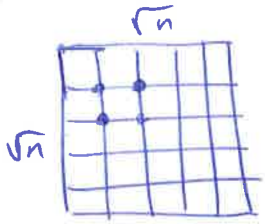
Swendsen-Wang (SW) algorithm

(for Ising/Potts model)

Last lecture:

- $G = (V, E)$



- Configurations $\sigma : V \to \{-1, +1\}$
- State space $\Omega := \{-1, 1\}^{|V|}$

Ising model:

$$\pi(\sigma) = \frac{1}{Z_I} e^{\beta \sum_{u \sim v} \sigma_u \sigma_v} \qquad Z_I = \sum_{\sigma \in \Omega} e^{\beta \sum_{u \sim v} \sigma_u \sigma_v}.$$

(*) $\beta > 0$ [ferromagnetic, favors equal spins btw. neighbors]

$\beta < 0$ [antiferromagnetic, " $\neq$ "]

Phase transition:

(*) $\beta < \beta_c$ (disorder) (subcritical) — configurations have roughly the same # of "+1" and "−1". (w.h.p.)

(*) $\beta > \beta_c$ (long range order) (supercritical) — configurations are either mostly "+1" or mostly "−1" w.h.p.

In the subcritical regime, it is easy to sample:

$(\beta < \beta_c)$: $T_{mix}(\text{Glauber dynamics}) = \Theta(n \log n)$ [Martinelli; Olivieri; Schlosman '94]

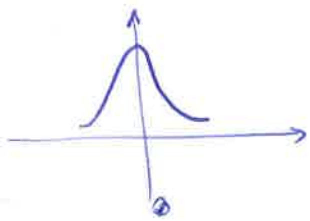
$(\beta > \beta_c)$: $T_{mix}(\text{Glauber dynamics}) = \exp(\Omega(\sqrt{n}))$.

Why?

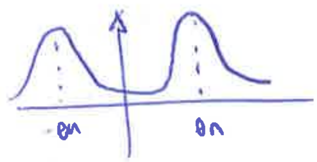
Let us look (for example) to the magnetization:

$$M \sim \sum_{v \in V} \sigma_v$$

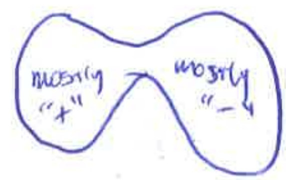
$\beta < \beta_c$



$\beta > \beta_c$



$\Rightarrow$



Question: How to sample from π when $\beta > \beta_c$?

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

  - Spins $:= \{1, \dots, q\}$ $\qquad \sigma: V \to \{1, \dots, q\}$
  - $\Omega_p := \{1, \dots, q\}^{|V|}$

$a(\sigma) = \#$ of agreements

$d(\sigma) = \#$ of disagreements

$$\pi_p(\sigma) = \frac{1}{Z_p} \cdot e^{\beta \cdot a(\sigma)}$$

The $q=2$, is just a renormalization / reparametrization of the Ising model as defined above,

$$\pi(\sigma) = \frac{1}{Z} \cdot \exp\left(\beta \sum_{u \sim v} \sigma_u \sigma_v\right)$$

$$= \frac{1}{Z_I} \cdot \exp\left(\beta\left(a(\sigma) - d(\sigma)\right)\right)$$

$$= \frac{1}{Z_I} \exp\left(2\beta a(\sigma) - \beta|E|\right)$$

$$= \frac{1}{Z_I e^{\beta|E|}} \cdot \exp\left(2\beta\, a(\sigma)\right) \quad \left[\begin{array}{l} 2\beta \to \beta \\ Z_I e^{\beta|E|} \to Z_P \quad q=2 \end{array}\right]$$

Phase transition

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

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

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

- Glauber dynamics have the same mixing behavior
  - $\beta < \beta_c(q)$ $\quad T_{mix} = \Theta(n \log n)$
  - $\beta > \beta_c(q)$ $\quad T_{mix} = \exp(\sqrt{n})$

SW algorithm ['87]

- Markov chain on ~~Ising~~ Potts configurations that converge to π_p.

Given $\sigma_t \in \Omega_p$, σ_{t+1} is obtained as follows:

1. Add each monochromatic edge independently with probability $p = 1 - e^{-\beta}$. to obtain $(\sigma_t, A_t \subseteq E)$.

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

3. Assing a new color uniformly at random to each connected component of A_t to obtain (σ_{t+1}, A_t)

4. Forget A_t to obtain σ_{t+1}.

Thm:

$\beta < \beta_c(q)$ $\quad t_{mix}(SW) = O(n)$ ('17)

$\beta > \beta_c(q)$ $\quad t_{mix}(SW) = \tilde{O}(n^3)$ ('11) [Ullrich]

$\Rightarrow$ Why is fast when Glauber dynamics is slow?

Intuition := It is is for the algorithm to jump from a mostly red configuration to a mostly green one, etc...

But, why...? Also, why is it correct?

- After step 2, only edge configuration.

Random-cluster model:

$$\mu(A) = \frac{1}{Z_{RC}} \left(\frac{p}{1-p}\right)^{|A|} q^{c(A)}$$

- $A \subseteq E$ configuration
- p, q model parameters
- $\Omega_{RC} = \{0,1\}^{|E|}$.

Phase transition:

- $p < p_c(q) = \frac{\sqrt{q}}{1-\sqrt{q}} = 1 - e^{-\beta_c(q)}$ [all components small]
- $p > p_c(q)$ [exactly one 'giant' component with most vertices]

Claim: If $\sigma \sim \pi_p$, and we apply step 1 and 2 from the SW algorithm, the resulting config. has distribution μ and moreover $Z_{RC} = Z_p$.

