Critical behavior for the model of random spatial permutations

Dissertation defense

John Kerl

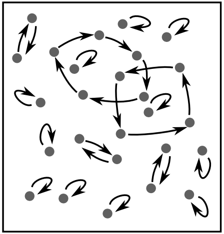
Department of Mathematics, University of Arizona

March 22, 2010

Outline

- The probability model
- Order parameters and criticality
- Markov chain Monte Carlo methods
- Finite-size scaling
- The worm algorithm
- Other work

The probability model



The probability model: definitions

State space: $\Omega_{\Lambda,N} = \Lambda^N \times \mathcal{S}_N$, where $\Lambda = [0, L]^3$ with periodic boundary conditions.

Point positions: $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ for $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \Lambda$.

Distance function (short-jump regime with periodic boundary conditions):

$$\|\mathbf{x} - \mathbf{y}\|_\Lambda = \|\mathbf{d}\|_\Lambda = \min_{\mathbf{n} \in \mathbb{Z}^3} \|\mathbf{d} + \mathbf{n}L\|.$$

Hamiltonian, where $T = 1/\beta$ and $r_\ell(\pi)$ is the number of ℓ-cycles in π:

$$H(\mathbf{X}, \pi) = \frac{T}{4} \sum_{i=1}^{N} \|\mathbf{x}_i - \mathbf{x}_{\pi(i)}\|_\Lambda^2 + \sum_{\ell=1}^{N} \alpha_\ell r_\ell(\pi).$$

- The first term discourages long permutation jumps, moreso for higher T.
- The temperature scale factor $T/4$, not $\beta/4$, is surprising but correct for the Bose-gas derivation of the Hamiltonian.
- The second term discourages cycles of length ℓ, moreso for higher α_ℓ. These interactions are not between points, but rather between permutation jumps.

The probability model: definitions

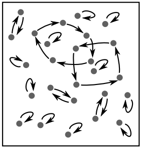
Fixed point positions (quenched model — includes all simulations done up to the present on the cubic unit lattice with $N = L^3$):

$$P_{\mathbf{X}}(\pi) = \frac{1}{Y(\Lambda, \mathbf{X})} e^{-H(\mathbf{X},\pi)}, \quad Y(\Lambda, \mathbf{X}) = \sum_{\sigma \in \mathcal{S}_N} e^{-H(\mathbf{X},\sigma)}.$$

Varying positions (annealed model — many theoretical results are available):

$$P(\pi) = \frac{1}{Z(\Lambda, N)} e^{-H(\mathbf{X},\pi)}, \quad Z(\Lambda, N) = \frac{1}{N!} \int_{\Lambda^N} Y(\Lambda, \mathbf{X}) \, d\mathbf{X}.$$

In either case, we write the expectation of an RV $S(\pi)$ as $\mathbb{E}[S] = \sum_{\pi \in \mathcal{S}_N} P(\pi) S(\pi)$.



Feynman (1953) studied long cycles in the context of Bose-Einstein condensation for interacting systems. See also Sütő (1993, 2002), and papers of Betz and Ueltschi.

The probability model: intuition

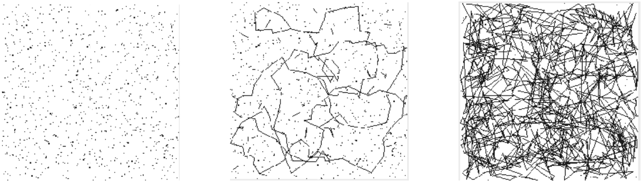
What does a typical random spatial permutation actually look like?
(Recall $H(\mathbf{X}, \pi) = \frac{T}{4} \sum_{i=1}^{N} \|\mathbf{x}_i - \mathbf{x}_{\pi(i)}\|_\Lambda^2 + \sum_{\ell=1}^{N} \alpha_\ell r_\ell(\pi)$.)

- As $T \to \infty$, the probability measure becomes supported only on the identity permutation. Large but finite T: there are tiny islands of 2-cycles, 3-cycles, etc.
- As $T \to 0$, length-dependent terms go to zero. The probability measure approaches the uniform distribution on $\mathcal{S}_N$: all π's are equally likely.

