Critical behavior for the model of random spatial permutations

John Kerl

Department of Mathematics, University of Arizona

September 9, 2009

Motivation

My research centers on Markov chain Monte Carlo methods in statistical mechanics. This includes side work on lattice percolation and self-avoiding walks; my thesis topic is random spatial permutations.

This model arises in the study of the Bose gas. It is also of intrinsic probabilistic interest. Theoretical history includes Bose-Einstein, Feynman, Penrose-Onsager, Sütő, Ueltschi-Betz.

Random permutations arise by symmetrizing the N-boson Hamiltonian with pair interactions and applying a multi-particle Feynman-Kac formula. System energy is now expressed in terms of point positions and permutations of positions, where permutations occur with non-uniform probability.

Interactions between permutations are interpreted as collision probabilities between Brownian bridges in Feynman time. Brownian bridges are integrated out, resulting in a model which lends itself readily to simulations without the need for CPU-intensive path-integral Monte Carlo (PIMC). This permits a new perspective on the venerable question: how does the critical temperature of Bose-Einstein condensation depend on inter-particle interaction strength?

The probability model

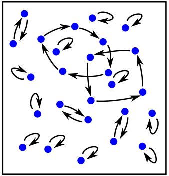
State space: $\Omega_{\Lambda,N} = \Lambda^N \times \mathcal{S}_N$, where $\Lambda = [0, L]^3$ with periodic boundary conditions.

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

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

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

- 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

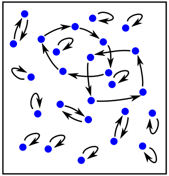
Fixed point positions (quenched model — includes all simulations done up to the present on the lattice $N = L^3$):

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

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



The probability model: intuition

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

- As $T \to \infty$, the probability measure becomes supported only on the identity permutation. For 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:

- The length of each permutation jump, $\|\pi(\mathbf{x}) - \mathbf{x}\|$, remains small.
- For T above a critical temperature T_c, all cycles are short: 2-cycles, 3-cycles, etc. $T_c \approx 6.8$, and positive α terms increase T_c.
- Phase transition at T_c: for $T < T_c$ jump lengths remain short but *long cycles form*.
- Figures: high T, medium but subcritical T, and low T.

Quantifying the onset of long cycles

We observe the following:

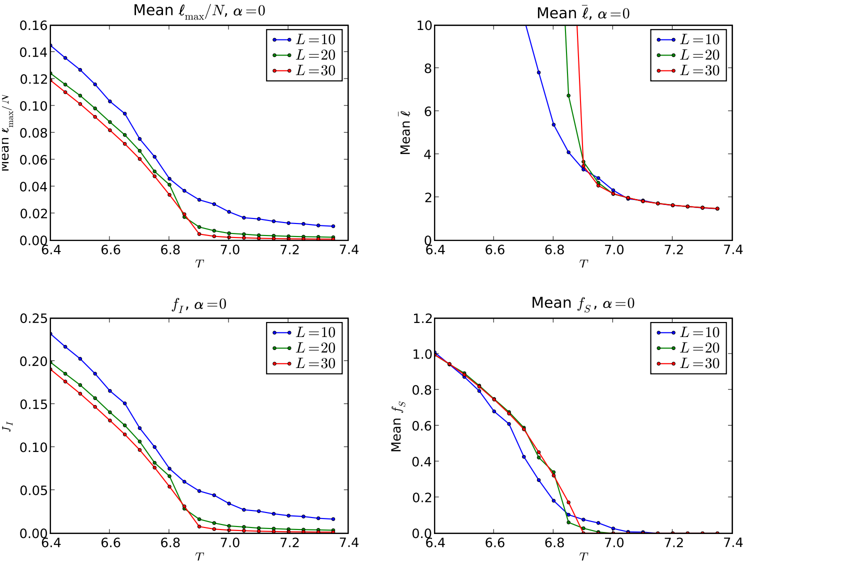
- For $T > T_c$, $\mathbb{E}[\ell_{\max}]$ is constant as $N \to \infty$: cycles remain finite.
- For $T < T_c$, $\mathbb{E}[\ell_{\max}]$ scales with N: there are arbitrarily long cycles, or infinite cycles, in the infinite-volume limit. Feynman (1953) studied long cycles in the context of Bose-Einstein condensation for interacting systems. See also Sütő (1993, 2002).

