Swendsen-Wang algorithm:

For $G=(V,E)$,

let $\Omega = \{+1,-1\}^V$

Configuration $\sigma \in \Omega$ is an assignment $\sigma : V \to \{+1,-1\}$.

Gibbs Distribution: $$\pi_I(\sigma) = \frac{e^{\beta \sum_{(v,w)\in E} \sigma_v \sigma_w}}{Z_I}$$

where partition function $Z_I = \sum_{\sigma \in \Omega} e^{\beta \sum \sigma_v \sigma_w}$

Note, $$\sum_{(v,w)\in E} \sigma_v \sigma_w = |M(\sigma)| - |D(\sigma)|$$

$|M(\sigma)|$: # of monochromatic edges, i.e., $(+,+)$ or $(-,-)$

$|D(\sigma)|$: non-monochromatic = Disagree edges, i.e., $(+,-)$ or $(-,+)$

Ferromagnetic: $\beta > 0$ so most likely configurations are all + & all − (called ground states)

but only 2 such whereas $\binom{n}{n/2}$ balanced configurations (half +, half −)

tradeoff b/w energy vs. entropy

all +, all − vs. $\binom{n}{n/2}$ balanced.

(called attractive in ML)

$\beta = 0$ (corresponding to infinite temperature as $\beta = \frac{1}{\text{temp}}$)

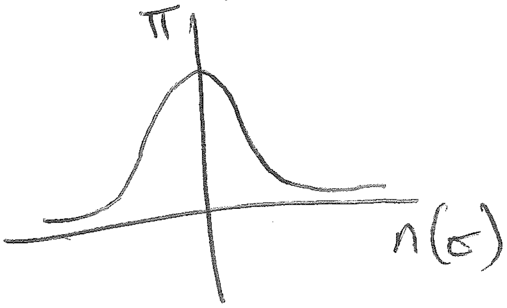
then balanced dominates as all configurations have same measure.

$\beta = \infty$ (zero temperature)

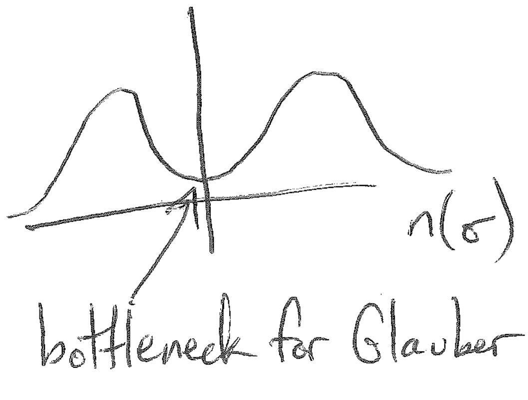
then all +, all − dominate (everything else has zero measure)

Look at magnetization $n(\sigma) = \sum_{v} \sigma(v)$

small β



large β



For $\sqrt{n} \times \sqrt{n}$ grid, recall $\exists\, \beta_c$ where:

$\forall \beta < \beta_c$: Glauber dynamics has $O(n \log n)$ mixing time

$\forall \beta > \beta_c$: Glauber has mixing time $\exp(\Omega(\sqrt{n}))$ i.e., $\geq e^{c\sqrt{n}}$ for constant $c > 0$.

update random vertex in every step.

Note, $|M(\sigma)| + |D(\sigma)| = |E|$

thus, $2|M(\sigma)| - |E| = |M(\sigma)| - |D(\sigma)|$

Hence,

$$\pi_I(\sigma) = \frac{e^{\beta(M(\sigma)-D(\sigma))}}{Z_I} = \frac{e^{2\beta M(\sigma)}}{e^{\beta|E|} Z_I} = \frac{e^{2\beta M(\sigma)}}{Z'}$$

Generalization from $q=2$ spins to integer $q \geq 2$ spins.

