Theory Working Group

29 Oct, 2008

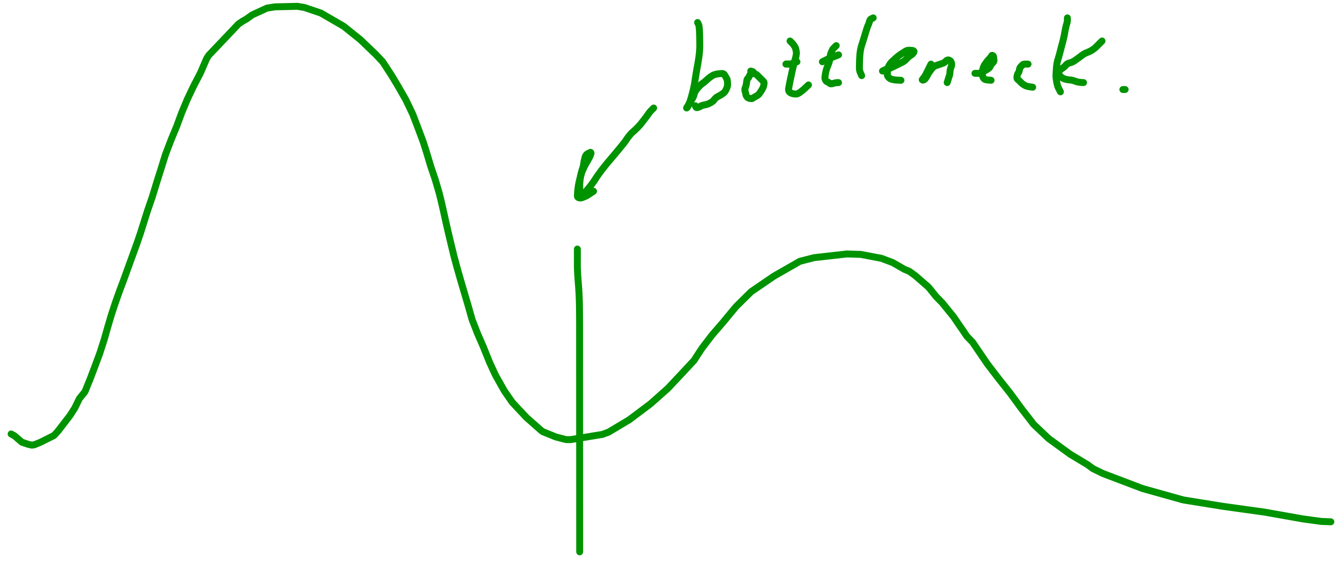
Parallel & Simulated Tempering

Plan for next few weeks

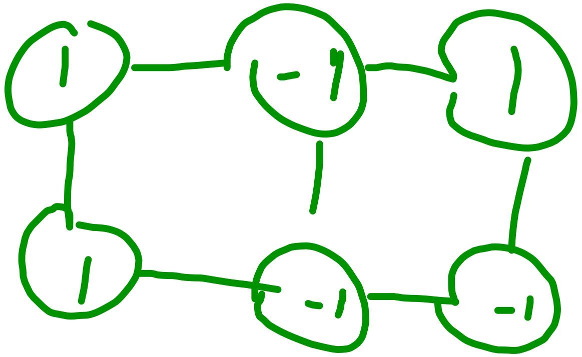
- Today Introduce Parallel & Simulated Tempering
- Present Conditions under which they are slow

- Present work of Eberle & Marinelli on Sequential Methods
- Open Question: Can show E&M runs quickly on problems where p. temp. is slow

Problem



Ising Model



$H(x) = \#$ of edges given different spin.

$w(x) = e^{-\beta H(x)}$

Cooling schedule

$$0 = \beta_0 < \beta_1 < \cdots < \beta_K = \infty$$

Simulated Annealing

Have Markov Transition kernels for every β

Run for a while using K_{β_0}, then K_{β_1}, K_{β_2}, ...

Pictorially

Biggest problem: How to choose cooling schedule!

Treat β as an auxiliary variable.

