DIRECTED SPATIAL PERMUTATIONS ON ASYMMETRIC TORI

ALAN HAMMOND AND TYLER HELMUTH

Abstract. We investigate a model of random spatial permutations on two-dimensional tori, and establish that the joint distribution of large cycles is asymptotically given by the Poisson–Dirichlet distribution with parameter one. The asymmetry of the tori we consider leads to a spatial bias in the permutations, and this allows for a simple argument to deduce the existence of mesoscopic cycles. The main challenge is to leverage this mesoscopic structure to establish the existence and distribution of macroscopic cycles. We achieve this by a dynamical resampling argument in conjunction with a method developed by Schramm for the study of random transpositions on the complete graph. Our dynamical analysis implements generic heuristics for the occurrence of the Poisson–Dirichlet distribution in random spatial permutations, and hence may be of more general interest.

1. Introduction

Random permutations arise in many contexts, with specific applications leading to different choices of laws. It was observed long ago that random *spatial* permutations are relevant for the study of low-temperature condensed matter physics, e.g., the superfluid transition of helium-4 [9]. The adjective spatial indicates that the law is biased by the geometry of the system under study. In physical contexts the spatial bias tends to zero as the temperature tends to zero.

To make this more concrete, consider the discrete setting of a finite graph $G = (V, E)$, and let $\beta \geq 0$ represent inverse temperature. One studies a law μ_β on permutations of V that is the uniform distribution if $\beta = \infty$, but otherwise is biased by the geometry of G. For example, adjacent vertices may be more likely to be in a common cycle than distant vertices. It is natural to wonder if μ_β with $\beta \gg 1$ retains features of the uniform measure on permutations of V. Of particular interest is the joint distribution of the largest cycles. The question is then whether or not the largest cycles follow $\mathsf{PD}(1)$, the Poisson–Dirichlet distribution with parameter 1. A precise definition of $\mathsf{PD}(1)$ will be recalled below, along with the fact that it describes the large cycles of a uniformly random permutation. Interest in this and related questions originates in physical considerations [9, 20, 2, 21]; recent years have seen an increase in related investigations, e.g., [10, 7, 4, 5, 3, 12, 13, 8].

There is a compelling heuristic argument for why one should expect $\mathsf{PD}(1)$ to arise in great generality for random spatial permutations, irrespective of any connections (or appeals) to physics. The main goal of this article is to rigorously implement this heuristic for a relatively simple model of random spatial permutations. We begin in Section 1.1 by introducing the model we study and formally stating our $\mathsf{PD}(1)$ convergence result, and then discuss $\mathsf{PD}(1)$ heuristics and a sketch of our proof in Section 1.2. It will be clear from this outline that our strategy should be viable in other contexts. The main contribution of this article is therefore an elucidation of how general $\mathsf{PD}(1)$ heuristics can turned into rigorous arguments. We indicate some directions for future study in Section 1.3.

1.1. Directed Spatial Permutations: Model and Results.

Let C_n denote the cyclic graph on n vertices. Let $\mathbb{T}_{n,m}$ denote the Cartesian product of C_n and C_m, i.e., $\mathbb{T}_{n,m} \cong \mathbb{Z}^2/(n\mathbb{Z} \times m\mathbb{Z})$ is an n by m subgraph of $\mathbb{Z}^2$ with periodic boundary conditions (a torus graph). Vertices in $\mathbb{T}_{n,m}$ will be denoted $\boldsymbol{x} = (x_1, x_2)$. Expressions in coordinates will always be interpreted modulo the dimensions of the torus:

$$\boldsymbol{x} = (x_1, x_2) = (x_1 \mod n, x_2 \mod m).$$

Date: 05 June 2023.

A bijection $\Phi\colon \mathbb{T}_{n,m} \to \mathbb{T}_{n,m}$ determines a discrete-time dynamical system. The orbits under Φ are cyclic subgraphs of $\mathbb{T}_{n,m}$ which will be called *cycles*. We write $|\mathsf{C}|$ for the number of vertices (equivalently, edges) in a cycle C. The aim of this article is to study the joint statistics of the large cycles of a family of random bijections defined as follows.

Let ϕ be the single-step distribution of a one-dimensional simple symmetric random walk, i.e., ϕ takes values in $\{-1, 0, 1\}$ and $\phi(1) = \phi(-1) = a$. Let $(\tilde{\phi}_{\boldsymbol{x}})_{\boldsymbol{x}\in\mathbb{T}_{n,m}}$ be a collection of IID random variables, each distributed as ϕ. Define $\tilde{\Phi}\colon \mathbb{T}_{n,m} \to \mathbb{T}_{n,m}$ by

$$\tilde{\Phi}(\boldsymbol{x}) = (x_1 + 1, x_2 + \tilde{\phi}_{\boldsymbol{x}} + 1). \tag{1}$$

When $n \neq m$ the presence of $+1$ in the second coordinate introduces a bias that will play an important role in what follows.

Note that $\tilde{\Phi}$ given by (1) will typically not be a bijection. Let Φ denote the random variable $\tilde{\Phi}$ conditioned on being a bijection. We call Φ a *directed spatial permutation*. Our symmetry hypothesis on ϕ implies that this random variable depends the parameter $a = \phi(1)$ as well as on n, m. We will study Φ as m tends to infinity under the following assumptions. The first is mild; the second will be discussed below.

(1) **Non-degeneracy**: $0 < a < 1/2$, and
(2) **Asymmetry of** $\mathbb{T}_{n,m}$: the horizontal dimension n is given by $n = m + C(m)$ with $C(m) = \lceil C'\sqrt{m \log m}\rceil$, C' an absolute constant.

Let $\mathcal{S}_N$ denote the set of permutations of N elements. For directed spatial permutations $N = nm$, but the next definition is more general.

Definition 1.1. *The* cycle structure $\mathcal{X}(\pi)$ *of a bijection $\pi \in \mathcal{S}_N$ is the non-increasing list of normalized lengths $(L_1, L_2, \dots)$ of the cycles of π, where the normalized length L of a cycle C is $N^{-1}|\mathsf{C}|$. If π consists of exactly k distinct cycles, then $L_j = 0$ for $j > k$.*

Before stating our main theorem we recall the definition of the *Poisson–Dirichlet distribution with parameter* 1, denoted $\mathsf{PD}(1)$. $\mathsf{PD}(1)$ is the probability distribution on the infinite-dimensional simplex

$$\Delta = \{(y_i)_{i\in\mathbb{N}_+} \mid y_i \geq 0, \sum_{i\in\mathbb{N}_+} y_i = 1, y_1 \geq y_2 \geq \dots\}$$

defined by the following procedure. Let $(U_i)_{i\in\mathbb{N}_+}$ be a sequence of IID random variables, uniform on $[0, 1]$. Set $x_1 = U_1$, and inductively $x_i = U_i(1 - \sum_{j=1}^{i-1} x_j)$. Let $(y_i)_{i\in\mathbb{N}_+}$ be the x_i in non-increasing order. Then $\mathsf{PD}(1)$ is the law of $(y_i)_{i\in\mathbb{N}_+}$. It is well-known that uniformly random permutations have $\mathsf{PD}(1)$ as their limiting cycle structure [16, 22].

Theorem 1.2. *Fix $0 < a < 1/2$, let $C(m) = \lceil C'\sqrt{m \log m}\rceil$, and suppose C' is a large enough absolute constant (independent of a). Let Φ be the law of directed spatial permutations on $\mathbb{T}_{n,m}$ with $n = m + C(m)$. As $m \to \infty$, the distribution of cycle lengths $\mathcal{X}(\Phi)$ converges weakly to $\mathsf{PD}(1)$.*

1.2. Heuristics and Proof Outline.

1.2.1. *Dynamical Heuristic for Poisson–Dirichlet.* The heuristic for why the Poisson–Dirichlet distribution should arise when studying random spatial permutations is based on a dynamical perspective on the uniform measure on $\mathcal{S}_N$. The relevant dynamics are the following. Write $[\![N]\!] = \{0, 1, \dots, N-1\}$.

Definition 1.3. *The* random transposition model *on $[\![N]\!]$ is the Markov chain $(\pi_j)_{j\geq 0}$ with state space $\mathcal{S}_N$ and transition probabilities*

$$\mathbb{P}[\pi_{j+1} = \sigma\pi_j] = \binom{N}{2}^{-1}, \qquad \sigma \text{ a transposition.}$$

By checking the detailed balance equations one can observe that the stationary distribution of the random transpositions model is the uniform distribution on $\mathcal{S}_N$. Since the limiting cycle structure of the uniform distribution is $\mathsf{PD}(1)$, the distribution of the cycle structure of the random transposition Markov chain will essentially be $\mathsf{PD}(1)$ when the chain has been run for long enough.

If we are only concerned with the sizes of cycles, an alternate view of the random transposition dynamics as a size-biased split-merge dynamics can be given. The $(j+1)^{\text{st}}$ step of the dynamics chooses two (not necessarily distinct) cycles C_1 and C_2 of π_j with probabilities $|\mathsf{C}_i|/N$, and (i) if $\mathsf{C}_1 \neq \mathsf{C}_2$, the cycles are merged, and (ii) if instead $\mathsf{C}_1 = \mathsf{C}_2$, the cycle is uniformly split into two cycles. We refer to case (i) as a *merge* and case (ii) as a *split*.

Heuristic conditions for when Poisson–Dirichlet statistics are likely to occur in models of random spatial permutations can now be given. Suppose the following conditions hold (possibly in an approximate sense, which we will not attempt to quantify):

[LC] Large cycles exist and occupy a positive fraction of space.

[M] Large cycles are well-mixed in space, in the following sense. Let X be a random variable that chooses two distinct points in space, independently of the cycle configuration. For example, X could be the endpoints of a uniformly chosen edge. Then we suppose that the probability the points of X are contained in two cycles C, C' is proportional to $|\mathsf{C}||\mathsf{C}'|$.

[D] The law on permutations is invariant under a local dynamics that updates by (i) choosing two points according to X and (ii) resampling where the points of X are mapped to with the correct marginal law.

Note that the resampling in [D] may result in no change in the configuration (the resampled configuration could be the same), but if a change occurs, then the selected cycles merge if they were distinct, and split if they were the same.

Condition [LC] ensures that [M] is a non-trivial statement. Conditions [M] and [D] indicate that the split-merge dynamics on large cycles is essentially the same as the random transposition dynamics, provided the marginal law does not have a systematic bias towards either splits or merges. Consequently one expects $\mathsf{PD}(1)$ to arise under a suitable normalization. If a consistent and systematic bias towards splits or merges is present in the local dynamics of [D], then the parameter 1 may change. A more explicit discussion of this heuristic in a particular context can be found in [11, Section 8], see also [5, Section 1.2].

1.2.2. *Proof Outline.* Our analysis of directed spatial permutations can be viewed as an implementation of the rough heuristic presented in Section 1.2.1. In fact our implementation is slightly different, in that we establish versions of the conditions [LC], [M] and [D] for large (but not macroscopic) pieces of cycles. The fact that one can argue for $\mathsf{PD}(1)$ via mesoscopic versions of these conditions should be a rather general fact, and showing how this can be done is one of the main contributions of this paper.

This section sketches our argument. At several steps in the full proof we will condition on events of high probability; we largely omit these technical matters in the following discussion.

First considerations and dynamics [D]. In Section 2 we examine the typical structure of the random bijections Φ. The key observation is that the correlation structure induced by conditioning the random map $\tilde{\Phi}$ to be a bijection is fairly simple. First, correlations only exist within each column of $\mathbb{T}_{n,m}$, as the condition to be a bijection is an independent condition on each column. By column we mean the set of vertices with fixed horizontal coordinate (see Section 1.4). Second, the correlations within each column can be explicitly described on an event of high probability: there is an explicit description of $\mathbb{P}[\phi_{\boldsymbol{x}} = 1, \boldsymbol{x} \in A]$ for a subset A of vertices in a column. These probabilities give a complete description of $(\phi_{\boldsymbol{x}})_{\boldsymbol{x}\in\mathbb{T}_{n,m}}$ since the variables form a bijection. The explicit description is in terms of the hard-core model from statistical physics.

There are two important consequences of these observations. First, the correlations in each column are rapidly (exponentially) decaying. This is a well-known fact about the hard-core model on a one-dimensional graph. Second, we can view the underlying randomness as arising from the hard-core model description. Doing so gives an explicit dynamics (Glauber dynamics for the hard-core model) that preserves the law of Φ. We use these dynamics to implement [D] from Section 1.2.1.

Cycles are not small, and a mesoscopic interpretation of large cycles [LC]. In Section 3 we start to explore the cycles of Φ by making use of the observation that the initial process of revealing a cycle has the law of a simple random walk. More precisely, c.f. (1), the first n vertices of a given cycle have the law of a simple random walk with step distribution $1 + \phi$. For tori $\mathbb{T}_{n,m}$ with suitable

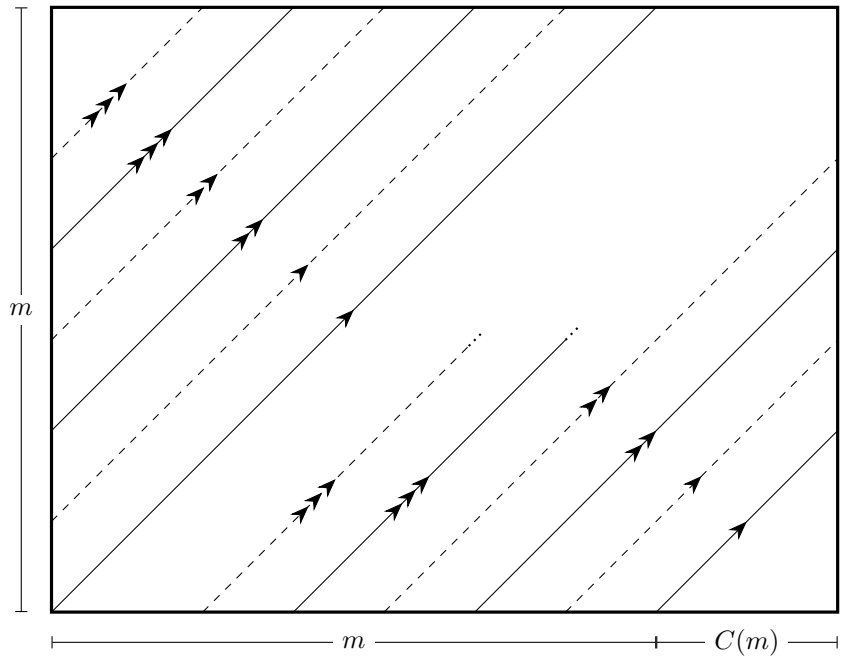

Figure 1. Plot of the expected location of portions of two cycles (solid and dashed). Horizontal traversals of the system, called strands in the main text, share the same number of arrows. After each strand the expected vertical displacement is $C(m)$. It takes $m/C(m)$ strands for the expected vertical displacement to be zero.

aspect ratios, the properties of these initial n steps imply that cycles cannot be too small. This is the first place where our assumption on $C(m)$ plays a role.

In more detail, the first n steps are those of a random walk with drift in the vertical direction. Our choice of $C(m)$ ensures that the first time a cycle returns to a column it is (with high probability) at vertical distance of order $C(m)$ from where it started. We may repeat this argument many times via a union bound, and this leads to a formalization of the idea that cycles are not too small in terms of *global traversals*. Roughly speaking, global traversals are mesoscopic objects that pass through the whole torus *vertically*, i.e., we follow a cycle until the first time its vertical displacement in a column is zero. This occurs after roughly $m/C(m)$ horizontal traversals of the system, see Figure 1. A precise definition of global traversals is given in Section 3; the preceding discussion neglects some technical aspects of the definition we use in practice. The essential point is that our assumption on $C(m)$ gives enough *a priori* control to deduce that each cycle will contain at least one global traversal. Note that this establishes [LC] on a mesoscopic (as opposed to macroscopic) scale — the macroscopic scale requires controlling order m (as opposed to $m/C(m)$) horizontal traversals.

Concentration of contacts and the mixing condition [M]. In Section 4 we leverage the geometry of global traversals to establish a form of [M] at a mesoscopic scale. This is done by showing that two global traversals encounter one another (*contact* one another) rather often, and in such a manner that the number of contacts between two global traversals is a concentrated random variable. Since the cycles of Φ are comprised of global traversals, this means that the cycles of Φ are well-mixed.

To understand this, consider Figure 1. Given one global traversal, any other global traversal starts in one of the "corridors" determined by the first. Ignoring correlations, the distance between the two global traversals evolves like a random walk with increments given by $X + X'$, X and X' distributed as ϕ. Due to the corridor structure, this distance between global traversals should be thought of as being a random walk run on an cycle of length $\Theta(C(m))$ for approximately $m^2/C(m)$ steps. The number

of contacts is the number of times this walk hits zero, and this is a concentrated random variable. Moreover, it is independent of the pair of global traversals chosen.

A precise formulation of contacts and concentration is given in Section 4.1. The (exponentially decaying) correlations that were neglected in the discussion above can be taken into account without significantly altering the conclusion. The details of this occupy Sections 4.2 and 4.3; they play no role in the sequel.

From [LC], [M], and [D] to $\mathsf{PD}(1)$. In Section 5 we consider the Glauber dynamics which leave Φ invariant, and use them to establish Theorem 1.2. There are two steps. First, in Section 5.1 we note that under these dynamics cycles split and merge with one another when Φ is updated at a contact between global traversals. Moreover, the dynamics choose contacts to update uniformly at random. Since the number of contacts between any pair of global traversals is concentrated around the same value, this implies the dynamics are *effectively* a size-biased split-merge dynamics on the cycles of Φ.

For random transpositions themselves we know $\mathsf{PD}(1)$ statistics are present in the invariant distribution. Our invariant distribution is not amenable to such a direct conclusion, so we must argue differently. We do this by utilizing work of Schramm that explained how size-biased split-merge dynamics lead to $\mathsf{PD}(1)$ *well before* equilibrium is reached in the random transposition context [19]. We show that the errors hidden in the word "effectively" in the previous paragraph do not cause any difficulties in implementing Schramm's argument. This is carried out in Section 5.2.

1.3. **Future Directions.** We have considered directed spatial permutations on asymmetric tori $\mathbb{T}_{n,m}$ with $n = m + C(m)$ with $C(m) = \lceil C'\sqrt{m \log m} \rceil$, C' large enough. It is natural to ask what happens for general values of $C(m)$. While our arguments certainly work for somewhat larger values of $C(m)$ (i.e., replacing $\sqrt{m \log m}$ with $m^{\frac{1}{2}+a}$, $a > 0$ not too large), new ideas will be needed at some point. For example, $C(m) = m$ is essentially the same situation as for $C(m) = 0$, and the *a priori* estimates we use to establish that cycles are not too small break down. It is with these *a priori* estimates in mind that we chose the scale $\sqrt{m \log m}$. It seems possible that our restrictions that the constant C' be large enough could be weakened at the expense of more technical arguments.

In the setting of a general choice of $C(m)$ it is not entirely clear what the correct scaling of the cycle structure is nor what the limiting distribution is. It is at least imaginable that there could be an analogue of the rational resonances phenomenon found in [14]. Note that if $C(m) \gg m^2$ then one will obtain a version of Theorem 1.2 (the first n steps of a cycle will be a well-mixed random walk on the m-cycle), but if $n = m + C(m)$ is constant, then one will not. It would be interesting to develop a more complete understanding of directed spatial permutations as $C(m)$ varies.

Our results concern the equilibrium distribution of directed spatial permutations on $\mathbb{T}_{n,m}$, and the proof uses a natural dynamics that preserve the equilibrium distribution. It seems natural to ask about the effect of these dynamics started away from equilibrium, in analogy with Schramm's proof of Aldous's conjecture that $\mathsf{PD}(1)$ emerges in the random transposition model well before the mixing time occurs [19]. E.g., if one starts from $\phi_{\boldsymbol{x}} = 0$ for all $\boldsymbol{x}$ and runs the Glauber dynamics underlying our proof, do $\mathsf{PD}(1)$ statistics arise prior to the mixing time of the chain?

