Lattice Permutations and Poisson-Dirichlet Distribution of Cycle Lengths

**Stefan Grosskinsky · Alexander A. Lovisolo ·
Daniel Ueltschi**

Abstract We study random spatial permutations on $\mathbb{Z}^3$ where each jump $x \mapsto \pi(x)$ is penalized by a factor $e^{-T\|x-\pi(x)\|^2}$. The system is known to exhibit a phase transition for low enough T where macroscopic cycles appear. We observe that the lengths of such cycles are distributed according to Poisson-Dirichlet. This can be explained heuristically using a stochastic coagulation-fragmentation process for long cycles, which is supported by numerical data.

Keywords Lattice permutations · Cycle lengths · Poisson-Dirichlet distribution

1 Introduction

Random permutations are common in probability theory and combinatorics [4]. They also occur in statistical mechanics, albeit with an additional spatial structure. With Λ denoting a finite box in $\mathbb{Z}^3$, we consider the set $\mathcal{S}_\Lambda$ of permutations of Λ, i.e., the set of bijections $\Lambda \to \Lambda$. The probability of a given permutation $\pi \in \mathcal{S}_\Lambda$ depends on the jump lengths in such a way that all sites are mapped in their neighborhoods. In this paper, we study the model with probability

$$\mathbb{P}_\Lambda(\pi) = \frac{1}{Z_\Lambda} \exp\left(-T \sum_{x \in \Lambda} \|x - \pi(x)\|^2\right). \tag{1}$$

Here $T > 0$ is a positive parameter and $\|x - \pi(x)\|$ denotes the Euclidean distance between x and $\pi(x)$. The normalization Z_Λ is defined by

$$Z_\Lambda = \sum_{\pi \in \mathcal{S}_\Lambda} \exp\left(-T \sum_{x \in \Lambda} \|x - \pi(x)\|^2\right). \tag{2}$$

S. Grosskinsky · A.A. Lovisolo
Centre for Complexity Science, University of Warwick, Coventry CV4 7AL, UK

S. Grosskinsky · D. Ueltschi (✉)
Department of Mathematics, University of Warwick, Coventry CV4 7AL, UK
e-mail: daniel@ueltschi.org

Fig. 1 A typical spatial
permutation with periodic
boundary conditions for small T.
The cycle that contains the origin
is depicted in *black* and it may be
long. Isolated sites belong to
1-cycles (i.e., they are mapped
onto themselves)
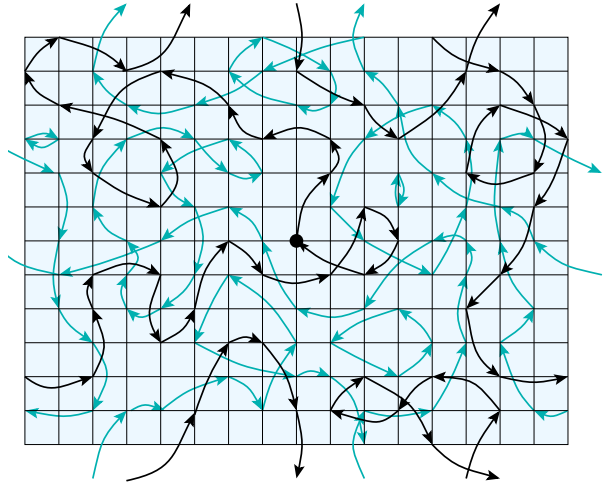
This model has its origin in Feynman's approach to the quantum Bose gas [13], where T
is proportional to the temperature. Bosons are described by Brownian trajectories with $\frac{1}{T}$
playing the rôle of time; this suggests the weights (1) for the permutation π. The presence
of a lattice is not really justified, but it does not affect the qualitative behavior, at least in
dimension 3 (we comment on dimension 2 at the end of the article).

Let us understand the qualitative behavior of the model when we vary the parameter T.
The most probable permutation is the identity, which has weight 1. Typical permutations
should be close to the identity when T is large, with small cycles here and there. The weight
in (1) penalizes large jumps and we expect that $\|x - \pi(x)\| \lesssim \frac{1}{\sqrt{T}}$. As T decreases, sites are
allowed to be mapped to more locations and the lengths of permutation cycles grow. One
expects that a phase transition takes place (in dimension 3 or more) that is accompanied by
the occurrence of infinite cycles. See Fig. 1 for a schematic spatial permutation with small T.

The phase transition was observed numerically in [15, 17]; it takes place at $T_c \approx 1.71$.
The fraction of sites that belong to macroscopic cycles was seen to converge to a non-
random function $\nu_\infty(T)$ as $|\Lambda| \to \infty$, which is continuous and monotone decreasing in T.
There are many macroscopic cycles and their sizes fluctuate. It was also observed that the
average length of the largest cycle scales like $0.624\nu_\infty(T)|\Lambda|$, which is identical to the
expectation of the largest cycle in a random permutation with uniform distribution [23]. The
latter observation was unexplained and puzzling at the time.

The situation is understood better now, and the explanation turns out to be surprisingly
general. The joint distribution of the length of the long cycles is given by the *Poisson-
Dirichlet distribution*. This distribution has been introduced by Kingman in [18] and has
cropped up in combinatorics, population genetics, number theory, Bayesian statistics and
probability theory, see [4, 12, 19–21] for details of applications and extensions.

Occurrence of the Poisson-Dirichlet distribution in models of statistical mechanics, i.e.
in models with spatial structure, seems to have been noticed only recently. It has been rig-
orously established in the "annealed" model of spatial permutations where the locations of
the sites are averaged upon [9]. The proof is inspired by [24] and it uses a representation
in terms of occupation numbers of Fourier modes, and non-spatial permutations within the
modes. Such a structure is not present here, however.

We recall the definition of the Poisson-Dirichlet distribution in Sect. 2.2 and provide
numerical evidence in Sect. 2.3 that it is present in our model. In order to explain this,

we show that the equilibrium state can be viewed as the stationary measure of an effective split-merge process. This strategy was recently applied successfully by Schramm to the random interchange model on the complete graph [22] (the result was first conjectured by Aldous, see [6]). The absence of spatial structure makes the situation much simpler, but it was nonetheless a *tour de force* to prove that long cycles occur, that they satisfy an effective split-merge process, and that their asymptotic distribution is Poisson-Dirichlet (see also [5] for subsequent simplifications and improvements). This strategy was also devised in [16] for the cycles and loops that arise in the Tóth and Aizenman-Nachtergaele representations of quantum Heisenberg models in three spatial dimensions [2, 25]. It allowed in particular to identify the parameter of the conjectured Poisson-Dirichlet distribution. Further situations that look similar include the random currents in the classical or quantum Ising models [1, 10, 14].

The key features are as follows: Long cycles are one-dimensional macroscopic objects and they are spread uniformly over the whole space. Introducing a suitable stochastic process with local changes, we observe that cycles are merged at a rate proportional to the number of "contacts" between them, and this number is proportional to the product of their lengths. Cycles are split at a rate proportional to the number of self-contacts, which is proportional to the square of the length. This is exactly analogous to a split-merge process on interval partitions [3, 7, 21]. As a consequence, the distribution of cycle lengths at equilibrium should be given by the invariant measure of the split-merge process, which is known to be the Poisson-Dirichlet distribution.

This explanation seems very attractive but it glosses over many technicalities. It assumes that a spatially uniform distribution of long cycles leads to a "mean-field" interaction and the correlations due to their spatial structure can be ignored. We provide mathematical background for these ideas in Sect. 3.1 and this allows us to state precise conjectures in Sect. 3.2. These conjectures are confronted with numerical results in Sect. 3.3. As it turns out, the above heuristics is fully confirmed.

2 Distribution of Long Cycles

2.1 Nature of Long Cycles

Let us first give precise definitions for "macroscopic", "mesoscopic" and "finite" cycles in the infinite volume limit. Given $x \in \Lambda$ and a permutation $\pi \in \mathcal{S}_\Lambda$, let $L_x(\pi)$ denote the length of the cycle that contains x, i.e., the number of sites in the support of this cycle.

