Scaling limit of ballistic self-avoiding walk interacting with spatial random permutations[*]

Volker Betz[†]   Lorenzo Taggi[‡]

Abstract

We consider nearest neighbour spatial random permutations on $\mathbb{Z}^d$. In this case, the energy of the system is proportional to the sum of all cycle lengths, and the system can be interpreted as an ensemble of edge-weighted, mutually self-avoiding loops. The constant of proportionality, α, is the order parameter of the model. Our first result is that in a parameter regime of edge weights where it is known that a single self-avoiding loop is weakly space filling, long cycles of spatial random permutations are still exponentially unlikely. For our second result, we embed a self-avoiding walk into a background of spatial random permutations, and condition it to cover a macroscopic distance. For large values of α (where long cycles are very unlikely) we show that this walk collapses to a straight line in the scaling limit, and give bounds on the fluctuations that are almost sufficient for diffusive scaling. For proving our results, we develop the concepts of spatial strong Markov property and iterative sampling for spatial random permutations, which may be of independent interest. Among other things, we use them to show exponential decay of correlations for large values of α in great generality.

Keywords: self-avoiding walk; random spatial permutations.
AMS MSC 2010: 82B20; 82B41; 82B41.
Submitted to EJP on June 15, 2018, final version accepted on May 27, 2019.

1 Introduction

Self-avoiding random walks are by now a classical topic of modern probability theory, although many questions still remain to be answered; we refer to the classic book [34] and the more recent survey [38]. A variant of self-avoiding walks are self-avoiding

[†]Technische Universität Darmstadt, Germany E-mail: betz@mathematik.tu-darmstadt.de
[‡]Weierstrass Institut for Applied Analysis and Stochastics, Germany
 E-mail: taggi@wias-berlin.de

polygons (see e.g. [26]), where the last step of the self-avoiding walk has to come back to the point of origin.

Spatial random permutations, on the other hand, are a relatively recent concept. They were originally introduced due to their relevance for the theory of Bose-Einstein condensation [5, 7, 41], but are of independent mathematical interest. The purpose of the present paper is to view spatial random permutations as systems of mutually self-avoiding polygons and compare the behaviour of a selected cycle of a spatial random permutation to the one of a self-avoiding walk or polygon. Put differently, a selected cycle of a spatial random permutation can be viewed as a self-avoiding polygon embedded into a background of other self-avoiding polygons, and we are interested in the effect that this embedding has. For this purpose, we restrict to nearest neighbour spatial random permutations, as they are most closely related to self-avoiding walks. For a finite subset Λ_n of $\mathbb{Z}^d$, the model is defined on the space $\mathcal{S}_\Lambda$ consisting of all bijective maps π on Λ with the property that $|\pi(x) - x| \in \{0, 1\}$. The probability measure (with order parameter α) is given by assigning the element $\pi \in \mathcal{S}_\Lambda$ the energy $\mathcal{H}(\pi) := \sum_{x \in \Lambda} |\pi(x) - x|$, and the probability

$$\mathbb{P}_\Lambda(\{\pi\}) = \frac{1}{Z(\Lambda)} \, \mathrm{e}^{-\alpha \mathcal{H}(\pi)} \,,$$

where $Z(\Lambda)$ is the partition function (normalising constant).

The most important questions in spatial random permutations concern the length of their cycles, in particular the existence of macroscopic cycles. It is known [5] that, when α is large enough, the probability that the origin is in a cycle of length larger than k decays exponentially with k, uniformly in the volume Λ. It is expected that for dimensions $d = 3$ and higher, there exists a critical value α_c of the order parameter so that for $\alpha < \alpha_c$, the cycle containing the origin (or any selected point) is of macroscopic length with positive probability. In $d = 2$, on the other hand, only a Kosterlitz-Thouless phase transition is expected, meaning that the probability of the cycle being larger than k decays exponentially if $\alpha > \alpha_c$ but algebraically otherwise. While there is good numerical evidence for the existence of long cycles in $d \geq 3$ [21, 24] and the Kosterlitz-Thouless transition [2] in $d = 2$, actually proving any positive statement about the existence of long cycles is the great unsolved problem of the theory of spatial random permutations. The only case where such a statement is known is for an annealed version of the model (with a slightly different energy) [4, 6]. The argument there relies on explicit calculations using Fourier transforms; all attempts to get away from this exactly solvable situation have so far failed. In [7] a non-rigorous argument is made that the study of models of non-spatial permutations with cycle weights may be useful, and such models have received some attention recently [3, 8, 13, 14].

In the context of self-avoiding walks, the concept of a phase transition from short to long loops is present in the following results. Consider a single step-weighted self-avoiding walk, i.e. fix a sequence of growing subsets Λ_n (e.g. cubes) of $\mathbb{Z}^d$, for each n two points a and z at opposite ends of Λ_n, and consider the set of all self-avoiding walks starting in a and ending in z. Let $\alpha \in \mathbb{R}$, and assign to each such self-avoiding walk X the weight $\exp(-\alpha|X|)$, where $|X|$ is the number of steps that X takes. Write μ' for the connective constant of the d-dimensional cubic lattice, which is defined in (2.4). When $\alpha > \log \mu'$, it is known that the shape of X converges to a straight line as $n \to \infty$, when scaled by $1/n$. Actually, when scaled by $1/n$ in the direction of $a - z$ and by $n^{-1/2}$ in the directions perpendicular to this vector, X converges to a Brownian Bridge. These results are implicit in the works [12, 30] and have been worked out by Y. Kovchegov in his thesis [32].

For $\alpha < \log \mu'$, on the other hand, the results are entirely different. As Duminil-Copin, Kozma and Yadin have recently shown [15], in this case the rescaled self-avoiding walk
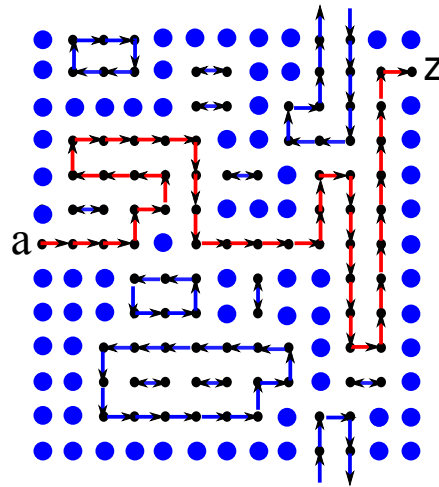
Figure 1: Representation of a bijection with a forced open cycle between a and z, when $\Lambda \subset \mathbb{Z}^2$ is a box with cylinder boundary conditions. If $\pi(x) = x$, then a circle is drawn at x, while if $\pi(x)$ is a neighbour of x, then an arrow is directed from x to $\pi(x)$.

becomes *weakly space filling*, meaning that it will only leave holes of logarithmic size in the graph. Their results also hold for the self-avoiding loop, i.e. in the case where a and z are chosen to be the same site. Note that for $\alpha > \log \mu'$, the self-avoiding *loop* will converge to a point in the scaling limit.

It is therefore a natural question what happens when the self-avoiding loop is embedded into a background of other self-avoiding loops, i.e. in the case of spatial random permutations. The proof of the existence of weakly space filling cycles would be particularly interesting, as this would imply that the expected length of the cycle is infinite, and thus a phase transition. Unfortunately, this is not what we can show. Instead, we give a somewhat negative result: we show that *if* there is a regime of space filling cycles, it must start at lower α than for the case of the self-avoiding polygon. More precisely, let μ be the *cyclic* connective constant of that graph, which is defined in (2.3), and let G be a finite sub-graph. We identify in Theorem 2.2 an $\alpha_0 < \log \mu$ so that for all $\alpha > \alpha_0$, and uniformly in the size of G, the length of a cycle through a given point has exponential tails. Thus in the interval $(\alpha_0, \log \mu)$ the single self-avoiding loop is weakly space filling while the self-avoiding loop embedded into an ensemble of other such loops is very short.

For our second result, we restrict to the case where $\Lambda_n = [0, n] \times [-n/2, n/2)^{d-1}$ and impose periodic boundary conditions on all except the first coordinate. We embed into the spatial random permutation a self-avoiding path starting in 0 and conditioned to end in a point z on the opposite side of Λ, see Figure 1. In this situation, we (almost) recover the results of [32], i.e. we show that for sufficiently large α, the self-avoiding walk starting in 0 collapses to a straight line in the scaling limit, as $n \to \infty$. Unlike in the case of the single self-avoiding walk, we do not have a good quantitative estimate on the threshold above which this behaviour holds, and we cannot quite control the fluctuations well enough to prove the convergence to a Brownian bridge. The reason is that the background of cycles introduces additional correlations that are very hard to control. This can be explained well by investigating the strategy of proof for the case of the self-avoiding walk, and discussing where it fails in our case.

The basic idea in the case of the self-avoiding walk (which we indeed adapt and extend) is to introduce regeneration points. The walk is forced to connect 0 to the other

side of the box; let us assume that it is the other side with respect to the first coordinate. A regeneration point is a point where the 'past' of the walk is entirely to the left (has smaller first coordinate) of that point, while the 'future' is entirely to the right. It is easy to see that if there are no regeneration points in an area of a given (horizontal) width w, then the self-avoiding path has to have at least $3w$ steps taking place in this area. Since every step that the walk takes introduces an additional factor of $e^{-\alpha}$ to its weight, and since α is large, it is possible to show that there must be many regeneration points, more precisely that the probability to not find any regeneration point in a vertical strip of width K steps decays exponentially in K, and this estimate holds uniformly in the size of the box Λ_n. On the other hand, regeneration points provide fresh starts for the walk (hence the name): when conditioned to start at a regeneration point, the resulting model is again a self-avoiding walk. The regeneration points themselves thus form random walk with iid steps, and by the exponential bound discussed above we have excellent control on the step size of that process. All scaling results now follow from standard limit theorems for random walks.

This argument breaks down at several places when going to spatial random permutations. First of all, the energy of the system is an extensive quantity for all values of α, i.e. it grows proportional to the volume. Since the energy of the embedded random walk only grows proportional to its length, in the case where it collapses to a straight line its energy will be a subdominant compared to energy provided by the environment. Therefore, it will be much harder to argue that a broad vertical strip with no regeneration points is unlikely just because the walk would need to take many steps. Secondly, even if we do have a regeneration point, just considering the area to the right of it will not decouple the past from the future. The reason is that with probability one (in the limit $n \to \infty$) a short cycle from the background will cross the vertical hyperplane containing the regeneration point, and introduce correlations.

We solve both problems by developing a method for estimating the decay of correlations for spatial random permutations. Since we expect this method to be of independent interest, and since it does not complicate matters much, we develop it for general graphs. The method is based on a strong version of the spatial Markov property, and on iterative sampling, and is strong enough to provide uniform mixing estimates in a rather general context. We then extend the concept of regeneration points and introduce (random) regeneration surfaces, that serve the same purpose as the (deterministic) hyperplanes that separate regions in the self-avoiding walk case. While our estimates are strong enough to give us the correct scaling limit, they are (just) not strong enough to get down to diffusive scaling. The main reason is that we can only show that consecutive regeneration points (and surfaces) have a distance of order $\log n$ with high probability, not a finite one as in the case of the self-avoiding walk.

If compared to cluster expansion ([23, 28, 29]), which is an alternative method for obtaining uniform mixing estimates in the perturbative regime, our *iterative sampling* technique has the advantage of being very simple, purely-probabilistic and non-entirely perturbative (more precisely, without the additional assumption (2.9) below, our uniform mixing estimate, Theorem 2.3, would just be a consequence of "exponentially small loops" and would thus be fully nonperturbative!). Nonperturbative results have been obtained for several models, for example for the Ising model [10], Bernoulli Percolation [11], self-avoiding walks [12, 30], and random cluster model [9], and in all these models they rely on the use of correlation inequalities of some kind. No correlation inequalities are known in spatial random permutations, and indeed we expect that finding such inequalities would have a significant impact on the subject area.

The significance of the model of spatial random permutations with a forced cycle goes far beyond the situation that we describe here. In [41] it is shown that (in a suitable

variant of spatial random permutations) the ratio of the partition functions of a system with a forced cycle and one without can be used to detect Bose-Einstein condensation: if this ratio stays positive uniformly in the volume and the spatial separation of the two endpoints of the forced cycle, this is equivalent to the presence of off-diagonal long range order [36], which itself is equivalent to Bose-Einstein condensation. Some progress has been made in understanding related models, e.g. the Heisenberg model and its connection to ensembles of mutually self- avoiding walks [40], or the random stirring models introduced by Harris [27] and further analysed in [1, 37, 31]. The big question about the existence of long loops, however, remains open.

On the probability side, our model is closely related to the connection is to the loop $O(n)$ model, which has been introduced in [18]. In this model a loop configuration ω is an undirected spanning subgraph of a graph G such that every vertex of this subgraph has degree zero or two. The weight of a loop configuration is proportional to $e^{-\alpha\, o(\omega)} n^{L(\omega)}$, where $o(\omega)$ is the number of edges in ω, $L(\omega)$ is the number of loops and n is a positive real. The case $n = 0$ corresponds formally to self-avoiding walk if one forces a path in the system in addition to the loops. If viewed as an ensemble of cycles, spatial random permutations are intimately related to the loop $O(n)$ model with $n = 2$, since each cycle of the permutation admits two possible orientations. The two models would be equivalent if in spatial random permutations cycles of length two were forbidden. On the hexagonal lattice, the loop $O(n)$ model has been conjectured to undergo a Kosterlitz-Thouless phase transition at the critical threshold $\log\left(\sqrt{2 + \sqrt{2 - n}}\right)$ when $n \leq 2$ [33]. This is compatible with our general finding that on every vertex-transitive graph the critical threshold of spatial random permutations, which corresponds more or less to the $n = 2$ case, is strictly less than the critical threshold for the self-avoiding walk, corresponding to the $n = 0$ case. Furthermore, it has been conjectured that only short cycles are observed at all values of α when $n > 2$. This has been rigorously proved only for n large enough in the article of Duminil-Copin, Peled, Samotij, and Spinka in [17], who also provide details on the structure of the typical configurations and provide evidence for the occurrence of a phase transition. Exploring the properties of the model at low values of n is of great interest and recently some progress has been made [16, 22, 39]. Most of the proofs that we present in this paper, can be reproduced for the loop $O(n)$ model for all values of n and α large enough, without the restriction of considering the hexagonal lattice (see Remark 3.2 below for further comments on this).

Our paper is organized as follows. We give precise definitions and state our results in Section 2. In Section 3 we prove our result on non-existence of long cycles, and provide various estimates on the partition functions over different domains which we will need later. In Section 4 we discuss the spatial Markov property and iterative sampling, and derive our results about exponential decay of correlations. In Section 5, we give the proof of our main result, Theorem 2.4, using the results of the previous sections.

2 Definitions and main results

We consider a finite simple graph $G = (V, E)$. A permutation on G is a bijective map $\pi : V \to V$ so that for all $x \in V$, either $\pi(x) = x$ or $\{x, \pi(x)\} \in E$. We write $\mathcal{S}_V$ for the set of all permutations on G, omitting the dependence on the edge set in the notation. Also, when $U \subset V$, we write $\mathcal{S}_U$ for the set of permutations on the subgraph (U, E_U) generated by U, i.e. where $E_U := \{(x, y) \in E : x \in U, y \in U\})$.

For a given $\pi \in \mathcal{S}_V$, we define its *energy* by

$$\mathcal{H}_V(\pi) = \sum_{x \in V} \mathbb{1}\{\pi(x) \neq x\}, \tag{2.1}$$

where the indicator above is 1 if $\pi(x) \neq x$ and 0 if $\pi(x) = x$. The probability of $\pi \in \mathcal{S}_V$ is

defined as,

$$\mathbb{P}_V(\pi) = \frac{\exp\left(-\alpha\,\mathcal{H}_V(\pi)\right)}{Z(V)}. \tag{2.2}$$

Above, $\alpha \in \mathbb{R}$ controls the preference of random permutations to have fixed points, and $Z(V)$ is called *partition function*.

Remark 2.1. The expression (2.1) is less general than it could be. By replacing the indicator function in that formula with edge weights $d : E \to E, \{x, y\} \mapsto d(x, y)$, we can generalize the definition of random permutations on graphs sufficiently so that classical cases including the quadratic jump penalization (see e.g. [2, 5]) are covered. While some of our results below (most notably the iterative sampling procedure) can be adapted to hold for this general situation, our main results compare the behavior of random permutations to the behavior of self-avoiding paths. Therefore, we prefer to stick with the narrower definition (2.1).

The most interesting objects in spatial random permutations are their cycles. For $\pi \in \mathcal{S}_V$ and $z \in V$, the *cycle of π containing z* is the directed graph on

$$\{ \pi^i(z) \in V \; : \; i \in \mathbb{N} \}$$

with edge set

$$\{ (\pi^i(z), \pi^{i+1}(z)) \in E \; : i \in \mathbb{N} \}.$$

We will denote it by $\gamma_z(\pi)$. We will regularly abuse the notation and use the symbol $\gamma_z(\pi)$ also to denote the vertex set of the cycle, viewed as a subset of V. We denote the total number of edges of $\gamma_z(\pi)$ by $\|\gamma_z\|$ and we refer to it as *length of γ_z*. In the special case where $\pi(z) = z$, $\gamma_z(\pi)$ has vertex set $\{z\}$ and empty edge set, therefore length 0. It is known [5] that there exists some $\alpha_0 > 0$ so that for all $\alpha > \alpha_0$, long cycles are exponentially unlikely. The first result of the present paper is to sharpen this statement by providing some information about the value of α_0. To state it, let us recall the definitions of connective constant and cyclic connective constant.

A *self-avoiding path* in a graph G is a finite directed subgraph (U, E') of G such that there is an enumeration $(x^1, x^2, \dots x^n)$ of U with the property that

$$E' = \{(x^i, x^{i+1}) : 1 \le i \le n - 1\}.$$

A *cycle* in G is a finite directed subgraph (U, E') of G such that there is an enumeration $(x^1, x^2, \dots x^n)$ of U with the property that

$$E' = \{(x^i, x^{i+1}) : 1 \le i \le n - 1\} \cup \{ (x^n, x^1) \}.$$

Note that by considering the graph (U, E') as directed, we give γ an orientation. For an infinite, vertex transitive graph ($\mathbb{Z}^d$ with the nearest neighbor edge structure being the most important example), we single out a vertex $0 \in V$ and call it the origin. We write SAW_n (respectively SAP_n) for the set of all self-avoiding paths (respectively cycles) starting from 0, with n edges. Thus, we have that $SAW_0 = SAP_0 = \{0\}$. Then the limits

$$\mu_G = \limsup_{n \to \infty} \sqrt[n]{|SAP_n|}, \tag{2.3}$$

and

$$\mu'_G = \lim_{n \to \infty} \sqrt[n]{|SAW_n|} \tag{2.4}$$

exist. For (2.4) this follows from a sub-additivity argument [25], while for (2.3) it follows from the fact that $|SAP_n| \le |SAW_n|$. The latter also immediately shows $\mu_G \le \mu'_G$ for all vertex-transitive graphs G. Hammersley [26] proved the remarkable fact that

$$\mu'_{\mathbb{Z}^d} = \mu_{\mathbb{Z}^d}.$$

We are now ready to state our first main result.

Scaling limit of ballistic self-avoiding walk interacting with spatial random permutations

Theorem 2.2. *Let G be any infinite vertex-transitive graph of bounded degree. Let α_0 be the unique solution of the equation*

$$\alpha + \tfrac{1}{2}\log(1 + \exp(-2\alpha)) = \log \mu_G. \tag{2.5}$$

Then for all $\alpha > \alpha_0$ there exist constants $C_0(\alpha), c_0(\alpha) > 0$ such that for any finite subgraph of G generated by $U \subset V$, for all $z \in U$, and for all $\ell \in \mathbb{N}$,

$$\mathbb{P}_U(\|\gamma_z\| > \ell) \leq C_0(\alpha) \exp\big(-c_0(\alpha)\,\ell\big). \tag{2.6}$$

Moreover, we can choose $C_0(\alpha)$ and $c_0(\alpha)$ so that $\lim_{\alpha \to \infty} c_0(\alpha)/\alpha = 1$ and $\limsup_{\alpha \to \infty} C_0(\alpha) < \infty$.

