Swendsen-Wang beats Heat-bath

Mario Ullrich

Friedrich Schiller University Jena

Borovets, Bulgaria, 29th August 2011
Overview

1 Introduction
   - Potts model
   - Random cluster model

2 Results
   - Main result
   - New algorithm
Problem:
Can one construct a rapidly mixing Markov chain for the Potts model at all temperatures?
For $q \in \mathbb{N}$ and a (finite) graph $G = (V, E)$, $N := |V|$, the $q$-state Potts model on $G$ is defined as the set of possible configurations $\Omega_P = \{1, \ldots, q\}^V$ together with the probability measure

$$\pi_\beta(\sigma) := \frac{1}{Z(G, \beta, q)} \exp\left\{ \beta \cdot \#\left\{ \{u, v\} \in E : \sigma(u) = \sigma(v) \right\} \right\}$$

for $\sigma \in \Omega_P$, where $Z$ is the normalization constant and $\beta \geq 0$ is the inverse temperature.
The most studied Markov chain is the *heat bath dynamics*. This is the Markov chain with transition probabilities

\[
P_{HB}(\sigma, \sigma^v, k) := P_{HB,\beta,q}(\sigma, \sigma^v, k) = \frac{1}{N} \frac{\pi_{\beta}(\sigma^v, k)}{\sum_{l=1}^{q} \pi_{\beta}(\sigma^v, l)},
\]

where \(\sigma^v, k(v) = k\) and \(\sigma^v, k(u) = \sigma(u), u \neq v\).
For the transition matrix $P$ of a Markov chain on state space $\Omega$ with stationary distribution $\pi$, we define the operator $P : L_2(\pi) \rightarrow L_2(\pi)$ by

$$Pf(x) := \sum_{y \in \Omega} P(x, y) f(y).$$

If $1 = \xi_1 \geq \xi_2 \geq \cdots \geq \xi_{|\Omega|} \geq -1$ are the (real) eigenvalues of the operator $P$, we define the spectral gap of the Markov chain by

$$\lambda(P) = 1 - \max\{\xi_2, |\xi_{|\Omega|}|\}.$$
Let \( \{P_n\}_{n \in \mathbb{N}} \) be a family of Markov chains with corresponding state spaces \( \Omega_n \). Then we call the Markov chains rapidly mixing (for \( \{\Omega_n\} \)), if

\[
\lambda(P_n)^{-1} = O\left((\log |\Omega_n|)^C\right)
\]

for some \( C \geq 0 \) and all \( n \in \mathbb{N} \).

In "our" case let \( \{G_n\}_{n \in \mathbb{N}} \) be a family of graphs, then \( \log |\Omega_n| = \log q \cdot |V_{G_n}| \).
Mixing for heat bath

An example of physical interest:
The two-dimensional square lattice \( \mathbb{L}_N = (V, E) \) with
\( V = \{1, \ldots, L\}^2 \), \( N = L^2 \), and
\( E = \{(v, w) \in V^2 : |v - w| = 1\} \).

It is well known that \( P_{\text{HB}} = P_{\text{HB}}^{\mathbb{L}_N} \), \( \beta, q \) satisfies

\[
\lambda(P_{\text{HB}})^{-1} = O(N), \quad \text{if } \beta < \beta_c(q) := \ln(1 + \sqrt{q}).
\]

\[
\lambda(P_{\text{HB}})^{-1} = e^{\Omega(N)}, \quad \text{if } \beta > \beta_c(q).
\]

\[
\lambda(P_{\text{HB}})^{-1} = O\left(N^C\right), \quad \text{if } q = 2 \text{ and } \beta = \beta_c(2),
\]

for some \( C > 0 \).

(Martinelli et al. ('94), Cesi et al. ('96), Lubetzky & Sly (2010),
Beffara & Duminil-Copin (2010))
Random cluster model

The *random cluster model* on $G = (V, E)$ has the state space $\Omega_{RC} = \{\omega \subseteq E\}$. So $\omega \in \Omega_{RC}$ induces a subgraph $(V, \omega)$ of $(V, E)$.

The probability measure on $\Omega_{RC}$ is given by

$$\mu_p(\omega) = \frac{1}{Z} p^{\lvert \omega \rvert} (1 - p)^{|E| - |\omega|} q^{C(\omega)},$$

where $C(\omega)$ is the number of connected components in $(V, \omega)$. 
Connection of the models

In the case \( p = 1 - e^{-\beta} \), the Potts model and the random cluster model are equivalent in the following sense.

We can define random mappings \( T : \Omega_P \leftrightarrow \Omega_{RC} \) and \( T^* : \Omega_{RC} \leftrightarrow \Omega_P \) such that

\[
X \sim \pi_\beta \Rightarrow T(X) \sim \mu_p
\]

and

\[
Y \sim \mu_p \Rightarrow T^*(Y) \sim \pi_\beta.
\]

We consider the Markov chain

\[
\sigma_{t+1} = T^* \circ T(\sigma_t).
\]
The most widely *used* algorithm is the *Swendsen-Wang dynamics*. For $\sigma \in \Omega_P$ let

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

The chain performs the following steps:

1. Given a Potts configuration $\sigma_t \in \Omega_P$ on $G$, delete each edge of $E(\sigma_t)$ independently with probability $1 - p = e^{-\beta}$. This gives $\omega \in \Omega_{RC}$.