- Macroscopic cycles occupy a non-zero fraction of the volume. The fraction of sites in macroscopic cycles is given by

$$\nu_{\text{macro}}(T) = \lim_{\varepsilon \to 0^+} \liminf_{|\Lambda| \to \infty} \frac{1}{|\Lambda|} \mathbb{E}_\Lambda \big(\#\{x \in \Lambda : L_x > \varepsilon |\Lambda|\} \big). \tag{3}$$

- Mesoscopic cycles are infinite cycles that are not macroscopic. The fraction of sites in mesoscopic cycles is given by

$$\nu_{\text{meso}}(T) = \lim_{K \to \infty} \liminf_{|\Lambda| \to \infty} \frac{1}{|\Lambda|} \mathbb{E}_\Lambda \left(\#\left\{ x \in \Lambda : K < L_x < \frac{|\Lambda|}{K} \right\} \right). \tag{4}$$

- Finally, the fraction of sites in finite cycles is given by

$$\nu_{\text{finite}}(T) = \lim_{K \to \infty} \liminf_{|\Lambda| \to \infty} \frac{1}{|\Lambda|} \mathbb{E}_\Lambda \big(\#\{x \in \Lambda : L_x < K\} \big)$$
$$= 1 - \nu_\infty(T). \tag{5}$$

Here, $\nu_\infty(T) = \nu_{\text{meso}}(T) + \nu_{\text{macro}}(T)$ is the fraction of sites in infinite cycles.
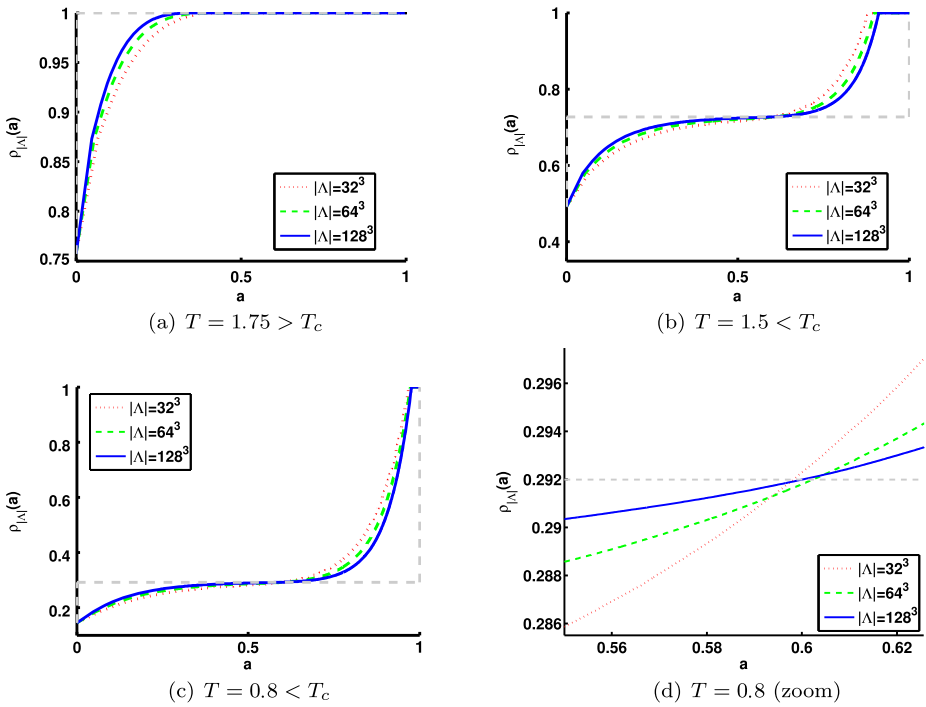
Fig. 2 Plots of the expected fraction of sites $\rho_{|\Lambda|}(a)$ in cycles of length smaller than or equal to $|\Lambda|^a$. The *horizontal dashed line* indicates $1 - \nu_\infty(T)$, the fraction of particles in finite cycles in the infinite volume limit. The *curves* have an intersection point independent of $|\Lambda|$ at $a \approx 0.6$ which is therefore used as the cutoff to distinguish long and short cycles. $\nu_\infty(0.8)$ is estimated to be 0.292. Averages were taken over 5×10^4 realizations

One expects that only finite cycles are present when T is large, that a phase with macroscopic cycles is present when T is smaller than a positive number T_c. This was proved in the annealed model in [8, 9]. We check this numerically in the lattice model. Let

$$\rho_{|\Lambda|}(a) = \mathbb{E}_\Lambda \left(\frac{\#\{x \in \Lambda : L_x \le |\Lambda|^a\}}{|\Lambda|} \right) \qquad (6)$$

denote the fraction of sites that belong to cycles of length less than or equal to $|\Lambda|^a$. Notice that $\rho_{|\Lambda|}(0)$ is the fraction of particles mapped onto themselves and that $\rho_{|\Lambda|}(1) = 1$. Numerical results for $\rho_{|\Lambda|}(a)$ are depicted in Fig. 2 for various parameters T.

In a finite domain, we need to define the cutoff that separates finite and "infinite" cycles. We can choose $|\Lambda|^a$ with *any* power $0 < a < 1$, since in the infinite volume limit, $\rho_\infty(a)$ will not depend on our choice. The graphs of $\rho_\Lambda(a)$ depend on the size of the domain, but we see in Fig. 2(d) that they cross the same point around $a \approx 0.6$ for $T < T_c$. This value of a is approximately independent of T and we choose it for the cutoff, since it significantly reduces finite size effects. For the numerical results, we define the fraction $\nu_{|\Lambda|}$ of sites that belong to infinite cycles by

$$\nu_{|\Lambda|}(T) = 1 - \rho_{|\Lambda|}(0.6). \qquad (7)$$

In accordance with results for the annealed model [8, 9] we expect that $\nu_{\text{meso}}(T) = 0$ and that

$$\lim_{|\Lambda| \to \infty} \nu_{|\Lambda|}(T) = \nu_\infty(T) = \nu_{\text{macro}}(T). \tag{8}$$

This is supported by the numerics in Fig. 2.

2.2 Griffiths-Engen-McCloskey and Poisson-Dirichlet Distributions

For a given permutation $\pi \in \mathcal{S}_\Lambda$ we call the cycle at x with length $L_x(\pi)$ *macroscopic* if $L_x(\pi) > |\Lambda|^{0.6}$, as discussed in the previous section. Let $L^{(1)}(\pi), L^{(2)}(\pi), \ldots, L^{(k)}(\pi)$ denote the cycle lengths in decreasing order, where $L^{(k)}$ is the smallest macroscopic cycle for the permutation π. If $\lambda^{(i)} = \frac{L^{(i)}}{|\Lambda|}$ is the fraction of the sites in the ith macroscopic cycle, we define

$$\nu(\pi) := \lambda^{(1)} + \cdots + \lambda^{(k)} \tag{9}$$

to be the fraction of sites in macroscopic cycles, and have $\mathbb{E}_\Lambda(\nu) = \nu_{|\Lambda|}(T)$. The sequence $(\lambda^{(i)})$ forms a random partition of the (random) interval $[0, \nu(\pi)]$. We now introduce the relevant measures on such partitions, that will allow us to describe the joint distribution of cycle lengths.

The Poisson-Dirichlet distribution (PD) is a one-parameter family but we only need the distribution with parameter 1, so we ignore the parameter altogether. It is best introduced with the help of the Griffiths-Engen-McCloskey distribution (GEM). The latter is also called the "stick-breaking" distribution. One can generate a random sequence of positive numbers $(\lambda_1, \lambda_2, \ldots)$ such that $\sum_i \lambda_i = \nu$ as follows:

- choose λ_1 uniformly in $[0, \nu]$;
- choose λ_2 uniformly in $[0, \nu - \lambda_1]$;
- choose λ_3 uniformly in $[0, \nu - \lambda_1 - \lambda_2]$;
- and so on. . . , always chopping a piece off the remaining portion of the "stick".