Potts model: spins $\{1, 2, \ldots, q\}$

state space $\Omega_P = \{1, 2, \ldots, q\}^V$

Configurations $\sigma \in \Omega_P$ is assignment $\sigma : V \to \{1, \ldots, q\}$

Gibbs distribution: $\pi_P(\sigma) = \dfrac{e^{\beta|M(\sigma)|}}{Z_P}$

$$M(\sigma) = \{(v,w) \in E : \sigma(v) = \sigma(w)\}.$$

$q=2$ is the Ising model with $\beta \mapsto \beta/2$.

$$\beta_c(q) = \ln(1 + \sqrt{q})$$ [Duminil-Copin Befara

$q=2$ was from [Onsager '44]

Random cluster model:

For $G=(V,E)$, RC has 2 parameters:

$$0<p<1 \quad \& \quad q>0$$

Configurations are subsets of edges $S \subset E$

Gibbs distribution for RC: $\mu_{RC}(S) = \dfrac{p^{|S|}(1-p)^{|E \setminus S|}\, q^{c(S)}}{Z}$

where $c(S) = \#$ of components in graph (V,S)
(isolated vertices count as a component)

When $q=1$: Percolation – keep edge e with prob. p
delete e w.p. $1-p$

For integer $q \geq 2$, set $p = 1-e^{-\beta}$ & use same q
then $Z_{RC} = Z_{Potts}$

Note, $\mu(S) = \dfrac{p^{|S|}(1-p)^{|E\setminus S|}\, q^{c(S)}}{Z} = \dfrac{(1-p)^{|E|}}{Z}\left(\dfrac{p}{1-p}\right)^{|S|} q^{c(S)}$

$(1-p = e^{-\beta})$ $\quad = \dfrac{p^{|S|}(1-p)^{-|S|}\, q^{c(S)}}{e^{\beta|E|}\, Z} = \dfrac{1}{Z_{RC}}\left(\dfrac{p}{1-p}\right)^{|S|} q^{c(S)}.$

Swendsen-Wang algorithm: For integer $q \geq 2$,

MC on $\Omega = \{1, \ldots, q\}^V$ = Potts configurations.

idea: From $X_t \in \Omega$,

1. transform to $A_t \subset E$ which is RC configuration;
2. Perform step $A_t \to A_{t+1}$ in RC model;
3. map A_{t+1} back to $X_{t+1} \in \Omega$ which is a Potts configuration.

Formally, from $X_t \in \Omega = \{1, \ldots, q\}^V$,

1. Let $M_t = \{(v,w) \in E : X_t(v) = X_t(w)\}$
   = set of monochromatic edges.
2. For each $(v,w) \in M_t$, independently
   add (v,w) to A_{t+1} w.p. $p = 1 - e^{-\beta}$
   & leave out (v,w) from A_{t+1} w.p. $1 - p = e^{-\beta}$.

   Let A_{t+1} be the remaining edges,
   ignore the colors in X_t so $A_{t+1} \subseteq E$
3. For each connected component C in (V, A_{t+1}),
   independently choose a color $c \in \{1, \ldots, q\}$ uar
   & assign every $v \in C$ with color c.
4. Let $X_{t+1} \in \{1, \ldots, q\}^V$ be the resulting coloring (ignore edges)

Recall, π_p = Gibbs distribution for Potts model
& μ_{RC} = Gibbs dist. for RC model.

<u>Claim 1:</u> If $X_t \sim \pi_p$ then after steps 1&2,
$$A_{t+1} \sim \mu_{RC}$$

<u>Claim 2:</u> If $A_{t+1} \sim \mu_{RC}$ then after steps 3&4,
$$X_{t+1} \sim \pi_p$$

Therefore, if $X_t \sim \pi_p$ then $X_{t+1} \sim \pi_p$

So π_p is a stationary distribution for the SW dynamics, and since SW is ergodic it's the unique stationary dist.

