Path Coupling and Aggregate Path Coupling

Yevgeniy Kovchegov
Oregon State University

(joint work with Peter T. Otto from Willamette University)

Mixing times for Glauber dynamics.

Important question: the relationship between the mixing time and the equilibrium phase transition of the corresponding statistical mechanical models.

• For models that exhibit a continuous phase transition: used standard path coupling (Bubley and Dyer '97). There rapid mixing is proved by showing contraction of the mean coupling distance between neighboring configurations.

• For models that exhibit a discontinuous phase transition, the standard path coupling method fails.

Our approach combines aggregate path coupling and large deviation theory to determine the mixing times of a large class of statistical mechanical models, including those that exhibit a discontinuous phase transition. The aggregate path coupling method extends the use of the path coupling technique in the absence of contraction of the mean coupling distances betweenl neighboring configurations.

Markov Chain Monte Carlo (MCMC).

Goal: simulating an Ω-valued random variable distributed according to a given probability distribution $\pi(z)$, given a complex nature of large discrete space Ω.

MCMC: generating a Markov chain $\{X_t\}$ over Ω, with distribution $\mu_t(z) = P(X_t = z)$ converging rapidly to its unique stationary distribution, $\pi(z)$.

Metropolis-Hastings algorithm: Consider a connected neighborhood network with points in Ω. Suppose we know the ratios of $\frac{\pi(z')}{\pi(z)}$ for any two neighbor points z and z' on the network.

Let for z and z' connected by an edge of the network, the transition probability be set to

$$p(z,z') = \frac{1}{M}\ \min\left\{1,\ \frac{\pi(z')}{\pi(z)}\right\} \qquad \text{for } M \text{ large enough.}$$

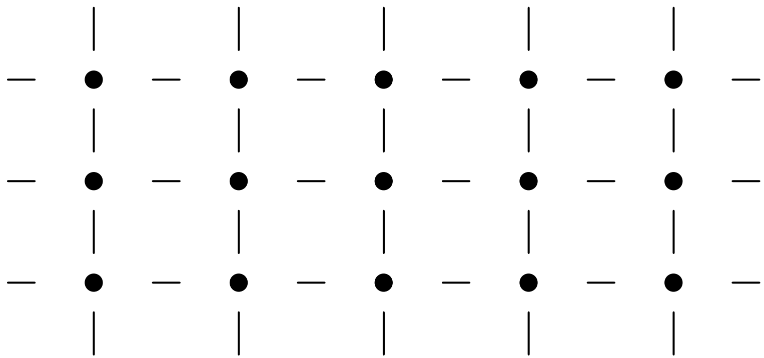
Gibbs Sampling: Ising Model.

Every vertex v of $G = (V, E)$ is assigned a spin

$$\sigma(v) \in \{-1, +1\}$$

The probability of a configuration $\sigma \in \{-1, +1\}^V$ is

$$\pi(\sigma) = \frac{e^{-\beta \mathcal{H}(\sigma)}}{Z(\beta)}, \quad \text{where} \quad \beta = \frac{1}{T}$$



Gibbs Sampling: Ising Model.

Every vertex v of $G = (V, E)$ is assigned a spin

$$\sigma(v) \in \{-1, +1\}$$

The probability of a configuration $\sigma \in \{-1, +1\}^V$ is

$$\pi(\sigma) = \frac{e^{-\beta \mathcal{H}(\sigma)}}{Z(\beta)}, \quad \text{where} \quad \beta = \frac{1}{T}$$

$$\begin{array}{ccccccccccc}
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & -1 & - & -1 & - & +1 & - \\
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & +1 & - & +1 & - & -1 & - \\
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & -1 & - & +1 & - & -1 & - \\
 & | & & | & & | & & | & & | &
\end{array}$$

Gibbs Sampling: Ising Model.

$\forall \sigma \in \{-1, +1\}^V$, the Hamiltonian

$$\mathcal{H}(\sigma) = -\frac{1}{2} \sum_{u,v:\ u \sim v} \sigma(u)\sigma(v) = -\sum_{edges\ \ e=[u,v]} \sigma(u)\sigma(v)$$

and probability of a configuration $\sigma \in \{-1, +1\}^V$ is

$$\pi(\sigma) = \frac{e^{-\beta \mathcal{H}(\sigma)}}{Z(\beta)}, \quad \text{where} \quad \beta = \frac{1}{T}$$

$Z(\beta) = \sum_{\sigma \in \{-1,+1\}^V} e^{-\beta \mathcal{H}(\sigma)}$ - normalizing factor.

Ising Model: local Hamiltonian

$$\mathcal{H}(\sigma) = -\frac{1}{2}\sum_{u,v:\ u\sim v}\sigma(u)\sigma(v) = \quad -\sum_{edges\ \ e=[u,v]}\sigma(u)\sigma(v)$$

The local Hamiltonian

$$\mathcal{H}_{local}(\sigma, v) = -\sum_{u:\ u\sim v}\sigma(u)\sigma(v)\ .$$

Observe: conditional probability for $\sigma(v)$ is given by $\mathcal{H}_{local}(\sigma, v)$:

$$\mathcal{H}(\sigma) = \mathcal{H}_{local}(\sigma, v) - \sum_{e=[u_1,u_2]:\ u_1,u_2\neq v}\sigma(u_1)\sigma(u_2)$$

Gibbs Sampling: Ising Model via Glauber dynamics.

$$\begin{array}{ccccccccccc}
 & \vert & & \vert & & \vert & & \vert & & \vert & \\
- & +1 & - & -1 & - & -1 & - & -1 & - & +1 & - \\
 & \vert & & \vert & & \vert & & \vert & & \vert & \\
- & +1 & - & -1 & - & \sigma(v) & - & +1 & - & -1 & - \\
 & \vert & & \vert & & \vert & & \vert & & \vert & \\
- & +1 & - & -1 & - & -1 & - & +1 & - & -1 & - \\
 & \vert & & \vert & & \vert & & \vert & & \vert &
\end{array}$$

