Bounding Chains for Swendsen-Wang

Outline

- The Potts and RCM models
- The Swendsen-Wang algorithm
- Coupling, bounding chains
- Theorems
- Simulations, plots

The Potts Model

$G=(V,E),\ Q=\{1,2,...,q\},\ X:V^{|V|}\rightarrow Q^{|V|},\ \mathrm{X}=(\mathrm{x(i)})_{\mathrm{i}\in\mathrm{V}}$

The probability distribution:

$$P_{Pt}(X)=\frac{1}{Z_{Potts}}\exp(\sum_{(i,j)\in E}J_{ij}(\delta_{x(i),x(j)}-1))$$

Simplified Ising model:

$$Q=\{0,1\},P_{Pt}(X)=\frac{1}{Z}\exp(-\beta\sum_{(i,j)\in E}(1-\delta_{x(i),x(j)}))$$

If β>0 → *ferromagnetic* model
β<0 → *antiferromagnetic* model

The Random Cluster Model

$N:E\rightarrow\{0,1\}^{|E|},N=(n_{ij})_{(i,j)\in E}$

n_{ij} are bond variables with the probability distribution:

$$P_R(N)=\frac{1}{Z_{RCM}}(\prod_{(i,j):n_{ij}=1}p_{ij})(\prod_{(i,j):n_{ij}=0}(1-p_{ij}))q^{C(N)}$$

where c(N) is the # of connected components in the graph whose edges are the bonds having n_{ij}=1.

Taking p_{ij}=p:

$$P_R(N)=p^{|A|}(1-p)^{|E|-|A|}q^{c(A)},A=\{(i,j)\in E\,|\,n_{ij}=1\}$$

Common Representation For The Potts And RCM Models

A joint probability distribution:

$$P(X,N)=\frac{1}{Z_{PRCM}}\prod_{(i,j)\in E}[(1-p_{ij})\delta_{n_{ij},0}+p_{ij}\delta_{n_{ij},1}\delta_{x(i),x(j)}]$$

Properties Of The Joint Model

i) $Z_{Potts} = Z_{RCM} = Z_{PRCM.}$

ii) The marginal over X is exactly $P_{Pt}(X)$.

iii) The marginal over N is exactly $P_R(N)$.

iv) P(N|X) is : if x(i)=x(j) then $n_{ij}=0$, else $n_{ij}\sim Be(p_{ij})$.

v) P(X|N) is : independently, for each connected cluster, a color is selected uniformly from {1,2,…,q}.

The Swendsen-Wang Algorithm

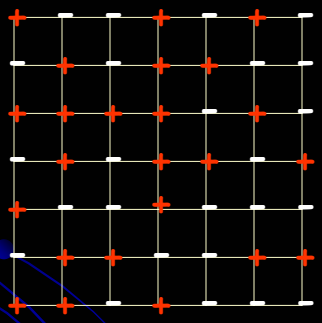
Two steps:

1) Edge creation phase: Between each two vertices colored identically an edge is created with probability $p=1-\exp(-\beta)$. Therefore, N_t is sampled from $P(N|X_t)$.

2) Coloring phase: given the current configuration of "on" edges, each cluster receives independently one color from Q. Or X_t gets sampled from $P(X|N_{t-1})$.

After $(X_t, N_t)_{t>0}$ converges to the stationary distribution given by P(X,N), one can get samples from both the Potts and the RCM models.

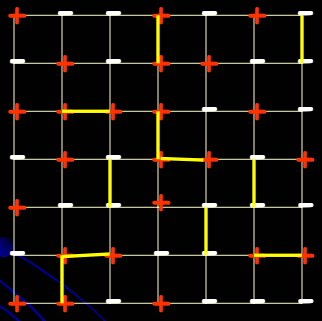
Swendsen-Wang Algorithm



An arbitrary Ising configuration according to

$$P(X) \sim e^{\beta \sum_{(i,j)} \delta_{x(i),x(j)}}$$

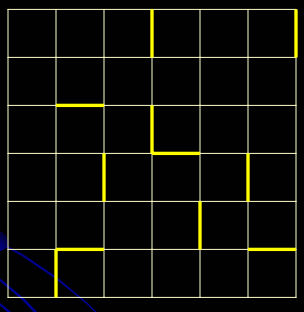
Swendsen-Wang Algorithm



Create a bond with probability $p = 1-e^{-\beta}$, if x(i)=x(j).