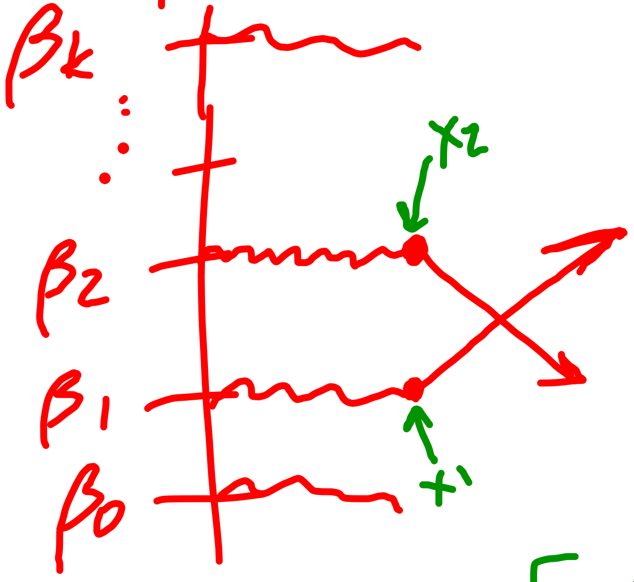
$$x \to x'$$

$$\beta \to \beta'$$

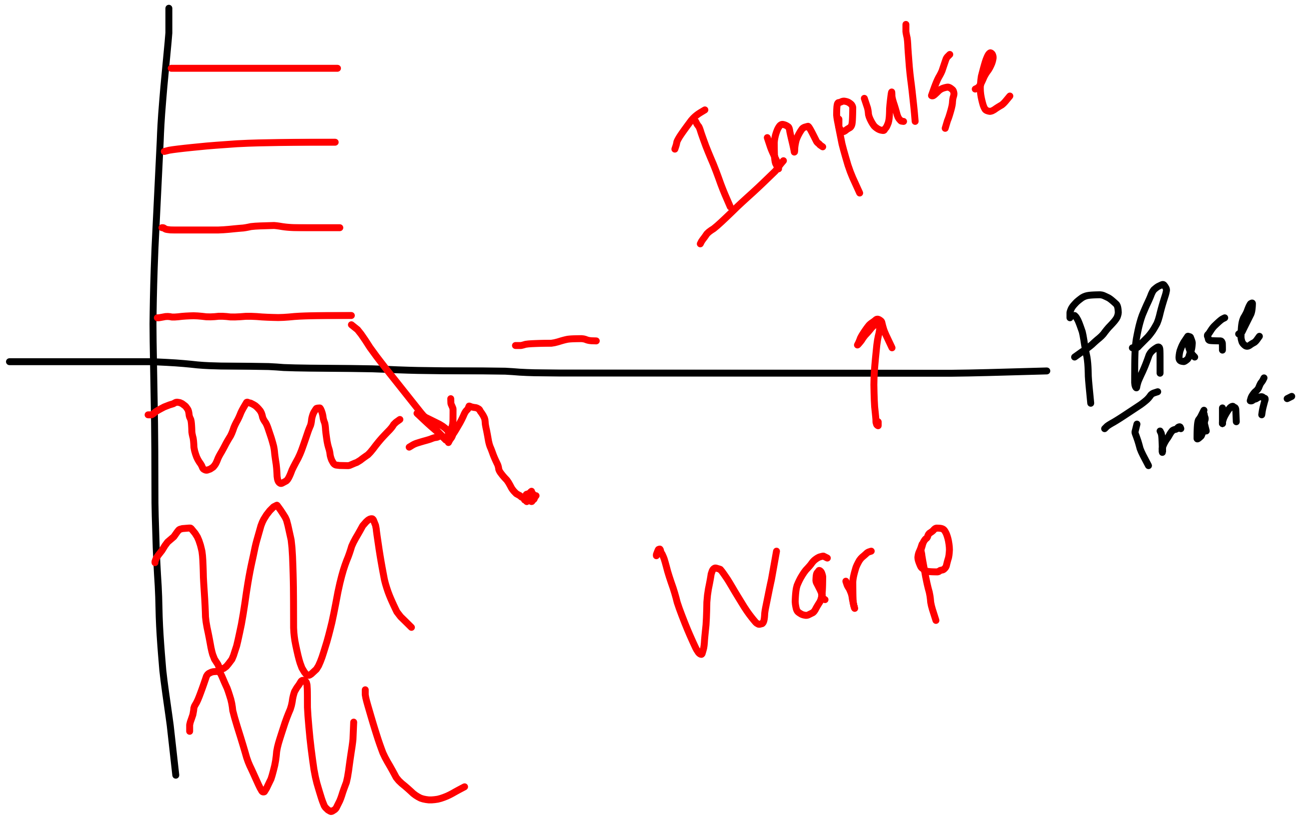
M. ratio $\beta \to \beta'$

$$\frac{w_{\beta'}(x) / Z_{\beta'}}{w_{\beta}(x) / Z_{\beta}}$$

Parallel Tempering



M-H ratio:
$$\frac{\left[w_{\beta_1}(X_2)/\ldots\right]\cdot\left[w_{\beta_2}(X_1)/\ldots\right]}{\left[w_{\beta_1}(X_1)/\ldots\right]\left[w_{\beta_2}(X_2)/\ldots\right]}$$


Impulse
Phase Trans.
Warp

Potts model

Like Ising

k different colors.

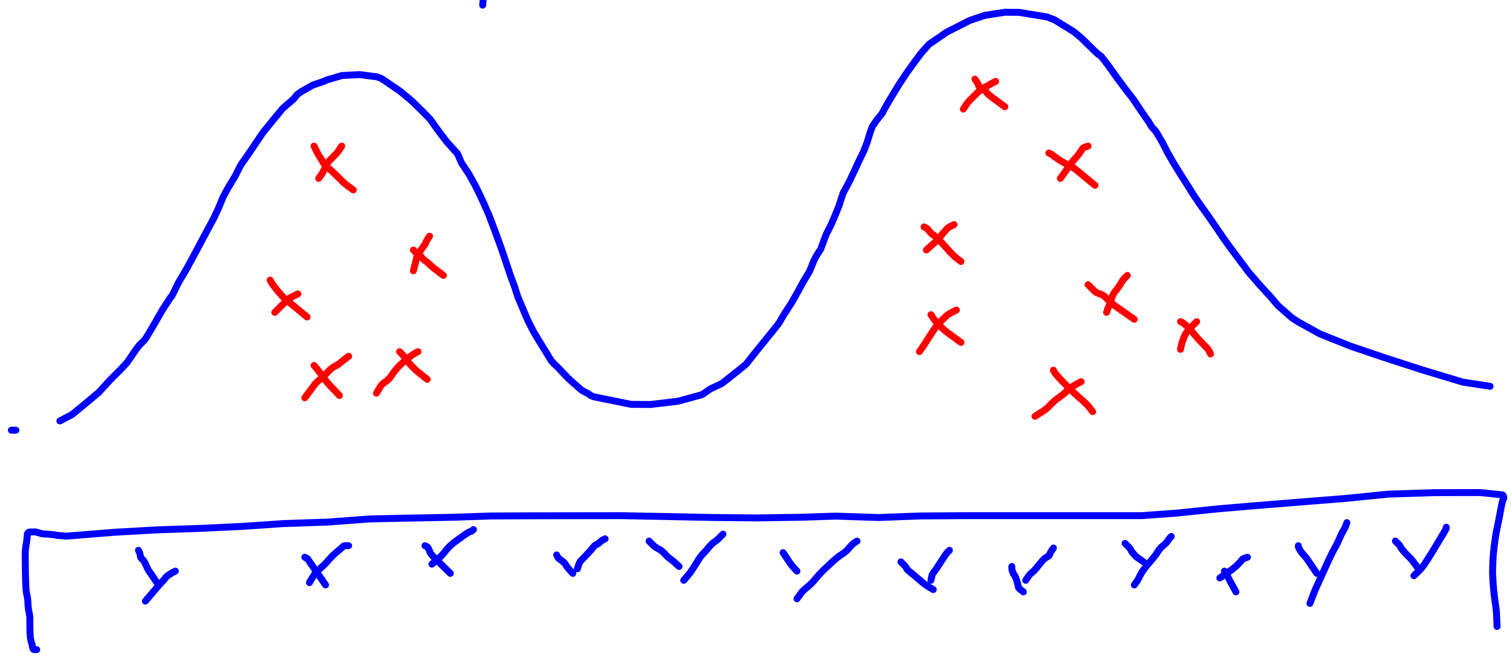
Mean-field model

G = complete graph

all edges included.