Other random variables ("order parameters") besides $\mathbb{E}[\ell_{\max}/N]$:

- Fraction of sites in long cycles, f_I, goes to zero in L above T_c, non-zero below.
- Correlation lengths $\xi(T)$ which are (spatial or hop-count) length of the cycle containing the origin: for $T < T_c$, these blow up in L.
- Winding numbers: number of x, y, z wraps around the 3-torus (Λ with p.b.c.). Scaled winding number: $f_S = \frac{\langle \mathbf{W}^2 \rangle L^2}{3\beta N}$. This behaves much like f_I, but is easier to compute with. Also, f_W: fraction of sites which participate in winding cycles.

Central goal of my dissertation work: quantify the dependence of T_c on α, where $\Delta T_c(\alpha) = \frac{T_c(\alpha) - T_c(0)}{T_c(0)}$. Known results and conjectures are formulated quantitatively in terms of $\lim_{\alpha \to 0} \Delta T_c(\alpha)$.

Behavior of order parameters as functions of L and T ($\alpha_\ell \equiv 0$)



Known results and conjectures

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

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

$$\Delta T_c(\alpha) \approx c\rho^{1/3}\alpha, \quad \text{for } \alpha \approx 0.$$

Known results and conjectures

Known results (continued):

- $\langle \ell_{\max} \rangle / N f_I$ is constant for $T < T_c$ for $\alpha_\ell \equiv 0$. (That is, the two order parameters f_I and $\langle \ell_{\max} \rangle / N$ have the same critical exponent.) For uniform-random permutations (Shepp and Lloyd 1966 solved Golomb 1964), $\langle \ell_{\max} \rangle / N \approx 0.6243$; unpublished work of Betz and Ueltschi has found $\langle \ell_{\max} \rangle / N f_I$ is that same number for the non-interacting case $\alpha_\ell \equiv 0$. Intuition: long cycles are "uniformly distributed" within the zero Fourier mode.

Conjectures:

- $\langle \ell_{\max} \rangle / N f_I$ is constant for $T < T_c$ for all interaction models. Questions: Why should this be true on the lattice? How does that constant depend on α?
- $\xi(T)$ is monotone in T: currently unproved either for the continuum or the lattice.
- $\rho_c(\alpha)$ formula holds not only for small cycle weights ($\alpha_\ell \to 0$ faster than $1/\log \ell$).

Open questions:

- To what extent does the $\rho_c(\alpha)$ formula hold true on the lattice?
- $\Delta T_c(\alpha)$ on the lattice should be similar to that on the continuum.
- $\Delta T_c(\alpha)$ is theoretically unknown for Ewens interactions (continuum or lattice).

Numerical/statistical techniques

- Run Markov chain Monte Carlo experiments for various values of L, T, interaction type, and interaction strength α.
- For each parameter combination, generate N typical permutations $\pi_1, \ldots, \pi_N$ from the stationary distribution. Compute random variables $X_i = X(\pi_i)$.
- Find the sample mean and estimate the variance of the sample mean (error bar). The correlation of the X_i's complicates the latter.
- Use finite-size scaling to compensate for finite-size effects: mathematically, we are interested estimating infinite-volume quantities based on finite-volume numerical experiments.

Metropolis sampling

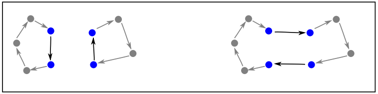
The expectation of a random variable θ (e.g. $\ell_{\max}/N$, f_I, f_S, ξ) is

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

The number of permutations, $N!$, grows intractably in N. The expectation is instead estimated by summing over some number M (10^4 to 10^6) typical permutations.

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

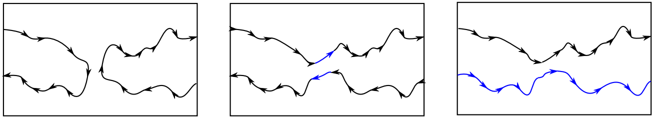
Metropolis step (analogue of single spin-flips for the Ising model): swap permutation arrows which end at nearest-neighbor lattice sites. This either splits a common cycle, or merges disjoint cycles:



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

Metropolis sampling and winding numbers: the GKU algorithm

- Figure part 1: a long cycle on the torus almost meets itself in the x direction.
- Part 2: after a Metropolis step, one cycle winds by $+1$, and the other by -1. Metropolis steps 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.