$$P(X,N) \sim \prod_{(i,j)} \left[(1-p)\delta_{n_{ij},0} + p\delta_{x(i),x(j)}\delta_{n_{ij},1} \right]$$

Swendsen-Wang Algorithm

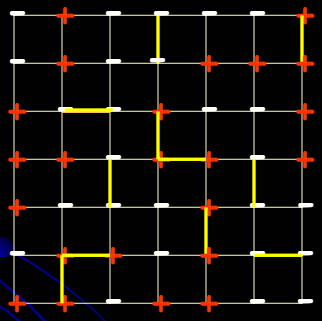


Remove the colors of the sites.

$$P(N) \sim \sum_{\{x\}} \prod_{(i,j)} \left[(1-p)\delta_{n_{ij},0} + p\delta_{x(i),x(j)}\delta_{n_{ij},1} \right]$$

$$= p^{|A|}(1-p)^{|E|-|A|} q^{N(A)}$$

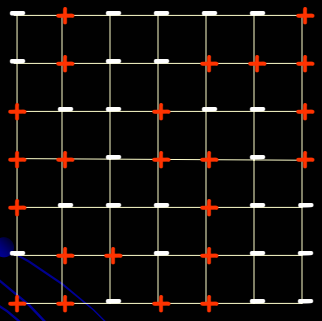
Swendsen-Wang Algorithm



Assign a new color for each cluster at random from Q. Isolated sites are also considered clusters.

Go back to P(X,N).

Swendsen-Wang Algorithm



Remove the bonds to finish one sweep.

Go back to P(X).

Perfect (Exact) Sampling Protocols

Coupling Markov Chains:

- Coupling from the past (CFTP) (Propp, Wilson 1996)
- Fill's algorithm (Fill, Murdoch, Rosenthal 1999)
- Read-Once CFTP (Wilson 2000)

Modified Acceptance/Rejection

- Popping Algorithms (Propp, Wilson 1998)
- Randomness Recycler (Fill, Huber 2001)

Techniques For Building A Good Coupler

Monotonicity (Propp, Wilson 1996)

Multigamma coupling (Murdoch, Green 1998)

Bounding chains (Häggström, Nelander 1999)
(Huber 1999, 2004)

Multishift coupling (Wilson 2000)

Mixing Time

- TV distance: $\| p - \pi \|_{TV} = \frac{1}{2} \sum_x | p(x) - \pi(x) |$
- Mixing time: $\tau(\varepsilon) = \max_x \min\{t : \| P^t(x,\cdot) - \pi \| \le \varepsilon\}$

Definition: X_t is rapidly mixing if $\tau(\varepsilon) \le \text{poly}[n, \ln(1/\varepsilon)]$

Note: We can bound the mixing time by coupling.

Coupling

Definition: A coupling is a MC on $\Omega \times \Omega$ s.t.

1) Each process $\{X_t\}_t$, $\{Y_t\}_t$ is a copy of the original MC,
2) If $X_t = Y_t$ then $X_{t+1} = Y_{t+1}$.

The (complete) coupling time T is:

$T = \max_{x,y} (E [T(x,y)])$, where

$T(x,y) = \min \{ t : X_t = Y_t \mid X_0 = x, Y_0 = y \}$.

Theorem: $\tau(\varepsilon) \le T \, e \ln(1/\varepsilon)$

Coupling

In a ***complete coupling*** a MC step → random function $f:\Omega \to \Omega$ satisfying $P(f(X_t) = x) = P(X_{t+1} = x \mid X_t)$.

At time t, $X_t = f_t \circ f_{t-1} \circ \ldots \circ f_0(X_0) := F_t(X_0)$.

The chains coupled if: $F_t(X_0) = F_t(Y_0) \to F_{t'}(X_0) = F_{t'}(Y_0)$, for all t'>t.

Definition: A chain has completely coupled if F_t is a constant.