Our arguments have made use of the geometric structure of $\mathbb{T}_{n,m}$ to give relatively simple arguments for the mixing of large cycles in the system. Applying our ideas in other settings would be very interesting. One possibility is to study models of directed spatial permutations in higher dimensions. Another is to study the properties of somewhat different models, e.g., random mappings (as opposed to bijections). Finally, there are well-known models like the random stirring model (see [8] for fantastic recent progress) or loop representations of quantum spin systems, see [11].

1.4. **Conventions and Notation.** We write $\mathbb{N}_+ = \{1, 2, 3, \dots\}$ and $\mathbb{N} = \{0, 1, 2, \dots\}$. Discrete intervals are denoted $[\![a, b]\!] = \{a, a+1, \dots, b-1, b\}$ for $a < b \in \mathbb{N}$, and we abbreviate $[\![b]\!] = [\![0, b]\!]$. For typographic convenience we sometimes write $a \wedge b$ for $\min\{a, b\}$. We use standard asymptotic notation: $f(x) = O(g(x))$ if there exists a $C > 0$ such that $|f(x)| \le Cg(x)$, and $f(x) = \Theta(g(x))$ if $f(x) = O(g(x))$ and $g(x) = O(f(x))$.

Let π_i denote the projection from $\mathbb{T}_{n,m}$ to the i^{th} coordinate, i.e., $\pi_i(\boldsymbol{x}) = x_i$ for $i = 1, 2$. The j^{th} *column* of $\mathbb{T}_{n,m}$ is $\pi_1^{-1}(j)$ for $j \in [\![n]\!]$; the j^{th} *row* is $\pi_2^{-1}(j)$, $0 \le j \le m-1$.

A *geometric random variable W with success probability p* satisfies $\mathbb{P}[W = n] = (1-p)^{n-1}p$, i.e., W is the trial on which success occurs. A random variable X is *sub-exponential* if there is a $c > 0$ such that $\mathbb{P}[|X| \geq t] \leq 2\exp\{-ct\}$ for all $t \geq 0$.

In many places our arguments will require that the parameter m controlling the size of $\mathbb{T}_{n,m}$ is large enough. We do not write this explicitly in our hypotheses.

2. The typical law of Φ

The collection of IID random variables $(\tilde{\phi}_{\boldsymbol{x}})_{\boldsymbol{x}\in\mathbb{T}_{n,m}}$ define a map $\tilde{\Phi}$ by $\tilde{\Phi}(\boldsymbol{x}) = (x_1 + 1, x_2 + \tilde{\phi}_{\boldsymbol{x}} + 1)$. Recall that $\tilde{\phi}$ are symmetric on $\{-1, 0, 1\}$, that we assume $\phi(1) = a \in \left(0, \frac{1}{2}\right)$, and that Φ denotes $\tilde{\Phi}$ conditioned to be a bijection. We similarly write $(\phi_{\boldsymbol{x}})_{\boldsymbol{x}\in\mathbb{T}_{n,m}}$ to denote $(\tilde{\phi}_{\boldsymbol{x}})_{\boldsymbol{x}\in\mathbb{T}_{n,m}}$ conditioned on $\tilde{\Phi}$ being a bijection. This conditioning induces correlations between the $\phi_{\boldsymbol{x}}$. This section describes these correlations. More precisely, by further conditioning on an event that occurs with high probability we will obtain an explicit description of the resulting correlation structure.

2.1. No global shifts.

By a slight abuse of notation, for $j \in [\![n]\!]$ we will write $\tilde{\Phi}_j$ for the y-coordinate of the restriction of $\tilde{\Phi}$ to the jth column, i.e., $\tilde{\Phi}_j(k) = k + \phi_{(j,k)} + 1$ records the y-coordinate of $\tilde{\Phi}((j,k))$. We think of $\tilde{\Phi}_j$ as a map from C_m to C_m. If $\tilde{\Phi}$ is a bijection of $\mathbb{T}_{n,m}$, then $\tilde{\Phi}_j$ induces a bijection of C_m for all $j \in [\![n]\!]$. Writing Φ_j for the restriction of Φ to column j, note that Φ_j and Φ_k are independent for $j \neq k$, as the condition that $\tilde{\Phi}$ is a bijection is a separate condition for each column.

Definition 2.1. *The map Φ_j is a* global up shift *if $\phi_{(j,k)} = 1$ for all $k \in C_m$. Column j is a* global down shift *if $\phi_{(j,k)} = -1$ for all $k \in C_m$.*

Having no global shifts *and* knowing that Φ_j is a bijection constrains Φ_j severely.

Lemma 2.2. *Suppose Φ_j is not a global shift. If $\phi_{(j,k)} = 1$, then $\phi_{(j,k+1)} = -1$. If $\phi_{(j,k)} = -1$, then $\phi_{(j,k-1)} = 1$.*

Proof. The proofs being similar, we derive only the first statement. Suppose that $\phi_{(j,k)} = 1$ for some $k \in [\![n]\!]$. We can make the following inferences from Φ_j being a bijection. The value $\phi_{(j,k+1)}$ cannot be zero. If $\phi_{(j,k+1)}$ equals one, then $\phi_{(j,k+2)}$ must also be one. By induction, $\phi_{(j,\ell)} = 1$ for all $\ell \in [\![n]\!]$, so that ϕ_j is a global shift, contrary to hypothesis. Thus, $\phi_{(j,k)} = -1$. ∎

It will sometimes be convenient to think of each $\phi_{(j,k)}$ as representing an *arrow* that points up (value 1), down (value -1), or horizontally (value 0). We say the arrows $\phi_{(j,k)}$ and $\phi_{(j,k+1)}$ *swap* if $\phi_{(j,k)}$ points up and $\phi_{(j,k+1)}$ points down. They are *parallel* if they are both horizontal.

Lemma 2.3. *Under Φ, the probability of a global shift occurring is at most ne^{-cm} for some $c = c(a) > 0$.*

Proof. Assume m is even; a similar argument applies to n odd. The probability that the unconditioned map $\tilde{\Phi}$ has a global shift in column j is $2a^m = 2(a^2)^{m/2}$, where we recall $a = \phi(1) \in \left(0, \frac{1}{2}\right)$. Consider the two partitions of the vertices in $\pi_1^{-1}(j)$ into adjacent pairs $\{(j,k), (j,k+1)\}$. For each choice of if these pairs of arrows swap or are parallel the resulting map is a bijection. Hence the probability of a bijection under $\tilde{\Phi}$ is at least $2((1-2a)^2 + a^2)^{m/2}$. This proves that conditionally on being a bijection, the probability of a global shift in column j is exponentially decaying in m if $0 < a < 1/2$. The lemma follows by a union bound over the n columns of $\mathbb{T}_{n,m}$. ∎

2.2. The underlying hard-core model.

We briefly recall the definition of the hard-core model, which we will shortly see is connected with the bijection Φ. Recall that an *independent set $I \subset V$* of a graph $G = (V, E)$ is a subset of vertices such that $u, v \in I$ implies that the edge $\{u, v\}$ is not in E. Write $\mathcal{I}$ for the set of independent sets of G. We sometimes identify elements of $\mathcal{I}$ with vectors $\sigma = (\sigma_x)_{x\in V} \in \{0,1\}^V$, with $\sigma_x = 1$ representing that x is in the independent set.

Definition 2.4. *The* hard-core model *with* activity $\lambda > 0$ *on a graph $G = (V, E)$ is the probability measure $\mathbb{P}$ on 2^V defined by*

$$\mathbb{P}[A] \propto \mathbb{1}_{\{A\in\mathcal{I}\}}\lambda^{|A|}, \tag{2}$$

where $|A|$ denotes the cardinality of $A \subset V$.

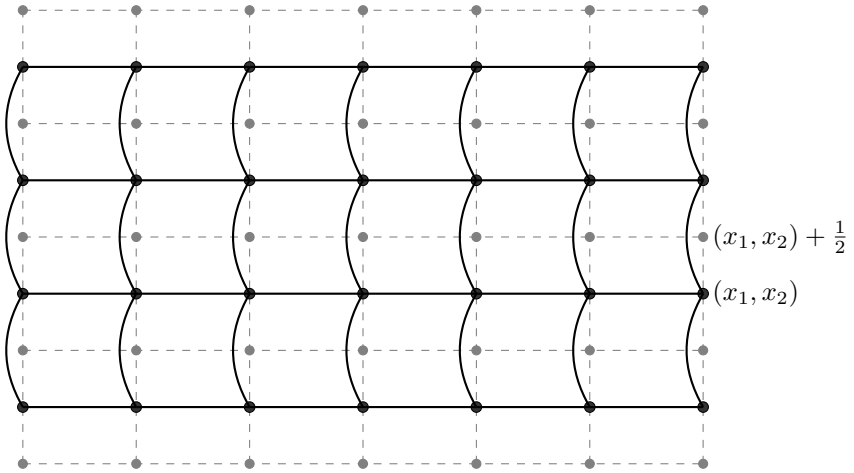

Figure 2. A subset of $\mathbb{T}_{n,m}$ is in grey, and a subset of $\mathbb{T}^\star_{n,m}$ is in black. The graph $\mathbb{T}^\star_{n,m}$ has vertices $v_{\boldsymbol{x}} = ((x_1, x_2), (x_1, x_2 + 1))$ given by the vertical edges of $\mathbb{T}_{n,m}$. Two vertices $v_{\boldsymbol{x}}$, $v_{\boldsymbol{x}'}$ form an edge in $\mathbb{T}^\star_{n,m}$ if either (i) $x_1 = x_1'$ and $x_2 - x_2' \in \{\pm 1\}$ or (ii) $x_2 = x_2'$ and $x_1 - x_1' \in \{\pm 1\}$.

Let $\mathbb{T}^\star_{n,m} \cong C_n \times C^\star_m$ be the dual of $\mathbb{T}_{n,m}$ (we write $C^\star_m \cong C_m$ for the dual of C_m). See Figure 2 and the accompanying caption for a depiction and precise description. It is convenient to identify the vertices of $\mathbb{T}^\star_{n,m}$ with the midpoints of the vertical edges of $\mathbb{T}_{n,m}$, and $\boldsymbol{x} + \frac{1}{2}$ will denote the vertex $((x_1, x_2), (x_1, x_2 + 1)) \in \mathbb{T}^\star_{n,m}$ for $\boldsymbol{x} \in \mathbb{T}_{n,m}$. See Figure 2. Recall the definition of swapping arrows from Section 2.1. It is convenient to extend this terminology to the vertices of $\mathbb{T}^\star_{n,m}$.

Definition 2.5. *A* swap *occurs on the edge $\boldsymbol{x} + \frac{1}{2}$ if $\big(\phi_{x_1}(x_2), \phi_{x_1}(x_2 + 1)\big)$ equals $(1, -1)$.*

Given a bijection of $\mathbb{T}_{n,m}$, consider the subset of vertices of $\mathbb{T}^\star_{n,m}$ at which a swap occurs on the corresponding edge of $\mathbb{T}_{n,m}$. This defines a map between bijections Φ of $\mathbb{T}_{n,m}$ and subsets of $V(T^\star_{n,m})$.

Proposition 2.6. *Consider the set of bijections Φ that do not contain a global shift. The induced law on subsets of $V(\mathbb{T}^\star_{n,m})$ is the product of independent hard-core model measures with activity $\lambda = \lambda(a) = a^2/(1-a)^2$ on each column $C^\star_m$ of $\mathbb{T}^\star_{n,m}$.*

Proof. By Lemma 2.2, every vertex $\boldsymbol{x}$ of $\mathbb{T}_{n,m}$ not assigned a horizontal arrow (i.e., $\phi_{\boldsymbol{x}} \neq 0$) is part of exactly one swap. Thus every configuration which is given non-zero probability is an independent set in each column $C^\star_m$ of $\mathbb{T}^\star_{n,m}$. Since any independent set corresponds to a possible set of swaps, every independent set occurs with positive probability.

Because Φ_j and Φ_k are independent for $j \neq k$, it is enough to show that the law on a single column $C^\star_m$ is given by (2). To this end, note that, for $A \subset V(C^\star_m)$ an independent set,

$$\mathbb{P}[A] \propto a^{2|A|}(1-a)^{n-2|A|} \propto \left(\frac{a^2}{(1-a)^2}\right)^{|A|}. \qquad \blacksquare$$

For the rest of this article we will denote the activity of the hard-core model induced by our model of random bijections by

$$\lambda = \lambda(a) = a^2(1-a)^{-2}. \tag{3}$$

Remark 1. *The proof of Proposition 2.6 used a bijection between configurations of Φ (conditioned on having no global shifts) and independent sets. Using the identification of independent sets with vectors $\sigma = (\sigma_x)_{x \in V}$, this bijection had the property that $\phi_{\boldsymbol{x}} = 1$ iff $\sigma_{\boldsymbol{x} + \frac{1}{2}} = 1$, $\phi_{\boldsymbol{x}} = -1$ iff $\sigma_{\boldsymbol{x} - \frac{1}{2}} = 1$, and*

$\phi_{\boldsymbol{x}} = 0$ *iff* $\sigma_{\boldsymbol{x}+\frac{1}{2}} = \sigma_{\boldsymbol{x}-\frac{1}{2}} = 0$. *In other words, there is a perfect coupling between the hard-core model and* Φ *conditioned on having no global shifts.*

Remark 2. *Lemma 2.3 implies that conditioning on the non-occurrence of a global shift is asymptotically irrelevant if* $n = n(m)$ *is sub-exponential in* m. *Under this conditioning the coupling of Remark 1 means that we can use the corresponding hard-core models as the underlying source of randomness of* Φ. *We will do this frequently in what follows.*

2.3. **Correlation decay in the hard-core model.** Since C_m is one-dimensional, the hard-core model on C_m has rapidly decaying correlations. For later use, the next lemma formalizes this. For $A, B \subset C_m$ let $d(A, B) = \min_{a\in A, b\in B} |a - b|$, where $|a - b|$ is the graph distance between a and b. Given a vector $\sigma = (\sigma_x)_{x\in V}$ and $A \subset V$ let $\sigma_A = (\sigma_x)_{x\in A}$.

Lemma 2.7. *Let* $\mathbb{P}$ *be the law of the hard-core model on* C_m, *and let* σ *denote the corresponding random independent set. There are positive constants* c_1, c_2, *depending on* λ *but not on* m *such that*

$$\left| \mathbb{P}\left[\sigma_A = \tilde{\sigma}_A, \sigma_B = \tilde{\sigma}_B\right] - \mathbb{P}_{HC}\left[\sigma_A = \tilde{\sigma}_A\right] \mathbb{P}_{HC}\left[\sigma_B = \tilde{\sigma}_B\right] \right| \leq c_1 e^{-c_2 d(A,B)}. \tag{4}$$

for all subsets $A, B \subset C_m$ *such that* A *is connected and all configurations* $\tilde{\sigma}$.

Proof. This is a well-known fact, and we omit a detailed proof. It can be derived by using the observation that the partition function (normalizing constant) $Z_n(\lambda)$ for the hard-core measure on a linear graph with n vertices satisfies the recurrence $Z_n(\lambda) = Z_{n-1}(\lambda) + \lambda Z_{n-2}(\lambda)$, which has an elementary solution. ∎

The next lemma records how correlation decay in the hard-core model translates into correlation decay for the random variables $\phi_{\boldsymbol{x}}$ that comprise Φ.

Lemma 2.8. *Condition on* Φ *having no global shifts. There exist positive constants* c_1, c_2 *depending on* λ *but not on* m *such that*

$$\left| \mathbb{P}\left[\phi_A = \tilde{\phi}_A, \phi_B = \tilde{\phi}_B\right] - \mathbb{P}\left[\phi_A = \tilde{\phi}_A\right] \mathbb{P}\left[\phi_B = \tilde{\phi}_B\right] \right| \leq c_1 e^{-c_2 d(A,B)}, \tag{5}$$

for all subsets A, B *belonging to single columns such that* A *is connected.*

Proof. The claim is immediate by independence if A, B are in disjoint columns. Otherwise, let $\mathbb{P}_{HC}$ be the law of independent hard-core models on each column of $\mathbb{T}^{\star}_{n,m}$, and let $\mathbb{P}$ be the induced law on $\{\phi_{\boldsymbol{x}}\}_{\boldsymbol{x}\in\mathbb{T}_{n,m}}$. Then by Remark 1, and letting σ_A denote the hard-core configuration corresponding to the arrows specified by $\tilde{\phi}_A$,

$$\mathbb{P}\big[\, \phi_{\boldsymbol{x}} = w \,\big|\, \sigma_A \,\big] = \begin{cases} \mathbb{P}_{HC}\big[\, \sigma_{\boldsymbol{x}-\frac{1}{2}} = 1 \,\big|\, \sigma_A \,\big] & w = -1 \\ \mathbb{P}_{HC}\big[\, \sigma_{\boldsymbol{x}+\frac{1}{2}} = 1 \,\big|\, \sigma_A \,\big] & w = 1 \\ \mathbb{P}_{HC}\big[\, \sigma_{\boldsymbol{x}-\frac{1}{2}} = \sigma_{\boldsymbol{x}+\frac{1}{2}} = 0 \,\big|\, \sigma_A \,\big] & w = 0 \end{cases}, \tag{6}$$

and similarly for computations of $\mathbb{P}\left[\phi_{\boldsymbol{x}} = w, \phi_{\boldsymbol{y}} = w' | \sigma_A\right]$. If A, B contain vertices in the same column the claim now follows from Lemma 2.7 since the probability of a configuration on the interval A is only affected by the configuration of B closest to each side of the interval. ∎

In Section 4.2 we will make use of the hard-core model on $\mathbb{Z}$. Let $\mathbb{P}_m$ denote the hard-core measure on C_m with a fixed activity λ. Correlation decay allows $\mathbb{P}_{\mathbb{Z}}$, the hard-core model on $\mathbb{Z}$, to be defined as the unique weak limit of the $\mathbb{P}_m$ as $m \to \infty$. The resulting measure is translation invariant and inherits the correlation decay properties of the finite-volume measures. In particular, $\mathbb{P}_{\mathbb{Z}}\left[\sigma_x = 1\right]$ exists, is positive if $\lambda > 0$, and is independent of x.

2.4. **Glauber dynamics for the hard-core model.** Let $G = (V, E)$ be a finite graph. The *Glauber dynamics* for the hard-core model are a Markov chain on $\{0,1\}^V$ whose equilibrium distribution is that of the hard-core model on G. Let $\sigma = (\sigma_v)_{v\in V}$ denote an independent set. A single step of the chain is defined as follows.

(1) Choose a vertex $v \in V$ uniformly at random.
(2) Conditionally on $\{\sigma_x \mid \{x, v\} \in E\}$, resample σ_v according to the hard-core measure.

The claim that the hard-core measure is the equilibrium distribution of the chain follows by checking the detailed balance equations.

Remark 3. *Specialize to the case $G = C_m^\star$, and recall the bijection between hard-core and arrow configurations from Remark 1. A Glauber dynamics update at v can change the value of σ_v only if $\sigma_{v+1} = \sigma_{v-1} = 0$. If $\sigma_v = 0$ as well, then the arrows at $v \pm \frac{1}{2}$ are parallel, while if $\sigma_v = 1$ there is a swap at v. If a Glauber update at v results in a change of σ_v then either parallel arrows become swapping arrows or vice versa. Viewing these arrows as being single steps in the cycle(s) of $v \pm \frac{1}{2}$, it follows that this update merges the cycles if they are distinct, and splits the cycle otherwise. This induced dynamics on cycles will be our main tool for studying the cycle structure of Φ.*

3. Equilibrium properties I: strands and global traversals

The main aim of this section is to define convenient geometric structures that allow us reduce geometric statements about cycles to statements about one-dimensional random walks. The key outputs of this section will be the idea of a *global traversal*, that all cycles are comprised of at least one global traversal, and that the number of global traversals in a cycle is an accurate proxy for the size of a cycle.