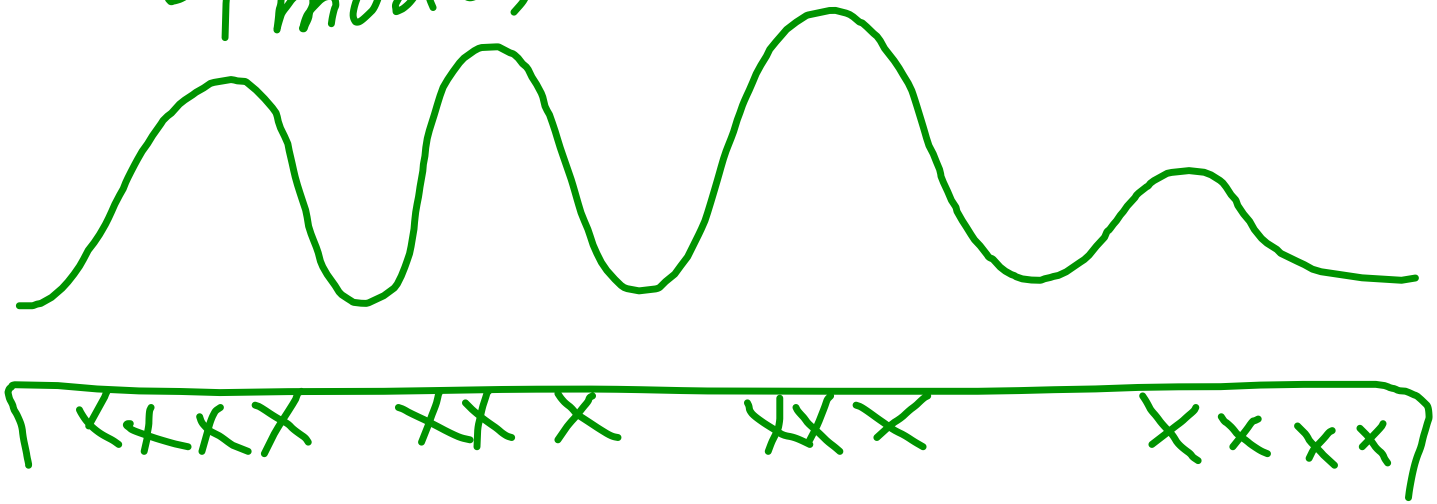
Fast or slow

k=2 (Ising mean field)
(Madras & Zheng 200?)
Tempering fast



k = 3

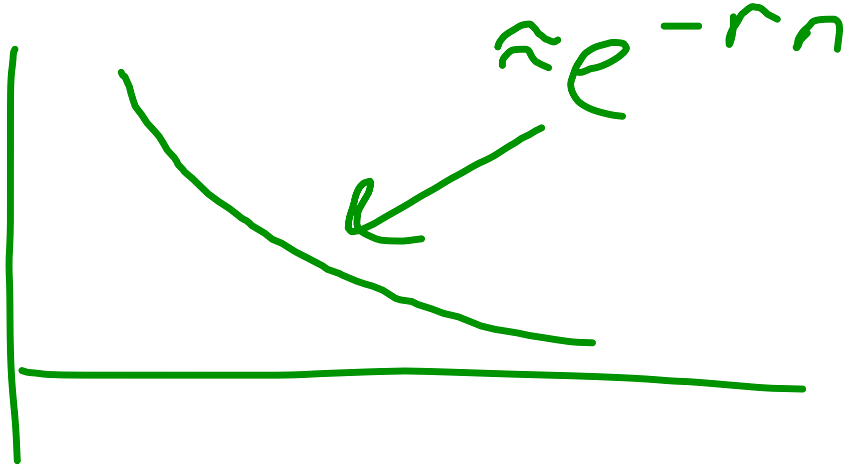
4 modes



Mixing Time

kernel K

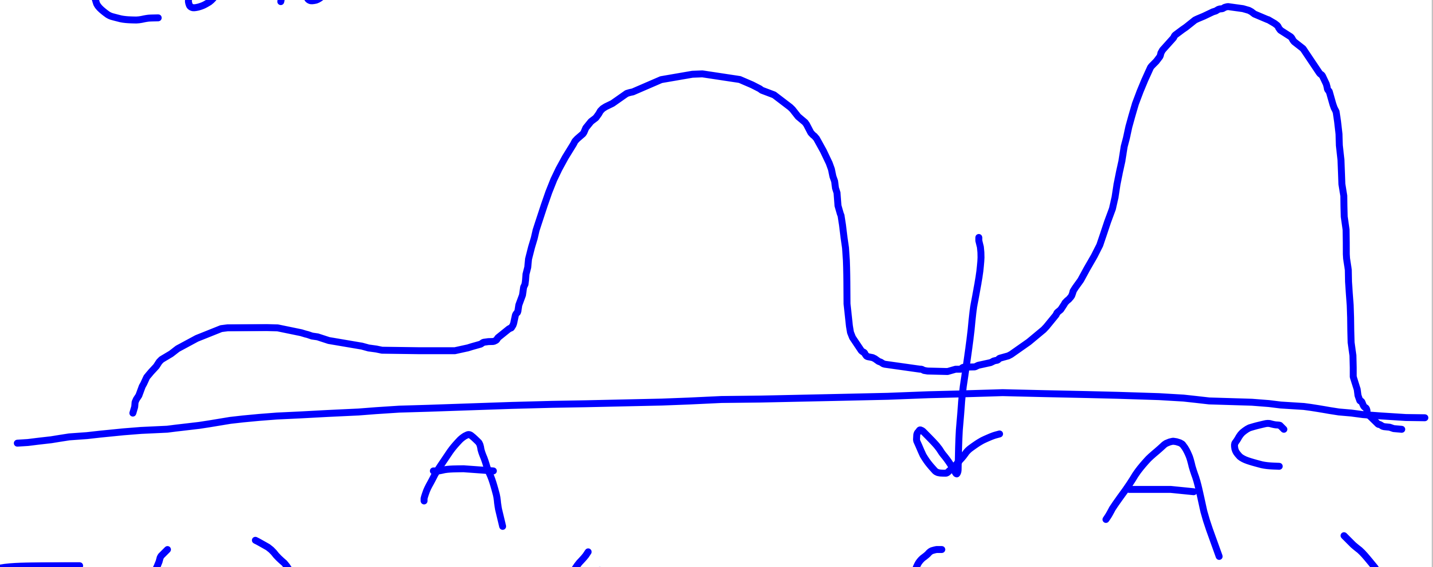
$$r := \inf_{\mu_0} \lim_{n\to\infty} \frac{-\ln(\|\mu_0 K^n - \pi\|_2)}{n}$$



K eigenvalue at 1
((corresponding to π)