Our current best algorithm (GKU) has two types of sweeps: (1) For each lattice site, do a Metropolis step as above (Gandolfo, K). (2) For each cycle in the permutation, reverse the direction of the cycle with probability $1/2$ (Ueltschi). This permits winding numbers of even parity in each of the three axes.

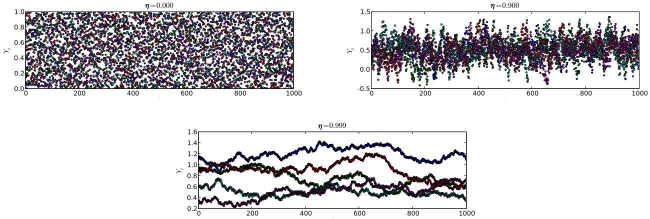
Methods for obtaining winding numbers of all parities: try (so far with mixed success) to adapt non-local updates (e.g. Swendsen-Wang for Ising) and worm algorithm. There are problems with low acceptance rate and stopping time for worm closure, respectively.

Variance of sample mean for correlated time series $c(k)$

The analysis applies to any stationary Markov chain X_t, $t = 0, 1, 2, \ldots$, with common mean μ_{X_t}, common variance $\sigma^2_{X_t}$, and exponential autocorrelation:

$$c(k) = \mathrm{Corr}(X_t, X_{t+k}) = \exp(-k/\tau_{\mathrm{exp}}) = \eta^k.$$

A toy-model Markov process Y_t , with fixed mean, fixed variance, and tunable autocorrelation exponent $\eta \in [0, 1)$, was constructed to test the analysis.



The memory induced by the autocorrelation results in a larger variance of the sample mean, which is already visible in the raw time-series data. We seek to quantify this.

Variance of sample mean for correlated time series: $\tau_{\rm int}$

The variance of the sample mean [Berg] is

$$\begin{aligned}\mathrm{Var}(\overline{X}_N) = \mathbb{E}[(\overline{X}_N - \mu_{X_t})^2] &= \frac{\sigma_{X_t}^2}{N}\left[1 + 2\sum_{t=1}^{N-1}\left(1 - \frac{t}{N}\right)\,\mathrm{Corr}(X_0, X_t)\right] \\ &\approx \frac{\sigma_{X_t}^2}{N}\left[1 + 2\sum_{t=1}^{\infty}\mathrm{Corr}(X_0, X_t)\right].\end{aligned}$$

The bracketed expression is the integrated autocorrelation time $\tau_{\rm int}$. Thus

$$\mathrm{Var}(\overline{X}_N) = \frac{\sigma_{X_t}^2}{N}\,\tau_{\rm int}$$

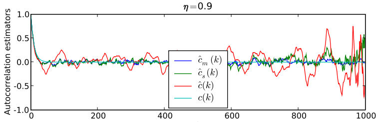
where $\sigma_{X_t}^2/N$ is the true variance of the sample mean only in the IID ($\eta = 0$) case. When the autocorrelation is $c(k) = \eta^k$, we have

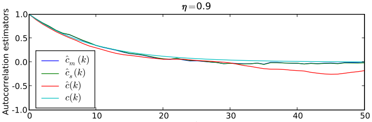
$$\tau_{\rm int} = \frac{1+\eta}{1-\eta}.$$

MCMC time series for the random-cycle model have $\eta \approx 0.99$ to 0.999: higher for T near T_c, lower farther away.

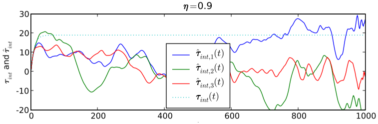
Estimation of $\hat{c}(k)$ and τ_{int}

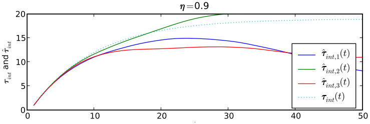
We estimate $c(k)$ by $\hat{c}(k)$ in the usual way. The estimator becomes poor for high k; even below that, it is fractionally underbiased.