Observe: conditional probability for $\sigma(v)$ is given by $\mathcal{H}_{local}(\sigma, v)$:

$$\mathcal{H}(\sigma) = \mathcal{H}_{local}(\sigma, v) - \sum_{e=[u_1,u_2]:\ u_1,u_2 \neq v} \sigma(u_1)\sigma(u_2)$$

Gibbs Sampling: Ising Model via Glauber dynamics.

$$
\begin{array}{ccccccccccc}
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & -1 & - & -1 & - & +1 & - \\
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & ? & - & +1 & - & -1 & - \\
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & -1 & - & +1 & - & -1 & - \\
 & | & & | & & | & & | & & | &
\end{array}
$$

Randomly pick $v \in G$, erase the spin $\sigma(v)$. Choose σ_+ or σ_-:

$$
\mathbf{Prob}(\sigma \to \sigma_+) = \frac{e^{-\beta\mathcal{H}(\sigma_+)}}{e^{-\beta\mathcal{H}(\sigma_-)}+e^{-\beta\mathcal{H}(\sigma_+)}}
$$

$$
= \frac{e^{-\beta\mathcal{H}_{local}(\sigma_+,v)}}{e^{-\beta\mathcal{H}_{local}(\sigma_-,v)}+e^{-\beta\mathcal{H}_{local}(\sigma_+,v)}} = \frac{e^{-2\beta}}{e^{-2\beta}+e^{2\beta}} \; .
$$

Glauber dynamics: Rapid mixing.

Glauber dynamics - a random walk on state space S (here $\{-1,+1\}^V$) s.t. needed π is stationary w.r.t. Glauber dynamics.

In high temperatures (i.e. $\beta = \frac{1}{T}$ small enough) it takes $O(n \log n)$ iterations to get "ε-close" to π. Here $|V| = n$.

Need: $\boxed{\max_{v \in V} deg(v) \cdot \tanh(\beta) < 1}$

Thus the Glauber dynamics is a fast way to generate π. It is an important example of **Gibbs sampling**.

Close enough distribution and mixing time.

What is "ε-close" to π? Start with σ_0:

$$
\begin{array}{ccccccccccc}
 & | & & | & & | & & | & & | & \\
- & +1 & - & +1 & - & +1 & - & +1 & - & +1 & - \\
 & | & & | & & | & & | & & | & \\
- & +1 & - & +1 & - & +1 & - & -1 & - & -1 & - \\
 & | & & | & & | & & | & & | & \\
- & -1 & - & -1 & - & -1 & - & -1 & - & -1 & - \\
 & | & & | & & | & & | & & | &
\end{array}
$$

If $P_t(\sigma)$ is the probability distribution after t iterations, the total variation distance

$$\|P_t - \pi\|_{TV} = \frac{1}{2} \sum_{\sigma \in \{-1,+1\}^V} |P_t(\sigma) - \pi(\sigma)| \leq \varepsilon \ .$$

Mixing Times.

Total variation distance between two distributions μ and ν:

$$\|\mu - \nu\|_{\text{TV}} = \sup_{A \subset \Omega} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|$$

Maximal distance to stationary:

$$d(t) = \max_{x \in \Omega} \|P^t(x, \cdot) - \pi\|_{\text{TV}}$$

where $P^t(x, \cdot)$ is the transition probability starting iat x and π is its stationary distribution.

Definition. Given $\varepsilon > 0$, the **mixing time** of the Markov chain is defined by

$$t_{\text{mix}}(\varepsilon) = \min\{t : d(t) \leq \varepsilon\}$$

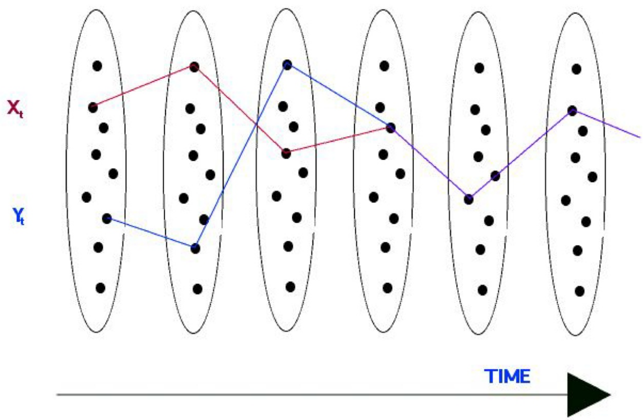
Mixing times categorized into two groups:

- rapid mixing = polynomial growth w.r.t. system size
- slow mixing = exponential growth w.r.t. system size

Coupling Method.

Ω - sample space, $\{p(i,j)\}_{i,j\in\Omega}$ - transition probabilities

Construct process (X_t, Y_t) on $\Omega \times \Omega$ such that
X_t is a $\{p(i,j)\}$-Markov chain and Y_t is a $\{p(i,j)\}$-Markov chain.



Once $X_t = Y_t$, let $X_{t+1} = Y_{t+1}$, $X_{t+2} = Y_{t+2}, \ldots$.

Coupling time: $\tau_c = \min\{t : \; X_t = Y_t\}$.

Successful coupling: $P(\tau_c < \infty) = 1$

Mixing Times, Coupling and Path Coupling.