This is equivalent to choosing a sequence of i.i.d. random variables $(\alpha_1, \alpha_2, \ldots)$ where each α_i is taken uniformly in $[0, 1]$, and then to form the sequence

$$\big(\alpha_1, (1 - \alpha_1)\alpha_2, (1 - \alpha_1)(1 - \alpha_2)\alpha_3, \ldots\big) \times \nu.$$

Our goal is to recognize that a given sequence has the distribution GEM. One can invert the above construction, and form an i.i.d. sequence out of a GEM sequence. Namely, if $(\lambda_1, \lambda_2, \ldots)$ is GEM on the interval $[0, \nu]$, the following sequence is i.i.d. with respect to the uniform distribution on $[0, 1]$:

$$(\alpha_1, \alpha_2, \alpha_3, \ldots) = \left(\frac{\lambda_1}{\nu}, \frac{\lambda_2}{\nu - \lambda_1}, \frac{\lambda_3}{\nu - \lambda_1 - \lambda_2}, \ldots \right). \tag{10}$$

PD is a distribution on ordered partitions, and a PD sequence can be obtained by rearranging a GEM sequence in decreasing order. On the other hand, given an ordered PD sequence, a GEM sequence can be obtained as a size-biased permutation of that sequence [21].

2.3 Numerical Observations of Cycle Lengths

The GEM distribution is easier to handle than the PD distribution, and it contains more information. We thus introduce an order on cycles allowing us to establish an order for the cycle lengths. This can be done as follows. First, choose an order for the sites of Λ. Then order the cycles according to the smallest sites in their support; namely, given two cycles $\gamma = (x_1, \ldots, x_{|\gamma|})$ and $\gamma' = (x'_1, \ldots, x'_{|\gamma'|})$, we say that $\gamma < \gamma'$ if and only if $\min_{1 \le i \le |\gamma|} x_i <$

Fig. 3 The lengths of macroscopic cycles, divided by the volume, give a random partition of $[0, v]$ which is expected to follow the GEM distribution
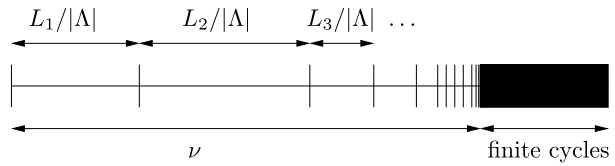
$\min_{1 \le i \le |\gamma'|} x'_i$. We then denote $L_1, L_2, \ldots$ the lengths of cycles larger than $|\Lambda|^{0.6}$ in this order, which is not to be confused with the notation L_x for lengths of cycles rooted in $x \in \Lambda$.

Let $v \equiv v(\pi)$ for a given permutation. Our aim is to show that $(\frac{L_1}{v|\Lambda|}, \frac{L_2}{v|\Lambda|}, \ldots)$ converges to GEM as $|\Lambda| \to \infty$, as illustrated in Fig. 3. This is equivalent to showing that

$$(\alpha_1, \alpha_2, \alpha_3, \ldots) = \left(\frac{L_1}{v|\Lambda|}, \frac{L_2}{v|\Lambda| - L_1}, \frac{L_3}{v|\Lambda| - L_1 - L_2}, \ldots \right) \tag{11}$$

converges to a sequence of i.i.d. uniform random variables in $[0, 1]$, see (10).

The Cumulative Distribution Function for α_i is defined as

$$F_{\alpha_i}(s) = P(\alpha_i \le s). \tag{12}$$

Numerical plots of F_{α_i} for the first three cycles can be found in Fig. 4. They clearly point to uniform random variables. Covariances can be found in Fig. 5, showing that they do indeed tend to 0 in the infinite volume limit. The discontinuities at $s = 1$ in Fig. 4 for $i = 1, 2, 3$ are due to the fact that in finite volumes it may happen that only 0, 1, 2 cycles larger than $|\Lambda|^{0.6}$, respectively, are present.

2.4 Markov Chain Monte-Carlo

To sample spatial permutations we use a Markov chain Monte-Carlo process which is ergodic and has $\mathbb{P}_\Lambda$ as its unique stationary distribution. Let B_Λ denote a suitable set of bonds, i.e., a set of unordered pairs $\{x, y\}$ of sites $x, y \in \Lambda$. Let $\tau_{xy} = \tau_{yx}$ denote the transposition of x and y. We say that two permutations $\pi, \pi' \in S_\Lambda$ are "in contact", noted $\pi \sim \pi'$, if there exists a bond $\{x, y\} \in B_\Lambda$ such that $\pi' = \pi \circ \tau_{xy}$. Let $Q(\pi, \pi')$, $\pi, \pi' \in \mathcal{S}_\Lambda$, be the transition matrix of a continuous-time Markov chain $(\pi_t : t \ge 0)$ on $\mathcal{S}_\Lambda$.

Proposition 2.1 *Suppose that B_Λ is large enough so that the graph (Λ, B_Λ) is connected, and that $Q(\pi, \pi') > 0$ whenever $\pi \sim \pi'$. The Markov chain with transition rates Q is ergodic.*

Proof This is done in [17], but we recall the argument here. The space $\mathcal{S}_\Lambda$ of lattice permutations is finite and irreducibility of the chain implies ergodicity. It then suffices to show that for all $\pi, \pi' \in \mathcal{S}_\Lambda$, $\pi' \ne \pi$ there exists $n < \infty$ such that $Q^n(\pi, \pi') > 0$. Every permutation π can be represented by a composition of transpositions. We still need to show that each of these transpositions can be written as a composition of transpositions along bonds of B_Λ.

Since the graph (Λ, B_Λ) is connected, there exists a connected path $(x_0, x_1, \ldots, x_m)$ such that $x_0 = x$, $x_m = y$, and $\{x_{i-1}, x_i\} \in B_\Lambda$ for all $1 \le i \le m$. One can check that the following composition gives $\tau_{x,y}$:

$$\tau_{x,y} = \tau_{x_0, x_1} \circ \cdots \circ \tau_{x_{m-2}, x_{m-1}} \circ \tau_{x_{m-1}, x_m} \circ \tau_{x_{m-2}, x_{m-1}} \circ \cdots \circ \tau_{x_0, x_1}. \tag{13}$$