magnitude rest strictly
bounded away from 1..

<u>Fact:</u> $r = -\ln(1 - \mathrm{Gap}(K))$

Conductance



$$\Phi_K(A) = \frac{\mathbb{P}_\pi(X_t \in A^c,\, X_{t+1} \in A)}{\pi(A^c)\,\pi(A)}$$

Fact: $\Phi_K(A) \geq \text{Gap}(K)$

Something like

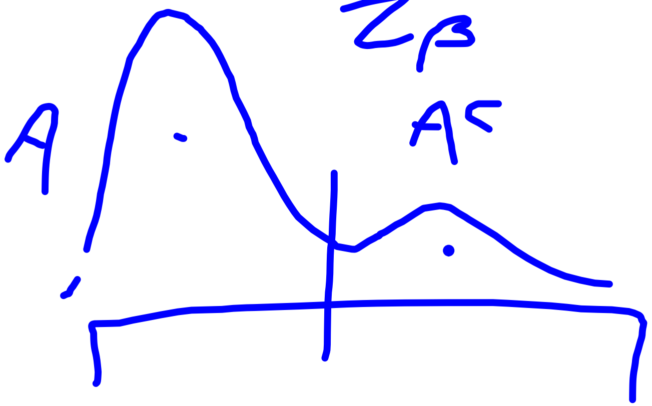
$$\text{Gap}(K) \geq \inf_A \frac{\Phi_K(A)^2}{2}$$