Proof:

$$\Pr[A \subseteq E] = \sum_{\sigma \in \Omega} \pi_p(\sigma) \cdot \Pr[\sigma \longrightarrow A]$$

$$= \sum_{\sigma : A \subseteq a(\sigma)} \frac{1}{Z_p} \cdot e^{\beta \cdot a(\sigma)} \cdot p^{|A|} (1-p)^{a(\sigma) \setminus |A|}$$

$$= \frac{1}{Z_p} \cdot \sum_{\sigma : A \subseteq a(\sigma)} \frac{1}{\cancel{(1-p)^{a(\sigma)}}} \cdot p^{|A|} \frac{\cancel{(1-p)^{a(\sigma)}}}{(1-p)^{|A|}}$$

$$\Pr[A \subseteq E] = \frac{1}{Z_p} \cdot \sum_{\sigma : A \subseteq a(\sigma)} \left(\frac{p}{1-p}\right)^{|A|}$$

$$= \frac{1}{Z_p} \left(\frac{p}{1-p}\right)^{|A|} \cdot q^{c(A)}$$

$$\sum_A \Pr[A] = 1 = \frac{1}{Z_p} \sum_A \left(\frac{p}{1-p}\right)^{|A|} \cdot q^{c(A)} = \frac{Z_{RC}}{Z_p}$$

$(Z_{RC} = Z_p)$. So

$$\Pr[A] = \mu(A).$$

<u>Claim 2</u> If $A \sim \mu$, ~~a~~after steps ③ and ④ of the dynamics, the resulting configuration have distribution π_p.

<u>Proof</u>:

$$\Pr[\sigma \in \Omega_p] = \sum_{A \in \Omega_{RC}} \mu(A) \cdot \Pr[A \to \sigma]$$

$$= \sum_{A : A \subseteq a(\sigma)} \frac{1}{Z_{RC}} \left(\frac{p}{1-p}\right)^{A} \cdot \cancel{q^{c(A)}} \; \frac{1}{\cancel{q^{c(A)}}}$$

$$\Pr[\sigma \in \Omega_p] = \frac{1}{Z_p} \sum_{A : A \subseteq a(\sigma)} \left(\frac{p}{1-p}\right)^{A}$$

$$= \frac{1}{Z_p} \cdot \sum_{k=1}^{a(\sigma)} \sum_{\substack{A : A \subseteq a(\sigma) \\ |A| = k}} \left(\frac{p}{1-p}\right)^{A} = \frac{1}{Z_p} \cdot \sum_{k=1}^{a(\sigma)} \binom{a(\sigma)}{k} \left(\frac{p}{1-p}\right)^{|A|}$$

$$\Pr[\sigma] = \frac{1}{Z_p}\left(\frac{p}{1-p}+1\right)^{a(\sigma)} = \frac{1}{Z_p}\left(\frac{1}{1-p}\right)^{a(\sigma)} = \frac{1}{Z_p}\cdot e^{\beta a(\sigma)}$$

■

Correctness of SW:

① Irreducible.

② Aperiodic

③ π_p is stationary. (by Claims 1 and 2

(*) The approch we used is due to [Edwards-Sokal '89.]

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

- $q = \{2, 3\}$ [Gheissari, Lubetzky '16] $T_{mix} = O(n^z)$ for some constant $z > 0$
- $q = 4$ [Lubetzky, Sly '10] $T_{mix} = O(n^{\log n})$ [we expect $T_{mix} = \Omega(n^z)$ also]
- $q \geq 5$ $T_{mix} = \exp(\Omega(\sqrt{n}))$

⇒ Why change of behavior at $q = 4$?

- Related to "continuity" of phase transition.

⇒ $1 \leq q < 4$ continuous ("smooth") phase transition.

$q > 4$ discontinuous ("sharp") "

$\Rightarrow$ Ok. $\mathbb{Z}_2$ we got. What about $\mathbb{Z}^d$? $d \geq 3$.

(We really don't know).

$\Rightarrow$ Other graphs

Meanfield case : $G = K_n$ complete graph on n vertices

$q = 2$ (Ising model) $\quad [\beta_c$

$\beta < \beta_c \quad {}^{\approx 2/n} \qquad T_{mix} = \Theta(1)$

$\beta > \beta_c \qquad T_{mix} = \Theta(\log n)$ $\quad \rightarrow$ [Long, Nachmias, Ning, Peres '11]

$\beta = \beta_c \qquad T_{mix} = \Theta(n^{1/4})$

$q \geq 3$ (Potts)

$\beta < \beta_L(q) \qquad T_{mix} = \Theta(1)$

$\beta > \beta_R(q) \qquad T_{mix} = \Theta(\log n)$

$\beta \in (\beta_L, \beta_R) \qquad T_{mix} = \exp(\Omega(n)) \qquad \beta_c \in (\beta_L, \beta_R)$

$\beta = \beta_L \qquad T_{mix} = \Theta(n^{1/3})$

$\beta = \beta_R \qquad T_{mix} = \Theta(\log n)$

$\rightarrow$ [2015]

$\boxed{\text{Conj: } \forall G, \forall \beta, q=2.\ \tau_{mix} = O(n^{1/4})}$

Thm [Guo & Jerrum '16]

$$\tau_{mix}(P_{SW}) = O(n^{10}) \text{ on any graph.}$$