It is interesting to compare the above result with the findings of Duminil-Copin, Kozma and Yadin [15], who study self-avoiding walks on $\mathbb{Z}^d$ with an energy proportional to their length. To connect to the present paper, it is best to view them as random permutations conditional on not having any cycles except the one through the origin. In [15] it is shown that when $\alpha < \log \mu'_{\mathbb{Z}^d}$, the resulting self-avoiding cycle is *weakly space filling*, in particular its expected length is infinite. Theorem 2.2 shows that without the conditioning on having just one cycle, the situation is drastically different: since $\mu'_{\mathbb{Z}^d} = \mu_{\mathbb{Z}^d}$ and since the solution α_0 of (2.5) is strictly smaller than $\log \mu_{\mathbb{Z}^d}$, in the interval $(\alpha_0, \log \mu_{\mathbb{Z}^d})$ the lonely self-avoiding cycle has infinite expected length (as the relevant subgraph becomes large), while the length of the cycle through the origin in ordinary spatial random permutations has exponential tails.

Our second main result states that correlations decay exponentially fast or, more precisely, that the effect of removing a finite set of sites from the graph vanishes exponentially fast with the graph distance from that set. Let ξ be a random variable on the positive integers. We say that our random permutation model has *cycle length bounded by ξ* in a finite or infinite graph (V, E) if, uniformly in all finite sub-volumes $U \subset V$ and all points $x \in U$, the length of the cycle γ_x containing x is stochastically dominated by ξ. In other words, for all $\ell \geq 1$ we assume that

$$\sup_{\substack{U \subset V \\ |U| < \infty}} \sup_{x \in U} \mathbb{P}_U(|\gamma_x| \geq \ell) \leq P(\xi \geq \ell). \tag{2.7}$$

Moreover, let $B \subset V$ be a finite set. We define the sigma algebra

$$\mathcal{F}_B = \sigma\big(\{\pi : \pi(x) = y,\, \pi^{-1}(x) = z\} : x \in B,\, y, z \in V\big). \tag{2.8}$$

Put differently, the value $f(\pi)$ of an $\mathcal{F}_B$-measurable function is determined by the set of values $\pi(x)$ and $\pi^{-1}(x)$, $x \in B$. For a set $U \subset V$, let $\mathbb{E}_U$ be the expectation with respect to the measure $\mathbb{P}_U$.

Theorem 2.3. *Let $G = (V, E)$ be a finite graph, consider two subsets $B \subset U \subset V$. Assume that a random permutation on (V, E) has cycle length bounded by a random variable ξ which admits exponential moments and such that its expectation satisfies*

$$\mathbb{E}(\xi) < 2. \tag{2.9}$$

There exist two positive constants $C, \kappa < \infty$, which depend only on ξ, such that, for any $\mathcal{F}_B$-measurable function f, we have

$$|\mathbb{E}_V(f) - \mathbb{E}_U(f)| \leq 2C\|f\|_\infty \sum_{x \in V \setminus U} e^{-\kappa d(x, B)}. \tag{2.10}$$

Above, d is the graph distance on G.

For our third main result, we restrict our attention to cylindrical subgraphs of $\mathbb{Z}^d$. For $n \in \mathbb{N}$, let

$$\Lambda_n := [0, n] \times (-n/2, n/2]^{d-1} \cap \mathbb{Z}^d.$$

Elements of Λ_n will be written in the form $x = (\bar{x}, \hat{x})$ with $\bar{x} \in \mathbb{Z}$ and $\hat{x} \in \mathbb{Z}^{d-1}$. We impose cylindrical boundary conditions on $[-n/2, n/2] \cap \mathbb{Z}^{d-1}$ (but not on $[0, n] \cap \mathbb{Z}$), and edges are then between nearest neighbours in Λ_n. We will denote the resulting graph with the same symbol Λ_n if no confusion can arise. For a subset $A \subset \Lambda_n$, we will also write A for the subgraph of Λ_n induced by A, i.e, the graph that retains all of the edges where both endpoints lie in A. For $a, z \in A$, we define $\mathcal{S}_A^{a \to z}$ as the set of maps $\pi : A \to A$ with the properties

(i): π is a bijection from $A \setminus \{z\}$ to $A \setminus \{a\}$.

(ii): $\pi(z) = z$

(iii): $|\pi(x) - x| \leq 1$ for all $x \in A$.

It is easy to see that $\mathcal{S}_A^{a \to z} = \emptyset$ if the vertex a is disconnected from z in the graph A, and that $z = \lim_{n \to \infty} \pi^n(a)$ otherwise. For given $\pi \in \mathcal{S}_A^{a \to z}$, we will always write $\gamma(\pi) = \mathrm{Orb}_\pi(\{a\})$. We have $\pi(A) = A \setminus \{a\}$, and $\gamma(\pi)$ is the trace of a self-avoiding walk starting at a and ending at z that is embedded in π. Thus the probability measure $\mathbb{P}_A^{a \to z}$ defined through

$$\mathbb{P}_A^{a \to z}(\{\pi\}) = \frac{1}{Z^{a \to z}(A)} \exp\left(-\alpha \sum_{x \in A} |\pi(x) - x|\right), \qquad (\pi \in \mathcal{S}_A^{a \to z}), \qquad (2.11)$$

describes a step-weighted self-avoiding walk interacting with a background of spatial random permutations. We also define

$$\ell_j := \{x \in \Lambda_n : \bar{x} = j\}, \qquad \text{and} \quad \mathcal{S}_A^{a \to \ell_n} := \bigcup_{z \in \ell_n} \mathcal{S}_{A \cup \{z\}}^{a \to z}.$$

The probability measure on $\mathcal{S}_A^{a \to \ell_n}$ will be (2.11), except that the normalisation is now given by $Z^{a \to \ell_n}(A) = \sum_{z : \bar{z} = n} Z^{a \to z}(A)$. We can now state the main result of this paper.

Theorem 2.4. *There exists $\alpha_0 > 0$, $D < \infty$ and $N \in \mathbb{N}$ so that for all $M > 0$,*

$$\sup_{n \geq N, \alpha > \alpha_0} \mathbb{P}_{\Lambda_n}^{0 \to \ell_n}\left(\max\{|\hat{y}| : y \in \gamma\} > M\sqrt{n \log n}\right) < D/M.$$

Thus for large α, the self-avoiding walk embedded into π converges to a straight horizontal line, and the vertical aberration can be proved to be just a bit larger than $\sqrt{n}$. Indeed, we expect that the true vertical aberration is exactly of the order $\sqrt{n}$ and that γ converges to a Brownian motion under diffusive scaling. This is known in the case of a self-avoiding walk in the infinite space without a background of spatial random permutations [32]. In our situation, the strong correlations prevent us from getting a presumably sharp upper bound on the fluctuations, and indeed also prevent a useful lower bound. We will comment on the places where we lose the necessary accuracy for diffusive behaviour at the end of the proof of Theorem 2.4.

3 Cycle length and partition function

In this section we prove Theorem 2.2 and provide further estimates comparing partition functions for different subsets of V. Our Theorem 2.2 treats only the case of vertex transitive graphs since we want quantitative estimates involving the connective

constant; but it is not difficult to modify our proof so that one can treat general graphs, including those with edge weights. In the latter case, some modifications will be necessary, as the graph distance is no longer a good quantity to measure the distance between sets. We do not pursue this any further in the present paper.

Recall that for a self-avoiding path or cycle γ, $\|\gamma\|$ denotes the number of its edges and that $|\gamma|$ denotes the number of its vertices. Our first comparison is

Proposition 3.1. *For any finite simple graph $G = (V, E)$, for any self-avoiding path or cycle $\gamma \subset G$, we have that*

$$\frac{Z(V \setminus \gamma)}{Z(V)} \leq \left(\frac{1}{1 + e^{-2\alpha}} \right)^{\frac{\|\gamma\|}{2}}. \tag{3.1}$$

Proof. Let $\gamma \subset V$ be a self-avoiding path or a cycle. Let us denote by $\{x^0, x^1, \ldots, x^{|\gamma|-1}\}$ the sequence of sites of γ, ordered such that (x^i, x^{i+1}) is an edge of γ. We let M be the largest even number such that $M \leq |\gamma|$. We claim that

$$Z(\gamma) \geq (1 + e^{-2\alpha})^{\frac{M}{2}}, \tag{3.2}$$

where we recall our abuse of notation: γ denotes the sites occupied by the cycle γ, and $Z(\gamma)$ is the partition function of the subgraph generated by those sites. To see (3.2), we partition γ into pairs (x^0, x^1), (x^2, x^3), ..., (x^{M-2}, x^{M-1}). Let $\hat{\gamma}$ be the graph obtained from γ by keeping only edges connecting vertices of the pair (x^i, x^{i+1}), for even $i \in [0, M-1]$, and removing all the other edges. Clearly, $Z(\gamma) \geq Z(\hat{\gamma})$, since $\hat{\gamma}$ is a subgraph of γ. Since the graph $\hat{\gamma}$ is composed of $\frac{M}{2}$ disjoint subgraphs containing two vertices connected by an edge each and since the contribution of each of these subgraphs is $(1 + e^{-2\alpha})$, (3.2) follows.

Now note that, by Proposition 4.1 (i),

$$Z(V) \geq Z(A) Z(V \setminus A). \tag{3.3}$$

Note also that if γ is a cycle, then $|\gamma| = \|\gamma\|$ is even and we can set $M = \|\gamma\|$, while if γ is a self-avoiding path, then $M \geq |\gamma| - 1 = \|\gamma\|$. Thus, we have that if γ is a self-avoiding path or a cycle, then

$$\frac{Z(V \setminus \gamma)}{Z(V)} \leq \frac{1}{Z(\gamma)} \leq \left(\frac{1}{1 + e^{-2\alpha}} \right)^{\frac{\|\gamma\|}{2}}. \qquad \square$$

The estimate of Proposition 3.1 is logarithmically sharp if the cycle γ is "stretched out" in the sense that the subgraph generated by the sites of γ contains no further edges beyond those of γ. If γ is "curly", meaning that many points in the relevant subgraph are connected by more than two edges, one could use these edges in order to get better lower bounds on $Z(A)$ and thus better upper bounds on the ratio $Z(V \setminus \gamma)/Z(V)$. The associated combinatorics do not look easy even in the case of $V \subset \mathbb{Z}^2$, though.

We are now ready to prove our first main theorem.

Proof of Theorem 2.2. Let G be an infinite vertex-transitive simple graph of bounded degree, let U be a finite subset of V. For $\pi \in \mathcal{S}_U$, $x \in U$, and a cycle $\tilde{\gamma} \subset U$, we have

$$\mathbb{P}_U(\gamma_x = \tilde{\gamma}) = e^{-\alpha \|\tilde{\gamma}\|} \frac{Z(U \setminus \tilde{\gamma})}{Z(U)}, \tag{3.4}$$

Let μ_G be the cyclic connective constant of G. The definition of cyclic connective constant (2.3) implies that for every $\delta > 0$ there exists $C_\delta > 0$ such that for any $n \in \mathbb{N}$,

$$|SAP_n| \leq C_\delta (\mu_G + \delta)^n.$$

By vertex transitivity, the same bound holds for $SAP_n(x)$, the set of self-avoiding cycles of length n starting in x. By Proposition 3.1, we then have that for all $\ell \in \mathbb{N}^+$,

$$\mathbb{P}_U(\|\gamma_x\| \geq \ell) = \sum_{n=\ell}^{\infty} \sum_{\tilde{\gamma} \in SAP_n(x) \cap U} \exp\big(-\alpha n\big) \frac{Z(U \setminus \tilde{\gamma})}{Z(U)} \tag{3.5}$$

$$\leq C_\delta \sum_{n=\ell}^{\infty} \exp\Big(-n\Big(\alpha + \frac{1}{2}\log\big(1 + e^{-2\alpha}\big) - \log\big(\mu_G + \delta\big)\Big)\Big) \tag{3.6}$$

Recall that α_0 is defined as the unique solution of (2.5) and that therefore it satisfies $0 < \alpha_0 < \log(\mu_G)$. For each $\alpha > \alpha_0$, we can find $\delta(\alpha, G) > 0$ small enough such that

$$c_0(\alpha) := \alpha + \frac{1}{2}\log\big(1 + e^{-2\alpha}\big) - \log\big(\mu_G + \delta(\alpha, G)\big) > 0. \tag{3.7}$$

Thus

$$\mathbb{P}_U(\|\gamma_x\| \geq \ell) \leq \frac{C_{\delta(\alpha, G)}}{1 - e^{-c_0(\alpha)}} e^{-\ell c_0(\alpha)}.$$

It is not difficult to see that for all large enough α, we can choose $\delta(\alpha, G) = 1$, and that then $\lim_{\alpha \to \infty} c_0(\alpha)/\alpha = 1$. This concludes the proof of the theorem. $\qquad\square$

Remark 3.2. Proposition 3.1 is the only point of this paper where the existence of cycles of length 2 is necessary. For the loop $O(2)$ model [17], corresponding to random permutations without cycles of length 2, the exponential bound (2.6) of Theorem 2.2 would still hold true, but only for $\alpha > \log \mu$. Exponential decay of cycle length for the loop $O(n)$ model has been proved in the paper [17] for any value of α when n is large.

The next proposition uses similar ideas as in the previous proof in order to give a complementary bound to the one of Proposition 3.1.

Proposition 3.3. *Let G be an infinite vertex-transitive simple graph, let α_0 be the unique solution of equation (2.5). For any $\alpha > \alpha_0$ there exists $c_1(\alpha) > 0$ such that for any finite $U \subset V$, and for all $A \subset U$,*

$$\frac{Z(U \setminus A)}{Z(U)} \geq c_1(\alpha)^{|A|}. \tag{3.8}$$

Moreover $c_1(\alpha)$ can be chosen such that $\lim_{\alpha \to \infty} c_1(\alpha) = 1$.

Proof. Fix $x \in U$. Using the notation as in the proof of Proposition 3.1, we get

$$Z(U) = \sum_{n=0}^{\infty} \sum_{\substack{\tilde{\gamma} \in SAP_n(x): \\ \tilde{\gamma} \subset U}} \sum_{\substack{\pi \in \mathcal{S}_U: \\ \gamma_x(\pi) = \tilde{\gamma}}} e^{-\alpha \mathcal{H}_U(\pi)} = Z(U \setminus \{x\}) \sum_{n=0}^{\infty} \sum_{\substack{\tilde{\gamma} \in SAP_n(x): \\ \tilde{\gamma} \subset U}} e^{-\alpha \|\tilde{\gamma}\|} \frac{Z(U \setminus \tilde{\gamma})}{Z(U \setminus \{x\})}$$

With $\tilde{\gamma} \in SAP_n(x)$ such that $\tilde{\gamma} \subset U$ for $n \geq 2$, we apply Proposition 3.1 to the self-avoiding path $\tilde{\gamma} \setminus \{x\}$ (of length $n-2$), and the graph $U \setminus \{x\}$, giving

$$\frac{Z(U \setminus \tilde{\gamma})}{Z(U \setminus \{x\})} \leq (1 + e^{-2\alpha})^{-(n-2)/2} \quad \text{for all } \tilde{\gamma} \in SAP_n(x), \ n \geq 2.$$

Thus,

$$\frac{Z(U)}{Z(U \setminus \{x\})} \leq 1 + (1 + e^{-2\alpha}) \sum_{n=2}^{\infty} |SAP_n| e^{-n\alpha} (1 + e^{-2\alpha})^{-n/2}$$

$$\leq 1 + \frac{(1 + e^{-2\alpha})C_{\delta(\alpha, G)}}{1 - e^{-c_0(\alpha)}} e^{-2c_0(\alpha)}.$$

Above, the constants are the same as in Theorem 2.2. By our knowledge of $C_{\delta(\alpha,G)}$ and $c_0(\alpha)$, clearly the right hand side of the above equation converges to one as $\alpha \to \infty$. Thus we have shown the claim for $A = \{x\}$, with

$$c_1(\alpha) = \left(1 + \frac{(1 + e^{-2\alpha})C_{\delta(\alpha,G)}}{1 - e^{-c_0(\alpha)}} e^{-2c_0(\alpha)}\right)^{-1}. \tag{3.9}$$

For general A, the claim follows from the telescopic product

$$\frac{Z(U \setminus A)}{Z(U)} = \frac{Z(U \setminus A)}{Z(U \setminus (A \setminus \{x^1\}))} \frac{Z((U \setminus (A \setminus \{x^1\})))}{Z(U \setminus (A \setminus \{x^1, x^2\}))} \cdots \frac{Z(U \setminus \{x^{|A|}\})}{Z(U)} \geq c_1(\alpha)^{|A|}. \quad \square$$

4 Iterative sampling and the spatial Markov property

Spatial random permutations have a fundamental spatial Markov property, which we will first discuss in a relatively simple form. Then we introduce a sort of strong Markov property, we present iterative sampling, which is a basic new technique that enables many of our proofs, and finally we prove Theorem 2.3, which is stated in an equivalent form as Theorem 2.3.

Let $G = (V, E)$ be a finite graph. Recall the definition of sigma algebra (2.8). For $\pi \in S_V$, we let

$$\mathrm{Inv}(\pi) := \{A \subset V : \pi(A) = A\}$$

be the family of π-invariant sets. The Markov property, statement (iii) in the theorem below, says that conditional on $A \in \mathrm{Inv}$, an $\mathcal{F}_A$ measurable random variable and an $\mathcal{F}_{A^c}$-measurable random variable are independent. In order to state it correctly, let us recall that $\mathbb{P}_A$ denotes the probability measure (2.2) on the subgraph generated by A, and write $\mathbb{E}_A$ for the corresponding expectation. For $\pi \in S_A$ we write $\pi \oplus \mathrm{id}$ for the permutation in S_V that is obtained by setting $\pi(x) = x$ for all $x \notin A$.

Proposition 4.1 (**Spatial Markov property**). *Let $A \subset V$, $A \neq \emptyset$. Then,*

(i): $\mathbb{P}_V(A \in \mathrm{Inv}) = Z(A)Z(A^c)/Z(V)$.

(ii): For $\mathcal{F}_A$-measurable f, we have $\mathbb{E}_V(f|A \in \mathrm{Inv}) = \mathbb{E}_A(f(. \oplus \mathrm{id}))$.