This shows that every $\pi \in \mathcal{S}_\Lambda$ is connected to the identity permutation under the Markov chain dynamics. $\qquad\square$
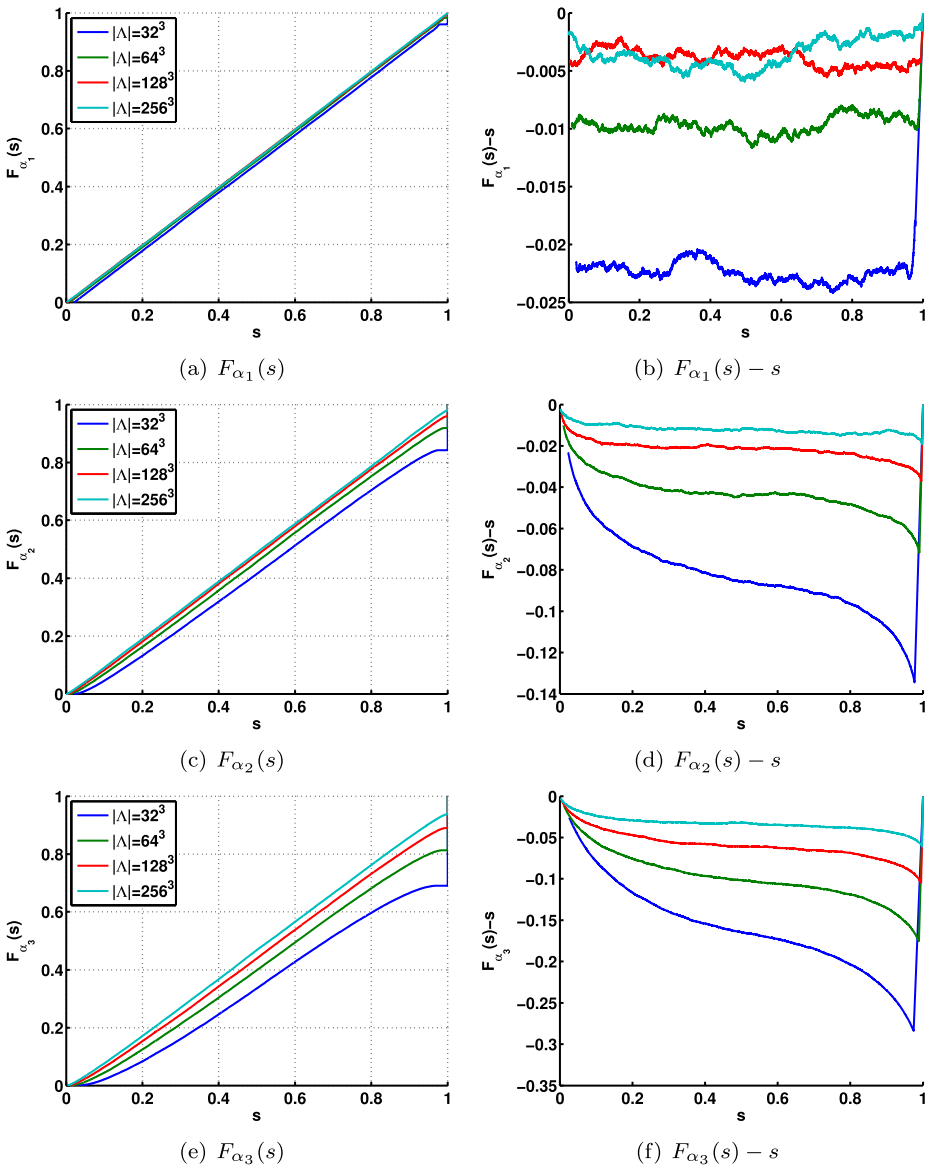
Fig. 4 (Color online) Cumulative distribution functions $F_{\alpha_i}(s)$ (12) with $i = 1, 2, 3$ for system sizes $|\Lambda| = 32^3, 64^3, 128^3, 256^3$ and $T = 0.8$. As the volume tends to infinity, they converge to the CDF of a uniform random variable. Averages were taken over 10^5 realizations

The composition of π with a transposition is explicitly given by

$$(\pi \circ \tau_{xy})(z) = \pi\left(\tau_{xy}(z)\right) = \begin{cases} \pi(z) & \text{if } z \neq x, y, \\ \pi(y) & \text{if } z = x, \\ \pi(x) & \text{if } z = y. \end{cases} \tag{14}$$

Fig. 5 Covariances of $\alpha_1, \alpha_2, \alpha_3$ as given in (11) shown for different system sizes $|\Lambda| = 32^3$, $64^3, 128^3, 256^3$ and $T = 0.8$
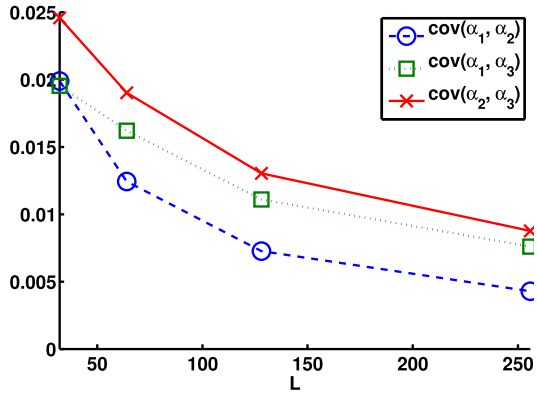
Let $H_\Lambda(\pi)$ denote the "energy" of $\pi \in \mathcal{S}_\Lambda$,

$$H_\Lambda(\pi) = T \sum_{x \in \Lambda} \|x - \pi(x)\|^2. \tag{15}$$

Here the Euclidean distance $\| \cdot \|$ is measured with periodic boundary conditions on a regular box $\Lambda \in \mathbb{Z}^3$. The distribution (1) then assumes the familiar form of the Gibbs state $e^{-H_\Lambda(\pi)}/Z_\Lambda$. For the transition rates we choose

$$Q(\pi, \pi') = \begin{cases} \frac{1}{|B_\Lambda|} \min(1, e^{-(H_\Lambda(\pi') - H_\Lambda(\pi))}) & \text{if } \pi' \sim \pi, \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Note that all rates are in [0, 1] and can therefore be used as acceptance probabilities for the standard Metropolis algorithm: pick a bond $\{x, y\} \in B_\Lambda$ uniformly at random and swap the images of x and y under π with probability 1 if this lowers the energy, and with probability $e^{-(H_\Lambda(\pi') - H_\Lambda(\pi))} < 1$ if the swap increases the energy.

It is clear that the measure $\mathbb{P}_\Lambda$ fulfills the detailed balance conditions, since

$$e^{-H_\Lambda(\pi)} Q(\pi, \pi') = e^{-H_\Lambda(\pi')} Q(\pi', \pi) \tag{17}$$

for all $\pi, \pi' \in \mathcal{S}_\Lambda$. This implies stationarity.

The particular algorithm we use is the "swap only" method described in [17]. The initial permutation is set to be the identity. The Metropolis steps are then as follows:

- Choose a bond $\{x, y\}$ of nearest-neighbors at random (we use periodic boundary conditions).
- The candidate permutation, $\pi' = \pi \circ \tau_{xy}$, replaces π with probability $\min(1, e^{-(H_\Lambda(\pi') - H_\Lambda(\pi))})$.

The Metropolis step is then computationally fast since $H_\Lambda(\pi \circ \tau_{xy}) - H_\Lambda(\pi)$ depends only on local terms:

$$H_\Lambda(\pi \circ \tau_{xy}) - H_\Lambda(\pi) = T\big(\|x - \pi(y)\|^2 + \|y - \pi(x)\|^2$$
$$- \|x - \pi(x)\|^2 - \|y - \pi(y)\|^2\big). \tag{18}$$

We use a standard approach based on the ergodic theorem for sampling, where we let the system equilibrate for a number of Metropolis steps of order $10^3 |\Lambda|$, and ensure that our measurements are spread over $5 \times 10^3 |\Lambda|$ steps. We have strong numerical evidence that the equilibration time of relevant observables is indeed of order $|\Lambda|$, as is shown in Fig. 6.

Fig. 6 (Color online) Expected value of the fraction of particles in macroscopic cycles for the permutation after t Metropolis steps starting from the identity permutation, for $|\Lambda| = 16^3, 32^3, 64^3, 128^3$ and $T = 0.8$. Rescaling time by $\frac{1}{|\Lambda|}$ shows the asymptotic behavior, showing that the equilibration time is of order $|\Lambda|$. The *dashed line* indicates the asymptote for the curves, $\nu_\infty(0.8)$. Averages were taken over 10^4 realizations
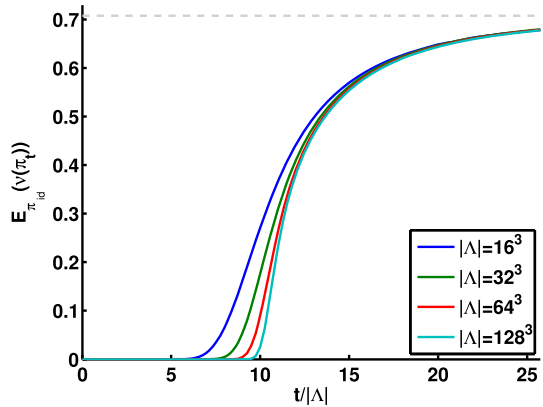
2.5 Periodic Boundary Conditions

We use periodic boundary conditions where $\Lambda \subset \mathbb{Z}^3$ is a 3-dimensional torus with equal side lengths L. This has the advantage of having less finite size effects than other choices such as closed boundary conditions. In the limit $|\Lambda| \to \infty$ we expect our results not to depend on that choice.

