Swendsen-Wang (SW) algorithm
(for Ising/Potts model)

Last lecture

- $G = (V,E)$
- Configurations $\sigma : V \to \{-1, 1\}$
- State space $\Omega = \{-1, 1\}^{|V|}$

Ising model:

$$\Pi(\sigma) = \frac{1}{Z_I} e^\beta \sum_{\sigma} e^{\beta \sum_{uv} \sigma_u \sigma_v}$$
$$Z_I = \sum_{\sigma} e^{\beta \sum_{uv} \sigma_u \sigma_v}$$

- $\beta > 0$ [ferromagnetic, favors equal spins btw. neighbors]
- $\beta < 0$ [antiferromagnetic]

Phase transition:

- $\beta < \beta_c$ (disorder) (subcritical) configurations have roughly the same # of "+1" and "-1".
- $\beta > \beta_c$ (long range order) (supercritical) configurations are either mostly "+1" or mostly "-1" w.h.p.
In the subcritical regime, it is easy to sample:

\[ \beta < \beta_c \]: \[ T_{\text{mix}} \left( \text{Glauber dynamics} \right) = \Theta(n \log n) \] [Martinelli, Olivieri, Stochsman '94]

\[ \beta > \beta_c \]: \[ T_{\text{mix}} \left( \text{Glauber dynamics} \right) = \exp \left( \Omega \left( \sqrt{n} \right) \right) \]

Why?

Let us look (for example) to the magnetization:

\[ \mu \sim \sum_{i \in V} \tau_i \]

\[ \beta < \beta_c \]

\[ \beta > \beta_c \]

\[ \Rightarrow \]

\[ \text{Essentially} \quad \Rightarrow \quad \text{mostly} \]

Question: How to sample from \( \Pi \) when \( \beta > \beta_c \)?

- We'll work in the more general framework of the Potts model.

- Spins: \( \xi_1, \ldots, \xi_V \)
- \( D_p = \{ \xi_1, \ldots, \xi_V \} \)

\[ \Pi_p (\xi) = \frac{1}{2^p} \cdot e^{\beta \cdot a(\xi)} \]
The $q=2$ case is just a renormalization / reparametrization of the Ising model as defined above.

\[
\Pi(\sigma) = \frac{1}{Z^2} \cdot \exp \left( \beta \sum_{uv} J_{uv} \sigma_u \sigma_v \right)
\]

\[
= \frac{1}{Z^2} \cdot \exp \left( \beta \left( a(\sigma) - d(\sigma) \right) \right)
\]

\[
= \frac{1}{Z^2} \cdot \exp \left( 2 \beta c(\sigma) - \beta |E| \right)
\]

\[
= \frac{1}{Z^2 e^{\beta |E|}} \cdot \exp \left( 2 \beta a(\sigma) \right)
\]

\[
\rightarrow \frac{2^\beta}{Z e^{\beta |E|}} \rightarrow Z^0 \quad q=2
\]

**Phase transition**

$\beta < \beta_c(q) \implies \frac{1}{q}$ vertices of each color w.h.p.

$\beta > \beta_c(q)$ \quad one dominant color class

$\beta_c(q) = \ln(1 + \sqrt{q}) \quad [\text{DC-B '10}]

*Glauber dynamics have the same mixing behavior:

$\beta < \beta_c(q) \quad \text{mix} = \Theta(n \log n)$

$\beta > \beta_c(q) \quad \text{mix} = \exp(\sqrt{n})$
SW algorithm [87]

° Markov chain on `\(\Pi_P\)` configurations that converge to `\(\Pi_P\)`.

Given `\(\Gamma_0 \in \Pi_P\)`, `\(\Gamma_{t+1}\)` is obtained as follows:

1. Add each monochromatic edge independently with probability `\(p = 1 - e^{-\beta}\)` to obtain `\((\Gamma_t, A, EE)\)`.

2. Forget the colors. (We just have `\(A_t\)`).

3. Assign a new color uniformly at random to each connected component of `\(A_t\)` to obtain `\((\Gamma_{t+1}, A)\)`.

4. Forget `\(A_t\)` to obtain `\(\Gamma_{t+1}\)`.

\[\text{Thm:}\]
\[
\beta < \beta_c (q) \quad \text{then} \quad T_{\text{mix}} (\text{SW}) = O(n) \quad (17)
\]
\[
\beta > \beta_c (q) \quad \text{then} \quad T_{\text{mix}} (\text{SW}) = O(n^3) \quad (11) \quad [\text{Ullrich}]
\]

\[\Rightarrow \text{Why is fast when Glauber dynamics is slow?} \]

\[\text{Intuition:} \quad \text{It is is for the algorithm to}\]
\[\text{Jump from a mostly red configuration}\]
\[\text{to a mostly green one, etc.}\]
But, why...? Also, why is it correct?

* After step 2, only edge configuration.

**Random-cluster model:**

\[ M(A) = \left( \frac{p}{1-p} \right)^{1 \lambda_1} q^{C(A)} \cdot A \in E \text{ configuration} \]

\[ r = \frac{1}{Z_{rc}} \]

**Phase transition:**

\[ P < P_c(q) = \frac{\sqrt{q}}{1-\sqrt{q}} = 1 - e^{-A_c(q)} \text{ [call components small]} \]

\[ P > P_c(q) \text{ [exactly one 'giant' component with most vertices]} \]

**Claim:** If \( \Sigma \sim \Pi_p \), and we apply step 4 and 2 from the SW algorithm, the resulting config has distribution \( M \) and moreover \( Z_{rc} = Z_p \).

