \ldots, b, \ldots, 9\)
  \(C': 1, 2, \ldots, a, \ldots, b, \ldots, 9\)

\[\mathbb{E}[d(X_1, X'_1)] \leq 1 - \frac{1}{n} \cdot \frac{q - \Delta}{q} + \frac{\Delta}{n} \cdot \frac{1}{q}\]

As long as \(q \geq 2\Delta + 1 \Rightarrow t_{mix} = O\left(\frac{nq}{q - \Delta} \cdot \lg n\right)\)

Metropolis mixes fast.
What about Glauber dynamics?
- Use comparison \(\rightarrow\) future lectures
- Use Dobrushin's conditions

Product Spaces

Suppose \(\mu\) is on \(\Omega = \Omega_1 \times \ldots \times \Omega_n\).

Glauber dynamic:
\(x = (x(1), \ldots, x(n)) \Rightarrow x' = (x(1), \ldots, z, \ldots, x(n))\)

\(z\) is sampled from \(i\)-th marginal of \(\mu\)
Conditioned on \(x(i) \not= x(i-1), x(i+1), \ldots, x(n)\).

Def: Call this \(\mu_i(. \mid X(i-1))\).
Goal: Show contraction of \(W\) based on
Hamming dist: \(d(x, x') = \#i : x(i) \not= x'(i)\).
**Dobrushin's Influence Matrix**

Adjacent $X,X'$: unique $j$ where $X(j)\neq X'(j)$. We write $X \sim_j X'$.

**Influence**: How much do marginals differ in Glauber when resampling coord $i$. 

$I[j\rightarrow i] = \max \left\{ \frac{d_{TV}(\mu_i(\cdot|X(-i)),\mu'_i(\cdot|X(-i)))}{d_{TV}(\mu_i(\cdot|X(-i)),\mu'_i(\cdot|X(-i)))} \mid X \sim_j X' \right\}$.

Note: The value of $X(i), X'(i)$ do not matter, so think of $X, X'$ as in $\Omega, \Omega \rightarrow \Omega, \Omega \rightarrow \Omega^q$.

Note: We have $I[j\rightarrow i] = 0$.

**Influence Matrix**: 

$$
\begin{bmatrix}
I[j\rightarrow i]
\end{bmatrix}
\text{row } i \rightarrow \text{col } j
$$

**Informal Thm**: If $I$ is "small" $\Rightarrow$ fast mixing.

Example. $\mu$ is uniform over $\{0,1\}^n$.

$\begin{align*}
I[j\rightarrow i] &= 0 \Rightarrow I = 0
\end{align*}$

Example. Coloring with a palette of $q$ colors.

$\mu_i(\cdot|X(i)) := \text{uniform over } \Omega_i \rightarrow \Omega^q - \{X(v) \mid v \text{ neighbor of } i\}.$

What is $I[j\rightarrow i]$:

$$
\begin{cases}
* j \neq i : 0 \\
* j = i : \frac{1}{q^2-A} \quad I \leq \frac{1}{q^2-A} A
\end{cases}
$$

Worst-case happens when neighbors of $i$ have diff colors in both $X, X'$. 

**Note**: We have $I[j\rightarrow i] = 0$.
**Thm:** If columns of $I$ sum to $\leq 1 - \frac{\delta}{n}$

\[ \Rightarrow W(yP, y'P) \leq (1 - \frac{\delta}{n}) W(y, y') \]

Glauber

**Proof:** Use path coupling: $X \sim X'$

\[ X_o \rightarrow X_1 \quad \text{? Couple} \]

- Pick same coord $i$
- Maximally couple replacement
  - the one defining $d_{TV}$

\[ d(X_o, X_o') = 1 \]

\[ \mathbb{E}[d(X_i, X_i')] \leq \frac{1}{n} x_0 + \sum_{i \neq j} \frac{1}{n} x (1 + I[j \rightarrow i]) \]

Pick $i = j$

\[ \leq \frac{n-1}{n} + \frac{\sum I[j \rightarrow i]}{n} \leq 1 - \frac{\delta}{n} . \]

**Corollary:** For coloring $I \leq \frac{1}{q-\Delta} A$, so column sums $\leq \frac{\Delta}{q-\Delta} = 1 - \frac{q-2\Delta}{q-\Delta}$.

For $q \geq 2\Delta + 1$ we have $1 - \frac{q-2\Delta}{n(q-\Delta)}$

contraction of $W \Rightarrow$

\[ t_{mix} = O \left( \frac{q-\Delta}{q-2\Delta} \cdot n \log n \right) \]

Dobrushin matrix is very useful for spin systems / graphical models / etc.

\[ \mu(x(i_1), x(n)) \propto \phi_1(x(i_1), x(2)), \phi_2(x(3), \ldots) \]

If $i$ and $j$ conditioned on everything else independent $\Rightarrow I[j \rightarrow i] = 0$.