Precisely, for $y \in \mathbb{Z}^d$ we define $v_i(y)$ to be the ith component of y modulo L, in such a way that $v_i(y) \in \{-\frac{L}{2} + 1, \ldots, \frac{L}{2}\}$ (we assume here that L is even, but the modifications for odd L are straightforward). The Euclidean distance on $\mathbb{Z}^d$ is then replaced by

$$\|y\| = \left(\sum_{i=1}^{d} |v_i(y)|^2 \right)^{1/2}. \tag{19}$$

Permutations on the torus can be characterized by their *winding number*, which is in reality a winding vector. The winding number of π in the ith direction, $i = 1, \ldots, d$, is the integer

$$W_i(\pi) = \frac{1}{L} \sum_{x \in \Lambda} v_i\big(\pi(x) - x\big). \tag{20}$$

In a large box, one should not expect any jumps of order L for positive $T > 0$ because of the Gaussian weights (1). The dynamics restricted to such permutations conserves the winding number:

Proposition 2.2 *Suppose that the permutation π satisfies*

$$\max_{x \in \Lambda} \|\pi(x) - x\| \leq \frac{L}{2} - 2.$$

Then $W_i(\pi \circ \tau_{xy}) = W_i(\pi)$ for all i and all pairs (x, y) of nearest-neighbors in Λ.

Proof Using (14), we have

$$W_i(\pi \circ \tau_{xy}) - W_i(\pi) = \frac{1}{L} v_i\big(\pi(y) - x\big) + \frac{1}{L} v_i\big(\pi(x) - y\big) - \frac{1}{L} v_i\big(\pi(x) - x\big)$$
$$- \frac{1}{L} v_i\big(\pi(y) - y\big). \tag{21}$$

Because of the modulo operation, we have

$$v_i(\pi(y) - x) = [\pi(y) - x]_i + k_1 L, \qquad v_i(\pi(x) - x) = [\pi(x) - x]_i + k_3 L,$$
$$v_i(\pi(x) - y) = [\pi(x) - y]_i + k_2 L, \qquad v_i(\pi(y) - y) = [\pi(y) - y]_i + k_4 L. \tag{22}$$

Here, $[\cdot]_i$ denotes the ith coordinate of the vector in $\mathbb{Z}^d$, and we always have $k_i \in \{-1, 0, 1\}$. It follows from the assumptions that $k_1 = k_4$ and $k_2 = k_3$, so that (21) vanishes. □

The consequence of this proposition is that we effectively lose ergodicity in very large systems, since the dynamics conserves the winding number on simulation time scales. On the other hand, when macroscopic cycles are present, we expect nonzero winding numbers to appear with positive probability in the equilibrium measure. The dynamics always start with the identity permutation which has zero winding number. By the proposition, the path to nonzero winding numbers must cross bottlenecks, i.e., permutations with large jumps which occur with probability less than $\frac{1}{Z} e^{-T(\frac{L}{2} - 1)^2}$. A big part of the phase space is not explored by the dynamics.

By introducing Monte-Carlo transitions that flip the orientations of cycles (which does not change their probability), it is easy to move between permutations with *even* winding numbers. On the other hand, it would be interesting to study the winding numbers of typical permutations at equilibrium, but this seems to be a very difficult task numerically since it requires dynamics that can move between odd an even winding numbers and still sample from the correct distribution.

Due to the metastability of the winding number, the actual mixing time of the Monte-Carlo dynamics is (at least) e^{cL^2} with $c > 0$. However, since we are only interested in cycle lengths and not in their orientation, we do not expect this to be relevant for the observables discussed in the present paper. In fact, as is shown in Fig. 6 and later in Fig. 13, we observe convergence on time scales of order $|\Lambda| = L^3$. This is still much faster than processes such as card shuffles leading to uniform distributions on permutations (see e.g. [27]), which are typically of order $|\Lambda|^3$ with logarithmic corrections. This is due to the fact that for positive temperature T the stationary distribution (1) is not uniform, and jumps in a typical permutation are local of order $1/\sqrt{T}$. Furthermore, the identity permutation we start with is actually the ground state (permutation with highest probability) of that measure.

3 Effective Split-Merge Process

In the previous section we presented numerical evidence that the lengths of macroscopic cycles are distributed according to the Poisson-Dirichlet distribution. The goal of the present section is to explain it with a split-merge process, defined below in Sect. 3.1. More precisely, the Markov chain Monte-Carlo process of Sect. 2.4, when restricted to the cycle structure, becomes an effective split-merge process with the correct rates. We formulate precise conjectures about macroscopic cycles in typical spatial permutations, which we then test numerically.

3.1 Split-Merge Process

Recall that a partition $\lambda = (\lambda^{(1)}, \lambda^{(2)}, \ldots)$ of the interval $[0, \nu]$ is a sequence of decreasing positive numbers such that $\sum_i \lambda^{(i)} = \nu$. Here, ν is any positive real number. The split-merge process $(\lambda(t): t \geq 0)$, also called coagulation-fragmentation, is a continuous-time stochastic process on partitions where the ith and jth components ($i \neq j$) merge with rate

Fig. 7 Illustration for the split-merge process. The partition undergoes a merge followed by two splits and another merge
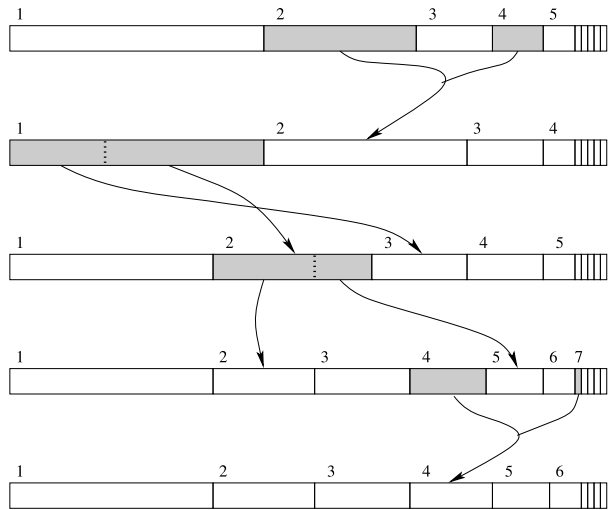
$$q_{ij} = 2\lambda^{(i)}\lambda^{(j)}/v^2, \tag{23}$$

and the ith component is split uniformly into two parts with rate

$$q_i = (\lambda^{(i)})^2/v^2. \tag{24}$$

Note that the rates are in $[0, 1]$ and they add up to 1, so they can be used directly for the following implementation of the process. If $\lambda(t) = (\lambda^{(1)}(t), \lambda^{(2)}(t), \ldots)$ denotes the partition at time t, one chooses the new configuration $\lambda(t + \text{Exp}(1))$ after an exponential waiting time with rate 1 as follows:

- Choose a first part of the partition with probability proportional to its size. That is, the index i is chosen with probability $\lambda^{(i)}(t)/v$. This is called "size-biased sampling".
- Choose a second part in the same manner, independently of the first. Let j the corresponding index.
- If $i \neq j$, merge $\lambda^{(i)}(t)$ and $\lambda^{(j)}(t)$. That is, the partition $\lambda(t + \text{Exp}(1))$ contains all parts $\lambda^{(k)}(t)$ with $k \neq i, j$, and a part of size $\lambda^{(i)}(t) + \lambda^{(j)}(t)$.
- If $i = j$, split $\lambda^{(i)}(t)$ uniformly. That is, the partition $\lambda(t + \text{Exp}(1))$ contains all parts $\lambda^{(k)}(t)$ with $k \neq i$, and two parts $u\lambda^{(i)}(t)$ and $(1 - u)\lambda^{(i)}(t)$, where u is a uniform random number in $[0, 1]$.
- The sequence is rearranged so that $(\lambda^{(k)}(t + \text{Exp}(1)))$ is decreasing.