(iii): For $\mathcal{F}_A$-measurable f and $\mathcal{F}_{A^c}$-measurable g, we have

$$\mathbb{E}_V(f\,g|A \in \mathrm{Inv}) = \mathbb{E}_V(f|A \in \mathrm{Inv})\,\mathbb{E}_V(g|A \in \mathrm{Inv}) = \mathbb{E}_A(f(. \oplus \mathrm{id}))\,\mathbb{E}_{A^c}(g(. \oplus \mathrm{id})).$$

(iv): For $\mathcal{F}_A$-measurable f and $\mathcal{B} \in \mathcal{F}_{A^c}$, we have

$$\mathbb{E}_V(f|A \in \mathrm{Inv}, \mathcal{B}) = \mathbb{E}_A(f). \tag{4.1}$$

Proof. By the additivity of $\mathcal{H}_V$, we have for $\mathcal{F}_A$-measurable f and $\mathcal{F}_{A^c}$-measurable g that

$$\sum_{\pi \in S_V : \pi(A) = A} f(\pi)g(\pi)\,e^{-\alpha\mathcal{H}_V(\pi)} = \sum_{\pi \in S_A} f(\pi \oplus \mathrm{id})\,e^{-\alpha\mathcal{H}_A(\pi)} \sum_{\tilde\pi \in S_{A^c}} g(\tilde\pi \oplus \mathrm{id})\,e^{-\alpha\mathcal{H}_{A^c}(\tilde\pi)}. \tag{4.2}$$

Inserting $f = 1$ and $g = 1$ into (4.2) and dividing by $Z(V)$ gives (i). Setting $g = 1$ in (4.2) and dividing by $Z(V)$ and using (i) gives $\mathbb{E}_V(f\mathbb{1}\{A \in \mathrm{Inv}\}) = \mathbb{E}_A(f)\mathbb{P}(A \in \mathrm{Inv})$, and thus (ii). For (iii), note that dividing (4.2) by $Z(A)Z(A^c)$ gives $\mathbb{E}_V(f\,g|A \in \mathrm{Inv})$ on the left hand side by using (i), while on the right hand side it gives $\mathbb{E}_A(f(. \oplus \mathrm{id}))\,\mathbb{E}_{A^c}(g(. \oplus \mathrm{id}))$. The remaining equality follows from (ii). For (iv), it suffices to use the equality $\mathbb{E}_V(f|A \in \mathrm{Inv}, \mathcal{B}) = \mathbb{E}_V(f\mathbb{1}\{\mathcal{B}\}|A \in \mathrm{Inv})/\mathbb{P}_V(B|A \in \mathrm{Inv})$, the fact that $A \in \mathrm{Inv}$ is equivalent to $A^c \in \mathrm{Inv}$, and (iii). $\square$

Following ideas from the theory of Markov chains, we can also formulate a strong Markov property for random permutations. A set-valued random variable $Q : S_V \to \mathcal{P}(V)$ will be called an *admissible random invariant set* if $Q(\pi) \in \mathrm{Inv}(\pi)$ for all $\pi \in S_V$, and if the event $\{Q = A\}$ is $\mathcal{F}_A$-measurable for all $A \subset V$. Moreover, we define the sigma algebra

$$\mathcal{F}_Q = \{R \in \mathcal{P}(S_V) : R \cap \{Q = A\} \in \mathcal{F}_A \text{ for all } A \subset V\}.$$

We then have

Proposition 4.2 (Strong Markov property). *Let Q be an admissible random invariant set. For $B \subset V$ and $\mathcal{F}_B$-measurable $f : S_V \to \mathbb{R}$ we have*

$$\mathbb{E}_V(f|\mathcal{F}_Q)\mathbb{1}\{Q \cap B = \emptyset\} = \mathbb{E}_V(f\mathbb{1}\{Q \cap B = \emptyset\}|\mathcal{F}_Q) = \mathbb{E}_{Q^c}\big(f(. \oplus \mathrm{id})\big)\mathbb{1}\{Q \cap B = \emptyset\}. \quad (4.3)$$

Proof. Since Q is admissible, it is $\mathcal{F}_Q$-measurable, and so the right hand side above is also $\mathcal{F}_Q$-measurable. Moreover, for a $\mathcal{F}_Q$-measurable random variable g, we have

$$\mathbb{E}_V\Big(\mathbb{E}_{Q^c}(f(. \oplus \mathrm{id}))\mathbb{1}\{Q \cap B = \emptyset\}g\Big) = \sum_{A : A \cap B = \emptyset} \mathbb{E}_V\Big(\mathbb{E}_{A^c}(f(. \oplus \mathrm{id}))\mathbb{1}\{Q = A\}g\Big)$$

$$= \sum_{A : A \cap B = \emptyset} \mathbb{E}_{A^c}\Big(f(. \oplus \mathrm{id})\Big)\,\mathbb{E}_V\Big(\mathbb{1}\{Q = A\}g\Big|A \in \mathrm{Inv}\Big)\,\mathbb{P}(A \in \mathrm{Inv})$$

$$= \sum_{A : A \cap B = \emptyset} \mathbb{E}_{A^c}\Big(f(. \oplus \mathrm{id})\Big)\,\mathbb{E}_A\Big(\mathbb{1}\{Q = A\}g(. \oplus \mathrm{id})\Big)\mathbb{P}(A \in \mathrm{Inv})$$

$$= \sum_{A : A \cap B = \emptyset} \mathbb{E}_V\Big(fg\mathbb{1}\{Q = A\}\Big) = \mathbb{E}_V\Big(fg\mathbb{1}\{Q \cap B = \emptyset\}\Big).$$

Above, in the first equality we sum over the realizations of Q, which are denoted by A, the second equality is true since $Q = A$ implies $A \in \mathrm{Inv}$ and the third equality holds by Proposition 4.1 (ii) and the fact that $\mathbb{1}\{Q = A\}g$ is $\mathcal{F}_A$-measurable for each $\mathcal{F}_Q$-measurable g. The fourth equality is due to Proposition 4.1 (iii), which is applicable since f is $\mathcal{F}_B$-measurable and thus in particular $\mathcal{F}_{A^c}$-measurable as $A \cap B = \emptyset$. Since the previous identities hold for any $\mathcal{F}_Q$-measurable random variable g, we obtain the second identity in (4.3). Since the indicator $\mathbb{1}\{Q \cap B = \emptyset\}$ is $\mathcal{F}_Q$-measurable, the first identity in (4.3) holds. The claim is thus shown. $\qquad\square$

We demonstrate the usefulness of the strong Markov property by applying it to the question of decay of dependence on the boundary conditions of spatial permutations. Let $G = (V, E)$ be a (large) graph, think of $V = \mathbb{Z}^d \cap [-N, N]^d$ with the nearest neighbour edge structure. Let $B \subset U \subset V$, where we think of U as being large and differing from V only "near the boundary", and of B as being "near the center" of V. We are interested in the difference between $\mathbb{E}_V(f)$ and $\mathbb{E}_U(f)$ in cases where the graph distance between U^c and B becomes large and f is $\mathcal{F}_B$-measurable. In other words, we are interested in how much the precise shape of the graph at the boundary influences the expectation of local random variables. This is a classical question in any theory related to Gibbs measures. Our answer to this question is the crude, but useful estimate that is provided in the next proposition.

Let $\tilde{G}$ be the disjoint union of two copies of G. Let $\tilde{V} = V_1 \cup V_2$ be its vertex set, where V_1 and V_2 are disjoint copies of V. Let $\phi : \tilde{V} \to \tilde{V}$ be the natural symmetry on $\tilde{V}$, i.e. the map such that each $x \in V_1$ is mapped to the vertex in V_2 that corresponds to x when V_1 and V_2 are identified with V. For $\tilde{\pi} \in S_{\tilde{V}}$, let $Q_A(\tilde{\pi})$ be the minimal $\tilde{\pi}$-invariant set that contains A and is *compatible with* ϕ. The latter means that $\phi(Q_A) = Q_A$. Since the intersection of two $\tilde{\pi}$-invariant, ϕ-compatible sets containing A retains these properties, and since $\tilde{V}$ has them, it is clear that such a minimal set exists.

Proposition 4.3. *Under the assumptions made above,*

$$|\mathbb{E}_U(f) - \mathbb{E}_V(f)| \leq 2\|f\|_\infty \mathbb{P}_{\tilde{V}}\big(Q_A \cap B \neq \emptyset \,\big|\, \tilde{\pi}|_A = \mathrm{id}\big). \tag{4.4}$$

Proof. As before, the superscript ˜ will refer to objects that are defined in $\tilde{G}$, the disjoint union of two copies of G. Any graph permutation $\tilde{\pi}$ on $\tilde{V}$ can be written as $\pi_1 \oplus \pi_2$ with π_i a permutation on V_i. For a $\mathcal{F}_B$-measurable random variable f on S_V, we define $f_1(\tilde{\pi}) = f(\pi_1)$ and $f_2(\tilde{\pi}) = f(\pi_2)$. We also define $A = V_1 \setminus U_1$, where U_1 is U regarded as a subset of V_1. In particular, $A \cap V_2 = \emptyset$. Then

$$\mathbb{E}_U(f) = \mathbb{E}_{\tilde{V}}(f_1 \mid \{\tilde{\pi}|_A = \mathrm{id}\}), \tag{4.5}$$

and

$$\mathbb{E}_V(f) = \mathbb{E}_{\tilde{V}}(f_2) = \mathbb{E}_{\tilde{V}}(f_2 \mid \{\tilde{\pi}|_A = \mathrm{id}\}).$$

From the definition $Q_A(\tilde{\pi})$ we have that $Q_A(\tilde{\pi}) = D$ if and only if
(i): $D \supset A$, D is $\tilde{\pi}$-invariant and ϕ-compatible;
(ii): Every strict subset of D fails to have at least one of the three properties.
So $\{Q_A = D\}$ is $\mathcal{F}_D$-measurable, and therefore Q_A is an admissible random invariant set. Now (4.5) gives,

$$\mathbb{E}_{\tilde{V}}(\mathbb{1}\{\tilde{\pi}|_A = \mathrm{id}\})\mathbb{E}_U(f) = \mathbb{E}_{\tilde{V}}(f_1\mathbb{1}\{\tilde{\pi}|_A = \mathrm{id}\}) =$$
$$\mathbb{E}_{\tilde{V}}\big(f_1\mathbb{1}\{\tilde{\pi}|_A = \mathrm{id}\}\mathbb{1}\{Q_A \cap B = \emptyset\}\big) + \mathbb{E}_{\tilde{V}}\big(f_1\mathbb{1}\{\tilde{\pi}|_A = \mathrm{id}\}\mathbb{1}\{Q_A \cap B \neq \emptyset\}\big), \tag{4.6}$$

where B is regarded as a subset of V_1. The first term of the right hand side is equal to

$$|\mathbb{E}_{\tilde{V}}\Big(\mathbb{E}_{\tilde{V}}(f_1\mathbb{1}\{\tilde{\pi}|_A = \mathrm{id}\}\mathbb{1}\{Q_A \cap B = \emptyset\}\big|\mathcal{F}_{Q_A}\big)\Big)$$
$$= \mathbb{E}_{\tilde{V}}\Big(\mathbb{E}_{Q^c}\big(f_1(. \oplus id)\big)\mathbb{1}\{\tilde{\pi}|_A = \mathrm{id}\}\mathbb{1}\{Q_A \cap B = \emptyset\}\Big), \tag{4.7}$$

since $\mathbb{1}\{\tilde{\pi}|_A = \mathrm{id}\}$ and $\mathbb{1}\{Q_A \cap B = \emptyset\}$ are both $\mathcal{F}_{Q_A}$-measurable, and thus Proposition 4.2 applies. But by the symmetry of Q_A, we have $\mathbb{E}_{Q_A}(f_1) = \mathbb{E}_{Q_A}(f_2)$. Thus we can do the same calculation with f_2 instead of f_1, and subtract the results. If in addition we divide by $\mathbb{E}_{\tilde{V}}(\mathbb{1}\{\tilde{\pi}|_A = \mathrm{id}\})$ we arrive at

$$\mathbb{E}_U(f) - \mathbb{E}_V(f) = \mathbb{E}_{\tilde{V}}\Big((f_1 - f_2)\mathbb{1}\{Q_A \cap B \neq \emptyset\}\big|\{\tilde{\pi}|_A = \mathrm{id}\}\Big), \tag{4.8}$$

which implies (4.4) and concludes the proof of the proposition. $\qquad\square$

When we want to compare two different subsets U_1 and U_2 of V, we could slightly adapt the definition of A above to get a similar estimate. In most cases, simply using the triangle inequality on (4.4) is enough.

In order to estimate the right-hand side of (4.4), we now need a tool to show the existence of invariant subsets with certain prescribed symmetries, and containing certain prescribed subsets of V. This tool is iterative sampling. Iterative sampling is a procedure to build a $\mathbb{P}_V$-distributed random variable step by step, reminiscent of the way that the full path of a Markov chain can be sampled step by step. In words, what we do is the following: we first sample a random permutation from $\mathbb{P}_V$, but keep only the cycles intersecting a (possibly random) set K_1. The only restriction is that K_1 may not depend on the permutation we just sampled, it has to be chosen independently. We end up with a set D_1 where we have kept the cycles, and a set B_1 where we have discarded them, where D_1 and B_1 are disjoint and their union is V. We then resample the permutation, independently, inside the subgraph generated by B_1. Again, we only keep the cycles intersecting some set K_2 that may depend on what has happened before and some external randomness, but not on the permutation inside B_1. We carry on until we exhaust V. Formally, the definition is as follows.

Definition 4.4. *A sampling strategy* for spatial random permutations on a finite graph *$G = (V, E)$ consists of two families of random variables $(\sigma_A)_{A \subset V}$ and $(K_A)_{A \subset V}$ such that*

(i): *for all A, σ_A is a $\mathbb{P}_A$-distributed random variable, in particular has values in S_A. $(\sigma_A)_{A \subset V}$ is a family of independent random variables.*

(ii): *K_A takes values in the power set $\mathcal{P}(A)$ for all A, and $\mathbb{P}(K_A = \emptyset) = 0$ when $A \neq \emptyset$.*

(iii): *For each $A \subseteq V$, the random variable K_A is independent of the family $(\sigma_B)_{B \subseteq A}$.*

Note in particular that we require that K_A is independent of σ_A. Given a sampling strategy, we define a *recursive sampling procedure* as follows: We set $B_1 = V$, and for $i \geq 1$ we set

$$\pi_i = \sigma_{B_i}, \quad D_i = \mathrm{Or}_{\pi_i}(K_{B_i}), \quad B_{i+1} = B_i \setminus D_i. \tag{4.9}$$

Above $\mathrm{Or}_\pi(A) = \{\pi^j(x) : x \in A, j \in \mathbb{N}\}$ denotes the orbit of A under π. This way we end up with a sequence (B_i, D_i, π_i) where $D_i \in \mathrm{Inv}(\pi_i)$, and where the D_i form a partition of V. It follows that the map

$$\boldsymbol{\pi} : V \to V, \quad \boldsymbol{\pi}(x) = \pi_i(x) \text{ whenever } x \in D_i \tag{4.10}$$

is a random element of S_V.

Lemma 4.5. *For any sampling strategy, the distribution of $\boldsymbol{\pi}$ is $\mathbb{P}_V$.*

Proof. We prove the lemma by induction on the number of sets K_i. Let us call $(K_A)_{A \subseteq V}$ an n-step sampling strategy if $K_A(\omega) = A$ for all ω in the underlying probability space such that $A = B_{n+1}(\omega)$. Since the event $\{A = B_{n+1}\}$ is independent of $(\sigma_{\tilde{A}})_{\tilde{A} \subseteq A}$, any sampling strategy can be turned into an n-step sampling strategy, and the union over the n-step sampling strategies for all n gives all sampling strategies.

For a one-step sampling strategy, we have that $K_A = A$ for all $A \neq V$. Putting $D_A(\bar{\pi}) = \mathrm{Or}_{\bar{\pi}}(A)$, we compute

$$\mathbb{P}(\boldsymbol{\pi} = \bar{\pi}) = \sum_{A \subset V} \mathbb{P}(\boldsymbol{\pi} = \bar{\pi}, K_V = A)$$

$$= \sum_{A \subset V} \mathbb{P}(\sigma_V = \bar{\pi}|_{D_A(\bar{\pi})}, \sigma_{V \setminus D_A(\bar{\pi})} = \bar{\pi}|_{V \setminus D_A(\bar{\pi})}, K_V = A). \tag{4.11}$$

For fixed A, all three events in the probability above are independent and moreover

$$\mathbb{P}(\sigma_V = \bar{\pi}|_{D_A(\bar{\pi})}) = \frac{1}{Z(V)} \, \mathrm{e}^{-\alpha \mathcal{H}_{D_A(\bar{\pi})}(\bar{\pi})} \, Z\big(V \setminus D_A(\bar{\pi})\big)$$

and

$$\mathbb{P}(\sigma_{V \setminus D_A(\bar{\pi})} = \bar{\pi}|_{V \setminus D_A(\bar{\pi})}) = \frac{1}{Z\big(V \setminus D_A(\bar{\pi})\big)} \, \mathrm{e}^{-\alpha \mathcal{H}_{V \setminus D_A(\bar{\pi})}(\bar{\pi})} \, .$$

Since $\mathcal{H}_{D_A(\bar{\pi})}(\bar{\pi}) + \mathcal{H}_{V \setminus D_A(\bar{\pi})}(\bar{\pi}) = \mathcal{H}_V(\bar{\pi})$, the product of the two terms above does not depend on A and is equal to $\mathbb{P}_V(\bar{\pi})$, and summing (4.11) over A gives the result in the case of one-step sampling strategies.

Now assume that the claim holds for all n-step sampling strategies up to some $n \in \mathbb{N}$. Let $(\sigma_A, K_A)_{A \subset V}$ be an $n + 1$-step sampling strategy. For fixed $\bar{\pi} \in S_V$ and $\bar{K} \subset V$ with $\mathbb{P}(K_V = \bar{K}, \sigma_V = \bar{\pi}) > 0$ we define the conditional measure

$$\mathbb{P}^{\bar{\pi}, \bar{K}} = \mathbb{P}(.|K_V = \bar{K}, \sigma_V = \bar{\pi}).$$

We write $\bar{D} = \mathrm{Or}_{\bar{\pi}}(\bar{K})$ and $\bar{B} = V \setminus \bar{D}$, and claim that under the measure $\mathbb{P}^{\bar{\pi}, \bar{K}}$, the families $(\sigma_A)_{A \subset \bar{B}}$ and $(K_A)_{A \subset \bar{B}}$ form a sampling strategy for random permutations on

the graph generated by $\bar{B}$. To check this claim, note that the event $\{K_V = \bar{K}, \sigma_V = \bar{\pi}\}$ is independent of the family $(\sigma_A)_{A \subset \bar{B}}$, and thus for all choices of $\pi_A \in S_A$,

$$\mathbb{P}^{\bar{\pi}, \bar{K}}(\sigma_A = \pi_A \ \forall A \subset \bar{B}) = \mathbb{P}(\sigma_A = \pi_A \ \forall A \subset \bar{B}). \tag{4.12}$$

This shows independence of the family $(\sigma_A)_{A \subset \bar{B}}$. Furthermore, for $A \subset \bar{B}$, $\kappa \subset A$, and $A_1, \ldots A_m \subseteq A$, we have

$$\mathbb{P}^{\bar{\pi}, \bar{K}}(K_A = \kappa, \sigma_{A_i} = \pi_{A_i} \ \forall i)$$
$$= \frac{\mathbb{P}(K_A = \kappa, K_V = \bar{K}, \sigma_V = \bar{\pi}, \sigma_{A_i} = \pi_i \ \forall i)}{\mathbb{P}(K_V = \bar{K}, \sigma_V = \bar{\pi})}$$
$$= \mathbb{P}^{\bar{\pi}, \bar{K}}(K_A = \kappa) \prod_{i=1}^{m} \mathbb{P}(\sigma_{A_i} = \pi_i). \tag{4.13}$$

Thus using (4.12) from right to left, we see that K_A is independent of the family $(\sigma_{\tilde{A}})_{\tilde{A} \subset A}$ under $\mathbb{P}^{\bar{\pi}, \bar{K}}$, and thus we have indeed a sampling strategy, which in addition clearly is an $m - 1$-step sampling strategy. Now,

$$\mathbb{P}(\boldsymbol{\pi} = \bar{\pi}) = \sum_{\bar{K} \subset V} \sum_{\bar{\sigma} \in S_V : \bar{\sigma}|_{\bar{D}} = \bar{\pi}|_{\bar{D}}} \mathbb{P}(\boldsymbol{\pi}|_{\bar{B}} = \bar{\pi}|_{\bar{B}}, K_V = \bar{K}, \sigma_V = \bar{\sigma}) =$$
$$= \sum_{\bar{K} \subset V} \sum_{\bar{\sigma} \in S_V : \bar{\sigma}|_{\bar{D}} = \bar{\pi}|_{\bar{D}}} \mathbb{P}^{\bar{\pi}, \bar{K}}(\boldsymbol{\pi} = \bar{\pi}|_{\bar{B}}) \mathbb{P}(K_V = \bar{K}, \sigma_V = \bar{\sigma})$$
$$= \sum_{\bar{K} \subset V} \sum_{\bar{\sigma} \in S_V : \bar{\sigma}|_{\bar{D}} = \bar{\pi}|_{\bar{D}}} \mathbb{P}_{\bar{B}}(\bar{\pi}) \mathbb{P}(K_V = \bar{K}) \, e^{-\alpha \mathcal{H}_V(\bar{\sigma})} \frac{1}{Z(V)}.$$

In the last line we used the induction hypothesis, and the independence of K_V and σ_V. For fixed $\bar{K}$, the sum over $\bar{\sigma}$ can now be carried out, and the resulting term is

$$\mathbb{P}_{\bar{B}}(\bar{\pi}) \mathbb{P}(K_V = \bar{K}) \, e^{-\alpha H_{\bar{D}}(\bar{\pi})} \frac{Z(\bar{B})}{Z(V)} = \mathbb{P}(K_V = \bar{K}) \mathbb{P}_V(\bar{\pi}).$$

The result now follows by summing over $\bar{K}$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Let us now come back to the task of constructing invariant sets with prescribed symmetries. We will be slightly more general than in the discussion leading up to (4.4), as this generality will be needed further on. Let $G = (V, E)$ be a finite graph. A *symmetry* of G is a bijective map ϕ on V such that for all $x, y \in V$, $\{\phi(x), \phi(y)\} \in E$ if and only if $(x, y) \in E$. For a group Φ of symmetries on G and an element $\pi \in S_V$, we say that $U \subset V$ is a Φ-compatible π-invariant set if U is invariant under π as well as under all $\phi \in \Phi$. We will use a sampling strategy which is defined as follows. The random variables $(\sigma_B)_{B \subset V}$ are independent and $\mathbb{P}_B$-distributed. Moreover, for any given subset A, the maps K_B are defined as follows,

$$K_V = \Phi(A), \text{ and for } B \neq V : K_B = B \cap \Phi(B^c) \text{ if } B \cap \Phi(B^c) \neq \emptyset, \text{ and } K_B = B \text{ otherwise,} \tag{4.14}$$

where $\Phi(A) := \bigcup_{\phi \in \Phi} \phi(A)$ is the symmetrization of A. For any configuration ω we put,

$$N(\omega) = \inf\{n \in \mathbb{N} : K_{B_n(\omega)}(\omega) = B_n(\omega)\}, \text{ and } \hat{A}(\omega) = B_{N(\omega)}^c(\omega). \tag{4.15}$$

Then $\hat{A}(\omega)$ is a $\boldsymbol{\pi}(\omega)$-invariant subset, compatible with Φ, and containing A. Recall that, from Lemma 4.5, $\boldsymbol{\pi}$ is distributed like $\mathbb{P}_V$. We say that U separates $A \subset V$ from $B \subset V$ if $A \subset U$ and $B \cap U = \emptyset$. Thus, if in addition $\hat{A}(\omega) \cap B = \emptyset$ for some subset B, we have found a Φ compatible, $\boldsymbol{\pi}(\omega)$-invariant subset separating A from B.
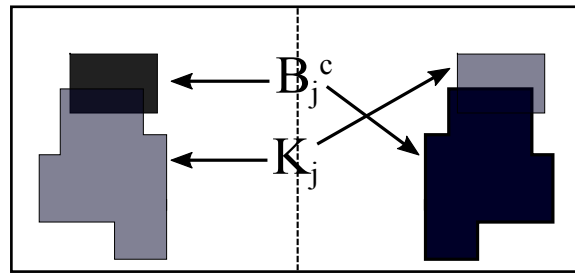
Figure 2: When the group of symmetries Φ contains the identity and the reflection with respect to the dotted vertical line and B_j^c corresponds to the black set in the figure, then K_j corresponds to the grey set.