Henceforth we set $C(m) = \lceil C'\sqrt{m \log m}\rceil$ for a constant C' that will be determined later (see Corollary 3.6 and Remark 6).

3.1. **Random walk interpretation: strands.** The cycles of a random bijection Φ can be thought of as closed random walk paths. The lack of independence between the random variables $(\phi_{\boldsymbol{x}})_{\boldsymbol{x}\in\mathbb{T}_{n,m}}$ that comprise Φ prevents direct control of these closed paths. This section introduces *strands* as a probabilistically convenient tool that will be used to study the local geometric structure of cycles.

Definition 3.1. *Fix $\boldsymbol{x} \in \mathbb{T}_{n,m}$, and let $\boldsymbol{x}_k = \Phi^k(\boldsymbol{x})$ for $k = 0, \dots, n-1$. The* strand $\mathsf{St}(\boldsymbol{x})$ starting at $\boldsymbol{x}$ *is the sequence $(\boldsymbol{x}_k)_{k=0}^{n-1}$. Explicitly,*

$$\mathsf{St}(\boldsymbol{x}) = \left(\left(x_1 + j, x_2 + \sum_{k=0}^{j-1}(\phi_{\boldsymbol{x}_k} + 1)\right)\right)_{j=0}^{n-1}. \tag{7}$$

The probabilistic content of a strand is contained in the increments $\phi_{\boldsymbol{x}_k}$, which encode a one-dimensional random walk trajectory $\tilde{\mathsf{St}}(\boldsymbol{x})$,

$$\tilde{\mathsf{St}}(\boldsymbol{x}) = \left(x_2 + \sum_{k=0}^{j-1}\phi_{\boldsymbol{x}_k}\right)_{j=0}^{n-1}, \tag{8}$$

with the convention that the sum from 0 to -1 is zero. The independence of the increments in distinct columns along a strand gives a simple but important observation.

Lemma 3.2. *Condition on the event that Φ contains no global shift. $\tilde{\mathsf{St}}(\boldsymbol{x})$ is equal in law to an n-step random walk with IID increments in $\{-1, 0, 1\}$. The law of the increments does not depend on $\boldsymbol{x}$, is symmetric, and has finite and non-zero variance.*

Proof. Without conditioning on Φ not containing a global shift the claim is immediate. Since the conditioning does not affect the distribution of a single arrow in a column, the conclusion follows. ∎

3.2. **Strand geometry.** The fact that $\tilde{\mathsf{St}}(\boldsymbol{x})$ is an n-step lazy simple random walk allows for strong *a priori* statements about the geometry of strands. We record the key lemma in its natural generality.

Lemma 3.3. *Let $S_0 = 0$, $S_k = \sum_{i=1}^k X_i$, with the X_i IID and distributed as X. Suppose X is symmetric about 0, $\mathrm{Var}X = \sigma^2 > 0$, and that X has a finite moment generating function. For $A > 0$ and $c > 0$ fixed constants the following estimate holds:*

$$\mathbb{P}\left[\max_{0\le j\le ct} |S_j| \ge A\sqrt{t\log t}\right] \le (1+o(1))t^{-\frac{A^2}{2c\sigma^2}}. \tag{9}$$

Proof. Let $a > 0$. Note that

$$\begin{aligned}\mathbb{P}\left[S_j \in [-a,a], 0\le j\le k\right] &= \mathbb{P}\left[\max_{0\le j\le k} S_j \le a\,,\ \min_{0\le j\le k} S_j \ge -a\right]\\ &= 1 - \mathbb{P}\left[\Big\{\max_{0\le j\le k} S_j > a\Big\}\bigcup\Big\{\min_{0\le j\le k} S_j < -a\Big\}\right]\\ &\ge 1 - 2\,\mathbb{P}\left[\max_{0\le j\le k} S_j > a\right]\\ &\ge 1 - 4\,\mathbb{P}\left[S_k > a\right]\end{aligned}$$

where the first inequality is by the union bound and the symmetry of the random walk, and the second inequality is by the reflection principle $\mathbb{P}\left[\max_{0\le j\le k} S_j > a\right] \le 2\mathbb{P}\left[S_k > a\right]$, see, e.g., [6, Theorem 5.2.7].

Upper bounding $\mathbb{P}\left[S_k > a\right]$ when $a = A\sqrt{t\log t}$ and $k = ct$ requires a large deviation estimate. We use [18, Theorem 5.23] (and a standard tail estimate for the normal distribution) to obtain

$$\mathbb{P}\left[S_{ct} > A\sqrt{t\log t}\right] \le (1+o(1))\exp(-(\frac{A}{\sqrt{c}}\sqrt{\log t})^2/2\sigma^2) = \big(1+o(1)\big)t^{-\frac{A^2}{2c\sigma^2}}. \tag{10}$$ ∎

Given a vertex $\boldsymbol{x}$, the cycle $\mathsf{C}(\boldsymbol{x})$ containing $\boldsymbol{x}$ can be decomposed into a sequence of strands. The second strand begins at the endpoint of the first, i.e., at the endpoint of $\mathsf{St}(\boldsymbol{x})$. This inductively decomposes $\mathsf{C}(\boldsymbol{x})$ into $(\mathsf{St}(\boldsymbol{x}_i))_{i=0}^k$ for some $k \in \mathbb{N}$, where $\boldsymbol{x}_i$ is the initial vertex of the i^{th} strand. Each of these strands are distinct. For $j \in [\![0,k]\!]$, the strands $\mathsf{St}(\boldsymbol{x}_j)$ and $\mathsf{St}(\boldsymbol{x}_{j+1})$ will be called *consecutive*, with $k+1$ interpreted as 0. The relation of being consecutive puts a cyclic order on strands that are contained in the same cycle.

Lemma 3.3 implies that consecutive strands are well-separated if the torus $\mathbb{T}_{n,m}$ is appropriately asymmetric, as the vertical distance between two consecutive strands in a single column is equal to the vertical distance between the first and last points of a single strand. To make a precise statement, we extend the distance $d(A,B)$ between subsets $A, B \subset C_m$ to distances between subsets of A and B of $\mathbb{T}_{n,m}$ as follows. Let A_j denote the subset of A in column j, and similarly for B_j. Define

$$d^-(A,B) = \min_{j\in[\![n]\!]} d(A_j,B_j), \quad d^+(A,B) = \max_{j\in[\![n]\!]} d(A_j,B_j).$$

Proposition 3.4. *Fix $A > 0$, and suppose $n = m + C(m)$ with $C' > A$. Let $\boldsymbol{x}_2$ be the endpoint of the strand $\mathsf{St}(\boldsymbol{x}_1)$. Then*

$$\mathbb{P}\left[d^-(\mathsf{St}(\boldsymbol{x}_1),\mathsf{St}(\boldsymbol{x}_2)) \le \lfloor C(m) - A\sqrt{m\log m}\rfloor\right] \le \big(1+o(1)\big)m^{-\frac{A^2-2}{2}}, \tag{11}$$

$$\mathbb{P}\left[d^+(\mathsf{St}(\boldsymbol{x}_1),\mathsf{St}(\boldsymbol{x}_2)) \ge \lfloor C(m) + A\sqrt{m\log m}\rfloor\right] \le \big(1+o(1)\big)m^{-\frac{A^2-2}{2}}, \tag{12}$$

Proof. We start with (11). Let $\mathsf{St}(\boldsymbol{x}_i) = \big(\boldsymbol{x}_0^{(i)}, \boldsymbol{x}_1^{(i)}, \dots, \boldsymbol{x}_{n-1}^{(i)}\big)$ for $i = 1, 2$. Note that

$$d^-(\mathsf{St}(\boldsymbol{x}_1),\mathsf{St}(\boldsymbol{x}_2)) = \min_{i=0,1,\dots,n-1}\left|\pi_2(\boldsymbol{x}_i^{(1)}) - \pi_2(\boldsymbol{x}_i^{(2)})\right|,$$

and, (recall (8))

$$\left|\pi_2(\boldsymbol{x}_i^{(1)}) - \pi_2(\boldsymbol{x}_i^{(2)})\right| = \left|\tilde{\mathsf{St}}(\boldsymbol{x}_0^{(i)})_{n-1} + C(m)\right| \mod m.$$

The preceding display and the fact that $C(m) = \lceil C'\sqrt{m\log m}\rceil \le m/3$ implies that the bound

$$\left|\pi_2(\boldsymbol{x}_i^{(1)}) - \pi_2(\boldsymbol{x}_i^{(2)})\right| \;\le\; \lfloor C(m) - A\sqrt{m\log m}\rfloor$$

can hold only if the endpoint $\tilde{\mathsf{St}}(\boldsymbol{x}_0^{(i)})_{n-1}$ of the increment random walk $\tilde{\mathsf{St}}(\boldsymbol{x}_0^{(i)})$ started at $\boldsymbol{x}_0^i$ has absolute value at least $A\sqrt{m\log m}$. Lemma 3.2 shows the hypotheses of Lemma 3.3 are satisfied, and applying Lemma 3.3 with $c = 1$ and $t = n$ yields

$$\mathbb{P}\left[\left|\pi_2(\boldsymbol{x}_i^{(1)}) - \pi_2(\boldsymbol{x}_i^{(2)})\right| \ge A\sqrt{m\log m}\right] \;\le\; \big(1+o(1)\big)m^{-\frac{A^2}{2}}\,,$$

where we have used that $n = (1+o(1))m$ and that $\sigma^2 < 1$ for the strand random walk. Hence by a union bound over the m consecutive strands in the system,

$$\mathbb{P}\left[d^-(\mathsf{St}(\boldsymbol{x}_1),\mathsf{St}(\boldsymbol{x}_2)) \le \lfloor C(m) - A\sqrt{m\log m}\rfloor\right] \;\le\; \big(1+o(1)\big)m^{-\frac{A^2-2}{2}}.$$

The bound (12) follows, *mutatis mutandis*, from the same argument. ∎

The remainder of this section collects useful properties of strands that hold with high probability.

Definition 3.5. *Let $S_L(r)$ be the event that all pairs of consecutive strands $\mathsf{St}(\boldsymbol{x})$, $\mathsf{St}(\tilde{\boldsymbol{x}})$ satisfy $d^-(\mathsf{St}(\boldsymbol{x}),\mathsf{St}(\tilde{\boldsymbol{x}})) \ge r\sqrt{m\log m}$. Formally, if $\boldsymbol{x}^+$ denotes the endpoint of $\mathsf{St}(\boldsymbol{x})$,*

$$S_L(r) = \bigcap_{\boldsymbol{x}\in\pi_2^{-1}(0)} \left\{d^-(\mathsf{St}(\boldsymbol{x}),\mathsf{St}(\boldsymbol{x}^+)) \ge r\sqrt{m\log m}\right\}. \tag{13}$$

Similarly, let $S_U(r)$ denote the event that all pairs of consecutive strands satisfy $d^+(\mathsf{St}(\boldsymbol{x}),\,\mathsf{St}(\tilde{\boldsymbol{x}})) \le r\sqrt{m\log m}$. Let $S^0(D) = S_L(D)\cap S_U(4+\frac{3D}{2})$.

Corollary 3.6. *Let $D > 0$ and $C' = 2 + 5D/4$. Then $\mathbb{P}\left[S^0(D)\right] \ge 1 - o(m^{-D/2})$.*

Proof. Let $A = 2 + D/4$. The result follows from Proposition 3.4 and the union bound since there are exactly m consecutive strands in any configuration. ∎

Remark 4. *Henceforth we will assume $C' = C'(D)$ has been chosen as in Corollary 3.6.*

Corollary 3.6 implies that we can study Φ conditionally on $S^0(D)$ occurring for some $D > 0$, i.e., we can work on the event that consecutive strands are neither very close nor very far from one another.

Let $\boldsymbol{x'}$ be the endpoint of $\mathsf{St}(\boldsymbol{x})$. The next lemma says that if $\boldsymbol{y}$ is in the same column as $\boldsymbol{x}$, and $\boldsymbol{y}$ is not too close to either $\boldsymbol{x}$ or $\boldsymbol{x'}$, then $\mathsf{St}(\boldsymbol{y})$ will not get near $\mathsf{St}(\boldsymbol{x})$ or $\mathsf{St}(\boldsymbol{x'})$.

Lemma 3.7. *Let $\eta' = D/2$. For D large enough there is a $c > 0$ such that the following holds with probability $1 - m^{-cD^2}$ on $S^0(D)$. Suppose $\boldsymbol{x}$ and $\boldsymbol{y}$ are in the same column, and $d^-(\boldsymbol{x},\boldsymbol{y}) > \eta'\sqrt{m\log m}$. If $d^-(\boldsymbol{x'},\boldsymbol{y}) \ge \eta'\sqrt{m\log m}$, then $d^-(\mathsf{St}(\boldsymbol{x'}),\mathsf{St}(\boldsymbol{y}))$ and $d^-(\mathsf{St}(\boldsymbol{x}),\mathsf{St}(\boldsymbol{y}))$ are at least $\Theta(\sqrt{m\log m})$.*

Proof. The event $S^0(D)$ entails that $d^-(\mathsf{St}(\boldsymbol{x}),\mathsf{St}(\boldsymbol{x'})) \ge D\sqrt{m\log m}$ for all $\boldsymbol{x}$. By (9) with $A = \frac{D}{5}$ and a union bound, there exists $c > 0$ such that each strand $\tilde{\mathsf{St}}(\boldsymbol{x})$ is contained in a corridor of width $\frac{D}{5}\sqrt{m\log m}$ with probability $1 - m^{-cD^2}$ if D is chosen large enough. Since $\eta' = D/2$ the conclusion follows. ∎

Remark 5. *In the sequel we set $S^1(D)$ to be the event that $S^0(D)$ occurs, that Φ does not contain a global shift, and that the conclusion of Lemma 3.7 holds. For D large, $S^1(D)$ occurs with probability $1 - o(m^{-\frac{D}{2}})$.*

3.3. **Global traversals and first properties.** This section introduces and discusses global traversals, the basic mesoscopic unit of our arguments.

3.3.1. *Definition of global traversals.* We would like to group strands into consecutive sequences that 'vertically span' the system. The next definition is one way to achieve this.

Definition 3.8. *Let* $\Gamma = \Gamma(m) = \lfloor \frac{m}{C(m)} - 2m^{1/4} \rfloor$. *The* global traversal $\mathsf{Gt}(\boldsymbol{x})$ *starting at* $\boldsymbol{x}$ *is the sequence of* Γ *consecutive strands starting with* $\mathsf{St}(\boldsymbol{x})$*, provided these strands are all distinct. If they are not distinct we say there is no global traversal at* $\boldsymbol{x}$*.*

The choice of Γ reflects that the expected vertical displacement along each strand is $C(m)$. This implies the expected vertical displacement after $\Gamma + 2m^{1/4}$ strands is zero. Roughly speaking, then, this is the number of strands after which a cycle has an appreciable chance to close. The omission of $2m^{1/4}$ strands in the definition of a global traversals is for technical convenience: it ensures that first and last strands of a global traversal do not interact with one another.

Given $\boldsymbol{x}$, recall that $\mathsf{C}(\boldsymbol{x})$ denotes the cycle of Φ that contains $\boldsymbol{x}$. Let $K(\boldsymbol{x})$ denote the integer number of strands contained in $\mathsf{C}(\boldsymbol{x})$, and set

$$\{\mathsf{C}(\boldsymbol{x})\}_{\mathsf{Gt}} = \frac{K(\boldsymbol{x}) \mod \Gamma}{\Gamma}, \quad \lfloor \mathsf{C}(\boldsymbol{x}) \rfloor_{\mathsf{Gt}} = \frac{K(\boldsymbol{x}) - \{\mathsf{C}(\boldsymbol{x})\}_{\mathsf{Gt}}}{\Gamma}. \tag{14}$$

These are the number of fractional and complete global traversals in $\mathsf{C}(\boldsymbol{x})$, respectively. To make this more explicit, note $K(\boldsymbol{x})$ is the number of vertices that $\mathsf{C}(\boldsymbol{x})$ contains in the column containing $\boldsymbol{x}$. Let $(\boldsymbol{x}_i)_{i=0}^{K(\boldsymbol{x})-1}$ denote the sequence of these vertices, starting from $\boldsymbol{x}$. Partition $\mathsf{C}(\boldsymbol{x})$ into a sequence of distinct global traversals $\mathsf{Gt}(\boldsymbol{x}_{j\Gamma})$ for $j = 0, \dots, \lfloor \mathsf{C}(\boldsymbol{x}) \rfloor_{\mathsf{Gt}}$ together with a sequence of $\{\mathsf{C}(\boldsymbol{x})\}_{\mathsf{Gt}}$ consecutive strands starting from $\boldsymbol{x}_{\lfloor \mathsf{C}(\boldsymbol{x}) \rfloor_{\mathsf{Gt}}\Gamma}$ and ending at $\boldsymbol{x}$. This decomposition of $\mathsf{C}(\boldsymbol{x})$ may depend on $\boldsymbol{x}$, but this ambiguity will cause no trouble: the key point is that $\{\mathsf{C}(\boldsymbol{x})\}_{\mathsf{Gt}}$ and $\lfloor \mathsf{C}(\boldsymbol{x}) \rfloor_{\mathsf{Gt}}$ are independent of $\boldsymbol{x}$.

3.3.2. *Existence and measurement of global traversals.* We will shortly show that each cycle contains at least one global traversal. This is a key step in reducing the study of cycles to the study of global traversals. To establish this we consider first consider long sequences of strands, which allows for a reduction to a statement about random walks. More precisely, set

$$Y_{nk+j}(\boldsymbol{x}) = \sum_{\ell=0}^{k-1} \tilde{\mathsf{St}}(\boldsymbol{x}_\ell)_{n-1} + \tilde{\mathsf{St}}(\boldsymbol{x}_k)_j, \qquad k \in [\![\frac{\Gamma}{100D}]\!], \quad j \in [\![n]\!]. \tag{15}$$

That is, Y is the random walk obtained by following the increments along $\Gamma/100D$ consecutive strands. Note that Y does not have independent increments due to the dependence between strands. Let $\tilde{Y}$ be the random walk with IID increments obtained by concatenating $\Gamma/100D$ independent copies of $\tilde{\mathsf{St}}(\boldsymbol{x})$.

Lemma 3.9. *Let* $\mathcal{E}$ *be the event that* $Y = \tilde{Y}$ *under the optimal coupling of* Y *and* $\tilde{Y}$*. There is an* $r > 0$ *such that* $\mathbb{P}\left[\mathcal{E} \mid S^1(D)\right] \geq 1 - e^{-r\sqrt{m \log m}}$.

Proof. We prove we can couple Y and $\tilde{Y}$ with probability $1 - Re^{-r\sqrt{m \log m}}$ for some $r, R > 0$ on $S^1(D)$. On $S^1(D)$ the maximal upward displacement of Y is $\frac{\Gamma}{100D}(4 + \frac{3D}{2})\sqrt{m \log m} \leq \frac{m}{2}$, so the last strand of Y is at least distance $D\sqrt{m \log m}$ from the first strand in every column. Thus the total variation distance between an increment of Y and an increment of $\tilde{Y}$ on $S^1(D)$ is at most $\exp(-\Theta(\sqrt{m \log m})$ by Lemma 2.8. A union bound over the $\frac{\Gamma}{100D} n \leq 2m^{3/2}$ increments gives the result. ∎

Lemma 3.10. *There are* $R, r > 0$ *such that*

$$\mathbb{P}\left[\textit{exists } \boldsymbol{x} \textit{ such that } \lfloor \mathsf{C}(\boldsymbol{x}) \rfloor_{\mathsf{Gt}} = 0 | S^1(D)\right] \leq Re^{-r(\log m)^{3/2}}. \tag{16}$$

Moreover, if $\boldsymbol{x}'$ *denotes the first vertex of the last strand of* $\mathsf{Gt}(\boldsymbol{x})$*, then we further have that*

$$m^{3/4}(\log m)^{1/2}(1 + o(1)) \leq d(\mathsf{St}(\boldsymbol{x}), \mathsf{St}(\boldsymbol{x}')) \leq 3m^{3/4}(\log m)^{1/2} \tag{17}$$

with probability $1 - Re^{-r(\log m)^{3/2}}$.