The process is illustrated in Fig. 7. Additional background can be found in [3, 7]. Tsilevich showed that the Poisson-Dirichlet distribution is invariant for the split-merge process [26]. It was proved in [11] that it is the unique invariant measure (see also [22]).

A key property of our Monte-Carlo process $(\pi_t: t \geq 0)$ described in Sect. 2.4 is that, at each step, either a cycle is split, or two cycles are merged. This is illustrated in Fig. 8. Let us state this precisely.

Proposition 3.1 *Let $\pi \in S_\Lambda$ and $x, y \in \Lambda$ with $x \neq y$.*

- *If x, y belong to the same cycle in π, then x, y belong to different cycles in $\pi \circ \tau_{xy}$.*
- *If x, y belong to different cycles in π, then x, y belong to the same cycle in $\pi \circ \tau_{xy}$.*
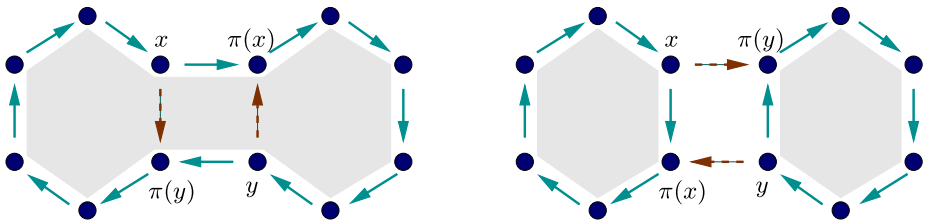
Fig. 8 The transposition τ_{xy} splits a cycle if x, y belong to the same cycle (*left*), or it merges two cycles if x, y belong to distinct cycles (*right*). These are the only two possibilities

See [17] for more details. It is clear that all cycles of π that do not involve x or y are also present in $\pi \circ \tau_{xy}$, and reciprocally. The length of the coalesced cycle is equal to the sum of the lengths of the two original cycles, and similarly for a fragmentation.

3.2 Effective Split-Merge Process for Macroscopic Cycles

We have seen in the previous section that each Monte-Carlo step results in either splitting a cycle, or merging two cycles. We have also seen in Sect. 2.1 that two kinds of cycles are present: The *finite* cycles, whose lengths do not diverge in the thermodynamic limit. And the *macroscopic* cycles, whose lengths are positive fractions of the volume. A Monte-Carlo step does one of the following:

(a) Merge two finite cycles.
(b) Merge a macroscopic cycle and a finite cycle.
(c) Merge two macroscopic cycles.
(d) Split a finite cycle (resulting in two finite cycles).
(e) Splits a macroscopic cycle, resulting in a finite and in a macroscopic cycles.
(f) Splits a macroscopic cycle, resulting in two macroscopic cycles.

One expects each of these options to take place with rates of order $O(1)$ in the limit $|\Lambda| \to \infty$. The ones that are relevant to the effective split-merge process are (c) and (f), since their effect is reflected in a change in $(\lambda^{(1)}(\pi(t)), \lambda^{(2)}(\pi(t)), \ldots)$, the ordered lengths of *macroscopic* cycles normalized by $|\Lambda|$. Accordingly, we introduce the rate R_{ij} at which the ith and jth largest cycles merge. It depends on the permutation π, and, with $i \neq j$, it is given by

$$R_{ij}(\pi) = 2 \sum_{x \in \gamma^{(i)}} \sum_{y \in \gamma^{(j)}} Q(\pi, \pi \circ \tau_{xy}). \tag{25}$$

Notice that $R_{ij}(\pi)$ scales to a constant as the volume diverges, since the sum over x, y is of order $|\Lambda|$, and $Q(\cdot)$ is of order $1/|B(\Lambda)|$, see (16). The rate R_i at which the ith largest cycle splits into two macroscopic cycles involves the cutoff K that distinguishes finite vs macroscopic cycles:

$$R_i(\pi) = \sum_{x \in \gamma^{(i)}} \sum_{k=K}^{L^{(i)} - K} Q(\pi, \pi \circ \tau_{x, \pi^k(x)}). \tag{26}$$

Here, $\pi^k(x)$ is the site at "distance" k of x along the cycle $\gamma^{(i)}$. If we set $K = 1$ in the right side, we get the rate at which $\gamma^{(i)}$ splits, irrespective of the sizes of the resulting cycles. The

Fig. 9 Schematic drawing of the situation in a mesoscopic box. It contains finite cycles in *light color*, and two legs that belong to each of the macroscopic cycles γ and γ'. The probability that γ and γ' merge in the next step is proportional to the number of 'contacts' between them, which in turn is proportional to $L_\gamma L_{\gamma'}$. The probability that γ splits into two macroscopic cycles is proportional to the self contacts *amongst different legs* of γ, which is in turn proportional to L_γ^2
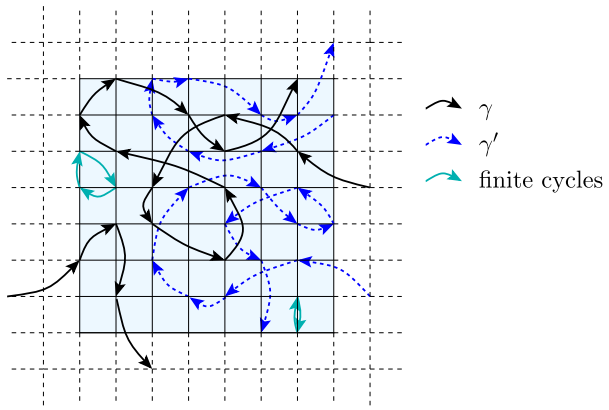
expression above gives the rate at which $\gamma^{(i)}$ splits in two cycles, each of which has length greater than K. As in previous sections we use the cutoff $K = |\Lambda|^{0.6}$.

We expect that, for almost all permutations in equilibrium, the rates R_{ij} and R_i are equal to those of the split-merge process modulo a constant time-scale R, resulting from the effective rate of macroscopic processes (c) and (f) above. Recall that $\lambda^{(i)}(\pi) = L^{(i)}(\pi)/|\Lambda|$ is a random variable.

Conjecture 1 *Let T such that $\nu_\infty(T) > 0$. There exists a number R (that depends on T but not on indices) such that for all i, j, and all $\varepsilon > 0$,*

$$\lim_{|\Lambda| \to \infty} \mathbb{P}_\Lambda \left(\left| R_{ij} - 2\lambda^{(i)}\lambda^{(j)}R \right| > \varepsilon \right) = 0,$$

$$\lim_{|\Lambda| \to \infty} \mathbb{P}_\Lambda \left(\left| R_i - \left(\lambda^{(i)}\right)^2 R \right| > \varepsilon \right) = 0.$$

The time scale of the effective split-merge process is determined by R, but the invariant measure is not, so the exact value of R is irrelevant. It is important, however, that it is identical for both the splits and the merges, and for all macroscopic cycles.

Let us explain the heuristics towards this remarkably simple behavior. Consider a mesoscopic box Λ' whose size is large enough so that boundary effects are irrelevant, yet small enough so that Λ is made up of a large number of mesoscopic boxes. The restriction of π on Λ' gives many finite cycles, and open legs that are parts of macroscopic cycles. See Fig. 9 for a schematic picture. Let us choose a pair of nearest-neighbors x, y at random, with the condition that x, y belong to distinct legs. The probability that τ_{xy} merges $\gamma^{(i)}$ and $\gamma^{(j)}$ is equal to the probability that x belongs to $\gamma^{(i)}$ and y belongs to $\gamma^{(j)}$, or conversely, which is equal to $2\lambda^{(i)}\lambda^{(j)}/\nu(T)^2$, up to vanishing finite-size effects. The probability that τ_{xy} splits $\gamma^{(i)}$ is equal to the probability that both legs belong to $\gamma^{(i)}$, which is equal to $(\lambda^{(i)}/\nu(T))^2$. This heuristics assumes that macroscopic cycles are spread uniformly in space, so they are present in all mesoscopic boxes in proportion to their size, and also that the local configurations do not depend on the situation in other boxes. Note that short range correlations in the cycle structure affect only the probability that a randomly chosen pair x, y belongs to distinct legs, which is absorbed in the constant R that determines the time scale of the effective split merge process.