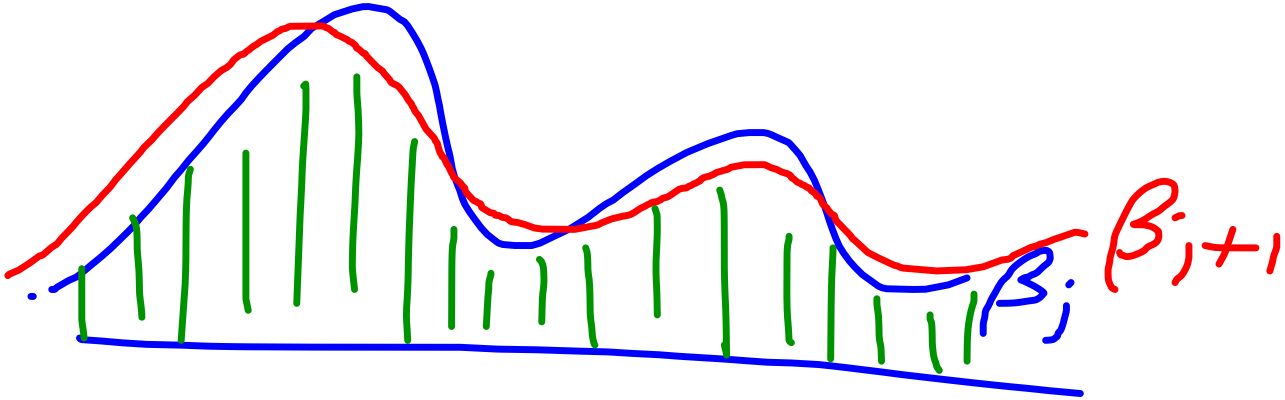
"Geometric Bounds on Gap"

Def'n: For $A \subseteq \Omega$, the persistance of A at level β is

$$\gamma(A, \Pi_\beta) = \min\left\{1, \frac{\Pi_\beta(A)}{\Pi(A)}\right\}.$$

$$\frac{e^{-\beta H(\cdot)}}{Z_\beta}$$





Overlap:

$$\delta(A, i, j) = \frac{1}{\pi_{\beta_j}(A)} \int_A \min\{\pi_{\beta_i}(A), \pi_{\beta_j}(A)\}$$

Theorem (Torpid Thm)

Let k^* be any level, $A \subseteq \Omega$

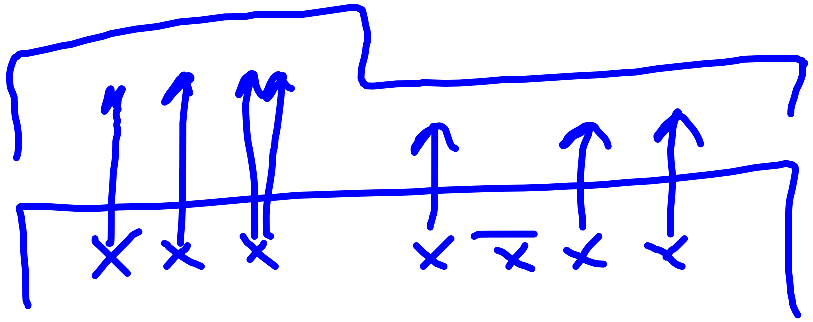
$$\text{Gap}(k) \le 12 \max_{k \ge k^*, \ell \le k^*} \left\{ \gamma(A, \pi_{\beta_k}) \cdot \max \left\{ \Phi_{\beta_k}(A) \cdot \delta(A,k,\ell)\,\delta(A^c,k,\ell) \right\} \right\}.$$

Theorem: (Fast Thm)

$$\tilde{\gamma}(A_1, \ldots, A_J) = \min_{k,j} \gamma(A_j, \pi_{\beta_k})$$

$$\tilde{\delta}(A_1, \ldots, A_J) = \min_{k,j} \delta(A_j, \pi_{\beta_k}, \pi_{\beta_{k+1}})$$

$$\mathrm{Gap}(K) \geq \frac{\tilde{\gamma}^{J+3}\,\tilde{\delta}^{3}}{2^{14(\#\text{levels})^5} J^3} \cdot \mathrm{Gap}(K_{\beta_0}) \cdot \min_{k,j} \mathrm{Gap}(\beta_k | A_j)$$



Question: Can we build a simple example w/ bad persistance where E&M method works well.