Coupling Inequality. Let (X_t, Y_t) be a coupling of a Markov chain where Y_t is distributed by the stationary distribution π. The **coupling time** of the Markov chain is defined by

$$\tau_c := \min\{t : X_t = Y_t\}.$$

Then, for all initial states x,

$$\|P^t(x,\cdot) - \pi\|_{TV} \leq P(X_t \neq Y_t) = P(\tau_c > t)$$

and thus

$$\tau_{mix}(\varepsilon) \leq \frac{\mathbb{E}[\tau_c]}{\varepsilon}$$

Thus

$$\tau_{mix}(\varepsilon) = \mathcal{O}(\mathbb{E}[\tau_c])$$

Path coupling method

Let $\{(X,Y)\}$ be a coupling of $P(x,\cdot)$ and $P(y,\cdot)$ for which $X_0 = x$ and $Y_0 = y$. Then

$$\|P(x,\cdot) - P(y,\cdot)\|_{TV} \leq P_{x,y}(X \neq Y)$$

Define a metric ρ on the space of configurations and let $(x = x_0, x_1, \ldots, x_r = y)$ be a minimal path joining configurations x and y such that each pair of configurations (x_{j-1}, x_j) are neighbors with respect to ρ. Then

$$P_{x,y}(X \neq Y) \leq \mathbb{E}_{x,y}[\rho(X,Y)] \leq \sum_{j=1}^{n} \mathbb{E}_{x_{j-1},x_j}[\rho(X_{j-1}, X_j)]$$

Path coupling method

Suppose the state space Ω of a Markov chain is the vertex set of a graph with path metric ρ. Suppose that for each edge $\{\sigma, \tau\}$ there exists a coupling (X, Y) of the distributions $P(\sigma, \cdot)$ and $P(\tau, \cdot)$ such that

$$\mathbb{E}_{\sigma,\tau}[\rho(X,Y)] \leq \rho(\sigma,\tau)e^{-\alpha} \qquad \text{for some } \alpha > 0$$

Then

$$t_{mix}(\varepsilon) \leq \left\lceil \frac{-\log(\varepsilon) + \log(\mathit{diam}(\Omega))}{\alpha} \right\rceil .$$

Contraction is required for ALL pairs of neighboring configurations.

Mean-field Blume-Capel model

Spin model defined on the complete graph on n vertices. The spin at site j is denoted by ω_j, taking values in $\Lambda = \{1, 0, -1\}$. The configuration space is the set Λ^n of sequences $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ with each $\omega_j \in \Lambda$.

In terms of a positive parameter $K > 0$ representing the interaction strength, the Hamiltonian is defined by

$$H_{n,K}(\omega) = \sum_{j=1}^{n} \omega_j^2 - \frac{K}{n} \left(\sum_{j=1}^{n} \omega_j \right)^2 .$$

Mean-field Blume-Capel model

For $n \in \mathbb{N}$, inverse temperature $\beta > 0$, and $K > 0$, the Gibbs measure or canonical ensemble for the mean-field B-C model is the sequence of probability measures

$$P_{n,\beta,K}(B) = \frac{1}{Z_n(\beta, K)} \cdot \int_B \exp[-\beta H_{n,K}]\, dP_n$$

where P_n is the product measure with marginals $\rho = \frac{1}{3}(\delta_{-1} + \delta_0 + \delta_1)$ and $Z_n(\beta, K)$ is the partition function

$$Z_n(\beta, K) = \int_{\Lambda^n} \exp[-\beta H_{n,K}]\, dP_n$$

Taking the system size n to infinity, called the "thermodynamic limit", yields the equilibrium state of the system.

Mean-field Blume-Capel model

Absorbing the noninteracting component of the Hamiltonian into the product measure P_n yields

$$P_{n,\beta,K}(d\omega) = \frac{1}{\tilde{Z}_n(\beta,K)} \cdot \exp\left[n\beta K\left(\frac{S_n(\omega)}{n}\right)^2\right] P_{n,\beta}(d\omega).$$

In this formula $S_n(\omega)$ equals the total spin $\sum_{j=1}^n \omega_j$, $P_{n,\beta}$ is the product measure on Λ^n with marginals

$$\rho_\beta(d\omega_j) = \frac{1}{Z(\beta)} \cdot \exp(-\beta\omega_j^2)\,\rho(d\omega_j),$$

and $Z(\beta)$ and $\tilde{Z}_n(\beta,K)$ are the appropriate normalizations.

Mean-field Blume-Capel model

With respect to the mean-field Blume-Capel model $P_{n,\beta,K}$, S_n/n satisfies the large deviations principle with speed n and rate function

$$I_{\beta,K}(z) = J_\beta(z) - \beta K z^2 - \inf_{y\in\mathbb{R}}\{J_\beta(y) - \beta K y^2\}$$

where

$$c_\beta(t) = \log \int_\Lambda \exp(t\omega_1)\, \rho_\beta(d\omega_1) = \log\left(\frac{1 + e^{-\beta}(e^t + e^{-t})}{1 + 2e^{-\beta}}\right)$$

and

$$J_\beta(z) = \sup_{t\in\mathbb{R}}\{tz - c_\beta(t)\}.$$

Mean-field Blume-Capel model

Large deviations principle.

(a) For each closed set C,

$$\limsup_{n\to\infty} \frac{1}{n} \log P_{n,\beta,K}\left(\frac{S_n}{n} \in C\right) \le -\inf_{z\in C} I_{\beta,K}(z)$$

(b) For each open set G,

$$\liminf_{n\to\infty} \frac{1}{n} \log P_{n,\beta,K}\left(\frac{S_n}{n} \in G\right) \ge -\inf_{z\in G} I_{\beta,K}(z)$$

Equilibrium macrostates:

$$\begin{aligned} \tilde{\mathcal{E}}_{\beta,K} &= \{x \in [-1,1] : I_{\beta,K}(x) = 0\} \\ &= \{x \in [-1,1] : x \text{ is a global min pt of } J_\beta(x) - \beta K x^2\} \end{aligned}$$