Proof. For convenience set $\boldsymbol{x} = (0,0)$. For the first statement, we must show that the cycle beginning at $\boldsymbol{x}$ will not be explored in fewer than Γ strands. We prove this by showing that $0 \le |\pi_2(\boldsymbol{x}_j)| \le m - m^{1/4}C(m)$ for all $j \in [\![0, \Gamma]\!]$ with the requisite probability, $\boldsymbol{x}_j$ representing the initial vertices of strands. On $S^1(D)$ consecutive strands are at vertical distance of order $\sqrt{m \log m}$, so this will suffice.

Decompose $\mathsf{C}(\boldsymbol{x})$ into sequences of $\Gamma/100D$ consecutive strands, and consider the first $100D$ such sequences. Write W for such a sequence of $\Gamma/100D$ strands. We will show that the collective fluctuations of the W cannot overcome the vertical gap created by the $-2m^{1/4}$ in the definition of Γ.

To do this, treat each W as a deterministic part plus Y, the random walk formed by the increments of W. Each strand has a deterministic vertical increase of $n = m + C(m)$, and hence a vertical displacement of $C(m)$. By displacement we mean $|\pi_2(\boldsymbol{x}_{j+1}) - \pi_2(\boldsymbol{x}_j)|$. The deterministic vertical displacement from the first 100D sequences is therefore

$$100D\frac{\Gamma C(m)}{100D} = \Gamma C(m) = m - 2m^{1/4}C(m)$$

A net fluctuation of size $2m^{1/4}C(m)$ is required for the first and last strands to meet. We will rule out a fluctuation of size $m^{1/4}C(m)$.

By Lemma 3.9 the fluctuation due to each sequence of strands can be treated as $\tilde{Y}$, a simple random walk of the same length. This replacement results in an error with probability at most $\exp(-r\sqrt{m \log m})$. Hence by a union bound, the net fluctuation in the first 100D sequences can be treated as that of a concatenation of 100D copies of $\tilde{Y}$, i.e., a simple random walk of length $\Gamma n = (1 + o(1))\Gamma m$. Let

$$x = \frac{m^{1/4}C(m)}{(\Gamma m)^{1/2}} \ge \frac{m^{1/4}(C(m))^{3/2}}{m} \ge (\log m)^{3/4}.$$

By [18, Theorem 5.23] (recall (10)), the probability of a fluctuation of size x is of order $\exp(-\Theta(x^2)) = \exp(-\Theta((\log m)^{3/2}))$. This proves the first claim. For the second, note that when a fluctuation of size at most x occurs the remaining gap is of size at least $(1 + o(1))m^{1/4}C(m)$ and at most $3m^{1/4}C(m)$. ∎

The next lemma says that $\lfloor \mathsf{C} \rfloor_{\mathsf{Gt}}$ is a good proxy for the size of a cycle.

Lemma 3.11. *Write $\cup_{\mathsf{C}}$ for a union over all cycles. There is an $r > 0$ such that*

$$\mathbb{P}\left[\bigcup_{\mathsf{C}} \left\{ \frac{\{\mathsf{C}\}_{\mathsf{Gt}}}{\lfloor \mathsf{C} \rfloor_{\mathsf{Gt}}} \ge m^{-1/8} \right\} \middle| S^1(D) \right] = \exp(-r\sqrt{m \log m}). \tag{18}$$

Proof. We work on the event that (17) holds for all global traversals. By a union bound over all $\boldsymbol{x}$ and Lemma 3.10, there is an $r > 0$ such that this event has probability at least $1 - \exp(-r\sqrt{m \log m})$.

The proof of (18) proceeds by a case analysis. By a union bound it is enough to consider a single cycle, as there are at most m cycles. Suppose a cycle C has exactly $k = \lfloor \mathsf{C} \rfloor_{\mathsf{Gt}} \le m^{1/8}$ completed global traversals. Concatenate these completed global traversals. Equation (17) implies the distance between the endpoint and initial point of the concatenation is of order $km^{3/4}(\log m)^{1/2}$. The number of strands that connect these points is of order $km^{3/4}(\log m)^{1/2}/C(m) = \Theta(km^{1/4})$, as each strand has a deterministic contribution of $C(m)$ plus a random fluctuation of at most $(4 + \frac{3D}{2})\sqrt{m \log m}$. Thus the fraction of strands in the fractional part is of order

$$\{\mathsf{C}\}_{\mathsf{Gt}}/\lfloor \mathsf{C} \rfloor_{\mathsf{Gt}} = \Theta(\frac{km^{1/4}}{k\Gamma}) = \Theta(\frac{m^{1/4}(\log m)^{1/2}}{m^{1/2}}) \le \frac{1}{m^{1/8}}.$$

On the other hand, if C has exactly $k = \lfloor \mathsf{C} \rfloor_{\mathsf{Gt}} > m^{1/8}$ completed global traversals, the claim follows since a fractional traversal consists of less than one global traversal. ∎

Let $|\mathsf{C}|_{\mathsf{Gt}} = \lfloor \mathsf{C} \rfloor_{\mathsf{Gt}} + \{\mathsf{C}\}_{\mathsf{Gt}}$. Lemma 3.11 shows that

$$m = \Gamma \sum_{\mathsf{C}} |\mathsf{C}|_{\mathsf{Gt}} = \Gamma(1 + o(m^{-1/8})) \sum_{\mathsf{C}} \lfloor \mathsf{C} \rfloor_{\mathsf{Gt}} \tag{19}$$

with high probability on $S^1(D)$. This says that the number of global traversals in cycles is within $o(1)$ of the maximal number m/Γ of possible global traversals in the system.

Remark 6. *Henceforth we define $S(D)$ to be the sub-event of $S^1(D)$ on which the events in Lemmas 3.10 and 3.11 occur. Note that $\mathbb{P}[S(D)] = 1 - o(m^{-D/2})$.*

4. Equilibrium properties II: Concentration of Contacts

This section studies pairs of global traversals, and establishes that the number of times they come into contact is a concentrated random variable. We make the notion of a contact precise in Section 4.1 and formalize concentration in Theorem 4.3. Subsequent subsections prove Theorem 4.3.

4.1. Definition of contacts, global concentration.

Recall the notions of swapping and parallel arrows from Section 2.1. Given $v \in V(\mathbb{T}^\star_{n,m})$, $v = \{(x_1, x_2), (x_1, x_2 + 1)\}$ (so v is a vertical edge of $\mathbb{T}_{n,m}$), write $v \pm \frac{1}{2}$ for (x_1, x_2) and $(x_1, x_2 + 1)$.

Definition 4.1. *Let $v \in V(\mathbb{T}^\star_{n,m})$ be a vertical edge. There is a* contact *at v* between $v + \frac{1}{2}$ *and $v - \frac{1}{2}$ if the arrows $\phi_{v\pm\frac{1}{2}}$ swap or are parallel.*

Given $A, B \subset V(\mathbb{T}_{n,m})$ and a realization of Φ, the *number of contacts between A and B* is the number of pairs $(a, b) \in A \times B$ such that there is a contact between a and b. Given a set $A \subset V(\mathbb{T}_{n,m})$, the *number of contacts contained in A* is the number of contacts between A and $V(\mathbb{T}_{n,m})$.

Lemma 4.2. *Let X be the number of contacts contained in a strand. There exist $c, \alpha > 0$ such that*

$$\mathbb{P}[|X - \alpha n| \geq \gamma \alpha n \mid S(D)] \leq e^{-c\gamma^2 n}, \qquad \gamma > 0.$$

Proof. Recall that on $S(D)$, we can use the underlying hard-core models on columns of $\mathbb{T}^\star_{n,m}$ as the source of randomness. A single strand is determined by n IID samples from the equilibrium measure of the hard-core model. Each sample results in i contacts with probability p_i for $i = 0, 1, 2$, $\sum_i p_i = 1$, $\sum_i i p_i = \alpha$, and $\alpha > 0$ since $\lambda > 0$. Hoeffding's inequality then implies the claim. ∎

The previous lemma indicated that contacts are plentiful. The next theorem indicates that they are rather uniformly distributed between pairs of disjoint global traversals. The proof of the theorem occupies Sections 4.2 and 4.3. The details of the proof will not be needed for subsequent developments.

Theorem 4.3. *Let $\boldsymbol{x} \neq \boldsymbol{y} \in \mathbb{T}_{n,m}$. Let X denote the number of contacts between $\mathsf{Gt}(\boldsymbol{x})$ and $\mathsf{Gt}(\boldsymbol{y})$ conditional on $\mathsf{Gt}(\boldsymbol{x}) \cap \mathsf{Gt}(\boldsymbol{y}) = \emptyset$. Then*

$$\mathbb{P}\left[|X - \mathbb{E}X| \geq m^{1/3} \middle| S(D)\right] \leq \exp\left(-\Theta(m^{1/7})\right), \quad \mathbb{E}X = \Theta\left(\frac{m}{\log m}\right). \tag{20}$$

4.2. The ideal gap chain.

The *ideal gap chain* $Z \colon \mathbb{N} \to \mathbb{N}_+$ is a time-homogeneous Markov chain that offers an idealized description of the vertical distance between two strands. In defining Z we make use of the hard-core model on $\mathbb{Z} + 1/2$ as introduced in Section 2.3. The shift by $1/2$ is made with Figure 2 and Remark 1 in mind: each hard-core configuration defines a corresponding arrow configuration ψ on $\mathbb{Z}$. The trajectories determined by a sequence of these arrow configurations will be our idealization of strands, and Z the distance between two such strands. We now make this precise.

Definition 4.4. *Let σ denote the hard-core model on $\mathbb{Z} + 1/2$ with activity λ given by (3). Let $p = \mathbb{P}[\sigma_{1/2} = 1] \in (0, 1/2)$ denote the probability of any particular vertex being occupied. Let $q_i(a) = \mathbb{P}[\sigma_{i+1/2} = 1 | \sigma_a = 1]$ for $a \in \{-1/2, 1/2\}$, and $q_i(0) = \mathbb{P}[\sigma_{i+1/2} = 1 | \sigma_{-1/2} = \sigma_{1/2} = 0]$. The* ideal gap chain $Z \colon \mathbb{N} \to \mathbb{N}_+$ *is the time-homogeneous Markov chain with transition probabilities ($i \geq 2$)*

$$\mathbb{P}\left[Z_1 - Z_0 = j \middle| Z_0 = i\right] = \begin{cases} p q_i(-1/2) & j = 2 \\ (1-2p) q_i(0) + p\big(1 - q_i(-1/2) - q_{i-1}(-1/2)\big) & j = 1 \\ (1-2p)\big(1 - q_i(0) - q_{i-1}(0)\big) + p\big(q_i(1/2) + q_{i-1}(-1/2)\big) & j = 0 \\ (1-2p) q_{i-1}(0) + p\big(1 - q_i(1/2) - q_{i-1}(1/2)\big) & j = -1 \\ p q_{i-1}(1/2) & j = -2 \end{cases}.$$

The last possibility does not occur if $i = 2$ since $q_1(1/2) = 0$. If $i = 1$ the transition probabilities are

$$\mathbb{P}\big[Z_1 - Z_0 = j \big| Z_0 = 1\big] = \begin{cases} pq_1(-1/2) & j = 2 \\ (1-2p)q_1(0) + p\big(1 - q_1(-1/2) - q_0(-1/2)\big) & j = 1 \\ (1-2p)\big(1 - q_1(0)\big) + p & j = 0 \end{cases}.$$

Let $\psi = (\psi_i)_{i\in\mathbb{Z}}$ denote the arrow configuration on $\mathbb{Z}$ associated to the hard-core model on $\mathbb{Z} + \frac{1}{2}$, and let $\Psi = (\psi^j)_{j\in\mathbb{N}}$ denote a sequence of independent copies of ψ. Thus $\psi_i^j \in \{-1, 0, 1\}$ is the arrow at i in the j^{th} copy of ψ, and for $(j, i) \in \mathbb{N} \times \mathbb{Z}$, $\Psi((j, i)) = \psi_i^j$. Set $\Psi^0((0, i)) = i$, and for $j \geq 0$ set $\Psi^{j+1}((0, i)) = \Psi^j((0, i)) + \psi^j_{\Psi^j((0,i))}$. Thus $(\Psi^j(0, i))_{j\in\mathbb{N}}$ is the analogue of a strand. More precisely, this is the analogue of $\tilde{\mathsf{St}}(\boldsymbol{x})$ (recall (8)), but the omission of the deterministic $+1$ will be irrelevant since we will be considering the distance between two strands.

Lemma 4.5. *Let $i \in \mathbb{N}_+$. Let $X\colon \mathbb{N} \to \mathbb{N}_+$ be given by $X_j = \big|\Psi^j((0, i)) - \Psi^j((0, 0))\big|$. Then X is equal in law to the ideal gap chain Z with $Z_0 = i$.*

Proof. We consider $i \geq 2$; the case $i = 1$ is similar and we omit the details. Since there is an independent hard core model σ for each $j \in \mathbb{N}$, it is enough to verify that $X_1 - X_0$ is distributed as $Z_1 - Z_0$.

Flip a three-sided coin with outcomes of probability p, p and $1 - 2p$ called heads, tails and blank. If it is heads, condition σ to contain $1/2$. Let $q_j(1/2)$ denote the conditional probability that $j + 1/2$ is in σ. Then the conditional distribution of $Z_1 - Z_0$ equals 0 with probability $q_i(1/2)$; -1 with probability $1 - q_i(1/2) - q_{i-1}(1/2)$; and -2 with probability $q_{i-1}(1/2)$.

If the outcome is tails, condition σ to contain $-1/2$. Let $q_j(-1/2)$ denote the conditional probability that $j + 1/2$ is in σ. Then the conditional distribution of $Z_1 - Z_0$ equals 2 with probability $q_i(-1/2)$; 1 with probability $1 - q_i(-1/2) - q_{i-1}(-1/2)$; and 0 with probability $q_{i-1}(-1/2)$.

If the outcome is blank, condition σ to contain neither $1/2$ nor $-1/2$. Let $q_j(0)$ denote the conditional probability that $j + 1/2$ is in σ. Then the conditional distribution of $Z_1 - Z_0$ equals 1 with probability $q_i(0)$; 0 with probability $1 - q_i(0) - q_{i-1}(0)$; and -1 with probability $q_{i-1}(0)$. ∎

Lemma 4.6. *There are independent events $(S_k)_{k\in\mathbb{N}}$ such that the ideal gap chain satisfies*

$$Z_{k+1} - Z_k = (X_k + X'_k)\mathbf{1}_{S_k} + E_k\mathbf{1}_{S_k^c}\,, \tag{21}$$

where

1. *X_k, X'_k, $k \in \mathbb{N}$, are independent random variables that share a symmetric and non-degenerate distribution μ taking values in $\{-1, 0, 1\}$;*
2. *E_k, $k \in \mathbb{N}$, are independent random variables taking values in $[\![-2, 2]\!]$; and*
3. *there are constants $c, C > 0$ such that the events S_j satisfy, for $k, k', j \in \mathbb{N}$,*

$$\mathbb{P}\big[S_k^c \big| Z_k = j\big] = \mathbb{P}\big[S_{k'}^c \big| Z_{k'} = j\big] \leq Ce^{-cj} \text{ for } j \in \mathbb{N}\,.$$

Proof. Lemma 2.7 implies that there exist positive constants C and c such that $|q_i(a) - p| \leq Ce^{-ci}$ for $a \in \{-1/2, 0, 1/2\}$. Items (1)–(3) of Lemma 4.6 then follow readily from Lemma 4.5. ∎

For $j \in \mathbb{N}_+$ let $\tau_j = \min\{k \in \mathbb{N} \mid Z_k = j\}$ denote the hitting time of j by Z, and $\tau_j^+ = \min\{k \in \mathbb{N} \mid Z_k \geq j\}$. The next proposition says that the ideal gap chain behaves like a simple random walk.

Proposition 4.7. *For $j_1, j_2 \in \mathbb{N}_+$ with $j_2 \geq j_1 > 1$,*

$$\mathbb{P}\big[\tau_{j_2}^+ < \tau_1 \,\big|\, Z_0 = j_1\big] = \Theta\big(j_1 j_2^{-1}\big). \tag{22}$$

Proof. We begin by deriving the upper bound, i.e., showing there exists a constant $C > 0$ for which

$$\mathbb{P}\big[\tau_{j_2}^+ < \tau_1 \,\big|\, Z_0 = j_1\big] \;\leq\; Cj_1 j_2^{-1}\,. \tag{23}$$

We will do this in three steps. First, we will construct a process $(Z'_k)_{k\geq 0}$ that stochastically dominates $(Z_k)_{k\geq 0}$. Informally, Z' is the process that results from forcing $E_k = 2$ in (21). Secondly, we will bound Z' above by a Markov chain that can be analyzed by standard methods. Third, we will use this Markov chain to push Z' down, which will imply that Z visits 1.

For the first step, define

$$Z'_{k+1} - Z'_k = (X_k + X'_k)\mathbf{1}_{S'_k} + 2 \cdot \mathbf{1}_{(S'_k)^c},$$

where the S'_k are independent events, with $\mathbb{P}\left[(S'_k)^c \mid Z'_k = j\right] = \max_{i \in [\![0,3]\!]} \mathbb{P}\left[S_k^c \mid Z_k = j - i\right]$; in this definition we set $\mathbb{P}\left[S_k^c \mid Z_k = i\right] = 0$ for $i \leq 0$. Set $Z'_0 = Z_0$, and suppose that $Z'_k \geq Z_k$. If $Z'_k - Z_k \geq 4$, then $Z'_{k+1} \geq Z_{k+1}$ since each process jumps by at most two. On the other hand, if $Z'_k - Z_k \in [\![0,3]\!]$, then we may couple Z'_{k+1} with Z_{k+1} such that $S_k^c \subset (S'_k)^c$, and hence $Z'_{k+1} \geq Z_{k+1}$.

For the second step, note that $(Z'_k)_{k \geq 0}$ with $Z'_0 = j_1$ is equal in law to the following auxiliary process $(Y_k)_{k \in \mathbb{N}}$ with $Y_0 = j_1$, observed only at even times $k \in 2\mathbb{N}$. Let $S(k,j)$ be events such that $\mathbb{P}\left[S(k,j)\right] = \mathbb{P}\left[S_k | Z_k = j\right]$, independent for different k. Suppose we are given the trajectory $(Y_i)_{i=0}^{2k}$. If $S(2k, Y_{2k})$ occurs, set $Y_{2k+1} = Y_{2k} + X$ and $Y_{2k+2} = Y_{2k+1} + X'$, where X and X' are independent random variables that share the symmetric law μ on $\{-1, 0, 1\}$ from Lemma 4.6. If $S(2k, Y_{2k})^c$ occurs, set $Y_{2k+1} = Y_{2k} + 1$ and $Y_{2k+2} = Y_{2k+1} + 1$.

The process Y is not Markov, but we may bound it above stochastically by a Markov process $Y^+ \colon \mathbb{N} \to \mathbb{N}_+$, $Y_0^+ = j_1$. We first explain this informally. The process Y flips a coin that lands heads with probability $\mathbb{P}\left[S(2k, Y_{2k})\right]$ at even times, takes two steps X, X' if the coin lands heads, and takes two steps $+1$ otherwise. Y^+ will instead flip a coin at each time, taking an independent step distributed as X for heads, and $+1$ for tails. To obtain the desired stochastic domination it suffices to show that we can arrange that if Y gets a tail at time $2k$, then Y^+ gets a tail at times $2k$ and $2k+1$, as if Y gets a head the increments of Y are dominated by those of Y^+.

To achieve this, define $p_k^+ = \max_{j \in [\![0,2]\!]} \sqrt{\mathbb{P}\left[S(k, Y_k^+ - j)^c\right]}$. In this formula, $\mathbb{P}\left[S(k,i)^c\right] = 0$ if $i \leq 0$ by convention. We make the Y^+ coins land heads with probabilities $1 - p_k^+$, and tails with the complementary probability p_k^+. With this definition Y^+ is Markov as its transition probabilities depend only on Y_k^+, and we can couple the coins as desired.

