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

The Problem

Problem:

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

The Potts Model

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* $\Omega_{\mathrm{P}} = \{1, \dots, q\}^V$ together with the probability measure

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

for $\sigma \in \Omega_{\mathrm{P}}$, where Z is the normalization constant and $\beta \geq 0$ is the inverse temperature.

The Dynamics I

The most studied Markov chain is the *heat bath dynamics*.
This is the Markov chain with transition probabilities

$$P_{\mathrm{HB}}(\sigma, \sigma^{v,k}) \;:=\; P^{G}_{\mathrm{HB},\beta,q}(\sigma, \sigma^{v,k}) \;=\; \frac{1}{N}\,\frac{\pi_\beta(\sigma^{v,k})}{\sum_{l=1}^{q} \pi_\beta(\sigma^{v,l})},$$

where $\sigma^{v,k}(v) = k$ and $\sigma^{v,k}(u) = \sigma(u)$, $u \neq v$.

Rapid Mixing

For the transition matrix P of a Markov chain on state space Ω with stationary distribution π, we define the operator $P: L_2(\pi) \to L_2(\pi)$ 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

$$\lambda(P) = 1 - \max\{\xi_2, |\xi_{|\Omega|}|\}.$$

Rapid Mixing

Let $\{P_n\}_{n\in\mathbb{N}}$ be a *family of Markov chains* with corresponding state spaces Ω_n. Then we call the Markov chains *rapidly mixing* (for $\{\Omega_n\}$), if

$$\lambda(P_n)^{-1} \;=\; \mathcal{O}\left((\log|\Omega_n|)^C\right) \qquad \text{for some } C \geq 0 \text{ and all } n \in \mathbb{N}.$$

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 $P_{\mathrm{HB}} = P_{\mathrm{HB},\beta,q}^{\mathbb{L}_N}$ satisfies

- $\lambda(P_{\mathrm{HB}})^{-1} = \mathcal{O}(N)$, if $\beta < \beta_c(q) := \ln(1 + \sqrt{q})$.
- $\lambda(P_{\mathrm{HB}})^{-1} = e^{\Omega(N)}$, if $\beta > \beta_c(q)$.
- $\lambda(P_{\mathrm{HB}})^{-1} = \mathcal{O}\left(N^C\right)$, if $q = 2$ 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 $\Omega_{\mathrm{RC}} = \{\omega \subseteq E\}$. So $\omega \in \Omega_{\mathrm{RC}}$ induces a subgraph (V, ω) of (V, E).

The probability measure on Ω_{RC} is given by

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

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

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 $T : \Omega_{\mathrm{P}} \mapsto \Omega_{\mathrm{RC}}$ and $T^* : \Omega_{\mathrm{RC}} \mapsto \Omega_{\mathrm{P}}$ such that

$$X \sim \pi_\beta \implies T(X) \sim \mu_p$$

and

$$Y \sim \mu_p \implies T^*(Y) \sim \pi_\beta.$$

We consider the Markov chain

$$\sigma_{t+1} = T^* \circ T(\sigma_t).$$

The Dynamics II

The most widely *used* algorithm is the *Swendsen-Wang dynamics*. For $\sigma \in \Omega_{\mathrm{P}}$ let

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

The chain performs the following steps:

1. Given a Potts configuration $\sigma_t \in \Omega_{\mathrm{P}}$ on G, delete each edge of $E(\sigma_t)$ independently with probability $1-p = e^{-\beta}$. This gives $\omega \in \Omega_{\mathrm{RC}}$.
2. Assign a random color independently to each connected component of (V, ω). Vertices of the same component get the same color. This gives $\sigma_{t+1} \in \Omega_{\mathrm{P}}$.

Mixing for Swendsen-Wang

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

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

where

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

where Δ is the maximal degree of G.

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

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

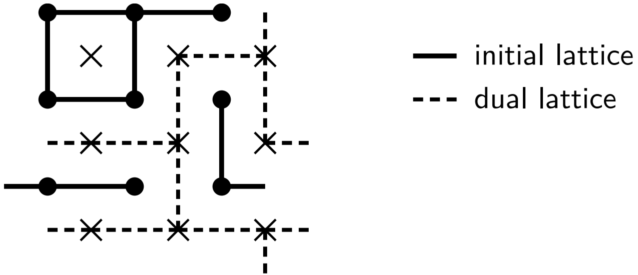
Comments

- 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.

Dual Lattices

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

$$e_D \in \omega_D \quad \Longleftrightarrow \quad e \notin \omega.$$



Duality

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

$$\mu^{G}_{p,q}(\omega) \;=\; \mu^{G_D}_{p^*,q}(\omega_D),$$

where the dual parameter p^* satisfies

$$\frac{p^*}{1-p^*} \;=\; \frac{q\,(1-p)}{p}.$$

The algorithm

Let $\widetilde{P} = \widetilde{P}_{p,q}^{G}$ be the SW algorithm on the random cluster model.

The dynamics performs the following steps:

1. Given a Potts configuration σ_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 $\widetilde{P}_{p^*,q}^{G_D}$ starting at $\omega_D \subset E_{G_D}$ to get a random cluster state $\widetilde{\omega}_D \subset E_{G_D}$.
3. Generate σ_{t+1} by $T^*(\widetilde{\omega})$.

Denote by M the transition matrix of this Markov chain.

Result

Proposition (U 2011)

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

$$\lambda(M) \;\geq\; \max\left\{\lambda(P^G_{p,q}),\, \lambda(P^{G_D}_{p^*,q})\right\}.$$

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

- $\lambda(M)^{-1} \;=\; \mathcal{O}\left(N\right),$ if $\beta \neq \beta_c(q) := \ln(1+\sqrt{q})$.
- $\lambda(M)^{-1} \;=\; \mathcal{O}\left(N^C\right),$ 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!