An estimator $\hat{\tau}_{\text{int}}$ for τ_{int} is found by summing values of $\hat{c}(k)$ until the sum becomes approximately flat (first turning point). The fractional underestimation of $\hat{c}(k)$ carries over to $\hat{\tau}_{\text{int}}$.

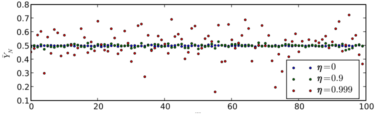


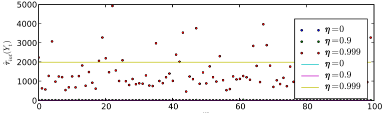


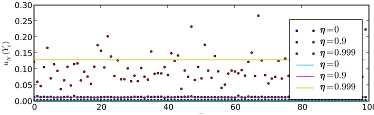
The true and estimated variance of the sample mean are

$$\mathrm{Var}(\overline{X}_N) = \frac{\sigma^2_{X_t}}{N}\,\tau_{\mathrm{int}} \qquad \text{and} \qquad u_N^2(X_t) = \frac{s_N(X_t)^2}{N}\,\hat{\tau}_{\mathrm{int}}.$$

The sample mean is unbiased for the true mean. The estimators $\hat{\tau}_{\mathrm{int}}$ and $u_N^2(X_t)$ are fractionally underbiased, and more widely varying with higher η. Here are results for $M = 100$ experiments of $N = 10000$ samples on the Y_t process.

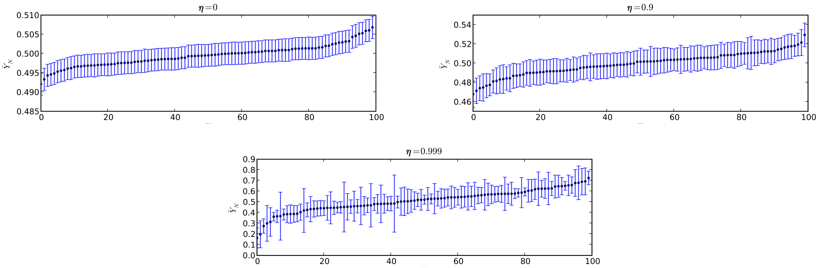






Estimation of the sample mean and its error bar

As a result, we may now clearly see the error of the error bar and its dependence on the autocorrelation exponent η.



Verdict: Compute the error-bar estimator as accurately as possible, keeping in mind that it is a rough estimator.

I have also shown that batched means, while facilitating IID analysis, improve neither the bias nor the variance of the error bar: batching N samples into N/B bins of size B reduces autocorrelation (good) but reduces sample size (bad). The two effects cancel.

We have an infinite-volume random variable $S_\infty(T)$, e.g. any of the order parameters above. The finite-volume quantity is $S_L(T)$. Define $t = (T - T_c)/T_c$. Examine, say, $0.99 < t < 1.01$.

The correlation length $\xi(T)$ follows a power law

$$\xi(T) \sim |t|^{-\nu}, \quad T \to T_c$$

For the infinite-volume quantity, we expect a power-law behavior

$$S_\infty(T) \sim t^\rho, (-t)^\rho, \quad \text{or} \quad |t|^\rho.$$

Finite-size scaling hypothesis: for T near T_c, $S_L(T)$ and $S_\infty(T)$ are related by a universal function Q which depends only on the ratio L/ξ:

$$S_L(T) = L^{-\rho/\nu} Q(L^{1/\nu} t) \sim L^{-\rho/\nu} Q((L/\xi)^{1/\nu}).$$

Finite-size scaling: parameter estimation and hypothesis testing

Collect MCMC experimental data, with error bars, for a range of L's, T's, and α's.

Estimation of critical exponents: given an order-parameter plot, vary the trial exponent $\hat{\rho}$. Raise the raw data to the $1/\hat{\rho}$ power. Find the $\hat{\rho}$ with least error in linear regression. Do the same for $\hat{\nu}$.