↑ this happens when $ij$ do not appear in the same factor
Example. (Coloring)

Define \( \phi : [q]^2 \to [0,1]^n \) as \( \phi(a, b) = \mathbb{1}[a \neq b] \)

\[ \max \prod_{u \sim v} \phi(x(u), x(v)) \]

one factor per edge

Example. (Hardcore model)

\( \mu : \) independent sets \( S \) of graph \( \mu \propto \lambda |S| \)

\[ \prod_{i \sim j} f(x(i)) \cdot \prod_{i \sim j} g(x(i), x(j)) \]

Max independent sets \( \max \prod_{u \sim v} \phi(x(u), x(v)) \)

Large \( \lambda \) is hard \( \leftarrow \) \( \max \) indep. sets

Small \( \lambda \) is easy:

\[ \Gamma[x \to i] = \sum \nu y : 0 \]

\[ \nu y : \lambda \frac{1}{1 + \lambda} \]

Worst case: neighbors unoccupied

\[ I \leq \frac{\lambda}{1 + \lambda} \cdot \lambda \]

So column sums are

\[ \lambda \Delta \leq \lambda \Delta \]

\[ \lambda \leq (1 - s) / \Delta \Rightarrow t_{\text{mix}} = O\left( \frac{n \log n}{\Delta} \right) \]

Remark: This is not the correct threshold!

\[ \lambda \leq (1 - s) \lambda_c (\Delta) \Rightarrow \text{fast mixing} \]

\[ \lambda \geq (1 - s) \lambda_c (\Delta) \Rightarrow \text{NP-hard!} \]
Example. (Ising model)

\[ \mu \text{ on } \{\pm 1\}^n \text{ with} \]

\[ \mu(x(n),-x(n)) \propto \exp \left( \sum_{ij} \beta_{ij} x(i)x(j) + \sum_i h_i x(i) \right) \]

If \( l_1 \text{ norm of } B \text{ matrix } \leq 1 - 8 \mu \text{ on } 5^{-1} \)

\[ \implies t_{\text{mix}} = O(\frac{n}{8 \ln n}) \]

Corollary. If \( \beta_{ij} \) are supported on \( \Delta \text{-max-deg} \)

graph and the same \( \beta \):

\[ |\beta| \leq \frac{1-\delta}{\Delta} \implies \text{fast mixing} \]

this is up to lower order terms

the correct threshold

Remark: Any spin system on \( \{\pm 1\}^n \)
with “soft” binary factors is an

Ising model.

Hardcore model is a limit
If vector $c \in \mathbb{R}^n_+$ has $cI \leq (1-\delta)c$
so far $\Rightarrow c$-Hamming has contraction

$$d(x, x') := \sum_i c_i 1[x(i) \neq x'(i)].$$

**Proof:** $x_0,x'_0$: Couple the same way as before.

$$|E[d(x, x')]| = \frac{1}{n} x_0 + \sum_{i \neq j} \frac{1}{n} x(c_j + c_i I_{ij}^{-1})$$

$$= \frac{n-1}{n} c_j + \frac{1}{n} (cI)_j \leq (1 - \frac{\delta}{n}) c_j = d(x_0, x'_0).$$

**Remark:** Every $\geq 0$ matrix has $\geq 0$ eigenvector.

$cI = \lambda_{\max} c$. This scheme "works" if $\lambda_{\max} \leq 1 - \delta$. [We need $\lambda_{\min}/\lambda_{\max}$ to be not so small]

**Intro to Spectral Analysis**

Perron-Frobenius [for ergodic chains]:

$P$ has a unique eigenvalue $\lambda$ and all others are $<1$ in magnitude.

**Eigenvalues of $P$** $\Rightarrow$ **Bounds on $t_{\text{mix}}$**

**Proof:**

$\sum \lambda_i = 1$ $\Rightarrow$ $Pt = \sum \lambda_i t\nu_i$

If $<1$ converges too.

$\nu \approx \sum \lambda_i t\nu_i$
**Lazification**

\[ AP A^{-1} = P \]

\[ P \Rightarrow \lambda P + (1-\lambda)I \]

**Intuition:** It only matters that \( \lambda_i \) are far from 1. If \( |\lambda_i| \approx 1 \) but far from 1, its norm shrinks in lazification.

**Time-reversible Chains**

When \( P \) is time-reversible w.r.t. \( \mu \) we have:

\[
\text{diag} \left( \mu \right) P = Q \text{ symmetric matrix}
\]

\[
\text{diag} \left( \mu \right)^{\frac{1}{2}} P \text{diag} \left( \mu \right)^{-\frac{1}{2}} = \text{diag} \left( \mu \right)^{-\frac{1}{2}} Q \text{diag} \left( \mu \right)^{-\frac{1}{2}}
\]

**Linear Algebra Fact:** Symmetric matrices have real eigenvalues, are diagonalizable, and have orthogonal eigenbasis.

**Corollary.** Time-reversible \( \Rightarrow \) real eigenvalues.

Can order: \( 1=\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda_n > -1 \)

**Def (spectral gap):** \( 1 - \lambda_2 \) or \( 1 - \max(\lambda_2, |\lambda_n|) \)

**Def (relaxation time):** \( \frac{1}{(1-\lambda_2)} \)

Intuitively how fast \( \lambda_2 \) decreases.

We will see later that in fact \( \lambda_2 \) dictates contraction of some quantity called \( \chi^2 \)-divergence.