Mean-field Blume-Capel model

Free energy functional:

$$G_{\beta,K}(x) = \beta K x^2 - c_\beta(2\beta K x)$$

$\tilde{\mathcal{E}}_{\beta,K} = \{x \in [-1,1] : x$ is a global min. point of $G_{\beta,K}(x)\}$

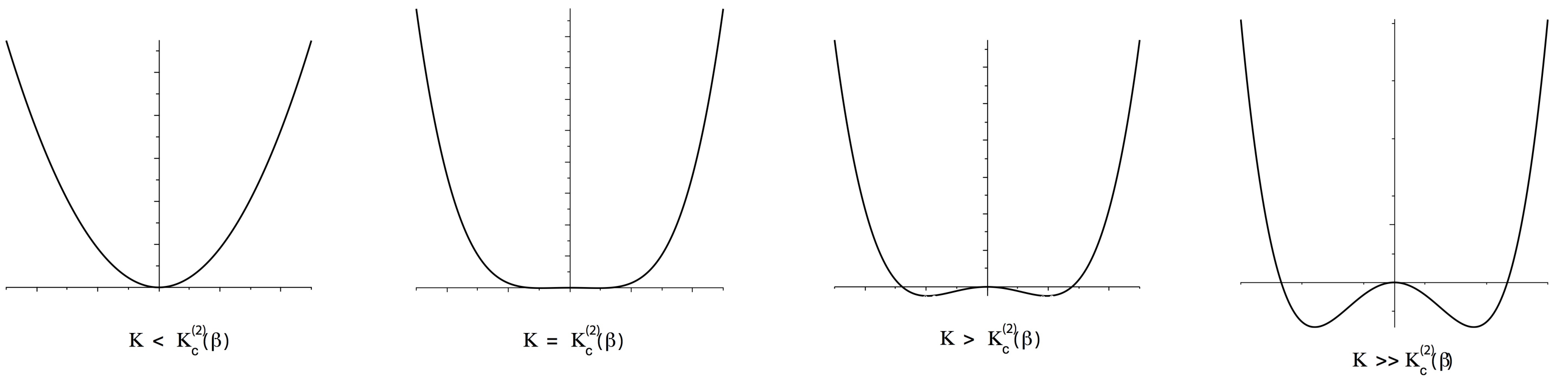
$G_{\beta,K}$ exhibits two distinct behaviors for

(a) $\beta \leq \beta_c = \log 4$ and (b) $\beta > \beta_c$.

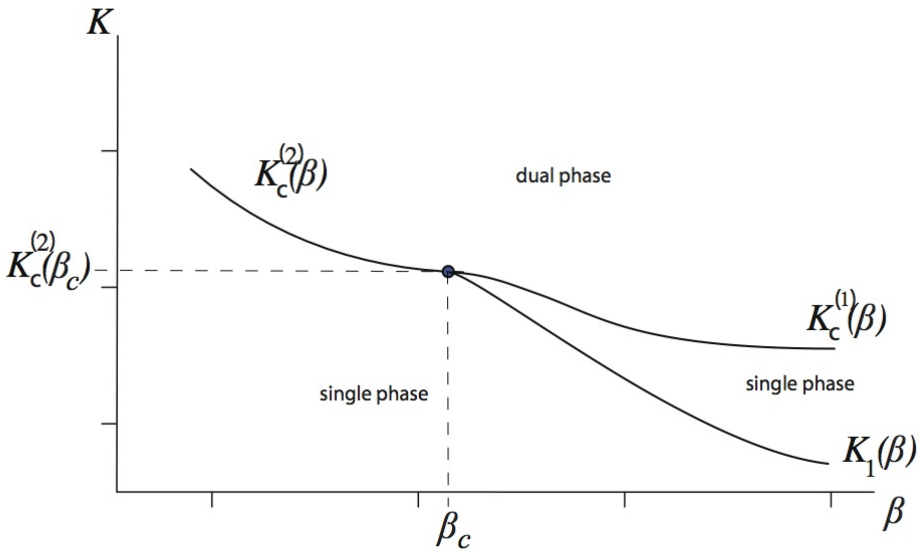
Mean-field Blume-Capel model

$\beta \leq \beta_c = \log 4$

$K = K_c^{(2)}(\beta)$ second-order, continuous phase transition point



Mean-field Blume-Capel model: Equilibrium phase diagram

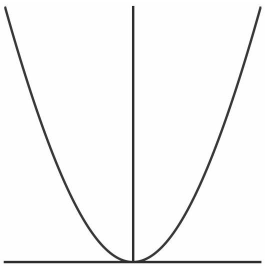


R.S. Ellis, P.T. Otto, and H. Touchette in AAP 2005

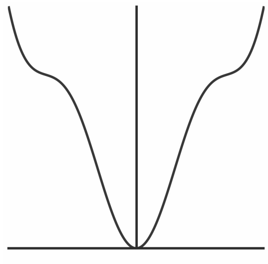
Mean-field Blume-Capel model

$\beta > \beta_c = \log 4$ $\quad K = K_1(\beta)$ metastable critical point

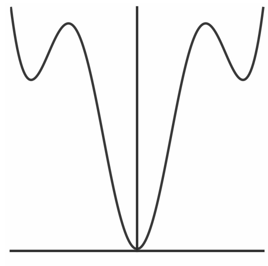
$K = K_c^{(1)}(\beta)$ discontinuous, first-order phase transition point