Proposition 4.6. *For $A, B \subset V$ and $\hat{A}$ defined as in* (4.15)*, we have*

$$\mathbb{P}_V(A \, \pi\text{-invariant, } \Phi\text{-compatible subset separating } A \text{ from } B \text{ exists}) \geq \mathbb{P}(\hat{A} \cap B = \emptyset).$$

Proof. We start with the full graph (V, E), $A \subset V$, and choose the symmetrization $\Phi(A) = \bigcup_{\phi \in \Phi} \phi(A)$ of A as our first set of which we want to keep the cycles. We draw a sample σ_V, and look at the cycles intersecting $K_1 := \Phi(A)$. If none of the cycles of σ_V intersecting K_1 leaves K_1, this means that K_1 is already a Φ-compatible, π-invariant subset of V, and we are done. Otherwise, we take $D_1 = \mathrm{Or}(K_1)$, and $B_2 = V \setminus D_1$. B_2^c is in general not a Φ-compatible subset of V: given a point $x \in D_1 \setminus K_1 \subset B_2^c$, all of its images under maps $\phi \in \Phi$ might be elements of B_2. Since x is an element of B_2^c, this means that B_2^c is not Φ-compatible in this case. We define $\tilde{K}_2 = B_2 \cap \Phi(B_2^c)$ (see also Figure 2). $\tilde{K}_2$ may still be empty, e.g. if we are lucky enough so that all cycles leaving K_1 jointly cover a Φ-compatible set. In that case, we are done, and B_2^c is the desired set, as it is Φ-compatible, π-invariant and contains A. Otherwise, we put $K_2 = \tilde{K}_2$. By the considerations we have just made, it follows that the cardinality of K_2 is bounded by $(|\Phi| - 1)|D_1 \setminus K_1|$. This estimate will be useful later on. We now sample a random permutation σ_{B_2} and keep all the cycles intersecting K_2 and arrive to a set D_2. Again, if $D_2 = K_2$, then $V \setminus K_2$ is a Φ-compatible, π-invariant subset of V, and we are done. Otherwise, we repeat the procedure, i.e. we set $B_3 = B_2 \setminus D_2$ and $\tilde{K}_3 = B_3 \cap \Phi(B_3^c)$. As before, we have $|K_3| \leq (|\Phi| - 1)|D_2 \setminus K_2|$. We continue in this way until we exhaust V (in which case we have found no non-trivial Φ-compatible subset), or until at some point $D_j = K_j$, in which case $B_{j+1}^c = \bigcup_{i \leq j} D_i$ is a Φ-compatible, π-invariant set containing A. $\qquad \square$

The previous proposition will only be useful if we have a way to control the probability that the random set $\hat{A}$ intersects B. At this point, we currently have no better tool than the following very crude estimate: if we write d for the graph distance on G, then

$$|\hat{A} \setminus \Phi(A)| < d(\Phi(A), B) \quad \text{implies} \quad \hat{A} \cap B = \emptyset. \tag{4.16}$$

Inequality (4.16) holds true since by construction, from every $x \in \hat{A}$ a path in G that is entirely contained in $\hat{A}$ leads to $\Phi(A)$. Fortunately, this crude estimate will be sufficient for our purposes. We therefore now consider bounds on the cardinality of $\hat{A}$.

We now introduce the notion of random permutation with bounded cycle length. Even if so far in this section we considered only finite graphs, we introduce this notion for graphs that might be finite or infinite. Recall that a random variable X is *stochastically dominated* by another random variable Y if $\mathbb{P}(X \geq k) \leq \mathbb{P}(Y \geq k)$ for all k. Let ξ be a random variable on the positive integers and recall the notion of "cycle length bounded by ξ", provided before the statement of Theorem 2.3. We use this notion in the next proposition.

Proposition 4.7. *Let ξ be an integer-valued random variable, and assume that a random permutation has cycle length bounded by ξ on the finite graph (V, E). Let $(\xi_n)_{n \in \mathbb{N}}$ be a sequence of independent copies of ξ. Then for each $A \subset V$ and each $\ell \in \mathbb{N}$, we have*

$$\mathbb{P}_V(|\mathrm{Or}(A)| \geq \ell) \leq \mathbb{P}(\sum_{i=1}^{|A|} \xi_i \geq \ell).$$

Proof. We consider the following sampling strategy: let $|A| = n$ and let $x^1, \ldots, x^n$ be the elements of A. We sample the cycles intersecting A one by one, in the order of the x^j. In other words, for $B \cap A \neq \emptyset$ we set $j_B = \min\{i \leq n : x^i \in B\}$, $x(B) = x^{j_B}$, and $K_B = \{x(B)\}$. When $B \cap A = \emptyset$, we set $K_B = B$. This way, we get a sampling procedure with sets $(B_i)_{i \in \mathbb{N}}$. Let $L_i = |\gamma_{x(B_i)}|$ be the number of vertices in the $i - th$ cycle that is sampled. When A is exhausted after τ steps, we set $L_i = 0$ for all $i > \tau$. Note that $\tau(\omega) \leq n$ for any ω, where ω denotes the realization in the probability space of the sampling procedure. The sequence $(L_i)_{i \in \mathbb{N}}$ of random variables is definitely not a Markov chain, and usually L_i will not even be measurable with respect to the σ-algebra generated by all the other L_k. However, it has the property that for given $\bar{B} \subset V$ and given cycle $\bar{\gamma}$,

$$\mathbb{P}(L_{k+1} > \ell | B_k = \bar{B} \text{ and } \gamma_{x(\bar{B})} = \bar{\gamma}) = \mathbb{P}_{\bar{B} \setminus \bar{\gamma}}(|\gamma_{x(\bar{B} \setminus \bar{\gamma})}| > \ell) \leq \mathbb{P}(\xi_{k+1} > \ell). \qquad (4.17)$$

Since $L_k = |\gamma_k|$, we conclude that when $\mathcal{F}_k$ is the σ-algebra generated by $L_1, \ldots, L_k$, then

$$\mathbb{P}(L_{k+1} > \ell | \mathcal{F}_k)(\omega) \leq \mathbb{P}(\xi_{k+1} > \ell)$$

for all ω. We have just checked Assumption 2 of our comparison lemma, Lemma 5.7, which we give in the Appendix as Lemma 5.7. In the notation of that lemma, $(M_j^1)_{j \geq 1} = (L_j)_{j \geq 1}$, and $(M_j^2)_{j \geq 0} = (\xi_j)_{j \geq 0}$. It is easy to see that Assumptions 1 and 3 also hold. The second statement of Lemma 5.7 then implies that for any $k \in \mathbb{N}$,

$$\mathbb{P}(\sum_{j=1}^{k} L_j \geq \ell) \leq \mathbb{P}(\sum_{j=1}^{k} \xi_j \geq \ell).$$

The statement of our proposition now follows when observing that $|\mathrm{Or}(A)| = \sum_{j=1}^{|A|} L_j$. $\quad\square$

Proposition 4.8. *Consider a sampling strategy with the property that there exists $M \in \mathbb{N}$ such that under the recursive sampling procedure (4.9), the inequality*

$$|K_{B_{i+1}}| \leq M |D_i \setminus K_{B_i}|$$

holds almost surely in the case $D_i \setminus K_{B_i} \neq \emptyset$, and $K_{B_{i+1}} = B_{i+1}$ otherwise. Let N be the random number of steps such that $K_{B_N} = B_N$, and let $\hat{A} = B_N^c$ be the set sampled before N (see also (4.15)). Assume further that the random permutation has cycle length bounded by an integer-valued random variable ξ on (V, E) and let $(Z_j)_{j \geq 0}$ be a Galton-Watson process where the offspring is distributed according to $M(\xi - 1)$, and with initial population $M|K_V|$. Let W be the (possibly infinite) total population of the Galton-Watson process. Then for each $\ell \in \mathbb{N}$, we have

$$\mathbb{P}_V(|\hat{A}| \geq \ell) \leq \mathbb{P}(W \geq \frac{M}{M+1} \ell).$$

Proof. By definition, $\hat{A}(\omega) = \bigcup_{i=1}^{N(\omega)-1} D_i(\omega)$. Moreover, $K_{B_i(\omega)}(\omega) \subset D_i(\omega)$ for all i, thus for all n,

$$\bigcup_{i=1}^{n} D_i(\omega) = \bigcup_{i=1}^{n} \left(D_i(\omega) \setminus K_{B_i(\omega)}(\omega) \right) \cup \bigcup_{i=1}^{n} K_{B_i(\omega)}(\omega).$$

Whenever $K_{B_i(\omega)}(\omega) \neq B_i(\omega)$, we have assumed $|K_{B_{i+1}(\omega)}(\omega)| \leq M|D_i(\omega) \setminus K_{B_i(\omega)}(\omega)|$. Therefore we find that for all ω,

$$|\hat{A}(\omega)| \leq (M+1) \sum_{i=1}^{N(\omega)-1} |D_i(\omega) \setminus K_{B_i(\omega)}(\omega)|.$$

Write $q_i(\omega) = |D_i(\omega) \setminus K_{B_i(\omega)}(\omega)|$ for brevity. Then for given sets $\bar{B}$ and $\bar{K}$,

$$\mathbb{P}(q_j \geq \ell | B_j = \bar{B}, K_{B_j} = \bar{K}) = \mathbb{P}_{\bar{B}}(\text{Or}(\bar{K}) - |\bar{K}| \geq \ell) \leq \mathbb{P}(\sum_{i=1}^{|\bar{K}|}(\xi_i - 1) \geq \ell),$$

where in the last step we used Proposition 4.7. Let $\mathcal{F}_j$ be the sigma-algebra generated by the $(q_k)_{k \leq j}$. Since $|K_{B_j(\omega)}(\omega)| \leq M q_{j-1}(\omega)$ for any ω, then we conclude that for any $\ell, m \in \mathbb{N}$, for any integer $j \geq 1$, and for any ω such that $q_{j-1}(\omega) = m$,

$$\mathbb{P}(q_j \geq \ell | \mathcal{F}_{j-1})(\omega) \leq \mathbb{P}\Big(\sum_{i=1}^{Mm}(\xi_i - 1) \geq \ell \Big)$$

$$= \mathbb{P}\Big(\sum_{i=1}^{Mm} M(\xi_i - 1) \geq M\ell \Big) \leq \mathbb{P}\Big(Z_j \geq M\ell \Big| Z_{j-1} = mM \Big). \quad (4.18)$$

Now we can again apply our comparison Lemma 5.7. In the notation of that lemma, $(M_j^1)_{j \geq 0} = (q_j)_{j \geq 0}$ and $(M_j^2)_{j \geq 0} = \frac{1}{M}(Z_j)_{j \geq 0}$, which is the Galton-Watson process rescaled by a constant M. Note that from Lemma 5.8, the Markov chain $(M_j^2)_{j \geq 0}$ satisfies the Assumption 1 of our comparison lemma, Lemma 5.7. Note also that from (4.18), we have that the Assumption 2 of Lemma 5.7 is fulfilled. Furthermore, since $\frac{1}{M}(Z_j)_{j \in \mathbb{N}}$ has initial population $\frac{1}{M}Z_0 = q_0 = |K_V|$, we also have that the Assumption 3 of the Lemma 5.7. Then, the Conclusion (b) of Lemma 5.7 gives that,

$$\mathbb{P}\Big(\sum_{i=0}^{N(.)-1} q_i \geq \ell \Big) \leq \mathbb{P}\Big(W \geq M\ell \Big).$$

Thus, we find that

$$\mathbb{P}(|\hat{A}| \geq \ell) \leq \mathbb{P}\Big((M+1) \sum_{i=0}^{N(.)-1} q_i \geq \ell \Big) \leq \mathbb{P}\Big((M+1)W \geq M\ell \Big).$$

The proof is finished. $\qquad \square$

Galton-Watson processes where the expected offspring per individual is strictly less than one have exponentially bounded total population. We can use this for

Proposition 4.9. *Consider the sampling strategy described before Proposition 4.6. Assume that the random permutation has cycle length bounded by ξ on the finite graph (V, E), and that $(|\Phi| - 1)(\mathbb{E}(\xi) - 1) < 1$. Then there exist constants $C_0 > 0$ and $\kappa_0 > 0$, depending on ξ and $|A|$ but not on V, so that*

$$\mathbb{P}_V(|\hat{A}| \geq \ell) \leq C_0 \exp(-\kappa_0 \, \ell), \quad (4.19)$$

where $\hat{A}$ has been defined in (4.15). *In particular,*

$$\mathbb{P}_V(A \; \pi\text{-invariant, } \Phi\text{-compatible subset separating } A \text{ from } B \text{ exits})$$
$$\geq 1 - C_0 \, e^{-\kappa_0 [d(\Phi(A), B) + |\Phi(A)|]}. \quad (4.20)$$

where $d(A, B)$ is the graph distance between A and B.

Proof. Observe that our sampling strategy (4.14) fulfils $|K_{B_{i+1}}| \leq (|\Phi| - 1)|D_i \setminus K_{B_i}|$ almost surely, so Proposition 4.8 can be applied with $M = |\Phi| - 1$. Thus, the first statement follows from subcriticality of the Galton-Watson process. For the second statement, we use that by (4.16),

$$\mathbb{P}(\hat{A} \cap B = \emptyset) \geq 1 - \mathbb{P}\left[|\hat{A} \setminus \Phi(A)| \geq d(\Phi(A), B)\right] = 1 - \mathbb{P}\left[|\hat{A}| \geq d(\Phi(A), B) + |\Phi(A)|\right]. \quad \square$$

We will be interested in cases where the set A itself is large. Then the above estimate is too weak. Indeed, since the constant C_0 there depends on A, we have too little control on its growth with $|A|$. We therefore need Theorem 2.3, whose proof is presented below. Note that Theorem 2.3 states exponential decay of correlations only in the context of comparing expectation of local function for different boundary conditions. An adaptation to the existence of Φ-compatible sets is straightforward.

Proof of Theorem 2.3. Let (V, E) be a finite graph, and $B \subset U \subset V$. We assume that a random permutation on (V, E) has cycle length bounded by a random variable ξ with $\mathbb{E}(\xi) < 2$. We let $C > 0$ and $\kappa > 0$ be two constants such that the total offspring W of a Galton-Watson process with initial population 1 and offspring distribution $\xi - 1$ fulfils $\mathbb{P}(W > n) \leq C\,\mathrm{e}^{-2\kappa n}$.

Let V_1 and V_2 be two disjoint copies of V and put $\tilde{V} = V_1 \cup V_2$, $\tilde{V}_0 = U_1 \cup V_2$ where U_1 is U regarded as a subset of V_1, and $A = \tilde{V} \setminus \tilde{V}_0$. Let ϕ be the natural graph isomorphism taking V_1 to V_2. For $\tilde{\pi} \in S_{\tilde{V}}$, let $Q_A(\boldsymbol{\pi})$ be the minimal $\boldsymbol{\pi}$-invariant, ϕ-compatible set containing A. In equation (4.4) we have seen that for $B \subset U$ and $\mathcal{F}_B$-measurable f,

$$|\mathbb{E}_V(f) - \mathbb{E}_U(f)| \leq 2\|f\|_\infty \mathbb{P}_{\tilde{V}}\left(Q_A \cap B \neq \emptyset \,\middle|\, \boldsymbol{\pi}|_A = \mathrm{id}\right). \quad (4.21)$$

On the right hand side above, we interpret B as a subset of $U_1 \subset V_1$.

For estimating the right hand side, we first note that when $A, A' \subset \tilde{V}$, then

$$Q_{A \cup A'} = Q_A \cup Q_{A'}.$$

Indeed, the implication $\supset$ holds since $A \subset B$ implies $Q_A(\boldsymbol{\pi}) \subset Q_B(\boldsymbol{\pi})$ for any $\boldsymbol{\pi}$. For the reverse implication, note that $Q_A(\boldsymbol{\pi}) \cup Q_{A'}(\boldsymbol{\pi})$ is $\boldsymbol{\pi}$-invariant and contains $A \cup A'$, and since ϕ is bijective we also have $\phi(Q_A \cup Q_{A'}) = \phi(Q_A) \cup \phi(Q_{A'}) = Q_A \cup Q_{A'}$. So, $Q_A \cup Q_{A'} \supset Q_{A \cup A'}$ by the minimality of the latter set. We conclude

$$\mathbb{P}_{\tilde{V}}\left(Q_A \cap B \neq \emptyset \,\middle|\, \boldsymbol{\pi}|_A = \mathrm{id}\right)$$
$$= \mathbb{P}_{\tilde{V}}\left(\bigcup_{x \in A}\{Q_{\{x\}} \cap B\} \neq \emptyset \,\middle|\, \boldsymbol{\pi}|_A = \mathrm{id}\right)$$
$$\leq \sum_{x \in A} \mathbb{P}_{\tilde{V}}\left(Q_{\{x\}} \cap B \neq \emptyset \,\middle|\, \boldsymbol{\pi}|_A = \mathrm{id}\right). \quad (4.22)$$

For estimating the latter probabilities, define $\phi_0(y)$ so that $\phi_0(y) = \phi(y)$ for $y \in U_1 \cup U_2$, and $\phi_0(y) = y$ if $y \in U_2^c$. Here, U_2 denotes U regarded as a subset of V_2. Let $Q^0_{\{x\}}$ be the minimal $\boldsymbol{\pi}$-invariant set containing x so that $\phi_0(Q^0_{\{x\}}(\boldsymbol{\pi})) = Q^0_{\{x\}}(\boldsymbol{\pi})$. First note that $\mathbb{P}_{\tilde{V}}(\boldsymbol{\pi}|_A = \mathrm{id}) = Z(\tilde{V}_0)/Z(\tilde{V})$, and thus

$$\mathbb{P}_{\tilde{V}}\left(Q_{\{x\}} \cap B \neq \emptyset \,\middle|\, \boldsymbol{\pi}|_A = \mathrm{id}\right) = \mathbb{P}_{\tilde{V}_0}(Q^0_{\{x\}} \cap B \neq \emptyset).$$

The relevant sampling strategy is then given by

$$K_{\tilde{V}_0} = \{x\}, \quad K_B = B \cap \phi_0(B^c) \text{ if } B \cap \phi_0(B^c) \neq \emptyset, \quad K_B = B \text{ otherwise.}$$

As before, we let $N = \inf\{j \in \mathbb{N} : K_{B_j} = B_j\}$ and $\hat{A} = B_N^c$. Then $\hat{A}(\boldsymbol{\pi})$ is ϕ_0-compatible and $\boldsymbol{\pi}$-invariant, so $Q_{\{x\}}^0(\boldsymbol{\pi}) \subset \hat{A}(\boldsymbol{\pi})$ for all $\boldsymbol{\pi}$. The sampling strategy fulfils the assumptions of Proposition 4.8 with $M = 1$, and therefore

$$\mathbb{P}_{\tilde{V}_0}(Q_{\{x\}}^0 \cap B \neq \emptyset) \leq \mathbb{P}_{\tilde{V}_0}(|\hat{A}| \geq d(x, B) + 1) \leq C\,\mathrm{e}^{-\kappa(d(x,B)+1)}\,.$$

C and κ are the constants defined in the statement of the proposition. Inserting this estimate in (4.22) and the result into (4.21) shows the claim. $\qquad\square$

Consider a vertex transitive graph $G = (V, E)$ that might be finite or infinite. Theorem 2.2 guarantees that there exists $\alpha_0' < \infty$ (which may be significantly larger than the α_0 given in that theorem) such that for all $\alpha > \alpha_0'$, the corresponding random permutation has cycle length bounded by a random variable ξ that has finite exponential moments and has $\mathbb{E}(\xi) < 2$. Theorem 2.3 then gives a bound on $|\mathbb{P}_{V_1}(\pi|_A = \mathrm{id}) - \mathbb{P}_{V_0}(\pi|_A = \mathrm{id})|$ for finite subsets $A \subset V_1 \subset V_0$ of V. However, this bound becomes meaningless when A is allowed to become large, since both of the involved probabilities then become very small. The following proposition improves the estimate of Theorem 2.3 to the strength that will be needed in the next section. We formulate it in terms of partition functions.