For intermediate T, things get more interesting:

- Lengths of each jump, $\|\pi(\mathbf{x}) - \mathbf{x}\|_\Lambda$, remain small: empirically, < 3.
- Above a critical temperature T_c, all cycles are short: 2-cycles, 3-cycles, etc. $T_c \approx 6.87$, and positive α terms increase T_c.
- Phase transition at T_c: below T_c, jump lengths remain short but long cycles form. Order-parameter RVs $1/\xi$, f_I, f_M, f_W, f_S (below) quantify this.

High T, medium but subcritical T, and low T:



Order parameters and criticality

Order parameters: $1/\xi$, f_S, f_W, f_I, f_M

The spatial cycle length and correlation length are

$$s_{\mathbf{x}}(\pi) = \sum_{j=1}^{\ell_{\mathbf{x}}(\pi)} \|\pi^j(\mathbf{x}) - \pi^{j-1}(\mathbf{x})\|_\Lambda \qquad \text{and} \qquad \xi = \overline{s}(\pi) = \frac{1}{N}\sum_{\mathbf{x}\in\Lambda} s_{\mathbf{x}}(\pi).$$

The winding number of π counts the integer number of wraps of π's cycles around the 3-torus in each of the three directions:

$$\mathbf{W}(\pi) = (W_x(\pi), W_y(\pi), W_z(\pi)) = \frac{1}{L}\sum_{i=1}^{N} \|\pi(\mathbf{x}_i) - \mathbf{x}_i\|_\Lambda$$

$$\mathbf{W}^2(\pi) = \mathbf{W}(\pi)\cdot\mathbf{W}(\pi) = W_x(\pi)^2 + W_y(\pi)^2 + W_z(\pi)^2.$$

The scaled winding number, f_S, arises in physics:

$$f_S = \frac{\mathbb{E}[\mathbf{W}^2]TL^2}{3N} = \frac{\mathbb{E}[\mathbf{W}^2]T}{3L}.$$

The fraction of sites in cycles which wind, f_W: self-explanatory.

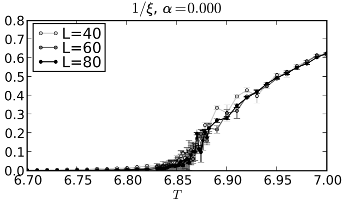
The fraction of sites in long cycles, f_I: defined in the dissertation. Intuition matches the name.

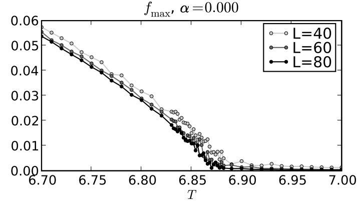
The scaled mean longest cycle length: $f_M = \mathbb{E}[\ell_{\max}]/N$.

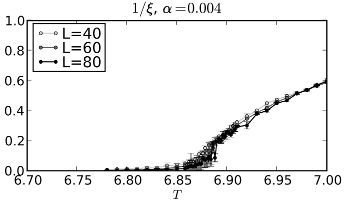
Behavior of order parameters as functions of L, T, and α.

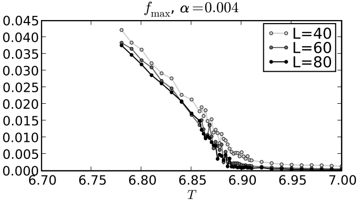
$1/\xi$ is right-sided; the rest are left-sided. All order-parameter plots tend to the right as α increases, i.e. $\Delta T_c(\alpha) = \frac{T_c(\alpha)-T_c(0)}{T_c(0)}$ is positive for small positive α.

Goal: quantify $\Delta T_c(\alpha)$'s first-order dependence on small α.









Known results and conjectures

Recall $H(\mathbf{X},\pi) = \frac{T}{4}\sum_{i=1}^{N} \|\mathbf{x}_i - \mathbf{x}_{\pi(i)}\|_\Lambda^2 + \sum_{\ell=1}^{N} \alpha_\ell r_\ell(\pi)$. We have the following models:

- Non-interacting model: $\alpha_\ell \equiv 0$.
- Two-cycle model: $\alpha_2 = \alpha$ and other cycle weights are zero.
- Ewens model: α_ℓ is constant in ℓ.
- General-cycle model: No restrictions on α_ℓ.