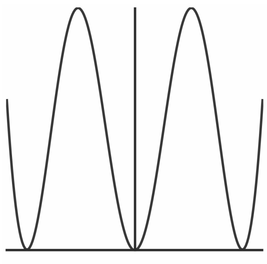
$K < K_1(\beta)$



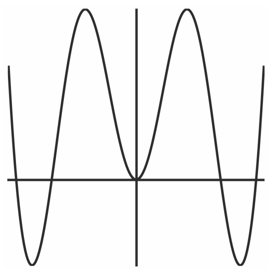
$K = K_1(\beta)$



$K_1(\beta) < K < K_c^{(1)}(\beta)$

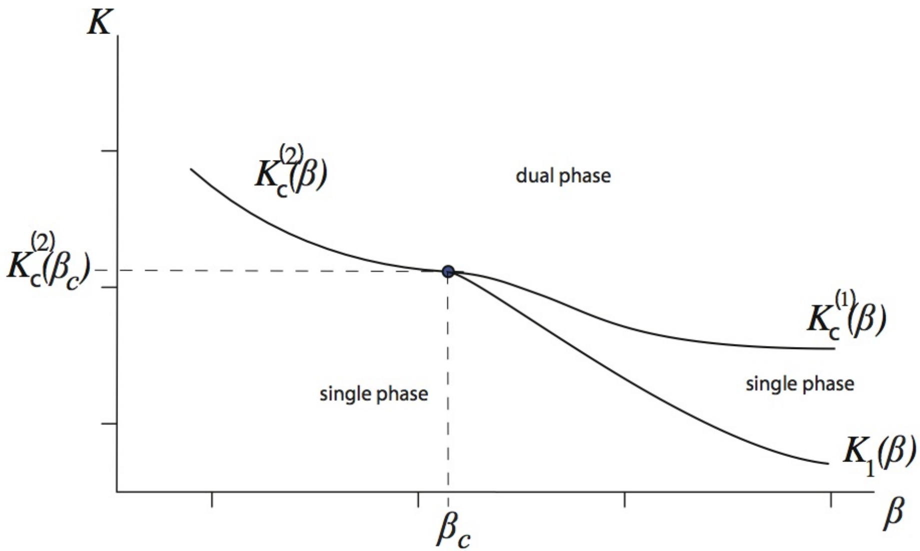


$K = K_c^{(1)}(\beta)$



$K > K_c^{(1)}(\beta)$

Mean-field Blume-Capel model: Equilibrium phase diagram



R.S. Ellis, P.T. Otto, and H. Touchette in AAP 2005

Glauber dynamics for BC model

Choose vertex of underlying complete graph uniformly then update the spin at the vertex according to $P_{n,\beta,K}$ conditioned on the event that the spins at all other vertices remain unchanged.

Reversible Markov chain with stationary distribution $P_{n,\beta,K}$.

Mean coupling distance for BC dynamics

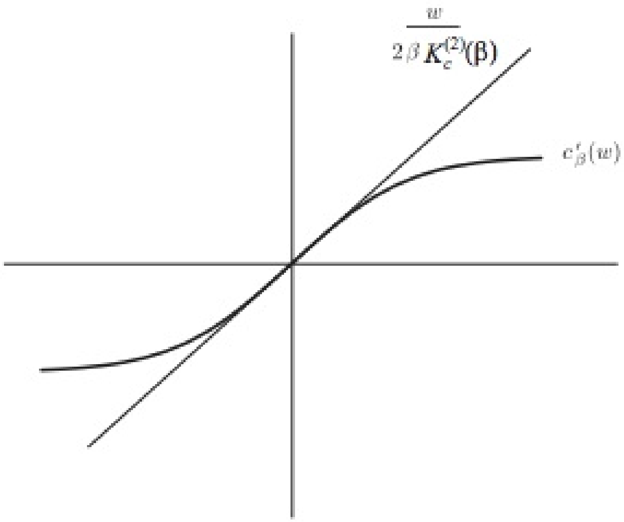
Path metric ρ on $\Omega^n = \{-1, 0, 1\}^n$ is defined by

$$\rho(\sigma, \tau) = \sum_{j=1}^{n} \left|\sigma_j - \tau_j\right|$$

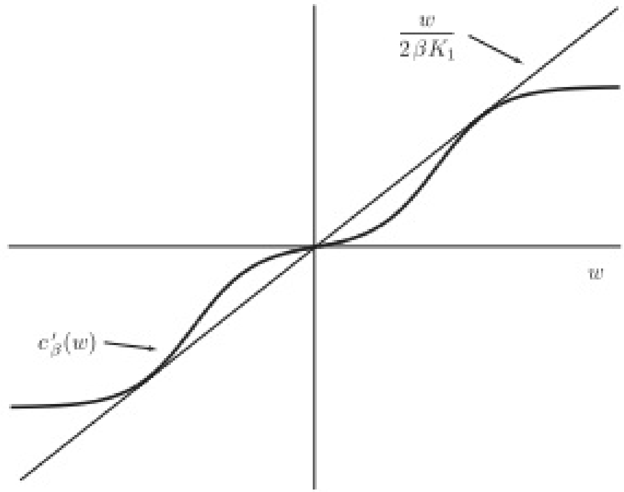
For a coupling (X, Y) of one step of the Glauber dynamics of the BC model starting in neighboring configurations σ and τ, asymptotically as $n \to \infty$,

$$\begin{aligned}
\mathbb{E}_{\sigma,\tau}[\rho(X,Y)] &\approx \frac{n-1}{n} + \frac{(n-1)}{n}\left[c_\beta'\left(2\beta K\frac{S_n(\tau)}{n}\right) - c_\beta'\left(2\beta K\frac{S_n(\sigma)}{n}\right)\right] \\
&\approx \frac{n-1}{n} + \frac{(n-1)}{n}2\beta K\left[\frac{S_n(\tau)}{n} - \frac{S_n(\sigma)}{n}\right] c_\beta''\left(2\beta K\frac{S_n(\sigma)}{n}\right)
\end{aligned}$$

Behavior of c'_β



$\beta \le \beta_c$



$\beta > \beta_c$

Rapid mixing for $\beta \leq \beta_c$

$$\mathbb{E}_{\sigma,\tau}[\rho(X,Y)] \approx \frac{n-1}{n} + \frac{(n-1)}{n} 2\beta K \left[\frac{S_n(\tau)}{n} - \frac{S_n(\sigma)}{n}\right] c_\beta'' \left(2\beta K \frac{S_n(\sigma)}{n}\right)$$