**Proof:**

\[ Pr[A \in E] = \sum_{A \in E} \Pi_p(A) \cdot Pr[\Sigma \rightarrow A] \]

\[ = \sum_{A \in A_{\Sigma \rightarrow A}} \frac{1}{Z_p} e^{\beta \cdot a(A)} \cdot p^{\lambda_1} (1-p)^{(\lambda_1)A_{\Sigma \rightarrow A}} \]

\[ = \frac{1}{Z_p} \cdot \sum_{A \in A_{\Sigma \rightarrow A}} \frac{1}{(1-p)^{a(A)}} \cdot p^{\lambda_1} \frac{(1-p)^{\lambda_1}}{(1-p)^{1A}} \]

\[ = \frac{1}{Z_p} \cdot \sum_{A \in A_{\Sigma \rightarrow A}} \frac{1}{(1-p)^{a(A)}} \cdot p^{\lambda_1} (1-p)^{\lambda_1} \]
\[ P_r[A \in E] = \frac{1}{Z_p} \cdot \sum_{\sigma \in \text{ac}(\pi)} (\frac{p}{1-p})^{1A} \]

\[ = \frac{1}{Z_p} \left( \frac{p}{1-p} \right)^{1A} \cdot q^{C(\pi)} \]

\[ \sum_A P_r[A] = 1 = \frac{1}{Z_p} \sum_A (\frac{p}{1-p})^{1A} \cdot q^{C(\pi)} = \frac{Z_{rc}}{Z_p} \]

\[ (Z_{rc}=Z_p), \text{ So} \]

\[ P_r[A] = M(A) \]

---

**Claim 2**

If \( A \rightarrow M \), after steps 3 and 4 of the dynamics, the resulting configuration have distribution \( T_\pi \).

**Proof**:

\[ P_r[\sigma \in \pi \rho_p] = \sum_{A \in \pi \rho_c} M(A) \cdot P_r[A \rightarrow \sigma] \]

\[ = \sum_{A \in \pi \rho_c} \frac{1}{Z_{rc}} \left( \frac{p}{1-p} \right)^{1A} \cdot q^{C(\pi)} \frac{1}{q^{C(\pi)}} \]

\[ P_r[\sigma \in \pi \rho_p] = \frac{1}{Z_p} \sum_{A \in \text{ac}(\sigma)} (\frac{p}{1-p})^{1A} \]

\[ = \frac{1}{Z_p} \sum_{A \in \text{ac}(\sigma)} \sum_{k=1}^{a(\pi)} \binom{a(\pi)}{k} \left( \frac{p}{1-p} \right)^{1A} \sum_{|\lambda|=k} \binom{a(\pi)}{k} \left( \frac{p}{1-p} \right)^{1A} = \frac{1}{Z_p} \sum_{k=1}^{a(\pi)} \binom{a(\pi)}{k} \left( \frac{p}{1-p} \right)^{1A} \]
\[ \Pr[A] = \frac{1}{2p} \left( \frac{P}{1-p} + 1 \right) a(\sigma) = \frac{1}{2p} \left( \frac{1}{1-p} \right) a(\sigma) = \frac{1}{2p} e^{-\lambda p} \]  

**Correctness of SW:**

1. Irreducible.
2. Aperiodic.
3. \( \Pi_p \) is stationary. (by Claims 1 and 2)

*The approach we used is due to Edwards-Sokal '84.*

---

**What happens when \( \beta = \beta_c(q) \)?**

- \( q = 2, 3 \) \( \Gamma_{\text{mix}} = O(n^2) \) for some constant \( \geq 0 \)
- \( q = 4 \) \( \Gamma_{\text{mix}} = O(n \log n) \) \([\text{we expect } \Gamma_{\text{mix}} = O(n^2)]\)
- \( q > 5 \) \( \Gamma_{\text{mix}} = \exp(D(\sqrt{n})) \)

**Why change of behavior at \( q = q_c \)?**

- Related to "continuity" of phase transition.

\[ 1 < q < q_c \] continuous ("smooth") phase transition.

\[ q > q_c \] discontinuous ("sharp")
Ok. $\mathbb{Z}^d$ we got. What about $\mathbb{Z}^d$, $d \geq 3$.

(We really don't know).

Other graphs

Mean-field case: $G = K_n$ complete graph on $n$ vertices

$q = 2$ (Ising model) \[\beta\]

$\beta < \beta_c$ \[T_{\text{mix}} = \Theta(1)\]

$\beta = \beta_c$ \[T_{\text{mix}} = \Theta(n^{\frac{1}{4}})\]

$\beta > \beta_c$ \[T_{\text{mix}} = \Theta(\log n)\]

$\beta = \beta_c$ \[T_{\text{mix}} = \Theta(n^{\frac{1}{4}})\]

$q > 3$ (Potts)

$\beta < \beta_c(q)$ \[T_{\text{mix}} = \Theta(1)\]

$\beta > \beta_c(q)$ \[T_{\text{mix}} = \Theta(\log n)\]

$\beta = \beta_c(q)$ \[T_{\text{mix}} = \Theta(n^{\frac{1}{4}})\]

$\beta = \beta_c$ \[T_{\text{mix}} = \Theta(\log n)\]

$\beta = \beta_c(q)$ \[T_{\text{mix}} = \exp(\Omega(n))\] $\beta_c \in (\beta_L, \beta_R)$

$\beta = \beta_L$ \[T_{\text{mix}} = \Theta(n^{\frac{1}{3}})\]

$\beta = \beta_R$ \[T_{\text{mix}} = \Theta(\log n)\]

$\Rightarrow [2015]$
Theorem [Goodfellow Ferrus '16]

$$\Gamma_{\text{mix}}(P_{\text{swap}}) = O(n^6)$$ on any graph.