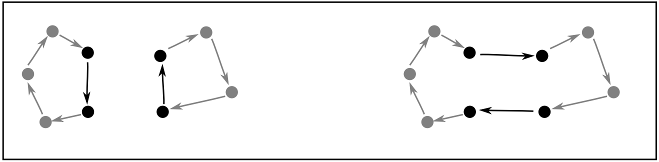
Known results for points on the continuum (obtained largely using Fourier methods):

- $\Delta T_c(\alpha)$ is known (to first order in α) for two-cycle interactions (Betz and Ueltschi, CMP 2008) and small cycle weights (Betz and Ueltschi 2008). (This taps into a long and controversial history in the physics literature: see Baym et al., EJP B 2001, or Seiringer and Ueltschi, PRB 2009, for surveys.) The critical (ρ, T, α) manifold relates ρ_c to T_c.

$$\rho_c(\alpha) \approx \sum_{\ell\geq 1} e^{-\alpha_\ell} \int_{\mathbb{R}^3} e^{-\ell\, 4\pi^2 \beta \|\mathbf{k}\|^2}\, d\mathbf{k} = \frac{1}{(4\pi\beta)^{3/2}} \sum_{\ell\geq 1} e^{-\alpha_\ell} \ell^{-3/2}$$

$$\Delta T_c(\alpha) \approx c\rho^{1/3}\alpha, \quad \text{for small } \alpha, \text{ with} \quad c \approx 2/3 \text{ when } \rho = 1.$$

Markov chain Monte Carlo methods



Metropolis sampling

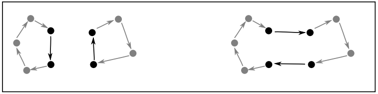
The expectation of a random variable S (e.g. f_W, f_M, f_I, f_S, ξ) is

$$\mathbb{E}[S] = \sum_{\pi \in \mathcal{S}_N} P(\pi) S(\pi).$$

$N!$ grows intractably in N. Instead, estimate expectations by summing over some number M (10^5 or 10^6) typical permutations. The sample mean is now a random variable with its own variance.

The usual technical issues of Markov chain Monte Carlo (MCMC) methods are known and handled in my simulations and dissertation: thermalization time, proofs of irreducibility, aperiodicity, and detailed balance (below), autocorrelation, batched means, and quantification of variance of sample means.

The fundamental Metropolis step (the analogue of single spin-flips for the Ising model) swaps permutation arrows which end at nearest-neighbor lattice sites. This either splits a common cycle, or merges disjoint cycles:

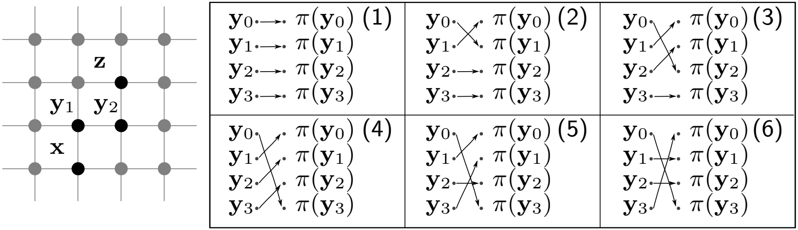


As usual, the proposed change is accepted with probability $\min\{1, e^{-\Delta H}\}$.

Correctness for the swap-only (SO) algorithm

Detailed balance, i.e. $P(\pi)M(\pi,\pi') = P(\pi')M(\pi',\pi)$ for all π, π', is easy to prove using standard Metropolis $M(\pi,\pi') \sim \min\{1, e^{-\Delta H}\}$. Here we prove irreducibility.