Contraction of mean coupling distance between neighboring configurations σ and τ if

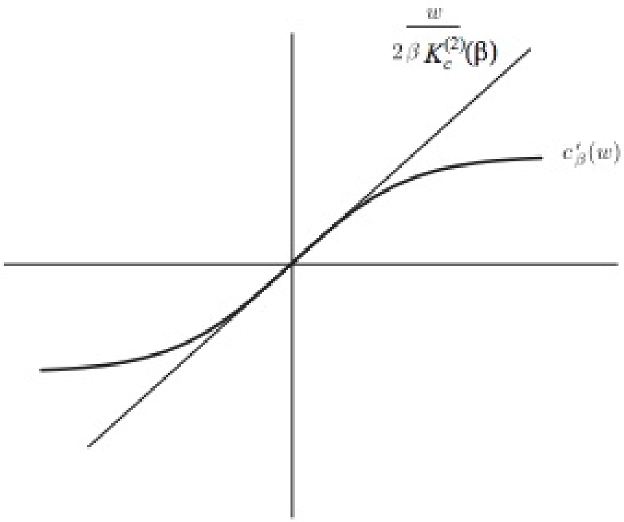
$$c_\beta''\left(2\beta K \frac{S_n(\sigma)}{n}\right) < \frac{1}{2\beta K}$$

For $\beta \leq \beta_c = \log 4$,

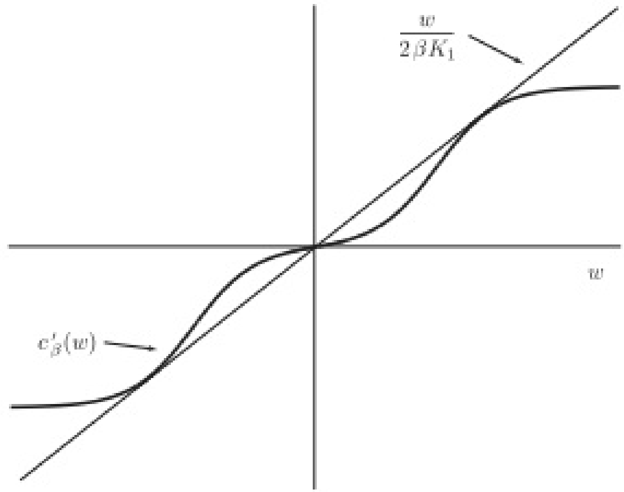
$$c_\beta''\left(2\beta K \frac{S_n(\sigma)}{n}\right) < c_\beta''(0) = \frac{1}{2\beta K_c^{(2)}(\beta)}$$

Rapid mixing when $K < K_c^{(2)}(\beta)$.

Behavior of c'_β



$\beta \leq \beta_c$



$\beta > \beta_c$

Rapid mixing for $\beta > \beta_c$.

Aggregate path coupling:

Let $(\sigma = x_0, x_1, \ldots, x_r = \tau)$ be a path connecting σ to τ and monotone increasing in ρ such that (x_{i-1}, x_i) are neighboring configurations.

$$\mathbb{E}_{\sigma,\tau}[\rho(X,Y)] \leq \sum_{i=1}^{r} \mathbb{E}_{x_{i-1},x_i}[\rho(X_{i-1},X_i)]$$

$$= \frac{(n-1)}{n}\rho(\sigma,\tau) + \frac{(n-1)}{n}\left[c'_\beta\left(\frac{2\beta K}{n}S_n(\tau)\right) - c'_\beta\left(\frac{2\beta K}{n}S_n(\sigma)\right)\right]$$

Assume $S_n(\sigma)/n \sim 0$.

$$\begin{aligned}
\mathbb{E}_{\sigma,\tau}[\rho(X,Y)] &\leq \frac{K}{K_1(\beta)}\left[\frac{S_n(\tau) - S_n(\sigma)}{n}\right] \\
&\leq \rho(\sigma,\tau)\left[1 - \frac{1}{n}\left(1 - \frac{K}{K_1(\beta)}\right)\right]
\end{aligned}$$

Rapid mixing for $\beta > \beta_c$

Let (X, Y) be a coupling of one step of the Glauber dynamics of the BC model that begin in configurations σ and τ, not necessarily neighbors.

Suppose $\beta > \beta_c$ and $K < K_1(\beta)$. Then for any $\alpha \in \left(0, \frac{K_1(\beta)-K}{K_1(\beta)}\right)$ there exists an $\varepsilon > 0$ such that, asymptotically as $n \to \infty$,

$$\mathbb{E}_{\sigma,\tau}[\rho(X, Y)] \leq e^{-\alpha/n}\rho(\sigma, \tau)$$

whenever $|S_n(\sigma)| < \varepsilon n$.

Rapid mixing for $\beta > \beta_c$

For $\beta > \beta_c$ and $K < K_1(\beta)$

$$P_{n,\beta,K}\{S_n/n \in dx\} \Longrightarrow \delta_0 \quad \text{as } n \to \infty.$$

For $Y_0 \stackrel{dist}{=} P_{n,\beta,K}$,

$$\begin{aligned}
\|P^t(X_0,\cdot) - P_{n,\beta,K}\|_{TV} &\leq P\{X_t \neq Y_t\} \\
&= P\{\rho(X_t,Y_t) \geq 1\} \\
&\leq \mathbb{E}[\rho(X_t,Y_t)] \\
&\leq e^{-\alpha t/n}\mathbb{E}[\rho(X_0,Y_0)] + ntP_{n,\beta,K}\{|S_n/n| \geq \varepsilon\} \\
&\leq ne^{-\alpha t/n} + ntP_{n,\beta,K}\{|S_n/n| \geq \varepsilon\}
\end{aligned}$$

Rapid mixing when $K < K_1(\beta)$.