Proposition 4.10. *Let (V, E) be finite or infinite. Consider finite subsets $A \subset V_1 \subset V_0$ of V. Let α_0' be large enough so that for all $\alpha > \alpha_0'$, the random permutation has cycle length bounded by a random variable ξ with $E(\xi) < 2$. Let $C > 0$ and $\kappa > 0$ be two constants such that the total offspring W of a Galton-Watson process with initial population 1 and offspring distribution $\xi - 1$ fulfils $\mathbb{P}(W > n) \leq C\,\mathrm{e}^{-2\kappa n}$. Let $c_1(\alpha)$ be defined as in (3.9). Define*

$$D = D(A, V_0, V_1, \alpha, \xi) = \exp\Big(\frac{2C}{c_1(\alpha)}\sum_{x \in A}\sum_{y \in V_0 \setminus V_1}\mathrm{e}^{-\kappa d(x,y)}\Big), \qquad (4.23)$$

where $d(x, y)$ denotes the graph distance between x and y. Then

$$\frac{1}{D}\frac{Z(V_1 \setminus A)}{Z(V_1)} \leq \frac{Z(V_0 \setminus A)}{Z(V_0)} \leq D\frac{Z(V_1 \setminus A)}{Z(V_1)}.$$

Proof. Let us put $n = |A|$, and write

$$A = \{x^1, x^2, \ldots x^n\}, \quad \text{and} \quad A_i = \{x^i, x^{i+1}, \ldots, x^n\} \subset A$$

for $i \leq |A|$. We set $A_{n+1} = \emptyset$. Then, for $j = 0, 1$,

$$\frac{Z(V_j \setminus A)}{Z(V_j)} = \prod_{i=1}^{n}\frac{Z(V_j \setminus A_i)}{Z(V_j \setminus A_{i+1})},$$

and we have

$$\log\frac{Z(V_1 \setminus A)}{Z(V_1)} = \sum_{i=1}^{n}\log\frac{Z(V_1 \setminus A_i)}{Z(V_1 \setminus A_{i+1})}$$

$$= \sum_{i=1}^{n}\log\frac{Z(V_0 \setminus A_i)}{Z(V_0 \setminus A_{i+1})} + \sum_{i=1}^{n}\log\Big(\frac{Z(V_1 \setminus A_i)}{Z(V_1 \setminus A_{i+1})}\frac{Z(V_0 \setminus A_{i+1})}{Z(V_0 \setminus A_i)}\Big). \quad (4.24)$$

The first sum on the right hand side above is equal to $\log\frac{Z(V_0 \setminus A)}{Z(V_0)}$. We use the inequality $\log(1 + |x|) \leq |x|$ on each term in the final sum above, and find

$$\Big|\log\frac{Z(V_1 \setminus A)}{Z(V_1)} - \log\frac{Z(V_0 \setminus A)}{Z(V_0)}\Big| \leq \sum_{i=1}^{n}\Big|\frac{Z(V_1 \setminus A_i)}{Z(V_1 \setminus A_{i+1})} - \frac{Z(V_0 \setminus A_i)}{Z(V_0 \setminus A_{i+1})}\Big|\frac{Z(V_0 \setminus A_{i+1})}{Z(V_0 \setminus A_i)}.$$

Now, by applying Proposition 3.3 with $U = V_0 \setminus A_{i+1}$ and $A = \{x^i\}$, we get

$$\left| \frac{Z(V_0 \setminus A_{i+1})}{Z(V_0 \setminus A_i)} \right| \leq \frac{1}{c_1(\alpha)}.$$

We now apply Theorem 2.3 with $f = \mathbb{1}\{\pi(x^i) = x^i\}$ and we obtain that,

$$\left| \frac{Z(V_1 \setminus A_i)}{Z(V_1 \setminus A_{i+1})} - \frac{Z(V_0 \setminus A_i)}{Z(V_0 \setminus A_{i+1})} \right| = |\mathbb{P}_{V_1 \setminus A_{i+1}}(\pi(x^i) = x^i) - \mathbb{P}_{V_0 \setminus A_{i+1}}(\pi(x^i) = x^i)|$$

$$\leq \sum_{y \in V_0 \setminus V_1} 2C \, \mathrm{e}^{-\kappa d(y,x^i)}. \quad (4.25)$$

We combine these estimates and obtain

$$\left| \log \frac{Z(V_1 \setminus A)}{Z(V_1)} - \log \frac{Z(V_0 \setminus A)}{Z(V_0)} \right| \leq \frac{2\,C}{c_1(\alpha)} \sum_{x \in A} \sum_{y \in V_0 \setminus V_1} \mathrm{e}^{-\kappa d(x,y)}.$$

Exponentiating this inequality gives $Z(V_1 \setminus A)/Z(V_1) \leq DZ(V_0 \setminus A)/Z(V_0)$ and thus the first claimed inequality. For the second one, notice that the only place where we used $V_1 \subset V_0$ is inequality (4.25). That estimate still holds when interchanging the roles of V_0 and V_1, and when we do this, we get the remaining inequality. $\qquad \square$

We will only use the following weaker form of the previous proposition.

Corollary 4.11. *In the situation of Proposition 4.10, write $B = V_0 \setminus V_1$. We then have*

$$\exp\left(-\frac{2C}{c_1(\alpha)} |A||B| \, \mathrm{e}^{-\kappa d(A,B)} \right) \frac{Z(V_1 \setminus A)}{Z(V_1)}$$

$$\leq \frac{Z(V_0 \setminus A)}{Z(V_0)} \leq \exp\left(\frac{2C}{c_1(\alpha)} |A||B| \, \mathrm{e}^{-\kappa d(A,B)} \right) \frac{Z(V_1 \setminus A)}{Z(V_1)}. \quad (4.26)$$

5 Proof of Theorem 2.4

5.1 Definitions and Notation

Recall that for $y = (y_1, \ldots, y_d) \in \mathbb{Z}^d$, we write $\bar{y} = y_1$ and $\hat{y} = (y_2, \ldots, y_d)$. For $y \in \Lambda_n$, we write

$$\mathcal{C}_y := \{x \in \Lambda_n : \bar{x} - \bar{y} \geq |\hat{x} - \hat{y}|\} \cup \{x \in \Lambda_n : \bar{x} \geq \bar{y} + \log n\}$$

for the forward cone starting in y that is broadened to full width after a logarithmic length (see also Figure 3a). A set $A \subset \Lambda_n$ will be called

$$\text{weakly } y\text{-admissible if } \{x \in \Lambda_n : \bar{x} \geq \bar{y} + \log n\} \subset A,$$
$$y\text{-admissible if in addition } \{x \in \Lambda_n : \bar{x} \geq \bar{y}, \hat{x} = \hat{y}\} \subset A,$$
$$\text{and strictly } y\text{-admissible if in addition } \mathcal{C}_y \subset A.$$

We write $\mathcal{A}_y^w$, $\mathcal{A}_y$ and $\mathcal{A}_y^s$ for the set of weakly admissible, admissible and strictly admissible sets, respectively.

For $y, z \in \Lambda_n$, $A \in \mathcal{A}_y$ and $\pi \in \mathcal{S}_A^{y \to z}$ recall that $\gamma(\pi) = \mathrm{Orb}_\pi(\{y\})$ denotes the trace of the self avoiding walk that is embedded in π, starting from y and ending in z. We order the elements of $\gamma(\pi)$ by order of their appearance in $\mathrm{Orb}_\pi(\{y\})$, i.e. $x \leq y$ if $y \in \mathrm{Orb}_\pi(\{x\})$. Together with this order, the set $\gamma(\pi)$ uniquely characterises a self-avoiding path from y to z. $\mathcal{L}_A^{y \to z}$ denotes the set of all ordered subsets of A such that their order makes them a self-avoiding nearest neighbour walk in A starting in y and ending in z. We define $\mathcal{L}_A^{y \to \ell_n} = \bigcup_{z \in \ell_n} \mathcal{L}_A^{y,z}$.
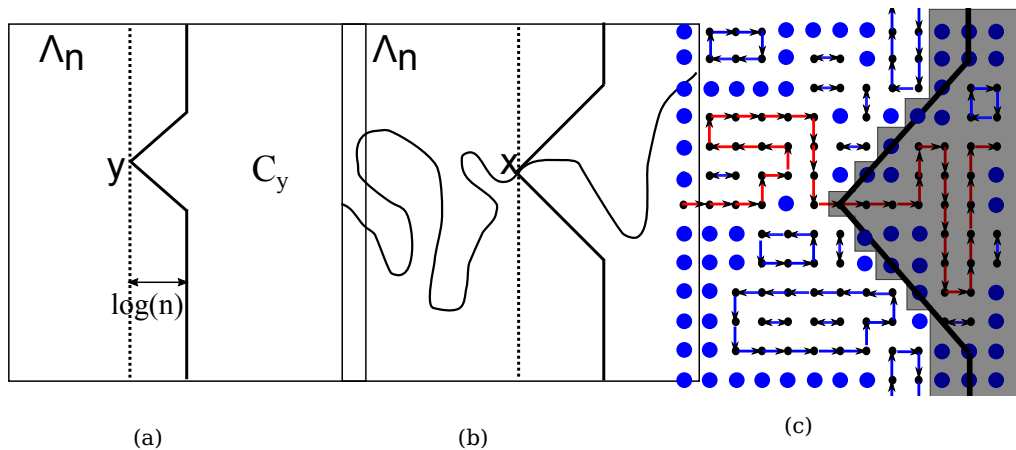
Figure 3: **(a)** A forward cone C_y. **(b)** A pre-regeneration point x. **(c)** The dark set represents the regeneration set of π through x, where x is the vertex in the centre of the figure.

For $\gamma \in \mathcal{L}_A^{y \to \ell_n}$, a point $x \in \gamma$ will be called *pre-regeneration point* of γ if $z \in C_x$ for all $z \geq x$, and $\bar{z} < \bar{x}$ for $z < x$. In words, the self avoiding path hits the set $\ell_{\bar{x}}$ precisely at x and stays in C_x thereafter (see also Figure 3b). Note that the latter requirement is more than what is usually required for regeneration points of self-avoiding walks. We will need it below in order to decouple the future of γ from part of the background consisting of loops in π.

Let $\pi \in \mathcal{S}_A^{y \to \ell_n}$, and let π_0 be the element of $\mathcal{S}_A$ with $\pi_0(x) = x$ whenever $x \in \gamma(\pi)$, and $\pi_0(x) = \pi(x)$ otherwise. A pre-regeneration point $x \in \gamma(\pi)$ will be called *regeneration point* if there exists a subset $R \subset A$ with the following four properties:

(R1) $R \subset \{z \in A : \bar{z} \geq \bar{x}\}$.

(R2) R is strictly x-admissible.

(R3) R is invariant under reflections leaving $\{z \in A : \hat{z} = \hat{x}\}$ invariant. In other words, $(\bar{z}, 2\hat{x} - \hat{z}) \in R$ whenever $z \in R$.

(R4) R is invariant under π_0.

If x is a regeneration point of π, the largest set with properties (R1)-(R4) is called *regeneration set* of π through x, and is denoted by $R(x, \pi)$. See also Figure 3c. Since (R1)-(R4) are invariant under unions of sets, the existence of any regeneration set implies the existence and uniqueness of $R(x, \pi)$. By convention, the point $0 \in \Lambda_n$ will be a regeneration point with $A(0, \pi) = \Lambda_n$ for all π, even though 0 might not be a pre-regeneration point of $\gamma(\pi)$. Also, $z = \max\{x : x \in \gamma(\pi)\}$ is a regeneration point by convention, with empty regeneration set. The set of all regeneration points of some $\pi \in \mathcal{S}_A^{y \to \ell_n}$ is denoted by $\mathcal{R}(\pi)$ and is non-empty by the above conventions. We order its elements by the order in which they appear in γ.

5.2 The number of regeneration points

In this subsection, we investigate how many regeneration points there are (at least) in a system of size n. Our first result is an estimate on the number of steps that γ can take before crossing a certain vertical hyperplane for the last time.

For $L \in \mathbb{N}$ and $\pi \in \mathcal{S}_A^{y \to \ell_n}$, set

$$x_L(\pi) := \max\{x \in \gamma(\pi) : \bar{x} = \bar{y} + L\}, \qquad \text{and} \qquad \rho_L(\pi) := \gamma(\pi) \setminus \mathrm{Orb}_\pi(x_L).$$

x_L is the last point at which the hyperplane $\{z \in A : \bar{z} = \bar{y} + L\}$ is crossed by γ, and ρ_L is the piece of γ that lies before that point. $|A|$ denotes the cardinality of a set A. We define

$$\mathcal{L}_A^{x \nearrow y} := \big\{\gamma \in \mathcal{L}_A^{x \to y} : \gamma \cap \ell_{\bar{x}} = \{x\}\big\}, \qquad \text{and} \quad \mathcal{L}_A^{x \nearrow \ell_n} := \bigcup_{y \in \ell_n} \mathcal{L}_A^{x \nearrow y}$$

In words, $\gamma \in \mathcal{L}_A^{x \nearrow y}$ never visits the hyperplane containing its starting point again after the first step.

Proposition 5.1. *For each $\delta > 0$, there exist $N \in \mathbb{N}$, $\alpha_0 < \infty$, $c > 0$ and $C < \infty$ such that for all $n > N$, $\alpha > \alpha_0$, $y \in \Lambda_n$ and $A, B \in \mathcal{A}_y$ with $B \subset A$, we have*

$$\sup\left\{\mathrm{e}^{cL}\, \mathbb{P}_A^{y \to \ell_n}\big(|\rho_L| \geq (1+\delta)L \big| \rho_L \subset B\big) : L \in (2\log n, n - \bar{y}]\right\} \leq C.$$

Proof. We start the proof by showing an a priori estimate on partition functions. We claim that there exist $\alpha_0 < \infty$ and $N \in \mathbb{N}$ such that for all $\alpha > \alpha_0$, $n > N$, $L > 2 \log n$, $y \in \Lambda_n$, $A \in \mathcal{A}_y$ $h \in A$ with $\bar{h} = \bar{y} + L$, and $\gamma \in \mathcal{L}_A^{h \nearrow \ell_n}$, we have

$$\frac{1}{2} \frac{Z(A)}{Z(\Lambda_n)} \leq \frac{Z(A \setminus \gamma)}{Z(\Lambda_n \setminus \gamma)} \leq 2 \frac{Z(A)}{Z(\Lambda_n)}. \tag{5.1}$$

Indeed, Theorem 2.2 guarantees the existence of $\alpha_0 > 0$ so that for $\alpha > \alpha_0$, the cycle length distribution of spatial random permutations (in an arbitrary domain) is stochastically bounded by the distribution of a geometric random variable ξ with success probability $1 - \mathrm{e}^{-\alpha/2}$. We can thus choose α_0 so large that the total population W of a Galton-Watson process with offspring distribution $(\xi - 1) \vee 0$ satisfies $\mathbb{P}(W > n) \leq \mathrm{e}^{-2\kappa n}$ for all $n \in \mathbb{N}$, where $\kappa = 2d + 1$. The assumptions of Corollary 4.11 are then fulfilled. Moreover, for any y, A, h and γ as above, we have $d(\gamma, \Lambda_n \setminus A) > \log n$ due to the fact that γ stays to the right of h. Since clearly $|\gamma||\Lambda_n \setminus A| \leq n^{2d}$, we have

$$|\gamma||\Lambda_n \setminus A|\, \mathrm{e}^{-\kappa d(\gamma, \Lambda_n \setminus A)} \leq \tfrac{1}{n},$$

and thus Corollary 4.11 allows us to choose N large enough so that (5.1) holds for all $n > N$, uniformly in y, A, h and γ with the stated properties.

For the next step, let α_0, N be as above and $\alpha > \alpha_0$, $n > N$, $y \in \Lambda_n$, $A, B \in \mathcal{A}_y$ with $B \subset A$, $L \in (2 \log n, n - \bar{y}]$, $h \in \ell_{\bar{y}+L}$ and $\rho \in \mathcal{L}_A^{y \to h}$ such that $\rho \subset B$. We define

$$\mathcal{S}_{A \setminus \rho}^{h \nearrow \ell_n} := \{\pi \in \mathcal{S}_{A \setminus \rho}^{h \to \ell_n} : \bar{x} > \bar{h} \text{ for all } x \in \gamma(\pi) \setminus \{h\}\}.$$

Since $\rho \subset B$, we have

$$\mathbb{P}_A^{y \to \ell_n}(\rho_L = \rho | \rho_L \in B) = \mathrm{e}^{-\alpha \|\rho\|} \frac{1}{\sum_{\pi \in \mathcal{S}_A^{h \to \ell_n}} \mathbb{1}\{\rho_L(\pi) \subset B\}\, \mathrm{e}^{-\alpha \mathcal{H}_{A \setminus \gamma}(\pi)}} \sum_{\pi \in \mathcal{S}_{A \setminus \gamma}^{h \nearrow \ell_n}} \mathrm{e}^{-\alpha \mathcal{H}_{A \setminus \gamma}(\pi)}$$

$$= \mathrm{e}^{-\alpha \|\rho\|} \sum_{\gamma \in \mathcal{L}_A^{h \nearrow \ell_n}} \mathrm{e}^{-\alpha \|\gamma\|} Z(A \setminus (\gamma \cup \rho)) \frac{1}{\sum_{\pi \in \mathcal{S}_A^{h \to \ell_n}} \mathbb{1}\{\rho_L(\pi) \subset B\}\, \mathrm{e}^{-\alpha \mathcal{H}_{A \setminus \gamma}(\pi)}} \tag{5.2}$$

For estimating the denominator, let $\rho_0 = \{x \in A : \bar{y} \leq \bar{x} \leq \bar{y} + L, \hat{x} = \hat{y}\}$ be the straight line from y to the hyperplane ℓ_h, and let $h_0 = \rho_0 \cap \ell_h$. Since B is admissible, $\rho_0 \subset B$, and

thus

$$
\sum_{\pi \in \mathcal{S}_A^{h \to \ell_n}} \mathbb{1}\{\rho_L(\pi) \subset B\} \, \mathrm{e}^{-\alpha \mathcal{H}_{A \setminus \gamma}(\pi)} \geq \sum_{\pi \in \mathcal{S}_A^{y \to \ell_n}} \mathbb{1}\{\rho(\pi)
$$

$$
\begin{aligned}
&= \rho_0\} \, \mathrm{e}^{-\alpha \mathcal{H}_A(\pi)} \; = \; \mathrm{e}^{-\alpha L} \, Z^{h_0 \to \ell_n}(A \setminus \rho_0) \\
&\geq \mathrm{e}^{-\alpha L} \, Z^{h_0 \nearrow \ell_n}(A \setminus \rho_0) \\
&= \mathrm{e}^{-\alpha L} \sum_{\gamma \in \mathcal{L}_A^{h_0 \to \ell_n}} \mathrm{e}^{-\alpha \|\gamma\|} \, Z(A \setminus (\rho_0 \cup \gamma)).
\end{aligned}
\tag{5.3}
$$

We now use Proposition 3.3 with $c_1(\alpha)$ as in (3.9) as well as (5.1) and find

$$
Z(A \setminus (\rho_0 \cup \gamma)) \geq c_1^{|\rho_0|} Z(A \setminus \gamma) \geq \frac{c_1^L}{2} \frac{Z(A) Z(\Lambda_n \setminus \gamma)}{Z(\Lambda_n)}.
$$

Using the other inequality of (5.1), we also find

$$
Z(A \setminus (\gamma \cup \rho)) \leq Z(A \setminus \gamma) \leq 2 \frac{Z(A) Z(\Lambda_n \setminus \gamma)}{Z(\Lambda_n)}.
$$

Using these estimates in (5.2) gives

$$
\mathbb{P}_A^{y \to \ell_n}(\rho_L = \rho | \rho_L \subset B) \leq 4 \, \mathrm{e}^{-\alpha(\|\rho\| - L)} \, c_1^L \frac{\sum_{\gamma \in \mathcal{L}_A^{h \nearrow \ell_n}} \mathrm{e}^{-\alpha \|\gamma\|} Z(\Lambda_n \setminus \gamma)}{\sum_{\gamma \in \mathcal{L}_A^{h_0 \nearrow \ell_n}} \mathrm{e}^{-\alpha \|\gamma\|} Z(\Lambda_n \setminus \gamma)} = 4 \, \mathrm{e}^{-\alpha(\|\rho\| - L)} \, c_1^L.
$$

The last equality is by translation invariance.