Proposition: Any π' is reachable from any other π using swaps.
Proof. Transpositions generate $\mathcal{S}_N$. We construct a sequence of (nearest-neighbor) swaps which results in a non-nearest-neighbor swap. We are given π, $\mathbf{x}$, and $\mathbf{z}$. Choose a nearest-neighbor path $\mathbf{y}_0 = \mathbf{x}$, $\mathbf{y}_1$, ..., $\mathbf{y}_{n-1}$, $\mathbf{y}_n = \mathbf{z}$. (See figure.)
Swaps: $(\mathbf{y}_0, \mathbf{y}_1)$, $(\mathbf{y}_0, \mathbf{y}_2)$, ... $(\mathbf{y}_0, \mathbf{y}_n)$; then $(\mathbf{y}_n, \mathbf{y}_{n-1})$, $(\mathbf{y}_n, \mathbf{y}_{n-2})$, ... $(\mathbf{y}_n, \mathbf{y}_1)$.
Then $\pi'(\mathbf{x}) = \pi(\mathbf{z})$, $\pi'(\mathbf{z}) = \pi(\mathbf{x})$, and $\pi'(\mathbf{y}) = \pi(\mathbf{y})$ for all other $\mathbf{y}$. □



Conclusion: given irreducibility, aperiodicity (also easy), and detailed balance, the Gibbs distribution is the invariant (and thus limiting) distribution for the SO chain.

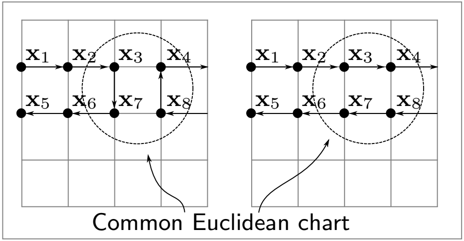
Conservation of winding number (with probability near 1)

Proposition: If jump lengths are less than $L/2$, swaps conserve winding number.
Proof. Swaps are done on pairs of arrows which end at nearest-neighbor sites. Due to short jump lengths, all four sites in a swap are in the same Euclidean chart lifted off the 3-torus. Thus

$$\mathbf{W}' - \mathbf{W} = \frac{1}{L} \sum_{i=1}^{N} \tilde{\mathbf{d}}(\pi'(\mathbf{x}_i), \pi(\mathbf{x}_i)) = \frac{1}{L} \left[\tilde{\mathbf{d}}(\pi'(\mathbf{x}), \pi(\mathbf{x})) + \tilde{\mathbf{d}}(\pi'(\mathbf{y}), \pi(\mathbf{y})) \right]$$

$$= \frac{1}{L} \left[\pi'(\mathbf{x}) - \pi(\mathbf{x}) + \pi'(\mathbf{y}) - \pi(\mathbf{y}) \right] = \frac{1}{L} \left[\pi(\mathbf{y}) - \pi(\mathbf{x}) + \pi(\mathbf{x}) - \pi(\mathbf{y}) \right] = 0.$$

□

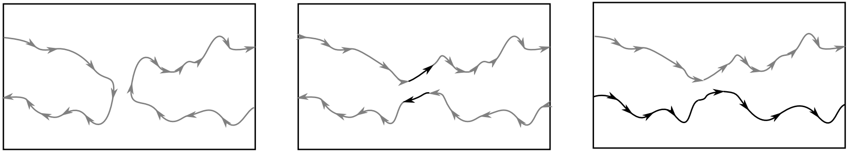


Partial solution: the swap-and-reverse (SAR) algorithm

Figure part 1: a long cycle on the torus almost meets itself in the x direction.

Part 2: after a swap-only step (above), one cycle winds by $+1$, and the other by -1. Metropolis steps in the short-jump-length regime create winding cycles only in opposite-direction pairs; total $W_x(\pi)$ is still zero.

Part 3: if we reverse one cycle (zero-energy move), $W_x(\pi)$ is now 2. This swap-and-reverse algorithm permits winding numbers of even parity in each of the three axes: one sweep proposes swaps at each lattice site. A second sweep reverses arrows on each cycle in the permutation with probability $1/2$.



Using SAR, it still takes a jump length $\approx L/2$ — which happens effectively never — to create an odd winding number. Band updates (see dissertation) are one idea; the worm algorithm (below) is another.

Finite-size scaling

Computational results: finite-size scaling method

Raw MCMC data yield $S(L, T, \alpha)$ plots as above, for each order parameter S.
Finite-size scaling (see Pelissetto and Vicari, `arXiv:cond-mat/0012164`, for a survey) determines the critical temperature $T_c(\alpha)$.