Bounding Chains

Definition: Let M_1 on C^V and M_2 on $P(C)^V$ using random sequences f_t and g_t. Then M_2 is a *BC* for M_1 if :

$$\text{i) } \forall\ t,\ x \in y \Rightarrow f_t(x) \in g_t(y),$$

$$\text{ii) } \forall\ v,\ |y(v)|=1 \Rightarrow g_t(y)(v) = f_t(y)(v)$$

In other words, if X_t is in Y_t, then X_{t+1} is in Y_{t+1} and if the color set is a single color at each vertex, then both chains evolve identically.

BC are a method for detecting complete coupling and consequently for obtaining exact samples.

Theorems

Theorem 1 (high temperature case) (Huber 2001):

If $p < \frac{1}{2(\Delta-1)}$, then the Swendsen-Wang bounding chain couples completely by time $O(\ln(2m))$ with probability at least 1/2.

Here Δ= maximum degree of a vertex and $m=|E|$.

Note: Similar results were obtained by Cooper and Frieze (1999)

Theorem 2 (low temperature case) (Huber 2001):

Suppose that $p \geq 1-\frac{1}{mQ}$. Then SW couples completely with probability at least 1/2 in time $2(mQ)^2$.

Theorems cont'd

Note: We believe that the second theorem is incorrect. The correct result should be the one below.

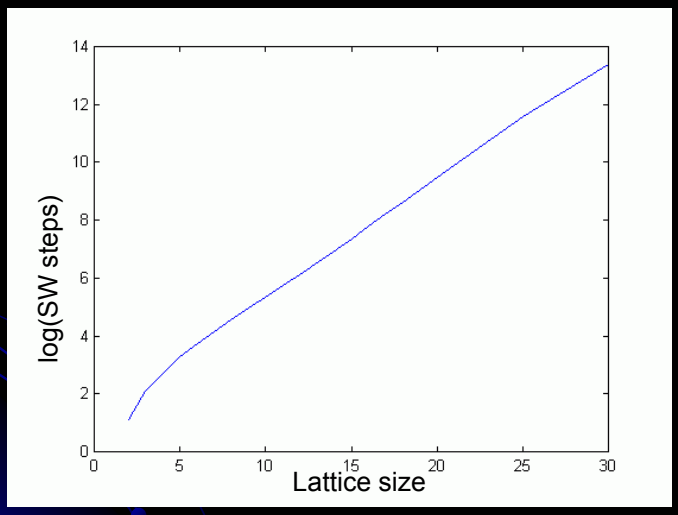
Theorem 2':

If $p \geq 1-\frac{1}{2mQ}$, then SW couples completely with probability at least 1/2 in time $O(\ln(2mQ))$.

Our simulations agree with this finding.

Simulations

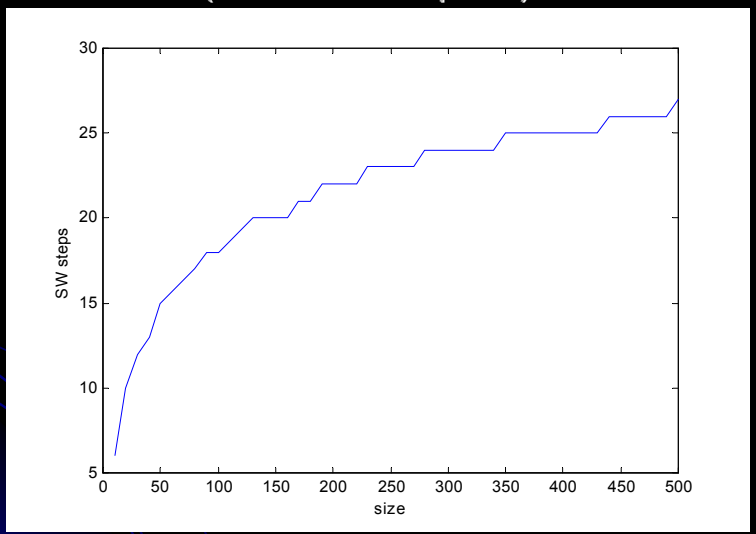
Version from first paper was exponential:



Then we implemented the second paper, which is good.