Finally, fix $\delta > 0$. Since $c_1(\alpha)$ converges to 1 as $\alpha \to \infty$, by making α_0 bigger if necessary we can achieve

$$
\mathbb{P}_A^{y \to \ell_n}(\rho_L = \rho) \leq \mathrm{e}^{-\alpha(\|\rho\| - (1 + \delta/2)L)}
$$

for all $\alpha > \alpha_0$, all $n > N$, and uniformly in all suitable A, B, L and ρ. Then, summing over all ρ (ordered by length) gives

$$
\begin{aligned}
\mathbb{P}(|\rho_L| > (1 + \delta)L | \rho_L \subset B) &\leq \sum_{m=(1+\delta)L}^{\infty} (2d)^m \, \mathrm{e}^{-\alpha(m - (1 + \delta/2)L)} \; , \\
&= \mathrm{e}^{-L(\alpha\delta/2 - (1+\delta)\log(2d))} \sum_{m=0}^{\infty} \mathrm{e}^{-m(\alpha - \log(2d))} \, .
\end{aligned}
$$

(The crude bound $(2d)^m$ on the number of all self-avoiding walks of length m could of course be improved using the connective constant, but in the present situation we do not gain any insight from that.) Now by further increasing α_0 if necessary, we see that the claim holds with $c = \alpha_0 \delta/2 - (1 + \delta)\log(2d)$, and $C = (1 - \mathrm{e}^{-\alpha_0 + \log(2d)})^{-1}$. □

The next statement, which is purely deterministic, shows that proposition (5.1) is useful for obtaining lower bounds on the number of pre-regeneration points.

Lemma 5.2. *Let $L \in \mathbb{N}$, $A \subset \Lambda_n$, $x, y \in A$ such that $\bar{y} = \bar{x} + L$, and $\gamma \in \mathcal{L}_A^{x \to y}$. Assume that $|\gamma| < (1 + \delta) L$. Then γ has at least $(1 - 3\delta) L$ pre-regeneration points.*

Proof. We recursively determine sequences of points in γ. We set $x_0 = x$. For $i \geq 1$, we define

$$
\begin{aligned}
y_i &:= \min\{y \in \gamma : y \geq x_{i-1}, y \text{ is not a pre-regeneration point}\}, \\
y_i' &:= \min\{y \in \gamma : y > y_i, y \notin \mathcal{C}_{y_i}\}, \\
m_i &:= \max\{\bar{x} : x \in \gamma, x \leq y_i'\}, \\
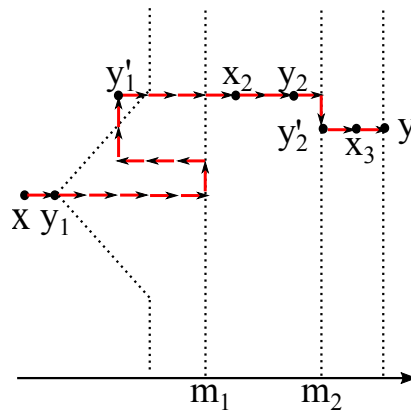x_{i+1} &:= \min\{x \in \gamma : \bar{x} > m_i\},
\end{aligned}
$$

Figure 4: A self-avoiding walk from x to y.

(see also Figure 4). Thus, when following the path γ, we have the following situation: At each x_i, the path crosses the hyperplane ℓ_{x_i} for the first time. Thereafter, we only see regeneration points until we hit y_{i+1}, although of course $y_{i+1} = x_i$ is possible, and at the point y'_{i+1} we actually notice that y_{i+1} was not a pre-regeneration point (see also the following paragraph!). We then follow γ further until we hit a vertical hyperplane that we have not seen before y'_{i+1}, and x_{i+1} is the place where we cross that hyperplane.

The crucial observation is that since x_i is the first time a certain hyperplane is crossed, and since all points between x_i and y_i are pre-regeneration points, the reason that y_i is not a pre-regeneration point can only be that the path leaves the enhanced forward cone; the other possibility, namely that some previous part of the path has already crossed the hyperplane containing y_i, is ruled out. This means that indeed y'_i is the place where we notice that y_i was not a pre-regeneration point. It also means that between y_i and x_{i+1}, the path stays strictly to the right of $\ell_{\bar{y}_i}$ and strictly to the left of $\ell_{\bar{x}_{i+1}}$. It is easy to see that in such a situation, at most a third of all the steps of the path can be 'to the right', i.e. between vertices z, w with $\bar{z} = \bar{w} + 1$. Since γ takes at least L steps to the right there are at most δL steps available for the other directions. Thus the union of all pieces of γ that lie between some y_i and the corresponding x_{i+1} has less than $3\delta L$ elements. Since all of the remaining at least $(1 - 3\delta)L$ points are regeneration points by construction, the claim is proved. $\qquad\square$

The next step is to show that for sufficiently large α, a pre-regeneration point that is sufficiently remote from the starting point of γ_0 is an actual regeneration point with uniformly positive probability. Recall that $\mathcal{R}(\pi)$ denotes the set of all regeneration points of π.

Proposition 5.3. *There exists $\alpha_0 < \infty$, $c > 0$ and $N \in \mathbb{N}$ such that for any $n \geq N$, $\alpha > \alpha_0$, $y \in \Lambda_n$, $A \in \mathcal{A}_y^w$ such that $y \in A$, $\gamma_0 \in \mathcal{L}_A^{y \to \ell_n}$, and any pre-regeneration point $x \in \gamma_0$ with $\bar{x} - \bar{y} \geq \log n$, we have*

$$\mathbb{P}_A^{y \to \ell_n}(x \in \mathcal{R} | \gamma = \gamma_0) > c.$$

Proof. Let $\tilde{c} < 1$. Pick first $\nu > d - 1$ so large that for all $n \in \mathbb{N}$,

$$\sum_{k=1}^{\infty} e^{-\nu k} |\{z \in \Lambda_n : \bar{z} = \bar{x} - 1, k - 1 \leq |\hat{z} - \hat{x}| < k\}| \leq \tilde{c}, \tag{5.4}$$

and then N so large that

$$\tilde{c} + N^{d-1-\nu} < 1.$$

Scaling limit of ballistic self-avoiding walk interacting with spatial random permutations

Fix $n \geq N$, $y \in \Lambda_n$, $A \in \mathcal{A}_y^w$, $\gamma_0 \in \mathcal{L}_A^{y \to \ell_n}$, and a pre-regeneration point x of γ_0 with $\bar{x} - \bar{y} \geq \log n$. The properties of A and x give $\{z \in \Lambda_n : \hat{z} = \hat{x}\} \subset A$. Given $\pi \in \mathcal{S}_A^{y \to \ell_n}$ with $\gamma(\pi) = \gamma_0$, we write $\mathsf{R}(x, \pi)$ for the set of all subsets R of A that satisfy properties (R1) to (R4) of regeneration sets. Since x is a pre-regeneration point of γ_0, we have $\mathrm{Orb}_\pi(x) \subset \mathcal{C}_x$ and $\gamma_0 \setminus \mathrm{Orb}_\pi(x) \subset \{z : \bar{z} < \bar{x}\}$ for all π such that $\gamma(\pi) = \gamma_0$. Since on the other hand $R \supset \mathcal{C}_x$ for all $R \in \mathsf{R}(x, \pi)$, we have

$$\mathbb{P}_A^{y \to \ell_n}(x \in \mathcal{R} | \gamma = \gamma_0) = \mathbb{P}_{A \setminus \gamma_0}(\mathsf{R}(x, .) \neq \emptyset).$$

Let Φ be the group of reflections in Λ_n that leave the line $\{z : \hat{z} = \hat{x}\}$ invariant.

We want to use Propositions 4.8 and 4.9, but we need to slightly adjust the sampling strategy leading to equation (4.15). The reason is that the set A, which plays the role of V in the context of the sampling strategy, is itself not necessarily invariant under Φ. We thus replace the strategy (4.14) by the strategy where $K_A = \Phi(A) \cap A$, and where for $B \subset A$

$$K_B = \begin{cases} B \cap \Phi(B^c) \cap A & \text{if } B \cap \Phi(B^c) \cap A \neq \emptyset \\ B & \text{otherwise} \end{cases}$$

Applying this sampling strategy with starting set $\ell_x^- := \{z \in \Lambda_n : \bar{z} < \bar{x}\}$ will produce the minimal π-invariant subset $Q_{\ell_x^-}(\pi)$ of A such that $\phi(Q_{\ell_x^-}(\pi)) \cap A \subset Q_{\ell_x^-}(\pi)$ for all $\phi \in \Phi$. Since $A \setminus \ell_x^-$ is invariant under Φ, so is $A \setminus Q_{\ell_x^-}(\pi)$, and thus

$$\mathsf{R}(x, \pi) \neq \emptyset \qquad \text{iff} \qquad Q_{\ell_x^-}(\pi) \cap \mathcal{C}_x = \emptyset,$$

and in this case $R(x, \pi) = A \setminus Q_{\ell_x^-}(\pi)$. On the other hand, it is easily checked that Propositions 4.8 and 4.9 stay valid for the new sampling strategy. Since $Q_{\ell_x^-}(\pi) \setminus \ell_x^- = (\bigcup_{z \in \Lambda_n : \bar{z} = \bar{x} - 1} Q_{\{z\}}(\pi)) \setminus \ell_x^-$ and $\mathcal{C}_x \cap \ell_x^- = \emptyset$, we find

$$\mathbb{P}_{A \setminus \gamma_0}(Q_{\ell_x^-}(\pi) \cap \mathcal{C}_x \neq \emptyset) \leq \sum_{z : \bar{z} = \bar{x} - 1} \mathbb{P}_{A \setminus \gamma_0}(Q_{\{z\}} \cap \mathcal{C}_x \neq \emptyset) \tag{5.5}$$

$$\leq \sum_{z : \bar{z} = \bar{x} - 1} \mathbb{P}_{A \setminus \gamma_0}(|Q_{\{z\}}| > \min\{|\hat{z} - \hat{x}|, \log n\}). \tag{5.6}$$

Theorem 2.2 and Propositions 4.8 and 4.9 guarantee that we can choose α_0, independently of A, so that for all $\alpha > \alpha_0$, all z with $\bar{z} = \bar{x} - 1$ and all $k \geq 1$, we have that

$$\mathbb{P}_{A \setminus \gamma_0}(|Q_{\{z\}}| > k) \leq e^{-\nu k}.$$

Then by condition (5.4), we find

$$\mathbb{P}_{A \setminus \gamma_0}(Q_{\ell_x^-}(\pi) \cap \mathcal{C}_x \neq \emptyset) \leq \tilde{c} + \sum_{z : \bar{z} = \bar{x} - 1} \mathbb{P}_{A \setminus \gamma_0}(|Q_{\{z\}}| > \log n) \leq \tilde{c} + n^{d-1} n^{-\nu} < 1. \tag{5.7}$$

Thus the claim holds for $c = 1 - \tilde{c} - n^{d-1-\nu} > 0$. $\qquad \square$

We are now ready to give the decisive statement of the first step of the proof, saying that with high probability, the first regeneration point being more than $\log n$ away from y is not further than a constant times $\log n$ away from y. For $y \in \Lambda_n$ and a strictly admissible set A, we define the $\mathbb{P}_A^{y \to \ell_n}$ random variable

$$X_1(\pi) = \min\{x \in \mathcal{R}(\pi) : \bar{x} \geq \bar{y} + \log n\}$$

Proposition 5.4. *Let $p \geq 1$. Then there exist $\alpha_0 < \infty$, $C < \infty$ and $N \in \mathbb{N}$ so that for all $y \in \Lambda_n$, $n > N$, $\alpha > \alpha_0$, $A \in \mathcal{A}_y^s$, $k \geq 1$ we have*

$$\mathbb{P}_A^{y \to \ell_n}\left(|X_1 - y| > k \log n \,\middle|\, \gamma \in \mathcal{C}_y\right) \leq C k^{-p}.$$

Proof. Clearly, it suffices to prove the claim for large k, and the claim trivially holds whenever $n - \bar{y} \leq k \log n$, since the last point of γ is a regeneration point by convention, and we conditioned on $\gamma \in \mathcal{C}_y$. Let thus $k \geq 2$, $\delta < 1/4$, and $L = k \log n$. Recalling the notation used in Lemma 5.1, we decompose

$$\mathbb{P}_A^{y \to \ell_n}(|X_1 - y| > k \log n | \gamma \subset \mathcal{C}_y) \leq \mathbb{P}_A^{y \to \ell_n}(|\rho_L| \geq (1 + \delta)L | \rho_L \subset \mathcal{C}_y) +$$
$$+ \mathbb{P}_A^{y \to \ell_n}(|X_1 - y| > k \log n, |\rho_L| < (1 + \delta)L | \rho_L \subset \mathcal{C}_y). \tag{5.8}$$

Here we used that $\rho_L \subset \mathcal{C}_y$ is equivalent to $\gamma \subset \mathcal{C}_y$ by the definitions of $\mathcal{C}_y$ and ρ_L and the fact that $L \geq 2 \log n$. Since $\mathcal{C}_y \subset \mathcal{A}_y$, we can apply Proposition 5.1 and find

$$\mathbb{P}_A^{y \to \ell_n}(|\rho_L| \geq (1 + \delta)L | \rho_L \subset \mathcal{C}_y) \leq C \, \mathrm{e}^{-ck \log n} \leq C \, \mathrm{e}^{-ck}.$$

uniformly in k, y and $A \in \mathcal{A}_y$ for all large enough α and n.

It remains to estimate the second term in (5.8). For this, we use the trivial equality

$$\mathbb{P}_A^{y \to \ell_n}(|X_1 - y| > k \log n, |\rho_L| < (1 + \delta)L | \rho_L \subset \mathcal{C}_y)$$
$$= \frac{\sum_{\rho \subset \mathcal{C}_y : |\rho| < (1+\delta)L} \mathbb{P}_A^{y \to \ell_n}(|X_1 - y| > K \log n | \rho_L = \rho) \, \mathbb{P}_A^{y \to \ell_n}(\rho_L = \rho)}{\sum_{\rho \subset \mathcal{C}_y} \mathbb{P}_A^{y \to \ell_n}(\rho_L = \rho)}$$

in order to see that the claim will be shown once we prove that there exists $\tilde{C} < \infty$ such that uniformly in ε, n and $\rho \subset A$ such that $|\rho| < (1 + \delta)L$ and such that ρ is the trace of some self-avoiding walk from y to some element of $\ell_{\bar{y}+L}$, we have

$$\mathbb{P}_A^{y \to \ell_n}(|X_1 - y| > k \log n | \rho_L = \rho) \leq \tilde{C} \, k^{-p}. \tag{5.9}$$

Lemma 5.2 tells us that such a ρ must have at least $(1 - 3\delta)k \log n \geq \frac{1}{4}k \log n$ pre-regeneration points. For $k > 8$, at least $(k/4 - 1) \log n \geq \frac{k}{8} \log n$ of them satisfy $\bar{x} - \bar{y} \geq \log n$. If any of these is a regeneration point, then X_1 is the minimal such point, and thus $|X_1 - y| \leq k \log n$.

Let therefore ρ be the trace of some self-avoiding walk from y to an element of $\ell_{\bar{y}+L}$, assume that it has $M > \log n$ pre-regeneration points that have horizontal distance larger than $\log n$ from y, and denote these points by $(x_i)_{i \leq M}$. By Proposition 5.3, each x_i is a regeneration point with uniformly positive probability. If the events $\{\pi : x_i \in \mathcal{R}(\pi)\}$ would be independent, this would already conclude the proof, by the argument of independent trials. Unfortunately, they are not independent, and so we will have to work harder. (Although we can not prove it, it is actually reasonable to conjecture that the events $\{x_i \notin \mathcal{R}\}$ are positively correlated since the fact that x_i is a pre-regeneration point but not a regeneration point means that there is some long cycle of π that prevents the existence of a regeneration surface; such a long cycle may still impact the next pre-regeneration point as well.)

The solution is to separate pre-regeneration points by invariant sets. Let $r_1 = x_1$, and

$$r_j = \min\{x_j : \bar{x}_j > \bar{r}_{j-1} + 2 \log n\}.$$

Since we started with at least $\frac{1}{8}k \log n$ pre-regeneration points (x_i), and since at most $2 \log n$ of them are between two consecutive r_i, we still retain at least $\tilde{k} \geq k/16$ points $(r_i)_{i \leq \tilde{k}}$. For $\pi \in \mathcal{S}_{A \setminus \gamma}$, denote again by $R(x, \pi)$ the set of all subsets of A satisfying properties (R1) to (R4). With

$$\mathcal{N}(x) := \{\pi \in \mathcal{S}_{A \setminus \gamma} : R(x, \pi) = \emptyset\}, \quad \mathcal{N}_i := \mathcal{N}(r_i),$$

for any $\gamma \subset A$ so that $\rho_L(\gamma) = \rho$, we then have

$$\mathbb{P}_A^{y \to \ell_n}(|X_1 - y| > K \log n | \rho_L = \rho) = \mathbb{P}_{A \setminus \gamma}\Big(\bigcap_{j \geq 1} \mathcal{N}(x_j)\Big) \leq \mathbb{P}_{A \setminus \gamma}\Big(\bigcap_{i=1}^{\tilde{k}} \mathcal{N}_i\Big).$$

We further define $w_i := \bar{r}_i + \log n$, and

$$\ell_{w_i}^- := \{z \in A : \bar{z} < \bar{w}_i\}, \qquad Q_i := \mathrm{Orb}_\pi(\ell_{w_i}^- \cap (A \setminus \gamma)), \qquad W_i := \{z \in A : \bar{z} \geq \bar{w}_i + \log n\}.$$

Let $m \leq \tilde{k}$. We decompose,

$$\mathbb{P}_{A \setminus \gamma}\Big(\bigcap_{i \geq 1}^m \mathcal{N}_i\Big) \leq \mathbb{P}_{A \setminus \gamma}(Q_{m-1} \cap W_{m-1} \neq \emptyset) + \mathbb{P}_{A \setminus \gamma}\Big(Q_{m-1} \cap W_{m-1} = \emptyset, \bigcap_{i \geq 1}^m \mathcal{N}_i\Big).$$

By the same reasoning that led to the bound on the second term in (5.7), the first term above is bounded by $n^{d-1-\nu}$, where we can choose ν as large as needed by making α large. For the second term we use the strong Markov property of SRP. Let $f(\pi) = \mathbb{1}\{\pi \in \mathcal{N}_m\}$, and note that $\bigcap_{i=1}^{m-1} \mathcal{N}_i \in \mathcal{F}_{Q_{m-1}}$ by the definition of the Q_i and the strict admissibility of regeneration sets. Thus

$$\mathbb{P}_{A \setminus \gamma}\Big(Q_{m-1} \cap W_{m-1} = \emptyset, \bigcap_{i \geq 1}^m \mathcal{N}_i\Big)$$
$$= \mathbb{E}_{A \setminus \gamma}\Big(\mathbb{E}_{A \setminus \gamma}(f | \mathcal{F}_{Q_{m-1}}) \mathbb{1}\{Q_{m-1} \cap W_{m-1} = \emptyset\} \mathbb{1}\{\bigcap_{i=1}^{m-1} \mathcal{N}_i\}\Big),$$

and Proposition 4.2 gives

$$\mathbb{E}_{A \setminus \gamma}(f | \mathcal{F}_{Q_{m-1}}) \mathbb{1}\{Q_{m-1} \cap W_{m-1} = \emptyset\} = \mathbb{P}_{Q_{m-1}}(\mathcal{N}_m) \mathbb{1}\{Q_{m-1} \cap W_{m-1} = \emptyset\}$$

almost surely. By Proposition 5.3 we find $\mathbb{P}_{Q_{m-1}(\pi)}(\mathcal{N}_m) \leq 1 - c$ uniformly in all allowed $Q_{m-1}(\pi)$ (note that by definition, $Q_{m-1}(\pi)$ is weakly admissible for all π such that $Q_{m-1}(\pi) \cap W_{m-1} = \emptyset$), and thus we conclude

$$\mathbb{P}_{A \setminus \gamma}\Big(\bigcap_{i=1}^m \mathcal{N}_i\Big) \leq (1-c)\mathbb{P}_{A \setminus \gamma}\Big(\bigcap_{i=1}^{m-1} \mathcal{N}_i\Big) + n^{d-1-\eta},$$

for all $m \leq \tilde{k}$. Therefore, by induction,

$$\mathbb{P}_{A \setminus \gamma}\Big(\bigcap_{i=1}^{\tilde{k}} \mathcal{N}_i\Big) \leq \sum_{j=0}^\infty (1-c)^j n^{d-1-\eta} = \frac{n^{d-1-\eta}}{c}.$$

It remains to choose α_0 so large that for $\alpha > \alpha_0$, we have $d - 1 - \eta < -p$. Then for sufficiently large n_0 and all $n > n_0$, we have

$$\mathbb{P}_A^{y \to \ell_n}(|X_1 - y| > k \log n | \gamma \subset \mathcal{C}_y) \leq C\, e^{-ck} + \frac{1}{c} n^{-p}.$$

For $k \leq n$, this proves the claim of Proposition 5.4; the last step in the proof consist in observing that for $k > n$, the required probability is trivially equal to zero when $k > d$. Thus Proposition 5.4 is proved. $\qquad \square$

5.3 The Markov chain of regeneration pairs

We define for each $\pi \in \mathcal{S}_{\Lambda_n}^{0 \to \ell_n}$ the following sequence of pairs $(x_i, R_i) \subset \Lambda_n \times \mathcal{P}(\Lambda_n)$: We set $X_0 = 0$ and $R_0 = \Lambda_n$, and then recursively define

$$X_{i+1}(\pi) = \begin{cases} \min\{x \in \mathcal{R}(\pi) : \bar{x} \geq \bar{X}_i(\pi) + \log n\} & \text{if } \bar{X}_i(\pi) + \log n \leq n, \\ z & \text{otherwise.} \end{cases}$$

In words, we pick a sequence of regeneration points that have mutual horizontal distance of at least $\log n$; as soon as we find, by this procedure, a regeneration point with horizontal distance less than $\log n$ to ℓ_n, we jump to $\max \gamma$ and stay there forever. In addition, we define $R_i(\pi) = R(X_i(\pi), \pi)$. We refer to the pairs (X_i, R_i) as regeneration pairs. Note that for $i = 1$, the definition of X_1 coincides with the one given in the previous part of the proof.