Slow mixing

Bottleneck ratio (Cheeger constant) argument.

For two configurations ω and τ, define the edge measure Q as

$$Q(\omega,\tau) = P_{n,\beta,K}(\omega)P(\omega,\tau) \quad \text{and} \quad Q(A,B) = \sum_{\omega\in A,\tau\in B} Q(\omega,\tau)$$

Here $P(\omega,\tau)$ is the transition probability of the Glauber dynamics of the BC model. The bottleneck ratio of the set S is defined by

$$\Phi(S) = \frac{Q(S,S^c)}{P_{n,\beta,K}(S)} \quad \text{and} \quad \Phi_* = \min_{S:P_{n,\beta,K}(S)\leq\frac{1}{2}} \Phi(S)$$

Then

$$t_{mix} = t_{mix}(1/4) \geq \frac{1}{4\Phi_*}$$

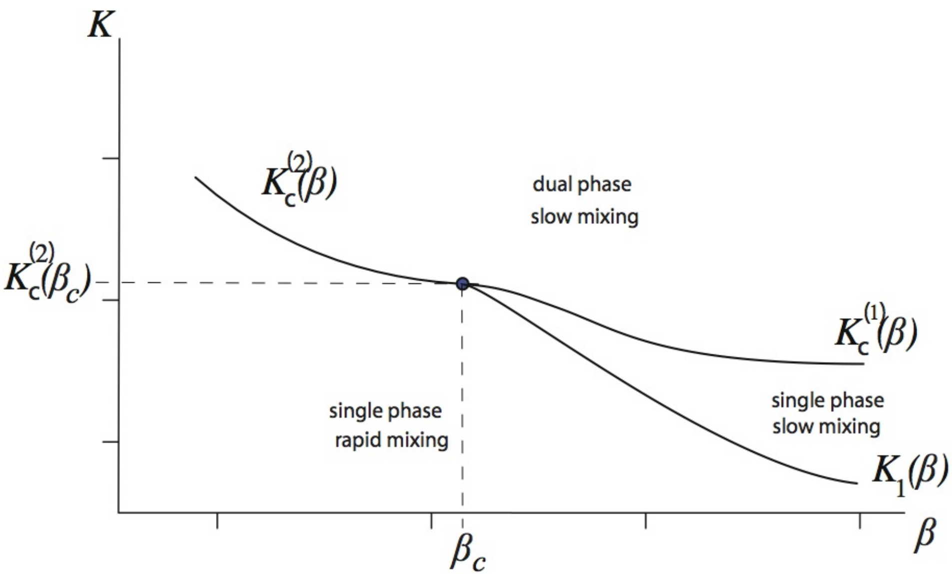
Slow mixing

Suppose $G_{\beta,K}$ has a minimum (either local or global) point at $\tilde{z} > 0$. Let z' be the corresponding local maximum point of $G_{\beta,K}$ such that $0 \leq z' < \tilde{z}$. Define the bottleneck set

$$A = \left\{ \omega : z' < \frac{S_n(\omega)}{n} \leq 1 \right\}$$

The bottleneck set A exists, and thus slow mixing, for (a) $\beta \leq \beta_c$ and $K > K_c^{(2)}(\beta)$, and (b) $\beta > \beta_c$ and $K > K_1(\beta)$.

Equilibrium structure versus mixing times



Y. K., P.T. Otto, and M. Titus in JSP 2011

Generalizing Aggregate Path Coupling Method

Configuration space:

Let q be a fixed integer and define $\Lambda = \{e^1, e^2, \ldots, e^q\}$, where e^k are the q standard basis vectors of $\mathbb{R}^q$, and $\omega = (\omega_1, \omega_2, \ldots, \omega_n) \in \Lambda^n$.

The magnetization vector (a.k.a empirical measure or proportion vector):

$$L_n(\omega) = (L_{n,1}(\omega), L_{n,2}(\omega), \ldots, L_{n,q}(\omega)),$$

where the kth component is defined by

$$L_{n,k}(\omega) = \frac{1}{n} \sum_{i=1}^{n} \delta(\omega_i, e^k)$$

- proportion of spins in configuration ω.

Generalizing Aggregate Path Coupling Method

Interaction representation function:

$$H(z) = H_1(z_1) + H_2(z_2) + \ldots + H_q(z_q)$$

Example: for the Curie-Weiss-Potts (CWP) model,

$$H(z) = -\frac{1}{2}\langle z, z\rangle = -\frac{1}{2}z_1^2 - \frac{1}{2}z_2^2 - \ldots - \frac{1}{2}z_q^2.$$

Hamiltonian: $H_n(\omega) = nH(L_n(\omega))$

Canonical ensemble:

$$P_{n,\beta}(B) = \frac{1}{Z_n(\beta)}\int_B \exp\left[-\beta H_n(\omega)\right] dP_n = \frac{1}{Z_n(\beta)}\int_B \exp\left[n\beta\, H\left(L_n(\omega)\right)\right] dP_n$$

where $P_n = \rho\times\ldots\times\rho$ and $Z_n(\beta) = \int\limits_{\Lambda^n} \exp\left[-\beta H_n(\omega)\right] dP_n$.

Generalizing Aggregate Path Coupling Method

Relative entropy: $R(\nu|\rho) = \sum_{k=1}^{q} \nu_k \log\left(\frac{\nu_k}{\rho_k}\right)$

Large deviations principle (R.S. Ellis, K. Haven, and B. Turkington, JSP 2000): The empirical measure L_n satisfies the LDP with respect to the Gibbs measure $P_{n,\beta}$ with rate function

$$I_\beta(z) = R(z|\rho) + \beta H(z) - \inf_t \{R(t|\rho) + \beta H(t)\}.$$

Equilibrium macrostates:

$$\mathcal{E}_\beta := \{\nu \in \mathcal{P} : \nu \text{ minimizes } R(\nu|\rho) + \beta H(\nu)\}$$

Generalizing Aggregate Path Coupling Method

Let ρ be the uniform distribution.