We proceed to the third step. Note that Lemma 4.6 implies

$$\mathbb{P}\left[Y_{k+1}^+ - Y_k^+ = j \middle| Y_k^+ = \ell\right] = \begin{cases} \mu(1) + \nu_1(\ell) & \text{for } j = 1 \\ \mu(0) + \nu_2(\ell) & \text{for } j = 0 \\ \mu(1) + \nu_3(\ell) & \text{for } j = -1 \end{cases},$$

where $|\nu_i(\ell)| \leq C \exp\{-c\ell\}$ for $i \in [\![1,3]\!]$. Choose a such that $|\nu_i(a')| < 1$ for all $a' \geq a$ and $i \in [\![1,3]\!]$. When $Y^+ \geq a$, Y^+ has the transitions of a birth-and-death chain (see, e.g. [17, Example 1.3.4]). A standard computation, which we give below in Lemma 4.8, shows the hitting times of Y^+ satisfy

$$\mathbb{P}^{Y^+}\left[\tau_{j_2}^+ < \tau_a \middle| Y_0^+ = j_1\right] \leq C j_1 j_2^{-1} \tag{24}$$

for some $C > 0$. We now deduce (23) from (24). Since $Y \leq Y^+$ almost surely, $Y_0 = Y_0^+ = j_1$, and Y^+ takes jumps of size at most one, the equality in law of Y and Z' at even times implies (24) also holds for the process Z' if τ_a is replaced by the hitting time of $\{a, a+1\}$. Since $Z \leq Z'$ almost surely, this implies the same statement for Z. Finally, the ideal gap chain has a positive probability of hitting 1 before j_2 if started at a or $a+1$, as can be seen by constructing an explicit trajectory. This completes the proof of (23).

To complete the proof of Proposition 4.7 we must prove the complementary lower bound, that there exists $c > 0$ such that $\mathbb{P}\left[\tau_{j_2}^+ < \tau_1 \,\middle|\, Z(0) = j_1\right] \geq c j_1 j_2^{-1}$. The argument is very similar to the one above, and so we only outline the steps. First, we repeat the above construction, but define Z' by setting $E_n = -2$ and taking the maximum of probabilities when $Z_k = Z'_k + i$ for $i \in [\![0,3]\!]$ (if such a jump would result in $Z'_k < 1$, we set $Z'_k = 1$). Second, introduce $Y^- \colon \mathbb{N} \to \mathbb{N}_+$ such that $Y \geq Y^-$ by following the specification of Y^+, but now replacing the $+1$ steps on tails with -1 steps, and again modifying the maximum. Third, the construction of an explicit trajectory to have Z hit 1 is replaced by the construction of an explicit trajectory to get Z to a state a above which the process Y^- is a birth-and-death process. The analogue of Lemma 4.8 for Y^- then yields the sought bound. ∎

Lemma 4.8. *In the context of the proof of Proposition 4.7, for $a < j_1 < j_2$ we have that*

$$\mathbb{P}^{Y^+}\left[\tau^+_{j_2} < \tau_a \middle| Y^+_0 = j_1\right] = \Theta(j_1 j_2^{-1})\,.$$

Proof. Let $\{\kappa_i : i \in \mathbb{N}\}$ be the increasing sequence given by $\kappa_{a-1} = 0$, $\kappa_a = 1$, and

$$\frac{\kappa_{\ell+1} - \kappa_\ell}{\kappa_\ell - \kappa_{\ell-1}} = \frac{\mu(1) + \nu_1(\ell)}{\mu(1) + \nu_3(\ell)} \text{ for } \ell \geq a\,. \tag{25}$$

Then (as $\tau^+_{j_2} = \tau_{j_2}$ and Y^+ is a standard birth-death chain, see, e.g. [17, Example 1.3.4])

$$\mathbb{P}^{Y^+}\left[\tau^+_{j_2} < \tau_a \middle| Y^+_0 = j_1\right] = \frac{\kappa_{j_1}}{\kappa_{j_2}}\,. \tag{26}$$

Note that

$$\frac{\kappa_{\ell+1} - \kappa_\ell}{\kappa_\ell - \kappa_{\ell-1}} = \big(1 + \nu_1(\ell)\mu(1)^{-1}\big)\big(1 + \nu_3(\ell)\mu(1)^{-1}\big)^{-1} = 1 + \Theta(e^{-c\ell})\,.$$

Hence, there exist constants $K \in (0,\infty)$ and $\chi \in (0,\infty)$ such that

$$\kappa_\ell = \ell\chi + K + \Theta(e^{-c\ell})\,, \tag{27}$$

and inserting this into (26) yields the lemma. ∎

Let $a \in \mathbb{N}$ be the state introduced above (24), i.e., the state such that $Y^\pm$ is a birth-and-death chain when at or above state a. The next lemma provides couplings of the bounding Markov chains $Y^\pm$ with Brownian motions when $Y^\pm \geq a$; these couplings will be useful for analysing hitting times of Z in the next section. More precisely, to circumvent the restriction $Y^\pm \geq a$, let $\mathcal{Y}^\pm$ denote $Y^\pm$ with jumps below level a suppressed (i.e., nothing happens if such a jump is attempted), and similarly, jumps of size 0 suppressed. The couplings require some notation. Write $\{\kappa^\pm_i\}_{i\geq 1}$ for the $\kappa^\pm_i$ given by (25) with $\kappa^\pm_{a-1} = 0$, $\kappa^\pm_a = 1$. The distinction between $\pm$ arises from the meaning of μ and ν in (25). Set

$$\bar\kappa^\pm_{a-i} = 2\kappa^\pm_a - \kappa^\pm_{a+i}, \quad i \geq 2, \qquad \text{and} \quad \mathcal{K}^\pm = \{\kappa^\pm_i\}_{i\geq a} \bigcup \{\bar\kappa^\pm_{a-i}\}_{i\geq 1}. \tag{28}$$

Lemma 4.9. *There is a coupling of $\mathcal{Y}^+$ with a Brownian motion $B\colon [0,\infty) \to \mathbb{R}$ such that*

1. *There is an increasing sequence of stopping times ρ^+_i with $\rho^+_0 = 0$ such that $B(\rho^+_i) \in \mathcal{K}^+$.*
2. *For each i, $\arg B(\rho^+_i) = \mathcal{Y}^+_i$, where $\arg x = a + i$ if $x = \kappa^+_{a+i} \in \mathcal{K}^+$ or $x = \bar\kappa^+_{a-i}$.*
3. *There are $\beta, c > 0$ such that if $b > 0$ and $E^+(i,b)$ is the event that $\rho^+_i - \beta i > bi$, then*

$$\mathbb{P}\left[E^+(i,b)\right] \leq e^{-cbi}. \tag{29}$$

An analogous coupling of $\mathcal{Y}^-$ with a Brownian motion exists, with $\mathcal{K}^+$ replaced with $\mathcal{K}^-$.

Proof. Suppose $\mathcal{Y}^+_0 = j \in \mathbb{N}_+$, $j \geq a$, and set $B(0) = \kappa_j$. We will construct the couplings for $\mathcal{Y}^+$ and $\mathcal{Y}^-$ in parallel (i.e., one should consistently choose the sign $+$ or the sign $-$). We first give the construction up to the first time τ that $\mathcal{Y}^\pm$ hits 1, and afterward we extend the construction to all times. To this end, iteratively define an increasing sequence of times $\{\rho^\pm_i : i \in \mathbb{N}\}$ such that $\mathbb{P}\left[\mathcal{Y}^\pm_{i+1} = k+1 \mid \mathcal{Y}^\pm_i = k\right]$ is equal to $\mathbb{P}\left[B(\rho^\pm_{i+1}) = \kappa^\pm_{k+1} \mid B(\rho^\pm_i) = \kappa^\pm_k\right]$. Recall $\rho^\pm_0 = 0$, and suppose an initial sequence $\rho^\pm_j$, $j \in [\![0, i-1]\!]$ has been specified such that for $i \in \mathbb{N}_+$, $B(\rho^\pm_j) \in \mathcal{K}^\pm$ for $j \in [\![0, i-1]\!]$. We now define $\rho^\pm_i$ to be the infimum of $t \geq \rho^\pm_{i-1}$ such that $B(t) \in \mathcal{K}^\pm$ with $B(t) \neq B\big(\rho^\pm_{i-1}\big)$. With these definitions, since the $\kappa^\pm_i$ form an increasing sequence,

$$\frac{\mathbb{P}\left[B(\rho^\pm_{i+1}) = \kappa_{j+1} | B(\rho^\pm_i) = \kappa^\pm_j\right]}{\mathbb{P}\left[B(\rho^\pm_{i+1}) = \kappa^\pm_{j-1} | B(\rho^\pm_i) = \kappa^\pm_j\right]} = \frac{\mathbb{P}\left[\mathcal{Y}^\pm_{i+1} = j+1 | \mathcal{Y}^\pm_i = j\right]}{\mathbb{P}\left[\mathcal{Y}^\pm_{i+1} = j-1 | \mathcal{Y}^\pm_i = j\right]} = \frac{\kappa^\pm_{j+1} - \kappa^\pm_j}{\kappa^\pm_j - \kappa^\pm_{j-1}}. \tag{30}$$

This gives our coupling up until the hitting time τ of a by $\mathcal{Y}^\pm$: $(\mathcal{Y}^\pm_i)_{i\in\mathbb{N}}$ is equal in law to $(\arg B(\rho^\pm_i))_{i\in\mathbb{N}}$.

To extend the coupling to all times, note that $B(\rho^\pm_{\tau+1})$ is equally likely to be $\bar\kappa^\pm_{a-1}$ and κ_{a+1}. The distances between the $\bar\kappa^\pm_{a-i}$ are exactly as for the κ^+_{a+i}, and hence the coupling can be continued by

making some small changes in indexing (the Brownian motion going to the left corresponds to $\mathcal{Y}^{\pm}$ going right when at the $\bar{\kappa}$). This establishes (1) and (2).

To establish (3), note that the law of $\rho_{i+1}^{\pm} - \rho_i^{\pm}$ is a mixture of the laws of hitting times of Brownian motion begun at one point in $\mathcal{K}^{\pm}$ on the set of adjacent points in $\mathcal{K}^{\pm}$. By (27) and (28) the distances between adjacent values of $\kappa^{\pm}$ lie in a compact subset of $(0,\infty)$. Hence by using the reflection principle to compute $\mathbb{P}[H < x]$, each such hitting time H satisfies $\mathbb{P}[H > x] \leq Ce^{-cx}$ for suitable $c, C > 0$, and these constants are independent of the value of $i \in \mathbb{N}$. Hence $\mathbb{P}\left[\rho_{i+1}^{\pm} - \rho_i^{\pm} > x\right] \leq Ce^{-cx}$. This implies that there is a $\beta \in (0,\infty)$ such that

$$\mathbb{P}\left[\rho_i^{\pm} - \beta i > bi\right] \leq C \exp\left\{-cbi\right\} \tag{31}$$

for $i \in \mathbb{N}_+$ and $b > 0$; the values of $c, C > 0$ may have changed. The deduction of this bound is by writing $\rho_{i+1}^{\pm} = \sum_{j=0}^{i} (\rho_{j+1}^{\pm} - \rho_j^{\pm})$, dominating the increments by a sequence of IID sub-exponential random variables, and applying Bernstein's inequality for sub-exponential random variables [23, Theorem 2.8.1]. ∎

4.3. **Concentration of Contacts.** This section completes the proof of Theorem 4.3. The strategy is to use the estimates of Proposition 4.7, which show that the ideal gap chain behaves like simple random walk, to implement the heuristic sketched in Section 1.2.2. There are three steps. Section 4.3.1 introduces an exploration procedure that measures the distance between two global traversals. Section 4.3.2 estimates how long it takes this exploration procedure to find contacts, and Section 4.3.3 uses these estimates and concentration inequalities to establish Theorem 4.3.

To lighten notation in this section we often omit rounding to nearest integer values from the notation, e.g., writing $\sqrt{m}$ in place of $\lfloor \sqrt{m} \rfloor$.

4.3.1. *Exploration Procedure.* To analyse contacts between two global traversals we introduce an exploration procedure. Set $\boldsymbol{x} = (0,x)$, $\boldsymbol{y} = (0,y)$, $x \neq y$. The choice of column 0 entails no loss of generality, and similarly we will assume $x < y$, i.e., $0 < y - x \leq m - (y - x)$. Define $\Phi_2^k(\boldsymbol{u}) = \pi_2(\Phi^k(\boldsymbol{u}))$, and (abusing notation) write $\Phi_2^k(x)$ in place of $\Phi_2^k(\boldsymbol{x})$ when $\boldsymbol{x} = (0,x)$. Set

$$\hat{Z}_{s,\ell} = \Phi_2^s(y) - \Phi_2^\ell(x). \tag{32}$$

Note that if $s = \ell \bmod n$, $\hat{Z}_{s,\ell}$ measures the vertical distance between two strands of $\mathsf{Gt}(\boldsymbol{x})$ and $\mathsf{Gt}(\boldsymbol{y})$. Our exploration procedure uses $\hat{Z}_{s,\ell}$, with $s = \ell = 0$ initially. We first describe the procedure informally. Let $\eta' = \frac{D}{2}$, and $\eta(m) = \eta'\sqrt{m \log m}$. We view $\eta(m)$ as a scale at which global traversals are (relatively) close. When the global traversals are close, the exploration procedure will examine two strands simultaneously, via $(\hat{Z}_{s+r,\ell+r})_{r \geq 0}$. When the global traversals are not close the procedure will instead examine a single strand from the "lower" global traversal, i.e., $(\hat{Z}_{s,\ell+r})_{0<r\leq n}$ if $\hat{Z}_{s,\ell}$ is positive, and $(\hat{Z}_{s+r,\ell})_{0<r\leq n}$ if $\hat{Z}_{s,\ell}$ is negative.

Fix $M \in \mathbb{N}_+$. To formally define the *exploration procedure* $(\tilde{Z}_k)_{k=0}^{Mn}$, set $\sigma_0 = 0$, $\tilde{Z}_0 = \hat{Z}_{0,0}$, and suppose that $(\tilde{Z}_k)_{0 \leq k \leq \sigma_j}$ and $\sigma_0 < \cdots < \sigma_j < Mn$ are given, with $\tilde{Z}_k = \hat{Z}_{s(k),\ell(k)}$ for specified $s(k)$, $\ell(k)$. In what follows we will write $\tilde{Z}_k = \hat{Z}_{s,\ell}$, omitting that $s = s(k)$ and $\ell = \ell(k)$. Then the exploration process continues as follows while $\sigma_j < Mn$:

(1) If $|\tilde{Z}_{\sigma_j}| < \eta(m) - 1$ and $\tilde{Z}_{\sigma_j} = \hat{Z}_{s,\ell}$, then $\tilde{Z}_{\sigma_j+1} = \hat{Z}_{s,\ell+1}$ and $\tilde{Z}_{\sigma_j+2} = \hat{Z}_{s+1,\ell+1}$. Repeat this until

$$\sigma_{j+1} = \inf_{r>0} \left\{ \tilde{Z}_{\sigma_j+2r} \in \{\sqrt{m} - 1, \sqrt{m}, \eta(m) - 1, \eta(m)\} \right\} \tag{33}$$

if this is at most Mn; otherwise set $\sigma_{j+1} = Mn$.

(2) For $k \geq 0$, if $|\tilde{Z}_{\sigma_j+kn}| \geq \eta(m) - 1$, then
   (a) if $\tilde{Z}_{\sigma_j+kn} = \hat{Z}_{s,\ell} > 0$, then $\tilde{Z}_{\sigma_j+kn+r} = \hat{Z}_{s,\ell+r}$ for $0 < r \leq n$.
   (b) If $\tilde{Z}_{\sigma_j+kn} = \hat{Z}_{s,\ell} \leq 0$, then $\tilde{Z}_{\sigma_j+kn+r} = \hat{Z}_{s+r,\ell}$ for $0 < r \leq n$.

   Let k' be the minimal k such that $|\tilde{Z}_{\sigma_j+kn}| < \eta(m) - 1$. Set $\sigma_{j+1} = \sigma_j + k'n$ if this is at most Mn, and otherwise set $\sigma_{j+1} = Mn$.

Thus $\tilde{Z}_k$ parametrizes $\hat{Z}_{s,\ell}$ in terms of the weakly increasing sequences $s(k)$ and $\ell(k)$. Since s and ℓ increase by at most one, this parametrization considers all vertices in the evolution of $\boldsymbol{x}$, $\boldsymbol{y}$ under Φ.

Remark 7. *Strands of $\mathsf{Gt}(\boldsymbol{x})$ are typically between two strands of $\mathsf{Gt}(\boldsymbol{y})$. As $\mathsf{Gt}(\boldsymbol{x})$ evolves, it may initially be closer to the lower strand of $\mathsf{Gt}(\boldsymbol{y})$, and later be closer to the upper strand (and vice versa). The type (2) evolution accounts for this, and realizes the "corridor" structure present in Figure 1.*

To analyze the exploration process it will be convenient to subdivide it into two types of *process steps*, defined by a subsequence t_j of the stopping times σ_j. Let $t_0 = \sigma_0 = 0$, $t_1 = \sigma_1$, and, for $j \geq 1$,

$$t_{j+1} = Mn \wedge \begin{cases} \operatorname{argmin}\{\sigma_k > t_j \mid |\tilde{Z}_{\sigma_k}| \in \{\eta(m)-1, \eta(m)\}\} & |\tilde{Z}_{t_j}| \leq \sqrt{m} \\ \operatorname{argmin}\{\sigma_k > t_j \mid |\tilde{Z}_{\sigma_k}| \leq \sqrt{m}\} & |\tilde{Z}_{t_j}| \in \{\eta(m)-1, \eta(m)\} \end{cases}.$$

We think of $[\![t_0, t_1 - 1]\!]$ as an *initial process step*; if $t_1 = 0$ this interval is the empty and there is no initial process step. Subsequent intervals $[\![t_i, t_{i+1} - 1]\!]$ are called *outer process steps* if $|\tilde{Z}_{t_i}| \in \{\eta(m)-1, \eta(m)\}$, and *inner process steps* otherwise.

We will primarily consider $\tilde{Z}_k$ up to the first time $T = \inf\{2k \mid s(2k) \geq \Gamma n \text{ or } \ell(2k) \geq \Gamma n\}$ that either $\mathsf{Gt}(\boldsymbol{x})$ or $\mathsf{Gt}(\boldsymbol{y})$ has been explored. Note $T \leq 2\Gamma n$. Our first lemma rules out contacts during outer process steps. Write $\mathsf{Gt}(\boldsymbol{x})_\ell$ for the ℓ^{th} vertex in $\mathsf{Gt}(\boldsymbol{x})$.

Lemma 4.10. *Suppose $[\![t_i, t_{i+1} - 1]\!]$ is an outer process step of $(\tilde{Z}_k)_{k \leq T}$. Then $d(\mathsf{Gt}(\boldsymbol{x})_{s(t)}, \mathsf{Gt}(\boldsymbol{y})) \geq \sqrt{m} - 1$ for all $t \in [\![t_i, t_{i+1} - 1]\!]$ on $S(D)$.*

Proof. During an outer process step, both types of evolution (1) and (2) in the definition of the exploration process occur. The claim follows by the definition of an outer process step during an evolution of type (1). During (2), strands are revealed one by one, and Lemma 3.7 implies $d(\mathsf{Gt}(\boldsymbol{x})_{s(t)}, \mathsf{Gt}(\boldsymbol{y})) = \Theta(\sqrt{m \log m})$ for the t considered while revealing a single strand. ∎

Combined with Lemma 4.10, the next lemma indicates that $(\tilde{Z}_k)_{k \leq T}$ is sufficient for detecting all contacts between $\mathsf{Gt}(\boldsymbol{x})$ and $\mathsf{Gt}(\boldsymbol{y})$.