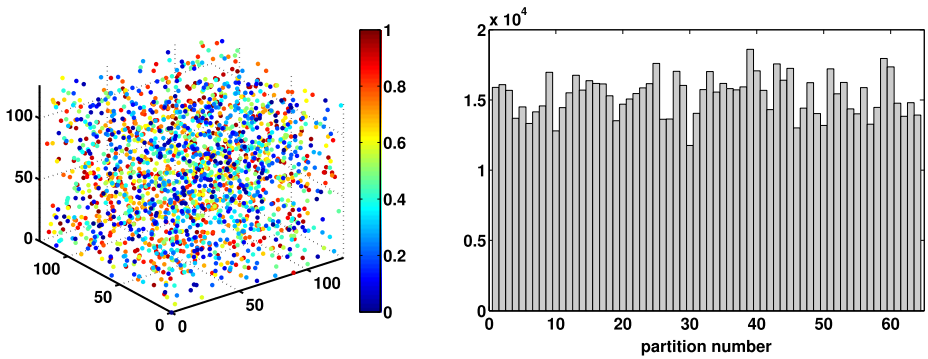In addition, we also conjecture that when a cycle is split, it is split uniformly.

Fig. 10 (Color online) A single permutation has been chosen with respect to the equilibrium measure for $|\Lambda| = 128$, $T = 0.8$. The longest cycle $\gamma^{(1)}$ is spread everywhere. (*Left*) Scatter plot of every 500th site along $\gamma^{(1)}$. The color (shades of gray) indicates the distance along $\gamma^{(1)}$, renormalized by the length of $\gamma^{(1)}$, starting from the site closest to the origin. (*Right*) The box Λ of volume 128^3 has been partitioned in 64 subsets of volume 32^3. The histogram depicts the number of sites of $\gamma^{(1)}$ that can be found in each subset. It is essentially constant

Conjecture 2 *Let T such that $\nu_\infty(T) > 0$ and define the CDF for a function of the split length $a \in [0, 1]$,*

$$\theta_i^{(a)}(\pi) = \frac{1}{R_i(\pi)} \sum_{x \in \gamma^{(i)}} \sum_{k=K}^{aL^{(i)}} P(\pi, \pi \circ \tau_{x, \pi^k(x)}). \tag{27}$$

Then

$$\lim_{|\Lambda| \to \infty} \mathbb{P}_\Lambda\left(\left|\theta_i^{(a)}(\pi) - a\right| > \varepsilon\right) = 0.$$

If these conjectures hold true, the effective process $(\lambda^{(1)}(\pi(t)), \lambda^{(2)}(\pi(t)) \ldots)_{t \geq 0}$ with stationary initial condition converges in the limit $|\Lambda| \to \infty$ to a split-merge process as defined in (23) and (24), running with total rate $R\nu_\infty^2$. Therefore the distribution of cycle lengths has to be invariant with respect to the split-merge process, so it has to be Poisson-Dirichlet. We now check these conjectures numerically.

3.3 Numerical Data About the Effective Split-Merge Process

In this section we give numerical evidence that the rates for splitting and merging macroscopic cycles converge to those of a split-merge process, thus confirming the Poisson-Dirichlet distribution of macroscopic cycles. The heuristics behind this argument, as described in the previous section, is that macroscopic cycles are distributed uniformly amongst lattice sites on a mesoscopic level, and are not confined to a bounded region. This is illustrated in Fig. 10.

In order to verify Conjecture 1, the rates of merging the two longest cycles, splitting the longest, and splitting the second longest in the next timestep were calculated using (25) and (26). We define

$$r_{ij}(\pi) = \frac{R_{ij}(\pi)}{2\lambda^{(i)}\lambda^{(j)}}, \qquad r_i(\pi) = \frac{R_i(\pi)}{(\lambda^{(i)})^2}. \tag{28}$$
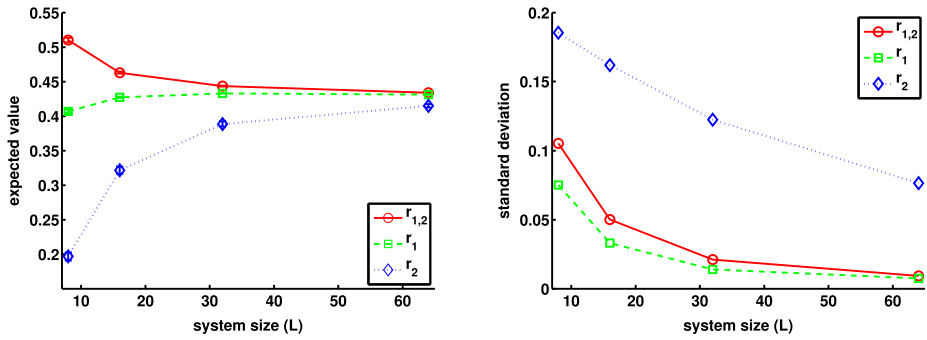
Fig. 11 Expectations (*left*) and standard deviations (*right*) of the rates defined in (28). $|\Lambda| = 8^3, 16^3,$ $32^3, 64^3, T = 0.8$. The standard error for the mean is within the marker size. Averages were taken over 10^3 realizations
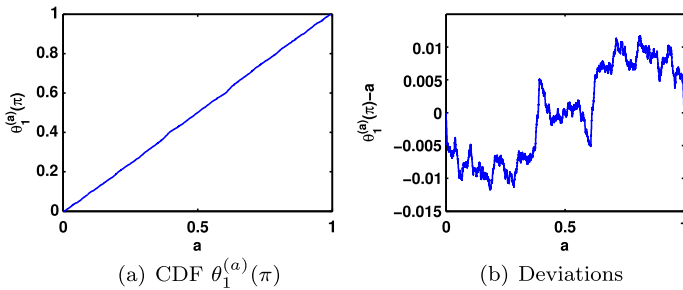

Fig. 12 Cycles split uniformly. Plot of $\theta_1^{(a)}(\pi)$ as defined in (27) for a typical permutation π under the equilibrium measure with $|\Lambda| = 64^3, T = 0.8$

Figure 11 shows that r_{ij} and r_i converge to a constant R as expected, supporting Conjecture 1. Note that in addition to convergence of the mean the variance is decreasing, confirming convergence in probability as stated in the conjecture.

If the cycle is split uniformly, $\theta_i^{(a)}(\pi)$ as defined in Conjecture 2 is the CDF of a uniform random variable on [0, 1]. The graph of $\theta_i^{(a)}(\pi)$ is shown in Fig. 12 for a given permutation π chosen randomly from the equilibrium measure, which confirms the expected behavior.

4 Further Prospects

The emergence of macroscopic cycles seems intriguing and it is worth being studied. Starting the Monte-Carlo Markov chain from the identity permutation, one expects the system to display only finite cycles for some time, before infinite objects built up. Let $\mathbb{E}_{\Lambda,t}$ denote the corresponding expectation, and $N_{|\Lambda|}(\pi)$ denote the number of cycles of length larger than $|\Lambda|^{0.6}$. Figure 13 shows $\mathbb{E}_{\Lambda,t}(N_{|\Lambda|})$ for $T = 0.8$ and various volumes. There is a peak at the transition to the phase with macroscopic cycles. It is certainly due to the presence of many mesoscopic cycles for a short time, that are going to merge afterwards. It would be interesting to get plots for numbers of cycles larger than $|\Lambda|^a$ for a other than 0.6.