Logarithmic moment generating function of X_1:

$$\Gamma(z) = \log\left(\frac{1}{q}\sum_{k=1}^{q} \exp\{z_k\}\right)$$

Free energy functional for the Gibbs ensemble $P_{n,\beta}$:

$$G_\beta(z) = \beta(-H)^*(-\nabla H(z)) - \Gamma(-\beta\nabla H(z)),$$

where $F^*(z) = \sup_{x\in\mathbb{R}^q}\{\langle x, z\rangle - F(x)\}$ denotes Legendre-Fenchel transform.

Then (M. Costeniuc, R.S. Ellis, and H. Touchette JMP 2005)

$$\inf_{z\in\mathcal{P}}\{R(z|\rho) + \beta H(z)\} = \inf_{z\in\mathbb{R}^q}\{G_\beta(z)\}$$

and, therefore, $\mathcal{E}_\beta = \{z \in \mathcal{P} : z \text{ minimizes } G_\beta(z)\}$

Generalizing Aggregate Path Coupling Method

$$\mathcal{E}_\beta = \{z \in \mathcal{P} : z \text{ minimizes } G_\beta(z)\}$$

We consider only the **single phase region** of the Gibbs ensemble; i.e. values of β where $G_\beta(z)$ has a unique global minimum, at z_β. There,

$$P_{n,\beta}(L_n \in dx) \to \delta_{z_\beta} \quad \text{as} \quad n \to \infty$$

For the Curie-Weiss-Potts model, the single phase region are values of β such that

$$0 < \beta < \beta_c := \frac{2(q-1)}{q-2}\log(q-1)$$

Generalizing Aggregate Path Coupling Method

Glauber dynamics: (i) Select a vertex i uniformly,

(ii) Update the spin at vertex i according to the distribution $P_{n,\beta}$, conditioned on the event that the spins at all vertices not equal to i remain unchanged.

For configuration $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$, let σ_{i,e^k} be the configuration that agrees with σ at all vertices $j \neq i$ and s.t. the spin at the vertex i is e^k; i.e.

$$\sigma_{i,e^k} = (\sigma_1, \sigma_2, \ldots, \sigma_{i-1}, e^k, \sigma_{i+1}, \ldots, \sigma_n)$$

Then if the current configuration is σ and vertex i is selected, the probability the spin at i is updated to e^k, denoted by $P(\sigma \to \sigma_{i,e^k})$, is equal to

$$P(\sigma \to \sigma_{i,e^k}) = \frac{\exp\left\{ -\beta n H(L_n(\sigma_{i,e^k}))\right\}}{\sum_{\ell=1}^{q} \exp\left\{ -\beta n H(L_n(\sigma_{i,e^\ell}))\right\}}$$

Generalizing Aggregate Path Coupling Method

Let $g^{H,\beta}(z) := \left(g_1^{H,\beta}(z), \ldots, g_q^{H,\beta}(z)\right)$, where

$$g_\ell^{H,\beta}(z) = \frac{\exp\left(-\beta\,[\partial_\ell H](z)\right)}{\sum_{k=1}^{q} \exp\left(-\beta\,[\partial_k H](z)\right)}$$

Main Result (**Y.K.** and Peter T. Otto, JSP 2015): Let z_β be the unique equilibrium macrostate. Suppose $\exists \delta \in (0,1)$ s.t.

$$\frac{\inf_{\pi: z \to z_\beta} \sum_{k=1}^{q} \int_\pi \left| \left\langle \nabla g_k^{H,\beta}(y), dy \right\rangle \right|}{\|z - z_\beta\|_1} \leq 1 - \delta$$

for all z in $\mathcal{P}$. Then the mixing time of the Glauber dynamics satisfies

$$t_{mix} = O(n \log n)$$

Generalizing Aggregate Path Coupling Method

Application: the generalized Curie-Weiss-Potts model (GCWP).

In GCWP model (B. Jahnel, C. Külske, E. Rudelli, and J. Wegener, MPRF 2015), for $r \geq 2$,

$$H(z) = -\frac{1}{r}\sum_{j=1}^{q} z_j^r$$

Then $g_k^{H,\beta}(z) = \frac{e^{\beta z_k^{r-1}}}{e^{\beta z_1^{r-1}}+\ldots+e^{\beta z_q^{r-1}}}$. Next, define

$$\beta_s(q,r) := \sup\left\{\beta \geq 0 : g_k^{H,\beta}(z) < z_k \text{ for all } z \in \mathcal{P} \text{ such that } z_k \in (1/q, 1]\right\}$$

Corollary (**Y.K.** and Peter T. Otto, JSP 2015):
If $\beta < \beta_s(q,r)$, then

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

Generalizing Aggregate Path Coupling Method

Remark: Rapid mixing region for classical Curie-Weiss-Potts model (GCWP with $r = 2$) was first obtained (among other things) in P. Cuff, J. Ding, O. Louidor, E. Lubetzy, Y. Peres and A. Sly, JSP 2012.

The rapid mixing region for GCWP was obtained in **Y.K.** and Peter T. Otto, JSP 2015 as a few page Corollary to the Main Result.

Note that

$$\beta_s(q,r) \leq \beta_c(q,r)$$

Here, the region of rapid mixing $\beta < \beta_s(q,r)$ is where

$$G_\beta(z) = \beta(-H)^*(-\nabla H(z)) - \Gamma(-\beta \nabla H(z))$$

has a unique local minimum (=global minimum).