[Guo-Jerrum '16] <u>For q=2:</u> For all G, all β, $T_{mix} = O(n^{10})$ for the SW dynamics.

big polynomial but it's Poly(n) for all G, all β for Ising

Proof of Claim 1: $X_t \sim \pi \Rightarrow A_{t+1} \sim \mu$

For SCE,

$$\Pr(A_{t+1}=S \mid X_t \sim \pi) = \sum_{\sigma \in \Omega = \{1,\ldots,q\}^V} \pi(\sigma) \Pr(A_{t+1}=S \mid X_t = \sigma)$$

$$= \sum_{\sigma : M(\sigma) \supseteq A(\sigma) = S} \frac{e^{\beta |M(\sigma)|}}{Z_p} p^{|S|} (1-p)^{|M(\sigma) \setminus S|}$$

$(1-p = e^{-\beta})$
$$= \sum_{\sigma} \frac{1}{Z_p} \cancel{\frac{1}{(1-p)^{|M(\sigma)|}}} p^{|S|} \frac{\cancel{(1-p)^{|M(\sigma)|}}}{(1-p)^{|S|}}$$

$$= \frac{1}{Z_p} \sum_{\sigma} \left(\frac{p}{1-p}\right)^{|S|}$$

$$= \frac{1}{Z_p} \left(\frac{p}{1-p}\right)^{|S|} \left|\{\sigma \in \{1,\ldots,q\}^V : M(\sigma) \supseteq S\}\right|$$

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

∎

Proof of claim 2: $A_{t+1} \sim \mu_{RC} \Rightarrow X_{t+1} \sim \pi_p$

For $\sigma \in \Omega_p = \{1, \dots, q\}^V$,

$$\Pr(X_{t+1} = \sigma \mid A_{t+1} \sim \mu_{RC})$$

$$= \sum_{S : S \subset M(\sigma)} \mu(S) \frac{1}{q^{c(S)}}$$

$$= \sum_{S : S \subset M(\sigma)} \frac{1}{Z_{RC}} \left(\frac{p}{1-p}\right)^{|S|} \frac{q^{c(S)}}{q^{c(S)}}$$

$$= \frac{1}{Z_{RC}} \sum_{k=1}^{|M(\sigma)|} \sum_{\substack{S : S \subset M(\sigma), \\ |S| = k}} \left(\frac{p}{1-p}\right)^k$$

$$= \frac{1}{Z_{RC}} \sum_k \binom{|M(\sigma)|}{k} \left(\frac{p}{1-p}\right)^k$$

$$= \frac{1}{Z_{RC}} \left(\frac{p}{1-p} + 1\right)^{|M(\sigma)|} = \frac{1}{Z_{RC}} \left(\frac{1}{1-p}\right)^{|M(\sigma)|}$$

$1 - p = e^{-\beta}$

$$= \frac{e^{\beta |M(\sigma)|}}{Z_{RC}} = \pi_p(\sigma). \quad \blacksquare$$

SW is poly(n) mixing for all G, all β
<u>when $q=2$</u>

What if $q \geq 3$?

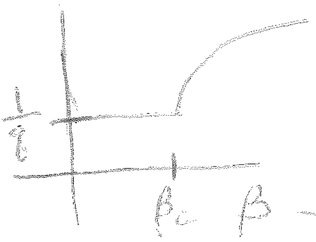
Examples where exponentially slow.

For $\sqrt{n} \times \sqrt{n}$ grid, $\beta = \beta_c$,

$q=2,3$: $T_{mix} = O(\text{poly}(n))$

$q=4$: $T_{mix} = n^{O(\log n)}$ (conjecture should be poly(n))

$q \geq 5$: $T_{mix} = \exp(\Omega(\sqrt{n}))$



$q=2$: [Lubetzky Sly '10]
$q \geq 3$: [Gheissari, Lubetzky '16]

Key: $q<4$: magnetization = # of spins of dominant color is continuous but derivative is discontin

$q \geq 5$: magnetization is discontinuous.