Define reduced temperature $t = \frac{T - T_c(\alpha)}{T_c(\alpha)}$, and correlation length ξ as above.

Hypotheses: (1) At infinite volume, $S \sim |-t|^\rho$ and $\xi \sim |t|^{-\nu}$ (power-law behavior).
(2) Finite-volume corrections enter only through a universal function Q_S of the ratio L/ξ:

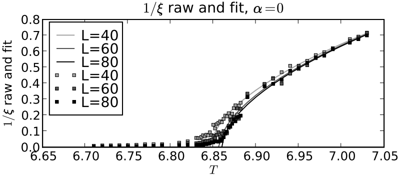
$$S(L, T, \alpha) = L^{-\rho/\nu} Q_S((L/\xi)^{1/\nu}) = L^{-\rho/\nu} Q_S(L^{1/\nu} t)$$

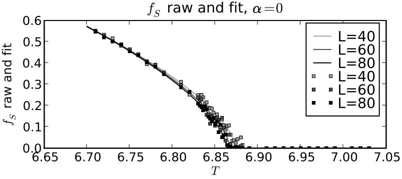
Method:

- Estimate critical exponents ρ, ν via power-law regression on MCMC data plots.
- Plot $L^{\hat{\rho}/\hat{\nu}} S(L, T, \alpha)$ as function of T. Since $t = 0$ at $T_c(\alpha)$, these plots for different L cross (up to sampling variability) at $T_c(\alpha)$.
- Having estimated $\hat{\rho}$, $\hat{\nu}$, and $\hat{T}_c(\alpha)$, plot $L^{\hat{\rho}/\hat{\nu}} S(L, T, \alpha)$ as function of $L^{1/\hat{\nu}} \hat{t}$. This causes all curves to collapse, confirming the FSS hypothesis.
- Regress $\Delta \hat{T}_c(\alpha)$ on α to estimate the constant c.

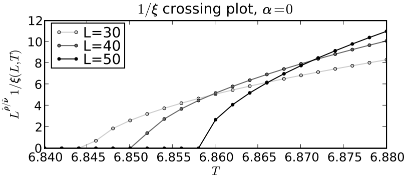
Computational results: power-law fit and crossing plots

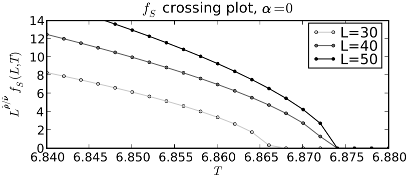
Raw data vs. power-law fit for $1/\xi$ and f_S with $\alpha = 0$:



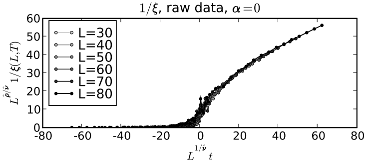


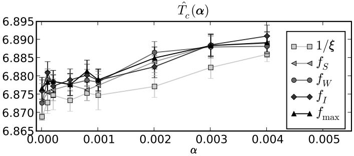
Plots for $1/\xi$, f_I, and f_M show crossing; plots for f_S and f_W do not. This is most clear at $L = 30, 40, 50$ where I did $M = 10^6$ MCMC samples for T near T_c, and most clear in the power-law-fit point of view:

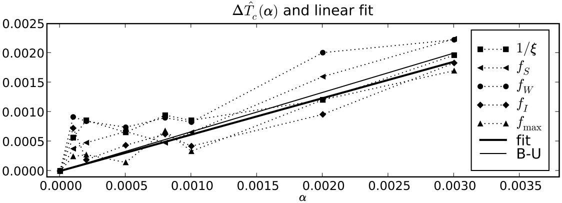




Collapse plots and $\Delta\hat{T}_c(\alpha) \approx \hat{c}\alpha$, given $\hat{\rho}$'s, $\hat{\nu}$, and $\hat{T}_c(\alpha)$