Lemma 4.11. *Suppose $\tilde{Z}_k = \hat{Z}_{s,\ell}$, and that $S(D)$ occurs. If $(\tilde{Z}_k)_{k \leq T}$ evolves by (1), then $|\tilde{Z}_{2k}|$ is the vertical distance $d(\mathsf{Gt}(\boldsymbol{x})_\ell, \mathsf{Gt}(\boldsymbol{y}))$.*

Proof. On $S(D)$, there is a distance at least $D\sqrt{m \log m}$ between strands of $\mathsf{Gt}(\boldsymbol{y})$. The lemma follows as $\eta(m) = \eta'\sqrt{m \log m}$, $\eta' = D/2$, and at even times during evolution by (1) $\tilde{Z}_{2k}$ is the vertical distance between two strands of $\mathsf{Gt}(\boldsymbol{x})$ and $\mathsf{Gt}(\boldsymbol{y})$. ∎

The next lemma reduces the analysis of inner process steps to an analysis of the ideal gap chain Z.

Lemma 4.12. *There exists a constant $c > 0$ such that the following holds on $S(D)$. Suppose that $[\![t_j, t_{j+1} - 1]\!]$ is an inner process step of $(\tilde{Z}_k)_{k \leq T}$ and $\tilde{Z}_{t_j} \neq 0$. There is a coupling of $\tilde{Z}$ and the ideal gap chain during this process step such that $(|\tilde{Z}_{2k}|)_{2k \in [\![t_j, t_{j+1}-1]\!]} = (Z_k)_{2k \in [\![t_j, t_{j+1}-1]\!]}$ with probability $1 - e^{-cm^{1/2}}$.*

Proof. On $S(D)$, during an inner process step the previously explored strands are at distance at least $\Theta(\sqrt{m \log m})$ away. This includes the initial strand of $\mathsf{Gt}(\boldsymbol{x})$ by Lemma 3.10; this lemma applies since we consider $\tilde{Z}_k$ only up until T. Thus Lemma 2.7 implies that computing transition probabilities of $\tilde{Z}$ by neglecting correlations from previous strands only introduces an exponentially small (in $\sqrt{m \log m}$) error. Lemma 2.7 also implies that the error in computing these transition probabilities on $\mathbb{Z}$ instead of C_m is also exponentially small (in m). The transition probabilities on $\mathbb{Z}$ are those that define the ideal gap chain, so the claim follows by a union bound, as $T \leq 2\Gamma n \leq m^2$. ∎

The following two remarks concern the disjointness of $\mathsf{Gt}(\boldsymbol{x})$ and $\mathsf{Gt}(\boldsymbol{y})$.

Remark 8. *Lemma 4.11 implies that if $y - x < D\sqrt{m \log m}$, then $\mathsf{Gt}(\boldsymbol{x})$ and $\mathsf{Gt}(\boldsymbol{y})$ are disjoint on $S(D)$: $\boldsymbol{y} \notin \mathsf{Gt}(\boldsymbol{x})$ since consecutive strands of $\mathsf{Gt}(\boldsymbol{x})$ are at least $D\sqrt{m \log m}$ apart on $S(D)$, and equation (17) of Lemma 3.10 implies that subsequent strands of $\mathsf{Gt}(\boldsymbol{x})$ do not get close to $\boldsymbol{y}$. Similarly,*

(17) *implies that* $\boldsymbol{x}$ *cannot be in* $\mathsf{Gt}(\boldsymbol{y})$. *Thus* $\mathsf{Gt}(\boldsymbol{x})$ *and* $\mathsf{Gt}(\boldsymbol{y})$ *are disjoint, as every other vertex of* $\mathsf{Gt}(\boldsymbol{y})$ *has an incoming arrow from a vertex not in* $\mathsf{Gt}(\boldsymbol{x})$.

Remark 9. *If* $y - x > D\sqrt{m \log m}$, *then* $\boldsymbol{y} \in \mathsf{Gt}(\boldsymbol{x})$ *is possible: this occurs if the exploration process starts its evolution according to (2) and hits zero. Note that this is consistent with Lemma 4.10, as the exploration process does not measure the vertical distance between strands when evolving according to (2). If* $\boldsymbol{y} \in \mathsf{Gt}(\boldsymbol{x})$ *we say the exploration process* fails, *as the process was designed for exploring disjoint global traversals (continuing the exploration process re-uses already revealed arrows). Otherwise the exploration process* succeeds, *and in this case each step of the exploration process reveals new arrows of* Φ. *Note that it cannot be that a subsequently revealed part of* $\mathsf{Gt}(\boldsymbol{x})$ *is in* $\mathsf{Gt}(\boldsymbol{y})$ *if* $\boldsymbol{y} \notin \mathsf{Gt}(\boldsymbol{x})$.

In what follows we implicitly only consider inner process steps when the exploration process succeeds, as in the case of failure the inner process degenerates to being identically zero.

At this stage we can indicate how the proof of Theorem 4.3 will proceed. We wish to show concentration for the number of contacts that occur during the inner process steps of a successful exploration procedure. By Lemma 4.12 these inner process steps can simulated by independent ideal gap chains. Concentration will follow provided there are many inner process steps, and that during each of them the number of contacts is reasonably concentrated. The first of these is addressed in Section 4.3.2, and the second in Section 4.3.3.

4.3.2. *Analysis of Exploration Procedure Steps.* This subsection gives estimates for the duration of initial, inner, and outer process steps. We also estimate the total number of process steps during an exploration procedure. First we analyse the duration of inner process steps.

Lemma 4.13. *Let* τ *denote the duration of an inner process step. There is a* $c > 0$ *such that for* $0 < j \leq \sqrt{m}$ *and* $x > 0$

$$\mathbb{P}\left[\tau > x(m \log m) \,\middle|\, |\tilde{Z}_0| = j\right] \leq \exp(-cx). \tag{34}$$

Moreover, $\mathbb{E}\tau = \Theta(m \log m)$.

Proof. Note that τ is the hitting time of $\{\eta(m) - 1, \eta(m)\}$ by $|\tilde{Z}|$. By Lemma 4.12 we can analyse $|\tilde{Z}|$ during inner steps using the ideal gap chain Z. To do this we will use the processes $Y^{\pm}$ introduced in the proof of Proposition 4.7; recall also the value a that was introduced so that $Y^{\pm}$ have non-degenerate jump distributions when $Y^{\pm} \geq a$. We will first show (34) using Y^{-}.

Recall that $Z_k \geq Y^{-}_{2k}$. Let $\mathcal{E}_t$ be the event that Z_k takes at least $c_1 t$ non-zero jumps while $Z_k \geq a$ and $0 \leq k \leq t$. Then there are $c_1, c_2 > 0$ such that $\mathbb{P}[\mathcal{E}_t^c] \leq e^{-c_2 t}$, as if $Z_k < a$, there is a uniformly positive probability that $Z_{k+a} > a$ by constructing an explicit path for the ideal gap chain.

To establish (34), recall the modification $\mathcal{Y}^{-}$ of Y^{-} that was introduced prior to Lemma 4.9. Then, given the above estimate on $\mathbb{P}[\mathcal{E}_t]$, we have

$$\mathbb{P}\left[\tau^{Z} \geq x \,\middle|\, Z_0 = j\right] \leq \mathbb{P}\left[\tau^{Y^{-}} \geq 2c_1 x \,\middle|\, \mathcal{Y}^{-}_0 = a\right] + e^{-c_2 x} \tag{35}$$

by having $\mathcal{Y}^{-}$ evolve only when $Z \geq a$. The factor of 2 accounts for that $\mathcal{Y}^{-}$ takes two steps for every single step of Z. We have written, e.g., τ^{Z} for the hitting time of Z.

To estimate $\mathbb{P}\left[\tau^{\mathcal{Y}^{-}} \geq 2c_1 x \,\middle|\, \mathcal{Y}^{-}_0 = a\right]$, we use the Brownian coupling of Lemma 4.9. Set $i = ym \log m$, $q = \beta$. Then

$$\begin{aligned}
\mathbb{P}\left[\tau^{\mathcal{Y}^{-}} > ym \log m \,\middle|\, \mathcal{Y}^{-}_0 = j\right] &\leq \mathbb{P}\left[\left\{\tau^{\mathcal{Y}^{-}} > ym \log m\right\} \cap (E^{-}(i,q))^{c} \,\middle|\, \mathcal{Y}^{-}_0 = j\right] + \mathbb{P}\left[E^{-}(i,q)\right] \\
&\leq \mathbb{P}\left[\sup_{t \in [0, 2\beta ym \log m]} |B(t)| \leq \kappa_{\eta(m)-1} \,\middle|\, B(0) = \kappa_{\sqrt{m}}\right] + Ce^{-cym \log m} \\
&\leq e^{-cy},
\end{aligned}$$

where the second inequality is by (31), and the third has used that $\kappa_{\eta(m)-1} = \Theta(\sqrt{m \log m})$. Using this with $y = 2c_1 x$ and (35) with x replaced by $xm \log m$ establishes (34).

The estimate for $\mathbb{E}\tau$ follows by an analogous computation using $\mathcal{Y}^+$ to show $\mathbb{P}[\tau > xm\log m \mid Z_0 = j] = \Theta(1)$ if x is small enough. ∎

We next consider the duration of outer process steps.

Lemma 4.14. *Let τ denote the duration of an outer process step. There is a $c > 0$ such that*

$$\mathbb{P}\left[\tau > xm\log m \,\middle|\, S(D)\right] \le e^{-cx}, \qquad 0 \le x \le \sqrt{m}. \tag{36}$$

Moreover, $\mathbb{E}\tau = \Theta(m\log m)$.

Proof. We begin by showing (36). Note that on $S(D)$ an outer process step starts at distance at most $K\sqrt{m\log m}$, $K = 4 + \frac{3D}{2}$, as this is the largest distance between two consecutive strands. Hence it is enough to prove that there is a $c_1 > 0$ such that

$$\mathbb{P}\left[\tau > xm\log m \,\middle|\, |\tilde{Z}_0| = j\right] \le \exp\{-c_1 x\}, \qquad j \le K\sqrt{m\log m}. \tag{37}$$

During an outer process step we can replace the random variable $\phi_{\boldsymbol{x}}$ of Φ by independent simple random walk increments with distribution ϕ at a cost of an error in probabilities of size at most $\exp\{-\Theta(m^{1/2})\}$ by Lemma 4.10 and Lemma 2.8. We do so in what follows; this is the source of the upper bound on x in (36).

Instead of determining $|\tilde{Z}_j|$ via the exploration procedure, we instead determine it by revealing $\mathsf{Gt}(\boldsymbol{y})$, and then revealing $\mathsf{Gt}(\boldsymbol{x})$. Let Z'_j be the distance from $\mathsf{Gt}(\boldsymbol{x})_j$ to $\mathsf{Gt}(\boldsymbol{y})$. By the preceding paragraph, Z' is a simple random walk with step distribution ϕ on an interval $[\![0, W]\!]$, where $W \in [\![D\sqrt{m\log m}, K\sqrt{m\log m}]\!]$ is random. Let τ' be the first time at which Z' is within distance $\sqrt{m} - 1$ of the boundary of $[\![0, W]\!]$. Then τ' is stochastically dominated by the first time Z' is within distance $\sqrt{m} - 1$ of the boundary of $[\![0, K\sqrt{m\log m}]\!]$. By a modification of the Brownian coupling from Lemma 4.9 (i.e., carrying out the same construction for simple random walk) and arguing as in the proof of Lemma 4.13, this implies $\mathbb{P}[\tau' > xm\log m | Z'_0 = j] \le \exp\{-c_1 x\}$ for all $j \in [\![0, K\sqrt{m\log m}]\!]$. To conclude the proof of (36), observe that τ' can differ from τ by at most a multiplicative constant, as outer process steps must evolve both $\mathsf{Gt}(Vx)$ and $\mathsf{Gt}(\boldsymbol{y})$ a positive proportion of the time.

To obtain the order of $\mathbb{E}\tau$, it suffices to prove $\mathbb{P}[\tau > xm\log m \mid S(D)] = \Theta(1)$ for some $x > 0$. This follows from the discussion above as Z' begins at distance $\Theta(\sqrt{m\log m})$ from the boundary of $[\![0, W]\!]$ and $W \ge D\sqrt{m\log m}$. ∎

The next lemma gives a similar estimate for the duration of the initial process step.

Lemma 4.15. *Let τ denote the duration of the initial process step started from $|\tilde{Z}_0| = j$, $K = 4 + \frac{3D}{2}$, and $k = (j/D\sqrt{m\log m} - K)_+$. There is a $c > 0$ such that*

$$\mathbb{P}[\tau > nk + xm\log m \mid S(D)] \le \exp(-cx), \qquad 0 \le x \le \sqrt{m}. \tag{38}$$

Proof. If $|\tilde{Z}_0| \le K\sqrt{m\log m}$, then $k = 0$ and this follows exactly from the arguments given for Lemmas 4.13 and 4.14. Now observe that since consecutive strands are separated by at least $D\sqrt{m\log m}$ on $S(D)$, if $j > K\sqrt{m\log m}$ then after (at most) k strands, $|\tilde{Z}_{kn}|$ will be at most $K\sqrt{m\log m}$. We can then apply the argument for $k = 0$. ∎

The exploration procedure alternates, after the initial process step, between inner and outer process steps. We block together pairs of inner and outer process steps into *full process steps*. In the next lemma the choices of $m^{3/8}$ and $m^{1/4}$ have been chosen for convenience.

Lemma 4.16. *Let N denote the number of full process steps for a successful exploration process $(\tilde{Z}_k)_{k\le T}$. Suppose $|\tilde{Z}_0| = O(\sqrt{m^{3/2}\log m})$. There are $\alpha, c_1 > 0$ and $\mu = \mu(m)$ such that*

$$\mathbb{P}\left[|N - \mu| > \alpha^{-1}m^{3/8} \mid S(D)\right] \le e^{-c_1 m^{1/4}}, \quad \mu = \Theta\left(\frac{m^{3/2}}{(\log m)^{3/2}}\right). \tag{39}$$

Proof. Let τ_i denote the duration of the i^{th} process step, where the initial process step is indexed by 0. Let $W_i = \tau_{2i-1} + \tau_{2i}$ be the duration of a full process step.

We first consider what happens for an exploration process run up to time Mn. We assume $m^{7/16} \leq M \leq 2\Gamma$, and aim to show (39) with $\mu = Mn/(\alpha m \log m)$. Note that $N = k$ is equivalent to Mn being in the interval $[\![\tau_0 + \sum_{i=1}^{k} W_i, \tau_0 + \sum_{i=1}^{k+1} W_i - 1]\!]$. We can bound $\sum_{i=1}^{k} W_i$ above and below by an IID sum of sub-exponential random variables, as conditional on $\tilde{Z}_{\tau_i}$, τ_{i+1} is a sub-exponential random variable by Lemmas 4.13 to 4.15. In the following we abuse notation slightly and continue to write W_i for these bounding variables.

The rest of the proof consists of estimating the probability that $N = k$, i.e., that W_{k+1} is the term that makes the sequence of partial sums $\sum_{i=1}^{r} W_i$ larger than $Mn - \tau_0 \approx Mn$. More precisely, by Lemmas 4.13 to 4.15, we have that

$$\mathbb{P}\left[\tau_0 \geq m^{11/8} \log m\right] \leq e^{-cm^{3/8}}, \quad \mathbb{P}\left[W_1 \geq m^{11/8} \log m\right] \leq e^{-cm^{3/8}}. \tag{40}$$

In applying Lemma 4.15 we have used the hypothesis $|\tilde{Z}_0| = O(\sqrt{m^{3/2} \log m})$. By (40) it suffices to show that there is an $\alpha > 0$ such that

$$\mathbb{P}\left[\sum_{i=1}^{k} W_i > Mn - 2m^{11/8} \log m\right] \leq e^{-cm^{1/4}}, \qquad k < \mu - \alpha^{-1} m^{3/8}, \tag{41}$$

$$\mathbb{P}\left[\sum_{i=1}^{k} W_i < Mn\right] \leq e^{-cm^{1/4}}, \qquad k > \mu + \alpha^{-1} m^{3/8}, \tag{42}$$

as the lemma then follows by a union bound over k since $N \leq m^2$. Define α by $\alpha m \log m = \mathbb{E}W_1$; $\alpha = \Theta(1)$ by Lemmas 4.13 and 4.14. By Bernstein's inequality for subexponential random variables [23, Theorem 2.8.2],

$$\mathbb{P}\left[\Big|\sum_{i=1}^{k} W_i - k\alpha m \log m\Big| \geq tm \log m\right] \leq 2 \max\{e^{-c_2 \frac{t^2}{k}}, e^{-c_3 t}\}.$$

The claim follows by re-arranging (41) and (42) and applying this estimate. This uses $M \leq 2\Gamma \leq 2\sqrt{m}$, and for (41) that $M \geq m^{7/16}$ so $m^{11/8} \log m = o(Mn)$. This completes the proof of the claim for an exploration procedure of length Mn.

To finish the proof of the lemma, write $T = M_T n$. T differs from $2\Gamma n$ only due to the phase of the initial process step that evolves a single global traversal. This phase takes time $O(m^{5/4})$ by the assumption $|\tilde{Z}_0| = O(\sqrt{m^{3/2} \log m})$; see the proof of Lemma 4.15. This alters M_T by at most $m^{1/4}$, and hence μ by at most $o(m^{3/8})$. This completes the proof. ∎

4.3.3. *Consequences.* Write $L(t) = |\{0 \leq s \leq t \mid Z_t = 1\}|$ for the local time of the ideal gap chain at state 1 up to time t. Note that the law of $L(t)$ depends on Z_0.

Lemma 4.17. *Let τ be the hitting time of $\{\eta(m) - 1, \eta(m)\}$ by the ideal gap chain with $Z_0 = 1$. Then*

$$L(\tau) = \sum_{i=1}^{R} V_i \tag{43}$$

where R is a geometric random variable with success probability $\Theta((\log m)^{-1/2})$ and the V_i are IID geometric random variables with success probability $\Theta(m^{-1/2})$. The variables R and V_i are independent.

Proof. Let $\tau_1 < \infty$ be the first hitting time of $\{\sqrt{m} - 1, \sqrt{m}\}$. For $r \geq 1$ let $h_r = \inf\{j > \tau_r \mid Z_j = 1\}$ and $\tau_{r+1} = \inf\{j > h_r \mid Z_j \in \{\sqrt{m} - 1, \sqrt{m}\}$. The number $R - 1$ of excursions to 1 from $\{\sqrt{m} - 1, \sqrt{m}\}$ is a geometric random variable with success probability $\Theta(\frac{1}{\sqrt{\log m}})$ by Proposition 4.7. This is by applying the strong Markov property at time τ_r, which yields an IID Bernoulli sequence of trials to hit 1 before $\{\eta(m) - 1, \eta(m)\}$; the variables are identically distributed as the hitting distribution on $\{\sqrt{m} - 1, \sqrt{m}\}$ is IID after a visit to 1.

If h_r occurs, the number of visits V_{r+1} to 1 before returning to $\{\sqrt{m}-1, \sqrt{m}\}$ is a geometric random variable with success probability $\Theta(m^{-1/2})$ by Proposition 4.7. This proves the claim, as there are also V_1 visits to 1 before hitting $\{\sqrt{m}-1, \sqrt{m}\}$. ■

The next technical lemma will allow us to control the visits of the exploration process to 1.

Lemma 4.18. *Let $N \in \mathbb{N}_+$, $\tilde{N} = \sum_{i=1}^{N} R_i$, and $\tilde{W} = \sum_{i=1}^{\tilde{N}} V_i$. Suppose the V_i an IID sequence of sub-exponential random variables and the R_i an IID sequence of $\mathbb{N}$-valued sub-exponential random variables. There is a $c > 0$ such that*

$$\mathbb{P}\left[|\tilde{W} - \mathbb{E}\tilde{W}| > 3N^{2/3}\right] \leq 4e^{-cN^{1/3}}.$$

Proof. By Bernstein's inequality [23, Theorem 2.8.1], there is a $c_1 > 0$ such that

$$\mathbb{P}\left[\left|\tilde{N} - \mathbb{E}\tilde{N}\right| > N^{2/3}\right] \leq e^{-c_1 N^{1/3}}, \tag{44}$$

so it suffices to estimate $|\tilde{W} - \mathbb{E}\tilde{W}| > 3N^{2/3}$ on the event $\mathcal{E} = \{|\tilde{N} - \mathbb{E}\tilde{N}| \leq N^{2/3}\}$. Write $\tilde{W} - \mathbb{E}\tilde{W}$ as

$$\sum_{i=1}^{\tilde{N}} V_i - \mathbb{E}V_i = \sum_{i=1}^{\mathbb{E}\tilde{N}} (V_i - \mathbb{E}V_i) + \mathbf{1}_{\tilde{N} > \mathbb{E}\tilde{N}} \sum_{i=\mathbb{E}\tilde{N}+1}^{\tilde{N}} (V_i - \mathbb{E}V_i) - \mathbf{1}_{\tilde{N} < \mathbb{E}\tilde{N}} \sum_{i=\tilde{N}+1}^{\mathbb{E}\tilde{N}} (V_i - \mathbb{E}V_i)\,. \tag{45}$$

Bernstein's inequality implies there is a $c_2 > 0$ such that

$$\mathbb{P}\left[\left| \sum_{i=1}^{\mathbb{E}\tilde{N}} (V_i - \mathbb{E}V_i) \right| > N^{2/3} \right] \leq e^{-c_2 N^{1/3}},$$

and that (interpreting $\sum_{i=k+r}^{k} a_i = 0$ if $r > 0$)

$$\begin{aligned}
\mathbb{P}\left[\left\{ \Big| \sum_{i=\mathbb{E}\tilde{N}+1}^{\tilde{N}} (V_i - \mathbb{E}V_i) \Big| > N^{2/3} \right\} \cap \mathcal{E} \right] &= \sum_{j=1}^{N^{2/3}} \mathbb{P}\left[\tilde{N} = \mathbb{E}\tilde{N} + j, \Big| \sum_{i=1}^{j} (V_i - \mathbb{E}V_i) \Big| > N^{2/3} \right] \\
&\leq \max_{j \in [\![1, N^{2/3}]\!]} \mathbb{P}\left[\Big| \sum_{i=1}^{j} (V_i - \mathbb{E}V_i) \Big| > N^{2/3} \right] \\
&\leq e^{-c_2 N^{1/3}}.
\end{aligned}$$

A similar argument bounds the probability that $\sum_{i=\tilde{N}+1}^{\mathbb{E}\tilde{N}} (V_i - \mathbb{E}V_i) > N^{2/3}$ on $\mathcal{E}$. The lemma follows by applying the triangle inequality to (45), using the above estimates, (44), and a union bound. ■

Proposition 4.19. *Let W be the number of visits of a successful exploration process $(\tilde{Z}_k)_{k \leq T}$ to 1. Suppose $|\tilde{Z}_0| = O(\sqrt{m^{3/2} \log m})$. There is a $c > 0$ such that*

$$\mathbb{P}\left[\left|W - \nu\right| \geq m^{1/3} \mid S(D)\right] \leq 5e^{-cm^{1/7}}, \quad \nu = \Theta\left(\frac{m}{\log m}\right). \tag{46}$$

Proof. Let N denote the number of full process steps of $(\tilde{Z}_k)_{k \leq T}$. By Lemma 4.16 there is a $c_1 > 0$ and $\mu = \mu(m)$ such that

$$\mathbb{P}\left[\left|N - \mu\right| > \alpha^{-1} m^{3/8}\right] \leq e^{-c_1 m^{1/4}}, \qquad \mu = \Theta\left(\frac{m^{1/2}}{\alpha (\log m)^{3/2}}\right).$$

Hence it suffices to verify (46) on the event $|N - \mu| \leq \alpha^{-1} m^{3/8}$.

By Lemma 4.10 visits to one only occur during inner process steps. To count these visits, couple each inner process step of $\tilde{Z}$ with an independent copy of the ideal gap chain by Lemma 4.12. This coupling may fail during the final inner process step, as the process $\tilde{Z}$ may end during this process step.

We obtain a lower bound on W by ignoring the final ideal gap chain contribution, and an upper bound by including it. By Lemma 4.17, we have obtained

$$W \in \Big[\!\Big[\sum_{i=0}^{N} \sum_{j=1}^{R_i - 1} V_{i,j}, \sum_{i=0}^{N+1} \sum_{j=1}^{R_i - 1} V_{i,j} \Big]\!\Big], \tag{47}$$

where the notation means the following. Let Q be a geometric random variable with success probability $\Theta((\log m)^{-1/2})$. In (47), the R_i, $i \geq 1$ are identically distributed random variables, and conditionally on $R_i \geq 1$, R_i is distributed as Q. The variable R_0 is stochastically bounded above by Q. The $V_{i,j}$ are identically distributed geometric random variables with success probability $\Theta(m^{-1/2})$. All variables are independent. The upper index $R_i - 1$ in the sums accounts for the fact that inner process steps do not start at 1, unlike the hypotheses of Lemma 4.17. The R_i are IID as the initial distribution of $\tilde{Z}$ for an inner process step is always the same. The variable R_0 is singled out as it accounts for visits to 1 in the initial process step; the domination claim holds as conditional on visiting 1 in the initial process step, R_0 has the same distribution as R_1.

The lemma now follows by a union bound and by applying Lemma 4.18 to both endpoints of the interval in (47), using that we can work on the event $|N - \mu| \leq m^{3/8}$. The power $m^{1/7}$ in the conclusion is for convenience, i.e., to drop logarithmic factors. The order of ν is found by computing $\mathbb{E}W = \mathbb{E}N\mathbb{E}R\mathbb{E}V = \Theta(\frac{m^{1/2}}{\alpha(\log m)^{3/2}}\sqrt{\log m}\sqrt{m}) = \Theta(\frac{m}{\log m})$. ∎

Proof of Theorem 4.3. We begin by giving the proof when $0 < y - x < D\sqrt{m \log m}$. In this case, by Lemma 4.11, the number of times $\mathsf{Gt}(\boldsymbol{x})$ and $\mathsf{Gt}(\boldsymbol{y})$ are at distance one is equal to the number of visits W of $(\tilde{Z}_k)_{k \leq T}$ to one, where $\tilde{Z}_0 = y - x$. Let A be the set of edges in a given column at which two global traversals are at distance one. On $S(D)$, the edges in A are at distance $\Theta(\sqrt{m \log m})$ from one another. The arrows at the endpoints of such an edge form a contact with positive probability, and up to an exponentially small error this is true for any such set A of contacts by Lemma 2.8. As a result we obtain that $\mathbb{P}\left[|X - \mathbb{E}X| \geq k^{2/3} \mid W = k, S(D)\right] \leq e^{-ck^{1/3}}$ by Bernstein's inequality. The theorem follows by Proposition 4.19.

Next we consider the general case. The additional ingredient to handle is that since global traversals do not quite vertically span the torus, it may not be possible to follow all of $\mathsf{Gt}(\boldsymbol{x})$ and $\mathsf{Gt}(\boldsymbol{y})$ starting from two vertices at distance less than $D\sqrt{m \log m}$. Instead, let $\boldsymbol{y}'$ be the closest point in $\mathsf{Gt}(\boldsymbol{y})$ to $\boldsymbol{x}$; by Lemma 3.10 (17), $\boldsymbol{y}'$ at distance $O(m^{3/4}(\log m)^{1/2})$ from $\boldsymbol{x}$. Similarly there is a $\boldsymbol{x}' \in \mathsf{Gt}(\boldsymbol{x})$ within distance $O(\sqrt{m \log m})$ of $\boldsymbol{y}$. To determine the number X of contacts between $\mathsf{Gt}(\boldsymbol{x})$ and $\mathsf{Gt}(\boldsymbol{y})$ we run an exploration process started from (i) $\boldsymbol{x}$ and $\boldsymbol{y}'$ and (ii) from $\boldsymbol{x}'$ and $\boldsymbol{y}$. In case (i) we run the process until $\mathsf{Gt}(\boldsymbol{y})$ ends, and in case (ii) until $\mathsf{Gt}(\boldsymbol{x})$ ends. The total time run by these exploration processes is $2\Gamma n - \Theta(m^{3/4}(\log m)^{1/2})$. The same argument as in the first paragraph relates $\tilde{Z}_k = 1$ to contacts, and hence the claim now follows by applying Proposition 4.19 (formally this lemma applies to a single exploration procedure, but the additional initial/terminal steps due to having two procedures do not alter the conclusion). ∎

5. Dynamics on cycles

Section 2.4 introduced Glauber dynamics that leave the equilibrium distribution of the hard-core model invariant. Since Φ is equivalent to a collection of hard-core models, running the Glauber dynamics on each column of $\mathbb{T}_{n,m}$ preserves Φ and its cycle structure. Section 5.1 describes the resulting effective dynamics on the cycle structure, which imitate the random transposition dynamics. This will allow us to prove Theorem 1.2 in Section 5.2 by making use of methods developed by Schramm [19].

Schramm's dynamical ideas are at this point relatively standard, and have been used (and exposited) in other works concerning random permutations, e.g., [1, 5]. As such we provide detailed proofs when our setting necessitates adjustments, but we have omitted proofs when (up to notation) they are as given by Schramm.

5.1. Glauber dynamics and the induced split-merge dynamics for Φ.

5.1.1. *Dynamics.* The Glauber dynamics from Section 2.4 naturally extend to give a dynamics on the n (disjoint) cycles C_m that comprise the torus $\mathbb{T}^\star_{n,m}$. To be precise, we define the *directed spatial permutation Glauber dynamics* by repeatedly

(1) choosing a uniformly random column $j \in [\![n]\!]$, and
(2) performing a Glauber update on the hard-core configuration on the j^{th} column of $\mathbb{T}^\star_{n,m}$.

This update to the hard-core configuration may change the corresponding arrow configuration, and hence the associated bijection of $\mathbb{T}_{n,m}$. We record two basic facts about these dynamics.

Lemma 5.1. *Suppose $(\Phi_i)_{i\geq 0}$ is the directed spatial permutations Glauber dynamics started from equilibrium, i.e., $\Phi_0 = \Phi$ conditioned to contain no global shift. Then:*

(1) *Φ_i is distributed as Φ_0 for all i, i.e., Φ_0 is the invariant measure of the dynamics.*
(2) *There is a $p_0 > 0$ such that $\mathbb{P}[\Phi_{i+1} \neq \Phi_i] \geq p_0$ for all $i \geq 0$.*

Proof. Proposition 2.6 implies that Φ conditioned on having no global shift is equivalent to a product of hard-core models on the columns of $\mathbb{T}^\star_{n,m}$. The first claim follows as the Glauber dynamics leave the hard-core models on each column invariant. For (2), it suffices to show that the directed spatial permutations Glauber dynamics have a uniformly positive probability to change the status of some vertex $\boldsymbol{v}$ at which a contact occurs. In hard-core model language, these transitions (i) set $\boldsymbol{v}$ to be unoccupied if it was previously occupied and (ii) set $\boldsymbol{v}$ to be occupied if it and its neighbours were previously unoccupied. The claim follows as the density of vertices satisfying the conditions in (i) and (ii) is deterministically positive, and so are the transition probabilities. ∎

5.1.2. *Induced split-merge dynamics.* This section describes the effective split-merge dynamics on cycles that are induced by the directed spatial permutation Glauber dynamics. Some results in this section make statements about properties of global traversals without discussing the arbitrary choices that must be made to decompose the cycles of Φ into global traversals. The conclusions are well-defined as they hold for any such decomposition; recall the discussion following Definition 3.8.

The next lemma constructs a good event on which we will be able to analyse the effective dynamics. The error estimates are convenient ones that suffice; we have not made an attempt to optimize them. Recall that C' is the constant in $C(m) = C'\sqrt{m \log m}$.

Lemma 5.2. *If C' is sufficiently large, there is an event $\mathcal{G} \subset S(D)$ with $\mathbb{P}[\mathcal{G}] = 1 - m^{-3}$ such that:*

(1) *there is a $c > 0$ such that $\{\mathsf{C}\}_{\mathsf{Gt}}/\lfloor \mathsf{C} \rfloor_{\mathsf{Gt}} < cm^{-1/8}$ for all cycles C;*
(2) *for all cycles C_i, C_j, the number of contacts between them is*

$$\lfloor \mathsf{C}_i \rfloor_{\mathsf{Gt}} (\lfloor \mathsf{C}_j \rfloor_{\mathsf{Gt}} - 1_{i=j}) \mathbb{E}X(1 + o(m^{-1/9})),$$

where X denotes the number of contacts between two disjoint global traversals; and
(3) *there is a $c_2 > 0$ such that each global traversal contains $c_2 \Gamma n$ contacts.*

Proof. Take $D > 6$. Choose C' to be at least as large as indicated in Corollary 3.6, so that $\mathbb{P}[S(D)] \geq 1 - o(m^{-3})$ by Remark 6. Then item (1) holds on $S(D)$ by Lemma 3.11 and Remark 6. Let $\mathcal{E}_3$ be the event that item (3) holds. Then $\mathbb{P}[\mathcal{E}_3] \geq 1 - e^{-cm}$ by Lemma 4.2 and a union bound. We will construct $\mathcal{G}$ as sub-event of $S(D) \cap \mathcal{E}_3$ by intersecting the latter with a further event $\mathcal{E}_2$.

Let $\mathcal{E}_2 = \bigcap_{i,j} \mathcal{E}_2(i,j)$, where $\mathcal{E}_2(i,j)$ indicates that item (2) holds for C_i and C_j. It suffices to prove $\mathbb{P}[\mathcal{E}_2(i,j)]$ occurs with all but stretched exponentially small probability. To show this, suppose C_i and C_j are comprised of k_i and k_j global traversals, respectively. On $S(D)$, $k_i, k_j \geq 1$. The number of contacts between the global traversals of C_i and C_j is a sum of $k_i(k_j - 1_{i=j})$ copies of X, X the number of contacts between two disjoint global traversals. By Theorem 4.3 and a union bound the number of contacts between the global traversals of C_i and C_j satisfies the desired estimate.

What remains is to show that the contacts that involve fractional global traversals of C_i and C_j are taken into account by the $o(m^{-1/9})$ error term. The number of contacts contained in the fractional global traversal of a cycle is at most twice the length of the fractional global traversal. On $S(D)$, item (1) thus bounds above the number of contacts in the fractional global traversal of C_i by $2k_i m^{-1/8} n = 2k\mathbb{E}X \cdot o(m^{-1/9})$. Thus the number of contacts between C_i and C_j that are

contained in fractional global traversals is at most $2(k_i + k_j)\mathbb{E}X \cdot o(m^{-1/9})$. The claim follows since $k_i + k_j \leq 2k_ik_j$. ∎

Lemma 5.3. *Conditional on $\mathcal{G}$, the probability that a uniformly chosen contact is contained in a fractional global traversal of a cycle C is at most $O(m^{-1/8}\lfloor \mathsf{C} \rfloor_{\mathsf{Gt}} / \sum_j \lfloor \mathsf{C}_j \rfloor_{\mathsf{Gt}})$. As a consequence, the probability that a uniformly chosen contact is between two global traversals is at least $1 - O(m^{-1/8})$.*

Proof. Fix a decomposition of the cycles into global traversals. If there is a contact at $\boldsymbol{v} \in \mathbb{T}^{\star}_{n,m}$, let $\{\boldsymbol{v}_1, \boldsymbol{v}_2\} = \{\boldsymbol{v} \pm \frac{1}{2}\}$. Let C_1 and C_2 denote the cycles containing v_1 and v_2 ($\mathsf{C}_1 = \mathsf{C}_2$ may occur).

Let E_i be the event that $\boldsymbol{v}_i$ is contained in the fractional part of C_i. By Lemma 5.2(1), at most a $O(m^{-1/8})$ fraction of the strands of C_i are in the fractional part of C_i, and the number of contacts in the fractional part of C_i is at most twice the length of the fractional part. Since the number of contacts in the global traversals of C_i is at least a constant fraction of the total length of the global traversals by Lemma 5.2(3), the probability of E_i is thus $O(m^{-1/8})$. The first conclusion follows. A union bound to estimate $\mathbb{P}[E_1^c \cap E_2^c]$ then yields the second conclusion. ∎

Lemma 5.4. *Let $E_{i,j}$ be the event that a uniformly chosen contact is between C_i and C_j. Let $\mathcal{E}_{\mathsf{Gt}}$ be the event that a uniformly selected contact is between two global traversals. Then $\mathbb{P}[E_{i,j} \mid \mathcal{E}_{\mathsf{Gt}} \cap \mathcal{G}]$ is proportional to $\lfloor \mathsf{C}_i \rfloor_{\mathsf{Gt}} \lfloor \mathsf{C}_j \rfloor_{\mathsf{Gt}}$ in the sense that*

$$\mathbb{P}[E_{i,j} \mid \mathcal{E}_{\mathsf{Gt}} \cap \mathcal{G}] = \frac{\lfloor \mathsf{C}_i \rfloor_{\mathsf{Gt}} (\lfloor \mathsf{C}_j \rfloor_{\mathsf{Gt}} - 1_{i=j})}{\sum_{k,\ell} \lfloor \mathsf{C}_k \rfloor_{\mathsf{Gt}} \lfloor \mathsf{C}_l \rfloor_{\mathsf{Gt}}} (1 + o(m^{-1/9})).$$

Moreover, in the case $i = j$ the contact is equally likely to be between any two distinct global traversals of C_i, up to an $o(m^{-1/9})$ error.

Proof. To estimate this probability we use that every contact is either between two cycles, or between a cycle and itself. The number of contacts between two cycles is given by Lemma 5.2(2). Thus the probability that a uniformly chosen contact that is contained in two global traversals is between C_i and C_j is (if $i \neq j$) $\frac{\lfloor \mathsf{C}_i \rfloor_{\mathsf{Gt}} \lfloor \mathsf{C}_j \rfloor_{\mathsf{Gt}}}{\sum_{k,\ell} \lfloor \mathsf{C}_k \rfloor_{\mathsf{Gt}} \lfloor \mathsf{C}_\ell \rfloor_{\mathsf{Gt}}}$ up to an error of size $o(m^{-1/9})$. If $i = j$ the numerator requires the subtraction of 1 as no global traversal comes into contact with itself. (For the same reason, the denominator is in fact a slight overcount: $\lfloor \mathsf{C}_k \rfloor_{\mathsf{Gt}}^2$ should be $\lfloor \mathsf{C}_k \rfloor_{\mathsf{Gt}} (\lfloor \mathsf{C}_k \rfloor_{\mathsf{Gt}} - 1)$. This can be absorbed into the $o(m^{-1/9})$ error term since the total number of global traversals is of order $\sqrt{m \log m}$.) ∎

5.2. **Proof of Theorem 1.2.** There are two steps in our proof, both using that the split-merge dynamics induced by the Glauber dynamics is very similar to the random transposition dynamics. The first step shows that most global traversals in Φ are in fact contained in large cycles. The second step, establishing the $\mathsf{PD}(1)$ limit, makes use of the fact that we can perfectly couple the two dynamics for a (relatively) short amount of time. This is sufficient as the dynamics acts very rapidly on the large cycles of Φ — this fact was one of the key observations used in [19] (see also [15]).

5.2.1. *Existence of large cycles.* The next two lemmas are analogues of Lemmas 2.3 and 2.4 of [19].

Definition 5.5. *The set of vertices in cycles C with $\lfloor \mathsf{C} \rfloor_{\mathsf{Gt}} \geq j$ is denoted $V_{\mathsf{C}}(j)$.*

Note that on $\mathcal{G}$, $V_{\mathsf{C}}(1)$ is the set of all vertices, as all cycles contain at least one global traversal.

Lemma 5.6. *Let $\varepsilon \in \left(0, \frac{1}{8}\right)$. Then*

$$\mathbb{E}[|V_{\mathsf{C}}(1) \setminus V_{\mathsf{C}}(\varepsilon C(m))|] \leq O(1)\varepsilon |\log \varepsilon| C(m). \tag{48}$$

Our proof of Lemma 5.6 follows the same strategy Schramm used to prove [19, Lemma 2.3] in the context of random transpositions. There are two main differences. First, Schramm's lemma was a non-equilibrium estimate, and some of the considerations required are not present in our equilibrium setting. Second, we have to take some care due to the fact that our control over the split-merge dynamics is based on global traversals. Cycles are not entirely comprised of global traversals, and this leads to errors when comparing the dynamics to random transposition split-merge dynamics. The proof shows these errors are negligible. Our exposition of the proof follows that presented in [19], both to facilitate comparison, and to emphasize that the key ideas are from this reference.

Proof of Lemma 5.6. Let $(\Phi_k)_{k\geq 0}$ denote the directed spatial permutation Glauber dynamics started from equilibrium, i.e., Φ conditional to contain no global shifts. We recall this is equivalent to Φ itself as $m \to \infty$ by Lemma 2.3. The chain $(\Phi_k)_{k\geq 0}$ is a lazy Markov chain, but by Lemma 5.1(2) there will be k non-lazy steps with probability $1 - o(1)$ after Ck total steps provided C is large enough. For this reason we measure time in terms of non-lazy steps in what follows.

The proof considers intervals of time on which a typical cycle will double in size. Set

$$\mu_s = \lceil a_s \rceil, \qquad a_s = 2\frac{C(m)}{2^s}\log_2 \frac{C(m)}{2^s}, \qquad s \geq 0,$$

$K = \lceil \log_2(\epsilon C(m))\rceil$, and $\tau_s = \sum_{i=0}^{s-1} \mu_i$ for $s \geq 1$. Our argument will end with $s = K$; after this interval a cycle that has successfully doubled in size at each step has size $2^K \geq \epsilon C(m)$. Set $t = \tau_K$. We will show that (48) holds for Φ_t with $t = \Theta(C(m)\log C(m)) = \Theta(\sqrt{m(\log m)^3})$. By stationarity this suffices. We begin with a few preparatory steps.

Since $\mathbb{P}[\mathcal{G}] = 1 - o(m^{-3})$ and each Φ_k is an equilibrium sample, the probability that $\mathcal{G}$ holds for $\Theta(t)$ steps is $1 - o(m^{-1})$ by a union bound. Thus by Lemmas 5.3 and 5.4 we may use the following description of the dynamics. With probability at least $1 - O(m^{-1/8})$ the dynamics is split-merge on the cycles, with their size measured in units of global traversals up to an error of size $m^{-1/9}$. With the complementary $O(m^{-1/8})$ probability either a cycle is split or two cycles are merged, but the cycles are selected with an unknown law and (if relevant) the distribution of the split is unknown. We call the first situation a *good transposition* and the second a *bad transposition*. A bad transposition corresponds to the event that a chosen contact is contained in a fractional global traversal. For the rest of the proof let $\gamma = 1/10$, and assume m is large enough such that both of these errors are at most $m^{-\gamma}$.

An important observation is that when a good tranposition occurs, the probability of splitting off a cycle of size s can be bounded above by essentially the same argument as for the random transposition model on $[\![N]\!]$ for $N \in \mathbb{N}_+$. For random transpositions, this probability is at most $2s/(N-1)$, as a random transposition is a choice of two random indices, and there are most $2s$ indices within distance s of the first index. To formulate an upper bound for directed spatial permutations, let $N_k = \sum_{\mathsf{C}(k)} \lfloor \mathsf{C}(k) \rfloor_{\mathsf{Gt}}$ denote the number of global traversals contained in the cycles of Φ_k, i.e., the sum is over the set of cycles of Φ_k. At each step of the dynamics a split or merge occurs, N_k may increase or decrease by one due to the joining of fractional global traversals / splitting of global traversals. Let good be the event that a good split occurs at time k. Then

$$\mathbb{P}[\mathsf{good},\ \text{split} \leq 2s] \leq \frac{2s}{N_k - 1}(1 + o(m^{-\gamma})),$$

where split $\leq 2s$ is shorthand for the event that a cycle containing at most $2s$ global traversals is created. This gives a uniform (in k) upper bound since $N_k \geq C(m)/2$ on $\mathcal{G}$ and $\lfloor \mathsf{C} \rfloor_{\mathsf{Gt}} \geq 1$ for all cycles.

We now move to the main analysis, which proceeds by determining the probability that typical cycles fail to grow. There are a few mechanisms to consider. The proof will often consider $|V|/(n\Gamma)$ for V a set of vertices. This is an estimate of the number of global traversals in V.

Failing due to a split: Let $s \geq 0$ and $k \in \{\tau_s + 1, \tau_s + 2, \ldots, \tau_{s+1}\}$. Let F^k be the set of vertices contained in a cycle of size at most 2^{s+1} that is created by a split at time k. At most two such cycles can be produced by a single step of the dynamics. Write good for the event of a good transposition occurring and bad for the complementary event. Then

$$\begin{aligned}
\mathbb{E}\frac{|F^k|}{n\Gamma} &\leq 2 \cdot (2^{s+1} + 1)\left(\mathbb{P}\left[\mathsf{good},\ \text{split} < 2^{s+1}\right] + \mathbb{P}\left[\mathsf{bad},\ \text{split} < 2^{s+1}\right]\right) \\
&\leq 2^{s+3}\left(\frac{2^{s+2}}{N_k - 1}(1 + o(m^{-\gamma})) + o(m^{-\gamma})\right) \\
&\leq \frac{2^{2s+7}}{C(m) - 2} + 2^{s+3}m^{-\gamma},
\end{aligned}$$

where the $+1$ in the first line accounts for the at most $n\Gamma$ vertices that can be in a cycle but not in a global traversal. Let $\tilde{F}^k = \cup_{\tau=1}^k F^\tau$. Then

$$\mathbb{E}\frac{|\tilde{F}^t|}{n\Gamma} \leq \sum_{s=0}^{K-1} \mu_s\left(\frac{2^{2s+7}}{C(m)-2} + 2^{s+3}m^{-\gamma}\right) = O(\epsilon|\log\epsilon|C(m)).$$

We have obtained the final estimate by using that the sum of the first term is $O(\epsilon|\log\epsilon|C(m))$ while the second is of order $o(C(m))$, as $m > C(m)$.

Failing to grow: Let $k \in \{\tau_s, \tau_s+1, \ldots, \tau_{s+1}-1\}$, $s \geq 0$. Cycles will not be likely to grow if there are too few reasonably sized cycles to merge with. To track this, set H^k to be $\emptyset$ if $|V_{\mathsf{C}}^k(2^s)| \geq \frac{C(m)}{2}$, and otherwise $H^k = V_{\mathsf{C}}(1)$. Set $\tilde{H}^k = \bigcup_{r=0}^k H^r$.

$H^k = V_{\mathsf{C}}(1)$ is a failure event, but if $H^k = \emptyset$ a cycle may still fail to grow. Write $V_{\mathsf{C}}^k(s)$ for the set of vertices in cycles of size at least s at time k. Define

$$B^s = V_{\mathsf{C}}^{\tau_s}(2^s) \setminus \left(\tilde{F}^{\tau_{s+1}} \cup \tilde{H}^{\tau_{s+1}-1} \cup V_{\mathsf{C}}^{\tau_{s+1}}(2^{s+1})\right),$$

and $\tilde{B}^s = \bigcup_{r=0}^s B^r$. Observe

$$V_{\mathsf{C}}(1) = V_{\mathsf{C}}^t(\epsilon C(m)) \cup \tilde{H}^t \cup \tilde{F}^t \cup \tilde{B}^{K-1}. \tag{49}$$

If $v \in B^s$, then for $k \in \{\tau_s, \ldots, \tau_{s+1}-1\}$ we have $|V_{\mathsf{C}}^k(2^s)| \geq \frac{C(m)}{2}$ since B^s and $\tilde{H}^{\tau_{s+1}-1}$ are disjoint by definition. Conditionally on v being in a cycle containing $r \in [\![2^s, 2^{s+1}-1]\!]$ global traversals at time k, and on $|V_{\mathsf{C}}^k(2^s)| \geq \frac{C(m)}{2}$, the probability the next step of the dynamics merges the cycle containing v with another cycle containing at least 2^s global traversals is at least

$$2^s\left(\frac{C(m)}{2} - 2^{s+1}\right)\binom{N_k}{2}^{-1}(1+o(m^{-\gamma})).$$

Since $N_k \leq C(m)$ and $\epsilon \leq \frac{1}{8}$, for $s \leq K-1$ this is bounded below by

$$2^s\left(\frac{C(m)}{2} - 2\epsilon C(m)\right)\frac{2}{C(m)(C(m)-1)}(1+o(m^{-\gamma})) \geq \frac{2^{s-1}}{C(m)}(1+o(m^{-\gamma})).$$

As a result,

$$\begin{aligned}\mathbb{P}[v \in B^s] \leq (1 - \frac{2^{s-1}}{C(m)}(1+o(m^{-\gamma})))^{\mu_s} &\leq \exp\left(-2^{s-1}\mu_s\frac{(1+o(m^{-\gamma}))}{C(m)}\right)\\ &\leq O(\frac{2^s}{C(m)})\exp\left(-o(m^{-\gamma})\log\frac{C(m)}{2^s}\right)\\ &= O(\frac{2^s}{C(m)})\end{aligned}$$

where the penultimate inequality used that $e > 2$. By summing over v this implies

$$\mathbb{E}\frac{|\tilde{B}^{K-1}|}{n\Gamma} = O(\epsilon C(m)).$$

Concluding: If $H^k = V_{\mathsf{C}}(1)$, then $|\tilde{F}^k \cup \tilde{B}^{s-1}| \geq \frac{n\Gamma C(m)}{2}$, so by Markov's inequality,

$$\mathbb{E}\frac{|\tilde{H}^t|}{n\Gamma} \leq C(m)\mathbb{P}\left[|\tilde{F}^t \cup \tilde{B}^{K-1}| \geq \frac{n\Gamma C(m)}{2}\right] \leq 2\mathbb{E}\frac{|\tilde{F}^t \cup \tilde{B}^{K-1}|}{n\Gamma}, \tag{50}$$

and this expectation is of order $O(\epsilon C(m)) + O(\epsilon|\log\epsilon|C(m)) = O(\epsilon|\log\epsilon|C(m))$ by the preceding two sections of the proof. Combined with (49) we are done. ∎

Lemma 5.7. *Let $0 < \epsilon, \alpha < \frac{1}{8}$, and let M be the minimal number of cycles that contain a $(1-\epsilon)$-fraction of the vertices of $\mathbb{T}_{n,m}$. Then there are $m_1(\epsilon)$ and a universal constant K such that for all $m \geq m_1(\epsilon)$*

$$\mathbb{P}\left[M > \alpha^{-1}|\log\alpha\epsilon|^2\right] \leq K\alpha. \tag{51}$$

Proof. Lemma 5.6 states $\mathbb{E}\left[|V_{\mathsf{C}}(1) \setminus V_{\mathsf{C}}(\epsilon C(m))|\right] \leq O(1)\epsilon|\log\epsilon|C(m)$. This equation implies the result exactly as discussed in the final paragraph of [19, Proof of Lemma 2.4]. ∎

5.2.2. *Existence of macroscopic cycles and convergence to* $\mathsf{PD}(1)$. Write $\Phi^{(m)}$ to emphasize the m-dependence of Φ. To establish Theorem 1.2 we must show that if Y has distribution $\mathsf{PD}(1)$, then for any $\epsilon > 0$, if m is sufficiently large then there is a coupling of $\mathcal{X}(\Phi^{(m)})$ and Y such that $\mathbb{P}\left[\|\mathcal{X}(\Phi^{(m)}) - Y\|_\infty < \epsilon\right] \geq 1 - \epsilon$. A theorem due to Schramm will be useful [19].

Theorem 5.8. *Suppose $(\pi_k^{(N)})_{k\geq 0}$ is the random transposition chain on $[\![N]\!]$ with an initial condition $\pi_0^{(N)}$. Suppose that for $\epsilon' > 0$ small enough,*

$$\mathbb{E}\left[N - |V_{\mathsf{C}}(\epsilon' N)|\right] < O(1)\epsilon'|\log\epsilon'|N \tag{52}$$

where the expectation is with respect to π_0. Let Y have distribution $\mathsf{PD}(1)$. For $\epsilon > 0$ let $q = \epsilon^{-1/2}$. If N is large enough, then there is a coupling of the random transposition dynamics such that

$$\mathbb{P}\left[\|\mathcal{X}(\pi_q^{(N)}) - Y\|_\infty < \epsilon\right] \geq 1 - \epsilon.$$

Remark 10. *Theorem 5.8 is not formally stated in* [19]*. We have extracted the hypotheses as discussed in* [19, Section 4]*, where Schramm proves his main theorem.*

We can now complete our proof of Theorem 1.2.

Proof of Theorem 1.2. Fix $\epsilon > 0$ and let $q = \epsilon^{-1/2}$. Given Φ, let N denote the number of global traversals contained in the cycles of Φ. By Lemma 5.2(1), $N \geq m/(2\Gamma)$ on the event $\mathcal{G}$, as each cycle contains at least one global traversal, and the total amount of the system in fractional global traversals cannot exceed that in global traversals. Note that (see Definition 3.8) $m/\Gamma \to \infty$ as $m \to \infty$.

Let π_0 be the partition of $[\![N]\!]$ that agrees with the cycle structure of Φ. That is, enumerate the global traversals, and if two global traversals are in a common cycle, we put them in a common part. Arbitrarily choose a permutation compatible with the given block structure. By Lemmas 5.6 and 5.7 and the first paragraph, π_0 satisfies the hypotheses of Theorem 5.8 when m is sufficiently large.

To complete the proof of the theorem we will use a coupling of the split-merge dynamics on the cycles of Φ induced by the directed spatial permutation Glauber dynamics with the split-merge dynamics induced by the random transposition dynamics started from π_0. Write Φ_t and π_t for the respective marginals. The coupling has to account for the fact that the number of global traversals N_t in Φ_t can change with t, and goes as follows. Note that $N_{t+1} - N_t \in \{-1, 0, 1\}$. By Lemmas 5.2 and 5.4 the induced split-merge dynamics selects two distinct uniformly chosen global traversals with probability $1 - o(m^{-1/9})$, and we can couple this perfectly with choosing two distinct vertices under the random transposition dynamics if $N_t = N_0$. To account for the possibility that $N_0 \neq N_t$, suppose $N_t - N_0 > 0$ and mark $N_t - N_0$ of the global traversals. The probability that one of these global traversals is selected by the dynamics is at most $2t/N_0$; if neither is selected we can couple a step of the dynamics as above. If $N_t - N_0 < 0$ we couple similarly, but this time instead adding $|N_t - N_0|$ fictitious global traversals. The probability that our coupling succeeds without ever selecting a marked or fictitious global traversal in the first q steps is of order $1 - \Theta(q^2)/N_0 = 1 - o(1)$. The probability of always selecting two distinct global traversals is of order $1 - qm^{-1/9} = 1 - o(1)$. Hence we may assume each step is coupled as above.

Under this coupling the number of global traversals in a cycle of Φ_t may not be exactly equal to the number of indices in the corresponding part of π_t, but the discrepancy is at most t. The discrepancy can arise due to merging or splitting of fractional global traversals. Since Φ is invariant under the Glauber dynamics, this implies that the size $\lfloor \mathsf{C}_i \rfloor_{\mathsf{Gt}}$ of the i^{th} largest cycle of Φ is within q of the number of vertices contained in the i^{th} largest cycle of π_q. Since q is of order one as $m \to \infty$, this discrepancy is negligible due to the normalization present in the definition of the cycle structure. This shows that we can couple the vector $(\lfloor \mathsf{C}_i \rfloor_{\mathsf{Gt}} / (\sum_j \lfloor \mathsf{C}_j \rfloor_{\mathsf{Gt}}))_{i\geq 1}$ with a $\mathsf{PD}(1)$ sample as desired by Theorem 5.8. To obtain the statement of Theorem 1.2, which measures cycle sizes by the number of vertices they contain, we use (19), which implies that measuring size by a count of global traversals is equivalent (up to errors of size $o(1)$) to measuring by a count of vertices. ∎

Acknowledgements. A.H. is supported by the National Science Foundation under DMS grants 1855550 and 2153359 and by the Simons Foundation as a 2021 Simons Fellow. T.H. was supported by an NSERC PDF when this work was initiated.

References

[1] Radosław Adamczak, Michał Kotowski, and Piotr Miłoś. Phase transition for the interchange and quantum Heisenberg models on the Hamming graph. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 57, pages 273–325. Institut Henri Poincaré, 2021.

[2] Michael Aizenman and Bruno Nachtergaele. Geometric aspects of quantum spin states. *Communications in Mathematical Physics*, 164:17–63, 1994.

[3] Omer Angel. Random infinite permutations and the cyclic time random walk. In *Discrete random walks (Paris, 2003)*, Discrete Math. Theor. Comput. Sci. Proc., AC, pages 9–16 (electronic). Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2003.

[4] Volker Betz and Daniel Ueltschi. Spatial random permutations and infinite cycles. *Communications in Mathematical Physics*, 285:469–501, 2009.

[5] Jakob E Björnberg, Michał Kotowski, Benjamin Lees, and Piotr Miłoś. The interchange process with reversals on the complete graph. *Electronic Journal of Probability*, 24:1–43, 2019.

[6] Rick Durrett. *Probability: Theory and Examples*, volume 49. Cambridge University Press, 2019.

[7] Dor Elboim and Ron Peled. Limit distributions for euclidean random permutations. *Communications in Mathematical Physics*, 369(2):457–522, 2019.

[8] Dor Elboim and Allan Sly. Infinite cycles in the interchange process in five dimensions. *arXiv:2211.17023*, 2022.

[9] Richard Phillips Feynman. Atomic theory of the λ transition in helium. *Physical Review*, 91(6):1291, 1953.

[10] Alexey Gladkich and Ron Peled. On the cycle structure of mallows permutations. *The Annals of Probability*, 46(2):1114–1169, 2018.

[11] Christina Goldschmidt, Daniel Ueltschi, and Peter Windridge. Quantum Heisenberg models and their probabilistic representations. *Entropy and the quantum II, Contemp. Math*, 552:177–224, 2011.

[12] Alan Hammond. Sharp phase transition in the random stirring model on trees. *Probab. Theory Related Fields*, 161(3-4):429–448, 2015.

[13] Alan Hammond and Milind Hegde. Critical point for infinite cycles in a random loop model on trees. *The Annals of Applied Probability*, 29(4):2067–2088, 2019.

[14] Alan Hammond and Richard Kenyon. Monotone loop models and rational resonance. *Probab. Theory Related Fields*, 150(3-4):613–633, 2011.

[15] Dmitry Ioffe and Bálint Tóth. Split-and-merge in stationary random stirring on lattice torus. *Journal of Statistical Physics*, 180(1):630–653, 2020.

[16] J Kingman. The population structure associated with the Ewens sampling formula. *Theoretical Population Biology*, 11(2):274–283, 1977.

[17] James R Norris. *Markov Chains*. Number 2. Cambridge University Press, 1998.

[18] Valentin V. Petrov. *Limit theorems of probability theory*, volume 4 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York, 1995. Sequences of independent random variables, Oxford Science Publications.

[19] Oded Schramm. Compositions of random transpositions. *Israel Journal of Mathematics*, 147:221–243, 2005.

[20] Bálint Tóth. Improved lower bound on the thermodynamic pressure of the spin 1/2 Heisenberg ferromagnet. *Lett. Math. Phys.*, 28(1):75–84, 1993.

[21] Daniel Ueltschi. Random loop representations for quantum spin systems. *Journal of Mathematical Physics*, 54(8):083301, 2013.

[22] AM Vershik and AA Shmidt. Limit measures arising in the asympyotic theory of symmetric groups. i. *Theory of Probability & Its Applications*, 22(1):70–85, 1977.

[23] Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge University Press, 2018.

UC Berkeley

Durham University