2. Assign a random color independently to each connected component of $(V, \omega)$. Vertices of the same component get the same color. This gives $\sigma_{t+1} \in \Omega_P$. 

Mario Ullrich

SW vs. heat-bath
Empirically, SW is much faster than HB, but there are not many results on the mixing properties of SW.

Positive results are only known for special classes of graphs (trees, cycles, complete graph etc.) or for high enough temperatures.

A negative result: Slow mixing on $\mathbb{Z}^d$, $d \geq 2$, at $\beta \approx \beta_c(q)$ for $q$ large enough. (Borgs, Chayes & Tetali (2010))
Main result

**Theorem (U 2011)**

Suppose that $P$ (resp. $P_{\text{HB}}$) is the transition matrix of the Swendsen-Wang (resp. heat-bath) dynamics, which is reversible with respect to $\pi^{G}_{\beta,q}$. Then

$$\lambda(P) \geq c_{\text{SW}} \lambda(P_{\text{HB}}),$$

where

$$c_{\text{SW}} = c_{\text{SW}}(G, \beta, q) := \frac{1}{2q^2} \left(q e^{2\beta}\right)^{-4\Delta},$$

where $\Delta$ is the maximal degree of $G$. 

Mario Ullrich

SW vs. heat-bath
Corollary

With this theorem we get new and old results on SW (up to a factor $N$). Especially, we get rapid mixing for the two-dimensional square lattice $\mathbb{L}_N$:

**Corollary (Square lattice $\mathbb{L}_N$)**

Let $P = P_{p,q}^{\mathbb{L}_N}$. Then

- $\lambda(P)^{-1} = O(N)$, if $\beta < \beta_c(q) := \ln(1 + \sqrt{q})$.
- $\lambda(P)^{-1} = O(N^C)$, if $q = 2$ and $\beta = \beta_c(2)$. 

Mario Ullrich

SW vs. heat-bath
The bounds are (probably) off by a factor $N$, because we compare a local and a highly non-local algorithm.

One may expect rapid mixing also at low temperatures. We were not able to prove this, but to modify the algorithm.
Let $G = (V, E)$ be a finite, planar graph and $G_D = (V_D, E_D)$ its dual graph. Then, to each RC configuration $\omega \in \Omega_{RC}$ there corresponds the dual configuration $\omega_D \subseteq E_D$, given by

$$e_D \in \omega_D \iff e \notin \omega.$$
Duality

It is easy to obtain (using Euler’s polyhedron formula)

\[ \mu_{p,q}^G(\omega) = \mu_{p^*,q}^{G_D}(\omega_D), \]

where the dual parameter \( p^* \) satisfies

\[ \frac{p^*}{1 - p^*} = \frac{q(1 - p)}{p}. \]
Let $\tilde{P} = \tilde{P}^G_{p,q}$ be the SW algorithm on the random cluster model.

The dynamics performs the following steps:

1. Given a Potts configuration $\sigma_t$ on $G$, generate a random cluster state $\omega \subset E_G$ by $\omega = T(\sigma_t)$.
2. Make one step of the Swendsen-Wang dynamics $\tilde{P}^{G_D}_{p^*,q}$ starting at $\omega_D \subset E_{G_D}$ to get a random cluster state $\tilde{\omega}_D \subset E_{G_D}$.
3. Generate $\sigma_{t+1}$ by $T^*(\tilde{\omega})$.

Denote by $M$ the transition matrix of this Markov chain.
Result

**Proposition (U 2011)**

Let $P_{p,q}^G$ be the Swendsen-Wang dynamics on a planar graph $G$, which is reversible with respect to $\pi_{\beta,q}^G$, and $M$ as above. Then

$$\lambda(M) \geq \max\{\lambda(P_{p,q}^G), \lambda(P_{p^*,q}^{G_D})\}.$$ 

We obtain for the square lattice $\mathbb{L}_N$:

- $\lambda(M)^{-1} = \mathcal{O}(N)$, if $\beta \neq \beta_c(q) := \ln(1 + \sqrt{q})$.
- $\lambda(M)^{-1} = \mathcal{O}(N^C)$, if $q = 2$ and $\beta = \beta_c(2)$. 
Partial answers

**Problem:**
Can one construct a rapidly mixing Markov chain for the Potts model at all temperatures?

For the square lattice $\mathbb{L}_N$:

- Yes, for $q = 2$.
- Yes, for $q > 2$ and $\beta \neq \beta_c(q)$.
- Probably not for $\beta = \beta_c(q)$ and $q$ large enough.
Thank you!