While the fraction of sites in large cycles, $\mathbb{E}_{\Lambda,t}(\nu)$, varies continuously with time on the scale $t/|\Lambda|$, the steps (c) and (f) of Sect. 3.2 take place at a very high rate (propor-

Fig. 13 Expected value of the number of cycles longer than $|\Lambda|^{0.6}$, $\mathbb{E}_{\Lambda,t}(N_{|\Lambda|})$, as a function of time. The Monte-Carlo chain starts from the identity permutation and has been recorded for various system sizes. Averages were taken over 10^4 realizations
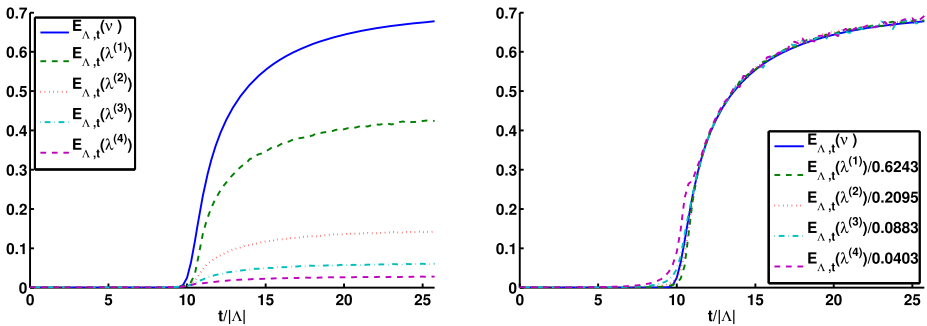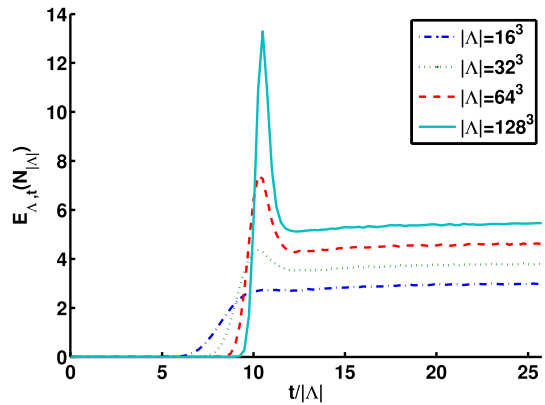
Fig. 14 The Poisson-Dirichlet distribution occurs already during equilibration. (*Left*) Expected values of the fraction of sites in macroscopic cycles $\mathbb{E}_{\Lambda,t}(\nu)$, and of the lengths of the four largest cycles (divided by the volume). (*Right*) The expectation of the i-th longest cycle has been divided by the average length of the i-th part in a random partition with Poisson-Dirichlet distribution. It is always close to $\mathbb{E}_{\Lambda,t}(\nu)$. $|\Lambda| = 128^3$, $T = 0.8$. Averages were taken over 10^4 realizations

tional to $|\Lambda|$) once the phase with macroscopic cycles has been reached. One then expects the lengths of macroscopic cycles to split and merge so fast that the Poisson-Dirichlet distribution appears immediately. The numerical results of Fig. 14 confirms this. Indeed, the expected lengths of the longest cycles, when divided by the number of sites in macroscopic cycles, is equal to the expected values obtained with respect to the Poisson-Dirichlet distribution, that can be found e.g. in [23].

Finally, let us comment on the physical dimension, taken here to be $d = 3$. It is safe to bet that everything is similar in all dimensions greater than 3. On the other hand, the dimension $d = 2$ remains mysterious. There are certainly no macroscopic cycles, as was observed in [15]. An open question is whether a phase occurs where a positive fraction of points belong to mesoscopic cycles. This has been ruled out in the "annealed" model that involves averaging over point positions [8, 9]. The present lattice model may be closer to a Bose gas with interactions, on the other hand, where a Kosterlitz-Thouless phase transition is expected. The presence of mesoscopic cycles could indeed be related to the slow decay of correlation functions.

References

1. Aizenman, M.: Geometric analysis of φ^4 fields and Ising models. Commun. Math. Phys. **86**, 1–48 (1982)
2. Aizenman, M., Nachtergaele, B.: Geometric aspects of quantum spin states. Commun. Math. Phys. **164**, 17–63 (1994)
3. Aldous, D.: Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. Bernoulli **5**, 3–48 (1999)
4. Arratia, R., Barbour, A.D., Tavaré, S.: Logarithmic Combinatorial Structures: A Probabilistic Approach. EMS Monographs in Mathematics. Eur. Math. Soc., Zürich (2003)
5. Berestycki, N.: Emergence of giant cycles and slowdown transition in random transpositions and k-cycles. Electron. J. Probab. **16**, 152–173 (2011)
6. Berestycki, N., Durrett, R.: Limiting behavior for the distance of a random walk. Electron. J. Probab. **13**, 374–395 (2008)
7. Bertoin, J.: Random Fragmentation and Coagulation Processes. Cambridge University Press, Cambridge (2006)
8. Betz, V., Ueltschi, D.: Spatial random permutations and infinite cycles. Commun. Math. Phys. **285**, 469–501 (2009)
9. Betz, V., Ueltschi, D.: Spatial random permutations and Poisson-Dirichlet law of cycle lengths. Electron. J. Probab. **16**, 1173–1192 (2011)
10. Crawford, N., Ioffe, D.: Random current representation for transverse field Ising model. Commun. Math. Phys. **296**, 447–474 (2010)
11. Diaconis, P., Mayer-Wolf, E., Zeitouni, O., Zerner, M.P.W.: The Poisson-Dirichlet law is the unique invariant distribution for uniform split-merge transformations. Ann. Probab. **32**, 915–938 (2004)
12. Feng, S.: The Poisson-Dirichlet Distribution and Related Topics. Probability and Its Applications, Springer, Berlin (2010)
13. Feynman, R.P.: Atomic theory of the λ transition in Helium. Phys. Rev. **91**, 1291–1301 (1953)
14. Grimmett, G.: Space-time percolation. In: In and Out of Equilibrium. 2. Prog. Probab., vol. 60, pp. 305–320. Birkhäuser, Basel (2008)
15. Gandolfo, D., Ruiz, J., Ueltschi, D.: On a model of random cycles. J. Stat. Phys. **129**, 663–676 (2007)
16. Goldschmidt, C., Ueltschi, D., Windridge, P.: Quantum Heisenberg models and their probabilistic representations. In: Entropy and the Quantum II. Contemporary Mathematics, vol. 552, pp. 177–224. Am. Math. Soc., Providence (2011). arXiv:1104.0983
17. Kerl, J.: Shift in critical temperature for random spatial permutations with cycle weights. J. Stat. Phys. **140**, 56–75 (2010)
18. Kingman, J.F.C.: Random discrete distributions. J. R. Stat. Soc. B **37**, 1–15 (1975)
19. Kingman, J.F.C.: Mathematics of Genetic Diversity. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 34. SIAM, Philadelphia (1980)
20. Pitman, J., Yor, M.: The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. Ann. Probab. **25**, 855–900 (1997)
21. Pitman, J.: Poisson-Dirichlet and GEM invariant distributions for split-and-merge transformations of an interval partition. Comb. Probab. Comput. **11**, 501–514 (2002)
22. Schramm, O.: Compositions of random transpositions. Isr. J. Math. **147**, 221–243 (2005)
23. Shepp, L.A., Lloyd, S.L.: Ordered cycle lengths in a random permutation. Trans. Am. Math. Soc. **121**, 340–357 (1966)
24. Sütő, A.: Percolation transition in the Bose gas. J. Phys. A **26**, 4689–4710 (1993)
25. Tóth, B.: Improved lower bound on the thermodynamic pressure of the spin 1/2 Heisenberg ferromagnet. Lett. Math. Phys. **28**, 75 (1993)
26. Tsilevich, N.V.: Stationary random partitions of a natural series. Teor. Veroyatnost. i Primenen. **44**, 55–73 (1999)
27. Wilson, D.: Mixing times of lozenge tiling and card shuffling Markov chains. Ann. Appl. Probab. **14**, 274–325 (2004)