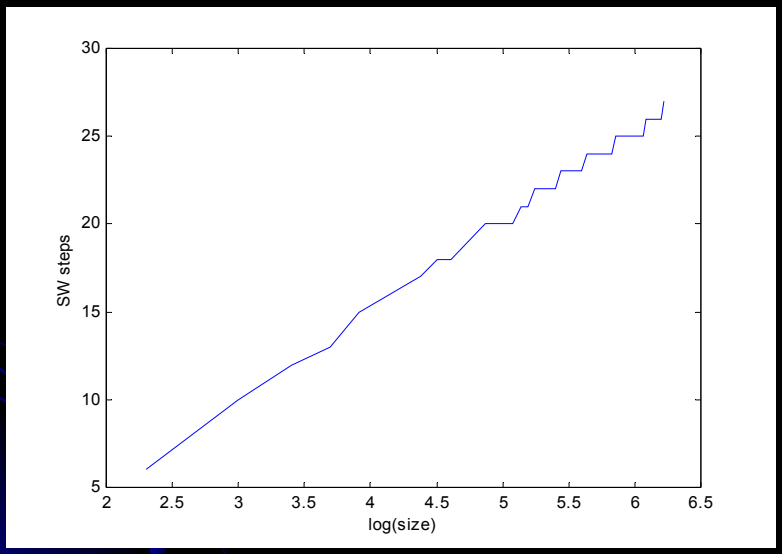
SW steps at p = 0.3

(below critical point)



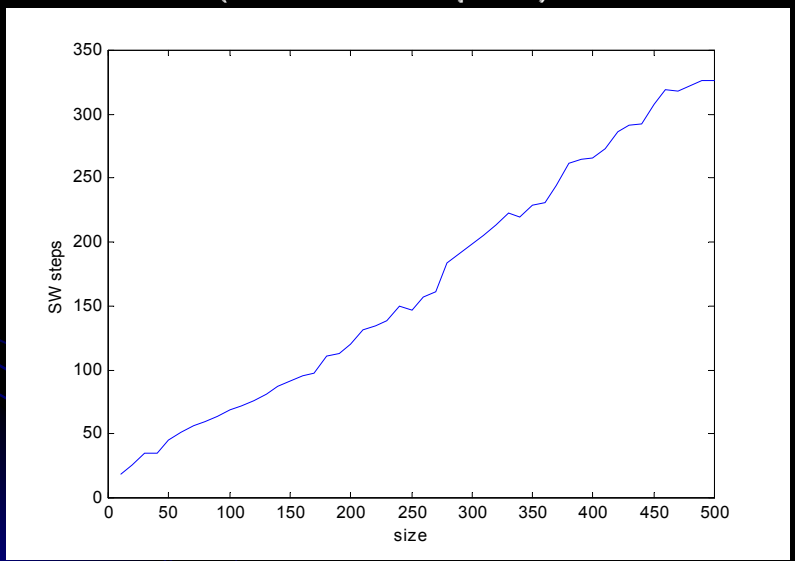
We suspect logarithmic dependency on size.

SW steps vs. Log(size) at p = 0.3



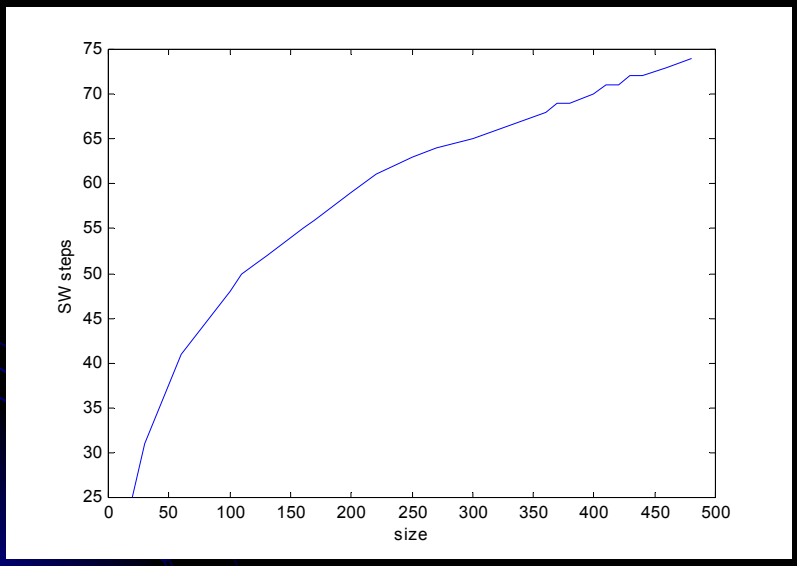
Indeed, dependence with log(size) seems linear