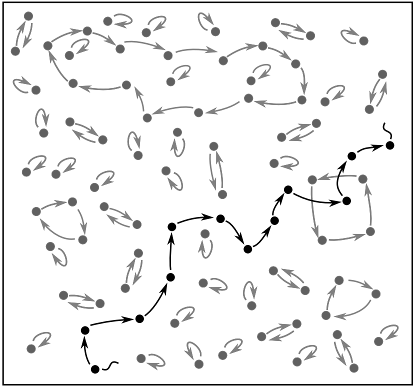






Omit $\alpha = 0.004$ since $\hat{T}_c(\alpha)$ begins to curve. Omit f_S and f_W due to non-crossing. Regressing on the $(\alpha, \Delta\hat{T}_c(\alpha))$ data, we find $\hat{T}_c(0) \approx 6.873 \pm 0.006$ and $\hat{c} \approx 0.618 \pm 0.086$ (2 σ error bars) for Ewens weights on the lattice. For small cycle weights on the continuum, Betz and Ueltschi have $T_c(0) \approx 6.625$ and $c \approx 0.667$. Conclusions: (1) Lattice structure modifies the critical temperature; (2) the α-dependent shift in critical temperature is unaffected.

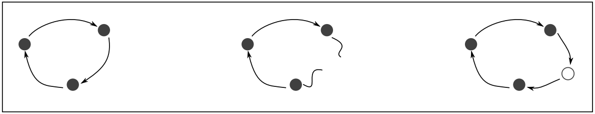
The worm algorithm



The worm algorithm: intuition

Random-cycle model with p.b.c. has multiple energy minima, indexed by winding numbers. Draw from path-integral Monte Carlo (PIMC) methods in physics. Tunnel through the energy barriers by opening a cycle, modifying it with swap-type moves at its tips, and closing it. Central point: this open cycle, or worm, can wander around the 3-torus (too) freely.

Not permutations anymore? In the figure: nothing $\mapsto \mathbf{x}_1$, $\mathbf{x}_1 \mapsto \mathbf{x}_2$, $\mathbf{x}_2 \mapsto \mathbf{x}_3$, $\mathbf{x}_3 \mapsto$ nothing. Rename the nothing to something, called the wormhole point, or w. It has no spatial coordinates and zero distance from any point. We now have $\pi \in \mathcal{S}_{N+1}$, an extended lattice, and an extended random-cycle model.



Same recipe applies as before: (extended) energy function and Metropolis moves; prove correctness. Invent any convenient extended energy for open π's agreeing with the original energy H for closed π's (proved next). Sample RVs only on closed π's.

The worm algorithm: marginality

Proposition: Let $\mathcal{S}_N \hookrightarrow \mathcal{S}_{N+1}$ by $\pi(w) = w$. Let H, H' be energy functions on $\mathcal{S}_N$ and $\mathcal{S}_{N+1}$ such that for all $\pi \in \mathcal{S}_N$, $H(\pi) = H'(\pi)$. Let P, P', Z, Z' be as above. Then for all $\pi \in \mathcal{S}_N$, $P'(\pi \mid \pi \in \mathcal{S}_N) = P(\pi)$.

Proof.: Let $\pi \in \mathcal{S}_N$. By definition of conditional expectation,

$$P'(\pi \mid \pi \in \mathcal{S}_N) = \frac{P'(\pi)\, 1_{\mathcal{S}_N}(\pi)}{P'(\mathcal{S}_N)}.$$

The numerator is Gibbs P for closed permutations, or 0 for open ones:

$$P'(\pi)\, 1_{\mathcal{S}_N}(\pi) = \frac{1}{Z'} e^{-H'(\pi)}\, 1_{\mathcal{S}_N}(\pi) = \frac{1}{Z'} e^{-H(\pi)}\, 1_{\mathcal{S}_N}(\pi)$$

since H and H' agree on closed π's. The denominator is total probability of closed permutations:

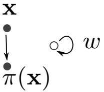
$$P'(\mathcal{S}_N) = \frac{1}{Z'} \sum_{\pi \in \mathcal{S}_N} e^{-H'(\pi)} = \frac{1}{Z'} \sum_{\pi \in \mathcal{S}_N} e^{-H(\pi)}.$$

Since $\pi \in \mathcal{S}_N$, the ratio is

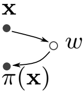
$$\frac{\frac{1}{Z'} e^{-H(\pi)}\, 1_{\mathcal{S}_N}(\pi)}{\frac{1}{Z'} \sum_{\pi \in \mathcal{S}_N} e^{-H(\pi)}} = \frac{e^{-H(\pi)}\, 1_{\mathcal{S}_N}(\pi)}{\sum_{\pi \in \mathcal{S}_N} e^{-H(\pi)}} = \frac{e^{-H(\pi)}\, 1_{\mathcal{S}_N}(\pi)}{Z} = P(\pi).$$

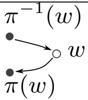
$\square$.