Crossing method for finding T_c: Once the exponents are known, plot $L^{\rho/\nu} S_L(T)$ as a function of T. Since at $T = T_c$ we have $t = 0$ and

$$L^{\rho/\nu} S_L(T) = Q(0),$$

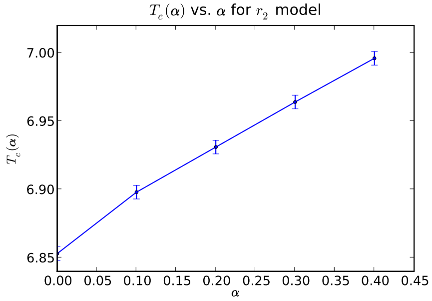
regardless of L, these curves will cross (approximately, due to sampling variability) at $T = T_c$.

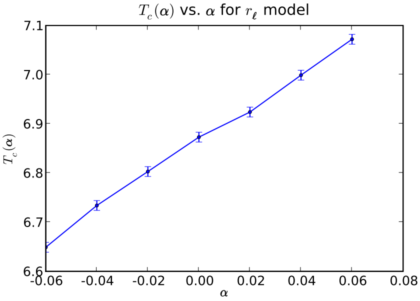
Testing of the FSS hypothesis: having estimated ρ, ν, and T_c, plot $L^{\rho/\nu} S_L(T)$ as a function of $L^{1/\nu} t$. This is a plot of the scaling function Q. If the hypothesis is correct, the curves for all L should coincide, or collapse.

Intermediate computational results: ΔT_c

These results are preliminary: no finite-size scaling; $L = 40$ here. For fixed L, one may sandwich $T_c(L)$ between the vertical asymptotes of $1/f_S$ and ξ. From such graphs, we obtain, with points on the lattice,

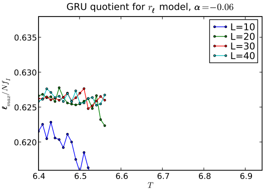
- $\Delta T_c(L)/\alpha = 0.0759 \pm 15\%$ for the r_2 model (vs. 0.088 theoretically for the continuum), and
- $\Delta T_c(L)/\alpha = 0.483 \pm 10\%$ for the Ewens model (theoretical value is unknown, but small-cycle-weight prediction for the continuum is 0.66).

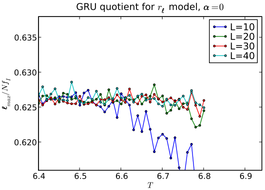


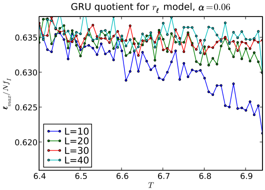


Computational results: GRU quotient $\langle \ell_{\max} \rangle / N f_I$

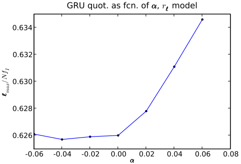
The GRU quotient varies with α in the Ewens model, but not in the r_2 model. For small L, it is non-constant for $T < T_c$; this bias seems to disappear as $L \to \infty$. (Needs a statistical confidence test.)







For r_2, GRU quotient is ≈ 0.626 regardless of α. For Ewens, averaging at all subcritical T's, we get the following dependence on α. This merits theoretical investigation.



Future work

Theory:

- Prove monotonicity of $\xi(T)$ for points on the continuum.
- Find theoretical expectations for the GRU quotient $\langle \ell_{\max} \rangle / N f_I$, as a function of α, on the continuum. Empirically, we know that there are negative-α and positive-α regimes with different α-dependence.

Experiment:

- Apply more careful finite-size scaling to simulation data.
- Conduct simulations with off-lattice quenched positions (Poisson point process). Lebowitz, Lenci, and Spohn 2000 showed that the point distribution for the Bose gas is not Poisson. Yet, this is a step away from the lattice and toward the true point distribution.
- Conduct simulations with varying (annealed) point positions on the continuum. This samples from the true point distribution. Software efficiency (namely, finding which points are near to which) requires a hierarchical partitioning of Λ.
- Develop an algorithm to permit odd winding numbers. (Hallway note: I would be delighted to discuss worm algorithms with a practitioner.)

Thank you for attending!