Proposition 5.5. *The sequence* (X_i, R_i) *is a* $\Lambda_n \times \mathcal{P}(\Lambda_n)$*-valued Markov chain, with transition matrix*

$$p\big((x, R), (x', R')\big) = \begin{cases} \mathbb{P}_{\Lambda_n}^{0 \to \ell_n}(X_1 = x', R_1 = R') & \text{if } (x, R) = (0, \Lambda_n), \\ \mathbb{P}_R^{x \to \ell_n}(y = \max \gamma) & \text{if } \bar{x} + \log n \geq n, \\ \mathbb{P}_R^{x \to \ell_n}\big(X_1 = x', R_1 = R' \big| \gamma \subset \mathcal{C}_x\big) & \text{otherwise.} \end{cases}$$

For the proof of this proposition, we need a variant of the spatial Markov property that we are now going to state. The proof of Proposition 5.5 is presented afterwards.

For any set $A \subset V = \Lambda_n$, we define $\tilde{\mathcal{F}}_A$ to be the σ-algebra over $\mathcal{S}_V^{0 \to \ell_n}$ generated by the forward evaluations $\pi \mapsto \pi(z)$ for $z \in A$. Note that in contrast to the situation in Proposition 4.1, we do not consider inverse images since these may not be defined in the situation with an open cycle. For any $\pi \in S_V^{0 \to \ell_n}$, we define the family of π-almost invariant sets

$$\mathrm{Inv}_0(\pi) := \big\{A \subset V : \pi(A) \subset A\big\}.$$

By this definition, the open cycle γ can enter an almost-invariant set A but not leave it, and if $A \cap \gamma \neq \emptyset$ then the image of A is A without the unique entrance point of γ into A. The reader should be warned that in the current situation, $A \in \mathrm{Inv}_0(\pi)$ does not imply $A^c \in \mathrm{Inv}_0(\pi)$.

Let $A \subset V$, $x \in A$, $a \in A^c$, and $\pi \in \mathcal{S}_V^{a \to \ell_n}$ such that $A \in \mathrm{Inv}_0(\pi)$ and $x = \min \gamma(\pi) \cap A$. Then there exist (unique) $\pi|_A \in \mathcal{S}_A^{x \to \ell_n}$ and $\pi|_{A^c} \in \mathcal{S}_{A^c \cup \{x\}}^{a \to x}$ such that $\pi|_A(z) = \pi(z)$ for all $z \in A$ and $\pi_{A^c}(z) = \pi(z)$ for all $z \in A^c$. In this situation, for $\tilde{\mathcal{F}}_{A^c}$-measurable $f : \mathcal{S}_V^{0 \to \ell_n} \to \mathbb{R}$, we use the same symbol to denote the function $f : \mathcal{S}_{A^c \cup \{x\}}^{a \to x} \to \mathbb{R}$ with $f(\pi) = f(\pi_{A^c})$. In the same way, we get $g(\pi) = g(\pi_A)$ for $\tilde{\mathcal{F}}_A$-measurable g.

Proposition 5.6. *Let* $A \subset V$, $a \in A^c$, $x \in A$. *Then,*

(i): $\mathbb{P}_V^{a \to \ell_n}(A \in \mathrm{Inv}_0, x = \min \gamma \cap A) = Z^{a \to x}(A^c \cup \{x\})Z^{x \to \ell_n}(A)/Z^{a \to z}(V).$

(ii): *For* $\tilde{\mathcal{F}}_{A^c}$*-measurable* f *and* $\tilde{\mathcal{F}}_A$*-measurable* g, *we have*

$$\mathbb{E}_V^{a \to \ell_n}(f|A \in \mathrm{Inv}_0, x = \min \gamma \cap A) = \mathbb{E}_{A^c \cup \{x\}}^{a \to x}(f)$$

and

$$\mathbb{E}_V^{a \to \ell_n}(g|A \in \mathrm{Inv}_0, x = \min \gamma \cap A) = \mathbb{E}_A^{x \to \ell_n}(g).$$

(iii): *For* $\tilde{\mathcal{F}}_{A^c}$*-measurable* f *and* $\tilde{\mathcal{F}}_A$*-measurable* g, *we have*

$$\mathbb{E}_V^{a \to \ell_n}(f \, g|A \in \mathrm{Inv}_0, x = \min \gamma \cap A) = \mathbb{E}_{A^c \cup \{x\}}^{a \to x}(f) \, \mathbb{E}_A^{x \to \ell_n}(g).$$

(iv): For $\tilde{\mathcal{F}}_A$-measurable g and $\mathcal{Q} \in \tilde{\mathcal{F}}_{A^c}$, we have

$$\mathbb{E}_V^{a \to \ell_n}(g | A \in \mathrm{Inv}_0, x = \min \gamma \cap A, \mathcal{Q}) = \mathbb{E}_A^{x \to \ell_n}(g). \tag{5.10}$$

Proof. Let $\mathcal{H}_B(\pi) := \sum_{x \in B} |\pi(x) - x|$. Then for all $A \subset V$, all $x \in A$ and all $\pi \in \mathcal{S}_V^{a \to \ell_n}$ such that $A \in \mathrm{Inv}_0(\pi)$ and $x = \min \gamma \cap A$, we have

$$\mathcal{H}_V(\pi) = \mathcal{H}_{A^c \cup \{x\}}(\pi|_{A^c}) + \mathcal{H}_A(\pi|_A).$$

This holds since $\pi|_{A^c}(x) = x$. Consequently, for $\tilde{\mathcal{F}}_{A^c}$ measurable f and $\tilde{\mathcal{F}}_A$-measurable g, we have

$$\sum_{\substack{\pi \in S_V^{a \to \ell_n}: \\ A \in \mathrm{Inv}_p(\pi), x = \min \gamma(\pi) \cap A}} f(\pi) g(\pi) \, \mathrm{e}^{-\alpha \mathcal{H}_V(\pi)}$$
$$= \sum_{\pi \in S_{A^c \cup \{x\}}^{a \to x}} f(\pi|_{A^c}) \, \mathrm{e}^{-\alpha \mathcal{H}_{A^c \cup \{x\}}(\pi_{A^c})} \sum_{\tilde{\pi} \in S_A^{x \to \ell_n}} g(\tilde{\pi}_A) \, \mathrm{e}^{-\alpha \mathcal{H}_A(\pi|_A)}. \tag{5.11}$$

The proof now follows the same steps as the proof of Proposition 4.1. □

Proof of Proposition 5.5. Fix $x_1, \dots x_{k+1} \in \Lambda_n$ and subsets $Q_1, \dots, Q_{k+1}$ of Λ_n which satisfy the deterministic conditions for regeneration points, in particular for all $i \le k+1$,

$$\bar{x}_i \ge \bar{x}_{i-1} + \log n, \qquad \text{and} \qquad \mathcal{C}_{x_i} \subset Q_i \subset \{z : \bar{z} \ge \bar{x}_i\}. \tag{5.12}$$

Let us write $\mathcal{J}_i = \{X_i = x_i, R_i = Q_i\}$ for brevity, and write

$$C(x, y) = \{\pi : y = \min\{w \in \mathcal{R}(\pi) : \bar{w} \ge \bar{x} + \log n\}\}.$$

Then

$$\mathbb{P}_{\Lambda_n}^{0 \to \ell_n}\left(X_{k+1} = x_{k+1}, R_{k+1} = Q_{k+1} \Big| \bigcap_{i=1}^k \mathcal{J}_i\right) = \mathbb{P}_{\Lambda_n}^{0 \to \ell_n}\left(C(x_k, x_{k+1}), Q_{k+1} = R(x_{k+1}) \Big| \bigcap_{i=1}^k \mathcal{J}_i\right)$$

$$= \mathbb{P}_{\Lambda_n}^{0 \to \ell_n}\left(C(x_k, x_{k+1}), R(x_{k+1}) = Q_{k+1} \Big| \mathrm{Orb}(x_k) \subset \mathcal{C}_{x_k}, \gamma \setminus \mathrm{Orb}(x_k) \subset \{z : \bar{z} < \bar{x}_k\}, R_k = Q_k, \bigcap_{i=1}^{k-1} \mathcal{J}_i\right)$$

$$= \frac{\mathbb{P}_{\Lambda_n}^{0 \to \ell_n}\left(C(x_k, x_{k+1}), R(x_{k+1}) = Q_{k+1}, \mathrm{Orb}(x_k) \subset \mathcal{C}_{x_k} \Big| \gamma \setminus \mathrm{Orb}(x_k) \subset \{z : \bar{z} < \bar{x}_k\}, R_k = Q_k, \bigcap_{i=1}^{k-1} \mathcal{J}_i\right)}{\mathbb{P}_{\Lambda_n}^{0 \to \ell_n}\left(\mathrm{Orb}(x_k) \subset \mathcal{C}_{x_k} \Big| \gamma \setminus \mathrm{Orb}(x_k) \subset \{z : \bar{z} < \bar{x}_k\}, R_k = Q_k, \bigcap_{i=1}^{k-1} \mathcal{J}_i\right)}.$$

The event $C(x_k, x_{k+1}) \cup \{R_{k+1} = R', \mathrm{Orb}(x_k) \subset \mathcal{C}_{x_k}\}$ is $\tilde{\mathcal{F}}_{\mathcal{C}_{x_k}}$-measurable, and since $\mathcal{C}_{x_k} \subset Q_k$, it is $\tilde{\mathcal{F}}_{Q_k}$-measurable. On the other hand, the event

$$\mathcal{Q} := \{\gamma \setminus \mathrm{Orb}(x_k) \subset \{z : \bar{z} < \bar{x}_k\}, R_k = Q_k, \bigcap_{i=1}^{k-1} \mathcal{J}_i\}$$

is $\tilde{\mathcal{F}}_{Q_k^c \setminus \{x_k\}}$-measurable, and contains the set $\{Q_k \in \mathrm{Inv}_{x_k}\} \cup \{x_k = \min \gamma \cap Q_k\}$. We can therefore apply part (iv) of Proposition 5.6 with $A = Q_k$ to both numerator and denominator above, and find

$$\mathbb{P}_{\Lambda_n}^{0 \to \ell_n}\left(X_{k+1} = x', R_{k+1} = Q_{k+1} \Big| \bigcap_{i=1}^k \mathcal{J}_i\right) = \mathbb{P}_{Q_k}^{x_k \to \ell_n}\left(X_1 = x_{k+1}, R_1 = Q_{k+1} \Big| \gamma \subset \mathcal{C}_{x_k}\right),$$

as claimed. □

We finally have all the pieces to conclude the proof of Theorem 2.4. By the symmetry condition (R3) of regeneration sets, the process $(\hat{X}_i)$, while not being a Markov process, is a $\mathbb{Z}^{d-1}$-valued martingale under $\mathbb{P}^{0\to\ell_n}_{\Lambda_n}$. By Doob's L^2-inequality, we have

$$\mathbb{E}(\max_{j\le n_0}|\hat{X}_j|^2) \le 4\mathbb{E}(|\hat{X}_{n_0}|^2) = 4\sum_{j=1}^{n_0}\mathbb{E}((|\hat{X}_j - \hat{X}_{j-1}|^2)). \qquad (5.13)$$

Using Proposition 5.4 with $p > 2$, we find that for suitable α_c, C and N,

$$\mathbb{E}^{0\to\ell_n}_{\Lambda_n}\Big(\frac{|X_j - X_{j-1}|^2}{(\log n)^2}\Big|X_j = y, R(X_j) = R\Big) \le \int_1^\infty \mathbb{P}^{y\to\ell_n}_R(|X_1| \ge k\log n)k\,\mathrm{d}k$$

$$\le C\int_1^\infty k^{1-p}\,\mathrm{d}k = C/(p-2) =: C_p$$

uniformly in $n > N$, $\alpha > \alpha_0$ and $R \in \mathcal{A}^s_y$. By integrating over possible y and R we conclude $\mathbb{E}(|\hat{X}_j - \hat{X}_{j-1}|^2) \le (\log n)^2 C_p$. Since $\bar{X}_{i+1} \ge \bar{X}_i + \log n$ by construction, the martingale takes at most $n/\log n$ steps before hitting the point $\lim_{n\to\infty}\pi^n(0)$, after which it does no longer move. This, Chebyshevs inequality and (5.13) give

$$\mathbb{P}(\max_{i\in\mathbb{N}}|\hat{X}_i| > M\sqrt{n\log n}) \le \frac{1}{M^2 n\log n}\mathbb{E}(\max_{i\in\mathbb{N}}|\hat{X}_i|^2) \le \frac{1}{M^2 n\log n}4\frac{n}{\log n}C_p(\log n)^2 = \frac{4C_p}{M^2}. \qquad (5.14)$$

In conclusion, we now know that with high probability none of the special regeneration points (X_i) is further than $\sqrt{n\log n}$ away from the line $\{\hat{x} = 0\}$.

In order to control the parts of γ in between the X_i, we first apply Proposition 5.4 with $k = n^{1/4}$ and $p = 8$, and find

$$\mathbb{P}^{0\to\ell_n}_{\Lambda_n}(|\bar{X}_{i+1} - \bar{X}_i| > n^{1/4}\log n|X_i = y, R_i = A) \le C_8 n^{-2}$$

for all $y \in \Lambda_n$ and all $A \in \mathcal{A}^s_y$. Since there are at most $n/\log n$ regeneration points, the union bound then gives (for n large enough)

$$\mathbb{P}^{0\to\ell_n}_{\Lambda_n}(\max_i|\bar{X}_{i+1} - \bar{X}_i| > n^{1/3}) \le C_8 n^{-1}.$$

We thus find

$$\mathbb{P}^{0\to\ell_n}_{\Lambda_n}(\max_{x\in\gamma}|\hat{x}| > 2M\sqrt{n\log n}) \le \frac{4C_p}{M^2} + \frac{C_8}{n} +$$

$$+ \mathbb{P}^{0\to\ell_n}_{\Lambda_n}\Big(\max_{x\in\gamma}|\hat{x}| > 2M\sqrt{n\log n}, \max_i|\bar{X}_{i+1} - \bar{X}_i| < n^{1/3}, \max_i|\hat{X}_i| < M\sqrt{n\log n}\Big). \qquad (5.15)$$

By definition, at a regeneration point γ crosses a certain vertical hyperplane for the last time. Applying Proposition 5.1 with $\delta = 1$, $B = \Lambda_n$ and $L = n^{1/3}$, we see that the probability that the piece of γ between X_i and X_{i+1} is longer than $2n^{1/3}$ is less than $C\,\mathrm{e}^{-cn^{1/3}}$ for each i. On the other hand, a path would need at least $2M\sqrt{n\log n}$ steps to exceed level $2M\sqrt{n\log n}$ when starting (and ending) below level $M\sqrt{n\log n}$. Thus the second line in (5.15) is bounded by $n\,\mathrm{e}^{-cn^{1/3}}$. Taking n so large that $n > M$ and $n\,\mathrm{e}^{-cn^{1/3}} < \frac{1}{M}$, Theorem 2.4 is proved. $\qquad\square$

Let us conclude by looking back over the proof and identifying the place where our estimates are not good enough to provide diffusive scaling. The most basic place where this happens is Proposition 5.1: here already, we can only prove a bound on the length of ρ_L (and thus many regeneration points) when L is of the order of $\log n$. This in turn is due to the inequality (5.1), which needs the separation of the part of γ to the right of $\ell_{\bar{h}}$ from

the set $\Lambda_n \setminus A$ by at least a distance of $\log n$. This is because (5.1) relies on Proposition 4.10, and in this result the sum that leads to the constant D needs to be controlled by sufficiently large distances between A and $V_0 \setminus V_1$, especially when the number of points in either of them diverges. The root cause of the problem is that while we do have exponential decay of correlations, the decay is not uniform in the size of the sets between which the correlations are measured. See also the discussion in the paragraph before Proposition 4.10. The latter proposition already goes some way towards solving this problem, but as it turns out it is not quite enough for obtaining optimal results.

Appendix

Given two finite or countable infinite sequences of non-negative integer-valued random variables, $\left(M_n^i\right)_{n \in I \subset \mathbb{N}}$, $i = 1, 2$, we say that $\left(M_n^2\right)_{n \in I}$ is stochastically larger than $\left(M_n^1\right)_{n \in I}$ and we write

$$\left(M_n^1\right)_{n \in I} \overset{d}{\preceq} \left(M_n^1\right)_{n \in I},$$

if for any finite sequence of non-negative integers n_0, n_1, ..., n_k, we have that,

$$P_1(M_0^1 > n_0, M_1^1 > n_1, \ldots M_k^1 > n_k) \leq P_2(M_0^2 > n_0, M_1^2 > n_1, \ldots M_k^2 > n_k),$$

where $P_i(\,\cdot\,)$ is the law of the i-th sequence, $i = 1, 2$. We are now ready to state our comparison Lemma.