The worm algorithm: Metropolis moves

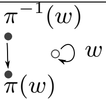


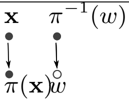
Open at $\mathbf{x}$ w.p. $\sim 1 \wedge e^{-\Delta H}$



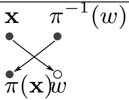


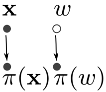
Close w.p. $\sim 1 \wedge e^{-\Delta H}$



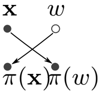


Head swap at $\mathbf{x}$ w.p. $\sim 1 \wedge e^{-\Delta H}$





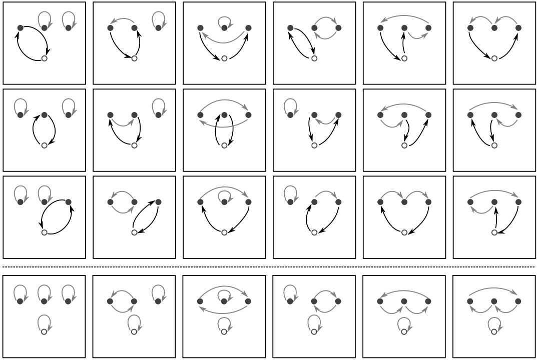
Tail swap at $\mathbf{x}$ w.p. $\sim 1 \wedge e^{-\Delta H}$



The worm algorithm: fibration and correctness

Key to proving correctness: fibration of $\mathcal{S}_{N+1}$ over $\mathcal{S}_N$.

- N open permutations lie over each closed permutation; fibers partition $\mathcal{S}_{N+1}$.
- Opens and closes stay within fibers.
- Head swaps and tail swaps move across fibers, and furthermore are transitive on fibers.
- Any SO swap can be constructed by an open, a head swap, and a close. (Hence irreducibility via SO, opens, and closes. Aperiodicity and detailed balance: also easy.)

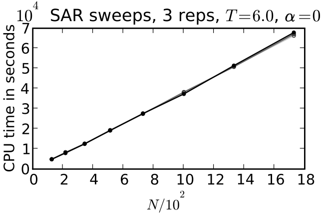


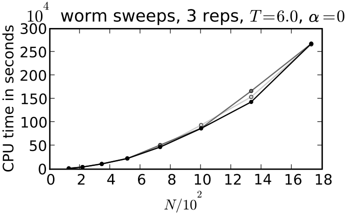
The worm algorithm: stopping time

Good news: examination of random-variable plots for $L = 10$, comparing SAR to worm, shows that similar results are produced — other than, of course, the winding-number histogram itself.

Problem: The the open worm tips wander around randomly within the L box, and fail to reconnect as L increases. Specifically, histograms show that the distribution of the wormspan $\|\pi(w) - \pi^{-1}(w)\|$ peaks around $L/2$.

SAR and worm CPU times are both $\sim aN + bN^2$. (Shown: $L = 5$ to 12.) SAR's b is tiny; worm's b is not. Interesting L (40-80 or so) are out of reach for the worm algorithm.





Other work

Other work

Dissertation items not presented today:

- Precise exposition of the theory of autocorrelation estimators for exponentially correlated Markov processes. Precise quantification of the advantages and non-advantages of batched means.
- Mean length of longest cycle as a fraction of the number of sites in long cycles recovers work of Shepp and Lloyd (1966) for non-spatial uniform permutations.

For the future:

- Use varying (annealed) point positions on the continuum. This samples from the true point distribution.
- Replace cycle-weight interactions in the Hamiltonian with those derived from the true Bose-gas model. Analytical as well as simulational work is needed in order to make this computationally tractable.

Thank you for attending!