SW steps at p = 0.99
(above critical point)



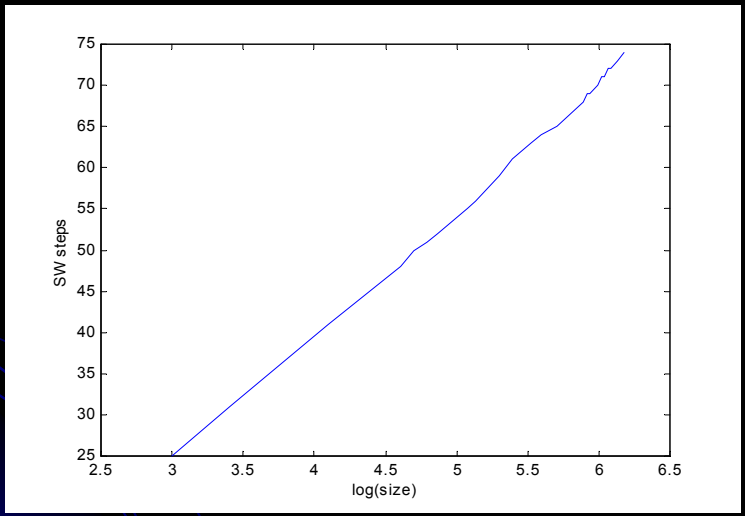
Dependence seems linear with size

SW steps at $p = 1-1/4n^2$
to Check Theorem 2 and 2'



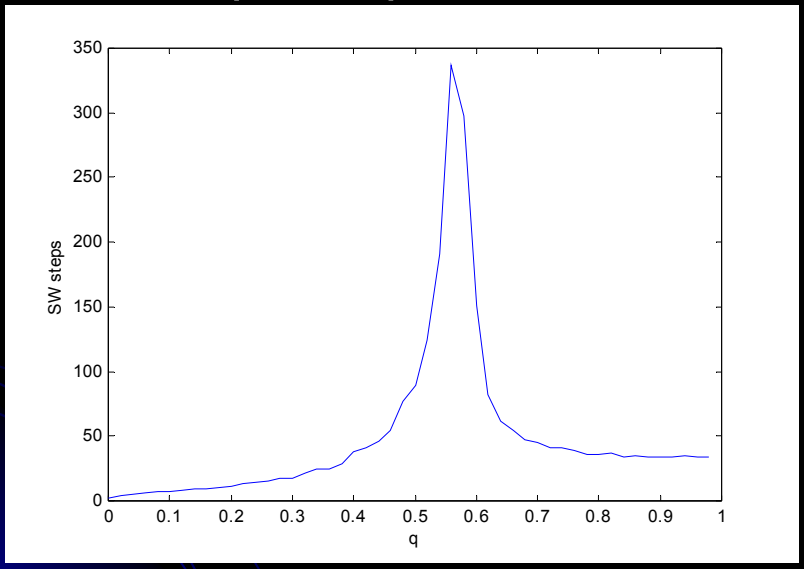
Dependence seems logarithmic with size n

SW steps vs. Log(size), $p=1-1/4n^2$
to Check Theorem 2 and 2'



Indeed, dependence seems linear with log(n), so Theorem 2' seems to hold.

SW steps vs. p, 200x200 lattice

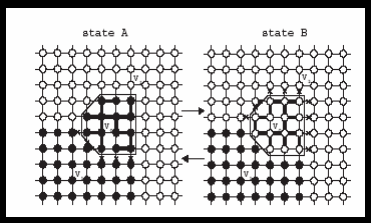


Increases dramatically around critical point ~0.58.

Swendsen-Wang
Simulations

Swendsen-Wang

(Barbu, Zhu 03)

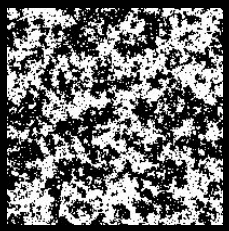


Swendsen-Wang: flips a cluster of spins at one time.

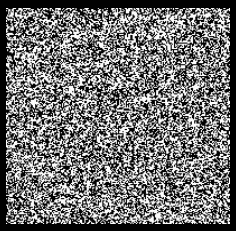
Experiments with 200x200 lattice:

Low T = 150
p = 0.67
coupling t = 18

T = 210
p = 0.55
coupling t = 19



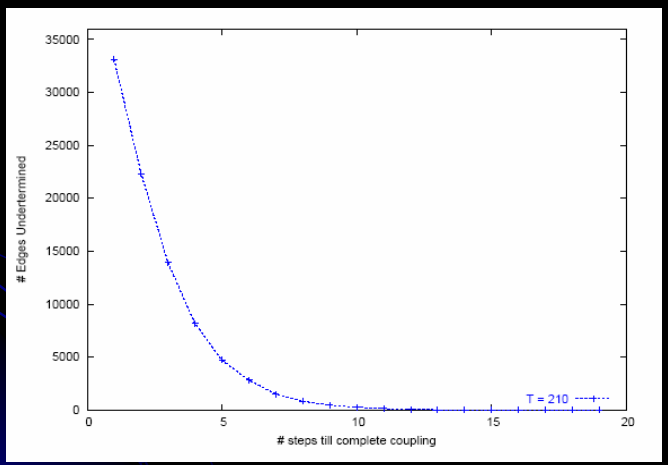
High T = 1000
p = 0.15
coupling t = 12



Bounding Chain

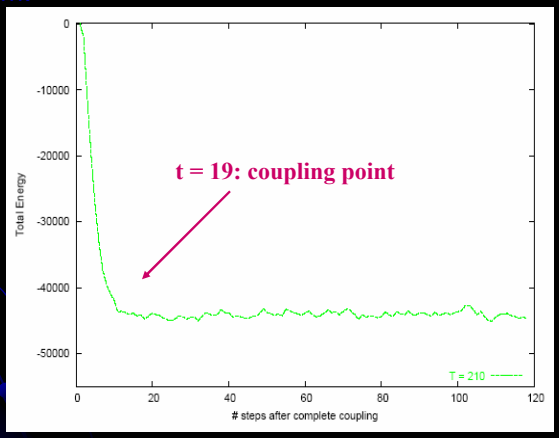
(Mark Huber 02)

- Bounding chains are means of detecting complete coupling.



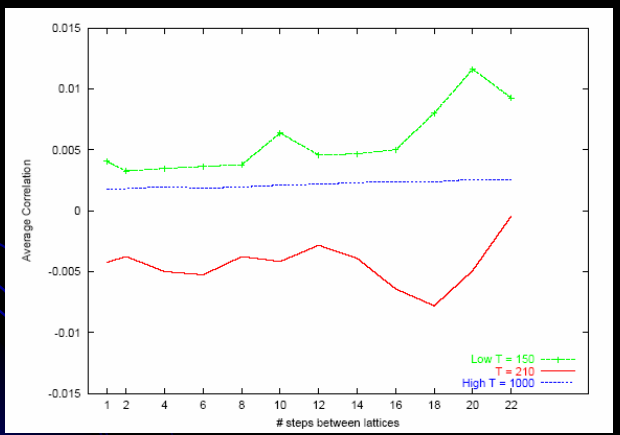
NO Critical Slowing Down

- At critical temperature, dynamics gets very slow if we use any local update method (e.g. Metropolis Importance Sampling).
- Being a cluster-based algorithm, Swendsen-Wang has less such problem.



Cluster (Non-Local) Algorithm

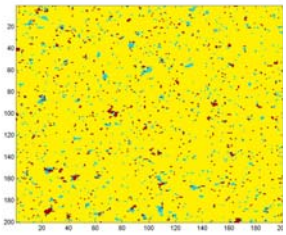
- Spins are less correlated at all Temperatures.
- Correlation: $\chi(r-r')=\langle s(r)s(r')\rangle-\langle s(r)\rangle\langle s(r')\rangle$.



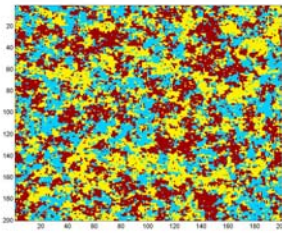
SW on Potts Model

Experiments with 200x200 lattice:

Low T = 150
p = 0.67
coupling t = 43



T = 180
p = 0.60
coupling t = 34



High T = 250
p = 0.49
coupling t = 31