Lemma 5.7 (Comparison Lemma). *Let $\left(M_n^i\right)_{n \in \mathbb{N}}$, $i \in \{1, 2\}$, be two non-negative, integer valued stochastic processes, defined in two different probability spaces, $(\Sigma_1, \mathcal{F}^1, P_1)$ and $(\Sigma_2, \mathcal{F}^2, P_2)$. Let $\left(\mathcal{F}_n^i\right)_{n \in \mathbb{N}}$, $i \in \{1, 2\}$ be their filtrations. Assume that $\left(M_n^2\right)_{n \in \mathbb{N}}$ is a Markov chain and let $\left(M_n^{2,\lambda}\right)_{n \in \mathbb{N}}$ be used to denote the Markov chain with initial state M_0^2 distributed according to the probability measure on the non-negative integers $\lambda(\,\cdot\,)$. Assume that*

- *(**Ass. 1**) The Markov chain is such that,*

$$\lambda \overset{d}{\preceq} \lambda' \implies \left(M_n^{2,\lambda}\right)_{n \in \mathbb{N}} \overset{d}{\preceq} \left(M_n^{2,\lambda'}\right)_{n \in \mathbb{N}}$$

- *(**Ass. 2**) For all integers $\ell, m, n \geq 0$, and $\omega \in \Sigma^1$ such that $M_n^1(\omega) = m$,*

$$P_1(M_{n+1}^1 > \ell \mid \mathcal{F}_n^1)(\omega) \leq P_2(M_{n+1}^2 > \ell | M_n^2 = m).$$

- *(**Ass. 3**) For a given initial distribution $\lambda^*(\,\cdot\,)$,*

$$M_0^1 \overset{d}{\preceq} M_0^{2,\lambda^*}.$$

Then, we have that

- *(**Concl. a**) The Markov chain $(M_n^{2,\lambda^*})_{n \in \mathbb{N}}$ is stochastically larger than $(M_n^1)_{n \in \mathbb{N}}$,*

$$(M_n^1)_{n \in \mathbb{N}} \overset{d}{\preceq} (M_n^{2,\lambda^*})_{n \in \mathbb{N}},$$

- *(**Concl. b**) For all $k \in \mathbb{N}$,*

$$\sum_{j=0}^{k} M_j^1 \overset{d}{\preceq} \sum_{j=0}^{k} M_j^{2,\lambda^*}.$$

Proof. Let us denote by $\overline{P}(\cdot) = P_1(\cdot) \times P_2(\cdot)$ and by $\overline{E}(\cdot)$ the expectation of $\overline{P}(\cdot)$. We start by proving the statement (a) of the lemma. Fix a finite sequence of non-negative integers, $n_1, n_2, \ldots, n_k$. We replace steps of the process 1 by steps of the process 2 one by one, showing that the probability of the event $\mathbb{1}\{ M_k^1 > n_k, M_{k-1}^1 > n_{k-1}, \ldots M_0^1 > n_0 \}$ cannot decrease at each replacement. Here is our first replacement,

$$P_1(M_k^1 > n_k, M_{k-1}^1 > n_{k-1}, \ldots M_0^1 > n_0) =$$
$$\overline{E}\left(\mathbb{1}\{M_k^1 > n_k, M_{k-1}^1 > n_{k-1}, \ldots M_0^1 > n_0\} \right) =$$
$$\overline{E}\left(\overline{E}(\mathbb{1}\{M_k^1 > n_k\} \,|\, \mathcal{F}_{k-1}^1) \, \mathbb{1}\{M_{k-1}^1 > n_{k-1}, \ldots M_0^1 > n_0\} \right) \leq$$
$$\overline{E}\left(\overline{E}(\mathbb{1}\{M_1^2 > n_k\} \,|\, M_0^2 = M_{k-1}^1) \, \mathbb{1}\{M_{k-1}^1 > n_{k-1}, \ldots M_0^1 > n_0\} \right) =$$
$$\overline{E}\left(\mathbb{1}\{M_1^{2,\delta_{M_{k-1}^1}} > n_k, M_{k-1}^1 > n_{k-1}, \ldots M_0^1 > n_0\} \right), \quad (5.16)$$

where for the second equality we used the definition of conditional expectation and for the first inequality we used our second assumption in the statement of the lemma. In the last equality we denote by $M_1^{2,\delta_{M_{k-1}^1}}$ the first step of the Markov chain which starts at time 0 from the value M_{k-1}^1 (then, the outer expectation integrates over the M_{k-1}^1). Here is our second replacement,

$$\overline{E}\left(\mathbb{1}\{M_1^{2,\delta_{M_{k-1}^1}} > n_k, M_{k-1}^1 > n_{k-1}, \ldots M_0^1 > n_0\} \right) =$$
$$\overline{E}\left(\overline{E}(\mathbb{1}\{M_1^{2,\delta_{M_{k-1}^1}} > n_k, M_{k-1}^1 > n_{k-1}\} \,|\, \mathcal{F}_{k-2}^1) \, \mathbb{1}\{M_{k-2}^1 > n_{k-2}, \ldots M_0^1 > n_0\} \right) \leq$$
$$\overline{E}\left(E^2(\mathbb{1}\{M_2^2 > n_k, M_1^2 > n_{k-1}\} \,|\, M_0^2 = M_{k-2}^1) \, \mathbb{1}\{M_{k-2}^1 > n_{k-2}, \ldots M_0^1 > n_0\} \right) =$$
$$\overline{E}\left(\mathbb{1}\{M_2^{2,\delta_{M_{k-2}^1}} > n_k, M_1^{2,\delta_{M_{k-2}^1}} > n_{k-1}, M_{k-2}^1 > n_{k-2}, \ldots M_0^1 > n_0\} \right). \quad (5.17)$$

In the expression on the left-hand side of the first inequality, $M_1^{2,\delta_{M_{k-1}^1}}$ can be interpreted as the step $n = 1$ of the Markov chain which starts from the initial state distributed according to M_{k-1}^1 conditioned on $\mathcal{F}_{k-2}^1$ (then, the outer expectation integrates over the $\mathcal{F}_{k-2}^1$). The inequality holds as we replace such an initial state for the Markov chain by a new state that, by our second assumption, is stochastically larger, namely by M_1^2 conditioned on $M_0^2 = M_{k-2}^1$. Then, our first assumption in the statement of the lemma guarantees that the whole Markov chain (not only the step 1, but also the following steps) is stochastically larger. After $k-1$ iterations we get the first inequality in the expression below.

$$P_1\left(M_k^1 > n_k, M_{k-1}^1 > n_{k-1}, \ldots M_0^1 > n_0 \right) \leq$$
$$\overline{E}\left(\mathbb{1}\{M_k^{2,\delta_{M_0^1}} > n_k, M_{k-1}^{2,\delta_{M_0^1}} > n_{k-1}, \ldots, M_1^{2,\delta_{M_0^1}} > n_2, M_0^1 > n_0\} \right) \leq$$
$$\overline{E}\left(\mathbb{1}\{M_k^{2,\lambda^*} > n_k, M_{k-1}^{2,\lambda^*} > n_{k-1}, \ldots, M_1^{2,\lambda^*} > n_2, M_0^{2,\lambda^*} > n_0\} \right), \quad (5.18)$$

For the last inequality we used our Assumptions 1 and 3. Indeed, on the left-hand side of the second inequality we have the Markov chain $(M_n^2)_{n \in \mathbb{N}}$ starting from initial state M_0^1, on the right-hand side we have the same Markov chain starting from initial state M_0^{2,λ^*}, which is stochastically larger than M_0^1 by our Assumption 3.

The proof of the part (b) of the lemma goes along the same replacement procedure as in the part (a). Let use define $W_k^i = \sum_{n=0}^{k} M_k^i$, $i \in \{1, 2\}$. We just illustrate our first and

second replacements below.

$$
\begin{aligned}
P_1\left(W_k^1 > n\right) &= E^1\Big(\mathbb{1}\{W_{k-1}^1 + M_k^1 > n\}\Big) \\
&= \overline{E}\Big(\; E^1\big(\; \mathbb{1}\{M_k^1 > n - W_{k-1}^1\} \mid \mathcal{F}_{k-1}^1, M_{k-1}^1 \;\big) \;\Big) \\
&\leq \overline{E}\Big(\; \overline{E}\big(\; \mathbb{1}\{M_1^2 > n - W_{k-1}^1\} \mid\; M_0^2 = M_{k-1}^1, \mathcal{F}_{k-1}^1 \;\big) \;\Big) \\
&= \overline{E}\Big(\; \mathbb{1}\{M_1^{2,\delta_{M_{k-1}^1}} + W_{k-1}^1 + > n\} \;\Big),
\end{aligned}
\tag{5.19}
$$

where we used our Assumption 2 for the inequality. Here is the second step, where we use our Assumptions 1 and 2.

$$
\begin{aligned}
&\overline{E}\Big(\; \mathbb{1}\{M_1^{2,\delta_{M_{k-1}^1}} + M_{k-1}^1 + W_{k-2}^1 > n\} \;\Big) \\
&= \overline{E}\Big(\; \overline{E}\big(\mathbb{1}\{M_1^{2,\delta_{M_{k-1}^1}} + M_{k-1}^1 > n - W_{k-2}^1\} \mid \mathcal{F}_{k-2}^1, M_{k-2}^1\big) \;\Big) \\
&= \overline{E}\Big(\; \overline{E}\big(\mathbb{1}\{M_2^{2,\delta_{M_{k-2}^1}} + M_1^{2,\delta_{M_{k-2}^1}} > n - W_{k-2}^1\} \mid \mathcal{F}_{k-2}^1, M_{k-2}^1\big) \;\Big) \\
&\leq \overline{E}\Big(\; \mathbb{1}\{M_2^{2,\delta_{M_{k-2}^1}} + M_1^{2,\delta_{M_{k-2}^1}} + W_{k-2}^1 > n\} \;\Big).
\end{aligned}
$$

The proof of the statement (b) of our Comparison Lemma follows by iteration. For the last step, we use our Assumption 3, the same as for the proof of our statement (a). □

The last result that we present in this section is that the simple Galton-Watson process is a Markov chain satisfying the Assumption 1 on the process $(M_j^2)_{j\in\mathbb{N}}$ in Lemma 5.7.

Lemma 5.8. *If Ξ and Ξ' are two independent random variables such that $\Xi \stackrel{d}{\preceq} \Xi'$, then $(Z_j^\Xi)_{j\in\mathbb{N}} \stackrel{d}{\preceq} (Z_j^{\Xi'})_{j\in\mathbb{N}}$, where $(Z_j^\Xi)_{j\in\mathbb{N}}$ is a simple Galton-Watson with offspring distribution ξ.*

We are now ready to prove the previous lemma.

Proof of Lemma 5.8. We couple the two Galton-Watson process in the same probability space. In this probability space we define an infinite array of independent random variables $\boldsymbol{\xi} = (\xi_{j,k})_{j\geq 1,\, k\geq 1}$ having the same distribution as ξ. For any $m \in \mathbb{N}$ and $\boldsymbol{\xi}$, we define the deterministic sequence $(Z_j(m,\boldsymbol{\xi}))_{j\in\mathbb{N}}$, which represents a Galton-Watson process starting from initial condition m, where the k-th individual of the j-th generation generates an offspring of size $\xi_{j,k}$. Let us define the sequence precisely. By definition, we set $Z_0(m,\boldsymbol{\xi}) = m$. We label individuals of the first generation according to an arbitrary order using the integers $1, 2, \ldots, m$. After having labelled individuals of the first generation, each individual k generates an offspring of size $\xi_{1,k}$. We set $Z_1(m,\boldsymbol{\xi})$ as the number of individuals of the generation $j = 1$. We label individuals of the second generation by using integer numbers increasing one by one, $1, 2, 3, \ldots$, in such a way that, for any two individuals A and B, if the label of the parent of A is smaller than the label of the parent of B, then the label of A is smaller than the label of B. Once labels are assigned, each individual k of the second generation generates an offspring of size $\xi_{2,k}$ and variables $\xi_{2,Z_1+1}, \xi_{2,Z_1+2}, \ldots$ remain unused. We set $Z_2(m,\boldsymbol{\xi})$ as the number of individuals of the generation $j = 2$.

At any step j, we assign labels according to such a rule and we use the variables $\xi_{j,k}$ for the children of each individual k. If for some j, $Z_j(m,\boldsymbol{\xi}) = 0$, then by definition $Z_{j+1}(m,\boldsymbol{\xi}) = Z_{j+2}(m,\boldsymbol{\xi}) = \ldots = 0$. The process can be seen as a tree where each

individual is connected to its children by a directed edge. By construction, we have that for any realization of the array $\boldsymbol{\xi}$, for all $j \in \mathbb{N}$,

$$Z_j\left(m, \boldsymbol{\xi})\right)_{j \in \mathbb{N}} \leq Z_j\left(m', \boldsymbol{\xi})\right)_{j \in \mathbb{N}} \tag{5.20}$$

whenever $m' \geq m$, as the tree corresponding to the Galton-Watson process with initial condition m is a sub-graph of the tree corresponding to the Galton-Watson process with initial condition m'.

Let us now fix a $k \in \mathbb{N}$ and a sequence of non-negative integers $n_0, n_1, \ldots, n_k$. For any fixed $\boldsymbol{\xi}$, we let $M(\boldsymbol{\xi}, n_0, \ldots n_k)$ be the smallest m such that

$$Z_j(m, \boldsymbol{\xi}) > n_j \text{ for all } j \text{ between } 0 \text{ and } k. \tag{5.21}$$

From the monotonicity property (5.20), we have that if (5.21) holds for m, then it holds for all $m' \geq m$. Let then $\tilde{P}(\,\cdot\,)$ be the law of the array $\boldsymbol{\xi}$, of Ξ and Ξ' together. We have that,

$$\tilde{P}\left(\Xi > n_0,\, Z_1(\Xi, \boldsymbol{\xi}) > n_1,\, \ldots,\, Z_k(\Xi, \boldsymbol{\xi}) > n_k\right) = \tilde{P}\left(\Xi \geq M(\boldsymbol{\xi}, n_0, \ldots n_k)\right) =$$
$$\sum_{\boldsymbol{\xi}} \tilde{P}\left(\Xi \geq M(\boldsymbol{\xi}, n_0, \ldots n_k) \,\big|\, \boldsymbol{\xi}\right) \tilde{P}(\,\boldsymbol{\xi}\,) \leq \sum_{\boldsymbol{\xi}} \tilde{P}\left(\Xi' \geq M(\boldsymbol{\xi}, n_0, \ldots n_k) \,\big|\, \boldsymbol{\xi}\right) \tilde{P}(\,\boldsymbol{\xi}\,) =$$
$$\tilde{P}\left(\Xi' \geq M(\boldsymbol{\xi}, n_0, \ldots n_k)\right) = \tilde{P}\left(\Xi' > n_0,\, Z_1(\Xi', \boldsymbol{\xi}) > n_1,\, \ldots,\, Z_k(\Xi', \boldsymbol{\xi}) > n_k\right), \tag{5.22}$$

For the first inequality we used that Ξ' is stochastically larger than Ξ by definition. As the two processes that we constructed in the space of $\tilde{P}(\,\cdot\,)$ have the same distribution as the Galton-Watson processes in the statement of the lemma, the previous inequality concludes the proof of the claim. $\qquad\square$

References

[1] N. Berestycki: *Emergence of giant cycles and slowdown transition in random transpositions and k-cycles.* Electron. J. Probab. 16, 152-173 (2011). MR-2754801

[2] V. Betz: *Random permutations of a regular lattice.* J. Stat. Phys., Volume 155, Issue 6, pp 1222-1248, (2014). MR-3207736

[3] V. Betz, D. Ueltschi, Y. Velenik: *Random permutations with cycle weights.* Ann. Appl. Probab. Volume 21, pp 312-331, (2011). MR-2759204

[4] V. Betz, D. Ueltschi: *Spatial random permutations and Poisson-Dirichlet law of cycle lengths.* Electron. J. Probab. Volume 16, Issue 41, (2011). MR-2820074

[5] V. Betz, D. Ueltschi: *Spatial random permutations and infinite cycles.* Comm. Math. Phys. 285, 469-501 (2009) MR-2461985

[6] V. Betz, D. Ueltschi: *Spatial random permutations with small cycle weights.* Probab. Th. Rel. Fields 149, 191-222 (2011). MR-2773029

[7] V. Betz, D. Ueltschi: *Critical temperature of dilute Bose gases.* Phys. Rev. A **81**, 023611 (2010).

[8] L. V. Bogachev, D. Zeindler: *Asymptotic statistics of cycles in surrogate-spatial permutations.* Comm. Math. Phys. 334, 1, p. 39-116. MR-3304271

[9] M. Campanino, D. Ioffe, and Y. Velenik: *Fluctuation theory of connectivities for subcritical random cluster models* Ann. Probab., Vol. 36, No. 4 (2008), 1287–1321 MR-2435850

[10] M. Campanino, D. Ioffe, and Y. Velenik: *Ornstein-Zernike theory for finite range Ising models above T_c.* Probability Theory and Related Fields 125(3):305-349 (2003) MR-1964456

[11] M. Campanino and D. Ioffe: *Ornstein-Zernike theory for the Bernoulli bond percolation on $\mathbb{Z}^d$.* Ann. Probab., Volume 30, Number 2 (2002), 652-682. MR-1905854

[12] J. T. Chayes and L. Chayes: *Ornstein-Zernike Behavior for Self-Avoiding Walks at All Noncritical Temperatures.* Comm. Math. Phys. 105, 2, 221-238 (1986)- MR-0849206

[13] D. Elboim, R. Peled: *Limit distributions for Euclidean random permutations.* arXiv:1712.03809. MR-3962003

[14] N. M. Ercolani, D. Ueltschi: *Cycle structure of random permutations with cycle weights.* Random Struct. Algor. 44, 109-133. MR-3143592

[15] H. Duminil-Copin, G. Kozma, and A. Yadin: *Supercritical self-avoiding walks are space-filling.* Ann. Inst. H. Poincaré Probab. Statist., Volume 50, Number 2, pp 315-326, 2014. MR-3189073

[16] H. Duminil-Copin, A. Glazman, R. Peled, Y. Spinka: *Macroscopic loops in the loop $O(n)$ model at Nienhuis' critical point.* ArXiv:1707.09335

[17] H. Duminil-Copin, R. Peled, W. Samotij, and Y. Spinka: *Exponential decay of loop lengths in the loop $O(n)$ model with large n.* ArXiv: 1412.8326v3. Accepted on Comm. Math. Phys. MR-3602816

[18] E. Domany, D. Mukamel, B. Nienhuis, and A. Schwimmer: Duality relations and equivalences for models with $O(n)$ and cubic symmetry, Nuclear Physics B 190 (1981), no. 2, 279-287.

[19] M. Draief and L. Massouli: *Epidemics and Rumours in Complex Networks.* London Mathematical Society Lecture Note Series: 369. MR-2582458

[20] R. G. Edwards and A. D. Sokal: Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm. Phys. Rev. D (3) 38 (1988), no. 6, 2009–2012 MR-0965465

[21] D. Gandolfo, J. Ruiz, D. Ueltschi: *On a model of random cycles.* J. Statist. Phys. 129, 663-676 (2007). MR-2360227

[22] A. Glazman, I. Manolescu: *Uniform Lipschitz functions on the triangular lattice have logarithmic variations.* ArXiv: 1810.05592.

[23] L. Greenberg, D. Ioffe: *On an invariance principle for phase separation lines*, in *Ann. Inst. H. Poincaré Probab. Statist.*, **41** (2005), no. 5, 871-885. MR-2165255

[24] S. Grosskinsky, A. A. Lovisolo, D. Ueltschi: *Lattice permutations and Poisson-Dirichlet distribution of cycle lengths.* J. Statist. Phys. 146, 1105-1121 (2012). MR-2903040

[25] J. M. Hammersley and D. C. Morton: *Poor man's Monte Carlo.* J. Roy. Stat. Soc. B, 16:23-38 (1954). MR-0064475

[26] J. M. Hammersley: *The number of polygons on a lattice.* Mathematical Proceedings of the Cambridge Philosophical Society, Volume 57, Issue 3, pp. 516-523 (1961). MR-0123489

[27] T. E. Harris: *Nearest-neighbor Markov interaction processes on multidimensional lattices.* Adv. Math. 9, pp. 66-89 (1972) MR-0307392

[28] D. Ioffe, S. Schlosman, F. Toninelli: *Interacting versus entropic repulsion for low temperature Ising polymers*, J. Stat. Phys. **158** (2015), no. 5, 1007-1050. MR-3313616

[29] D. Ioffe: *Multidimensional random polymers: a renewal approach*, in *Random Walks, random fields, and disordered systems*, 147-210, Lecture Notes in Math. **2144**, Springer, Cham, 2015. MR-3382174

[30] D. Ioffe: *Ornstein-Zernike Behaviour and Analyticity of Shapes for Self-Avoiding Walks On Zd.* Markov Processes and Related Fields 4 (1998), 324-350. MR-1670027

[31] R. Kotecky, P. Milos, D. Ueltschi: *The random interchange process on the hypercube.* Electron. Comm. Probab. 21, no. 4 (2016). MR-3485373

[32] Y. Kovchegov: *Brownian bridge in percolation, self-avoiding walks and related processes.* Ph.D thesis, Department of Mathematics, Stanford University, 2002. MR-2703295

[33] B. Nienhuis: *Exact Critical Point and Critical Exponents of $O(n)$ Models in Two Dimensions.* Phys. Rev. Lett., Volume 49, Number 15, (1982). MR-0675241

[34] N. Madras, G. Slade: *The Self-Avoiding Walk.* Birkhäuse (2013), reprint of the 1996 Edition. DOI 10.1007/978-1-4614-6025-1. MR-2986656

[35] L. S. Ornstein, F. Zernike: Proc. Acad. Sci. (Amst.) 17, 793–806 (1915).

[36] O. Penrose, L. Onsager: *Bose-Einstein condensation and liquid Helium.* Phys. Rev. 104, 576 (1956)

[37] O. Schramm: *Compositions of random transpositions.* Isr. J. Math. 147, 221-243 (2005). MR-2166362

[38] G. Slade: *The self-avoiding walk: a brief survey*. in *Surveys in Stochastic Processes*, J. Blath, P. Imkeller, S. Roelly (eds.), European Mathematical Society (2011). MR-2883859

[39] L. Taggi: *Shifted critical threshold for the loop $O(n)$ model at arbitrarily small n,* Electronic Comm. Probab. 2018, Vol. 23, paper no. 96, 1-9. MR-3896834

[40] B. Toth: *Improved lower bound on the thermodynamic pressure of the spin $1/2$ Heisenberg ferromagnet*. Lett. Math. Phys. 28, 75-84 (1993) MR-1224836

[41] D. Ueltschi: *Relation between Feynman Cycles and Off-Diagonal Long-Range Order*. Phys. Rev. Lett. 97, 170601 (2006). MR-2274342

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm