Electron. J. Probab. **24** (2019), no. 108, 1–43.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP366

The interchange process with reversals on the complete graph

Jakob E. Björnberg[*] Michał Kotowski[†] Benjamin Lees[‡]
Piotr Miłoś[§]

Abstract

We consider an extension of the interchange process on the complete graph, in which a fraction of the transpositions are replaced by 'reversals'. The model is motivated by statistical physics, where it plays a role in stochastic representations of xxz-models. We prove convergence to PD($\frac{1}{2}$) of the rescaled cycle sizes, above the critical point for the appearance of macroscopic cycles. This extends a result of Schramm on convergence to PD(1) for the usual interchange process.

Keywords: interchange process; XXZ model; Poisson-Dirichlet distribution.
AMS MSC 2010: 60K35; 60J27.
Submitted to EJP on April 25, 2019, final version accepted on September 18, 2019.
Supersedes arXiv:1808.08902.

1 Introduction

Recent years have seen a growing interest in the cycle structure of large random permutations. A major example is the interchange process, or random-transposition random walk. One motivation for studying this process is that it plays a key role in a stochastic representation of the most important quantum spin system, the ferromagnetic Heisenberg model. This representation was developed by Tóth in the early 1990's [23] (after an earlier observation by Powers [20]).

At about the same time, a closely related stochastic representation was discovered for the *anti*-ferromagnetic Heisenberg model, by Aizenman and Nachtergaele [2]. Very

[*]Department of Mathematics, Chalmers University of Technology and the University of Gothenburg, Sweden.
E-mail: jakob.bjornberg@gu.se
[†]Institute of Mathematics, Faculty of Mathematics, Informatics, and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland.
E-mail: michal.kotowski@mimuw.edu.pl
[‡]Fachbereich Mathematik, Technische Universität Darmstadt, Germany.
E-mail: lees@mathematik.tu-darmstadt.de
[§]Institute of Mathematics of the Polish Academy of Sciences, Warsaw, Poland.
E-mail: pmilos@mimuw.edu.pl

roughly speaking, in the ferromagnetic Heisenberg model the interaction between neighbouring electrons behaves like a transposition of the spins. In the antiferromagnetic model the interaction involves a 'reversal', which Aizenman and Nachtergaele depicted as on the right in Figure 1.

Figure 1: Pictorial representation of a transposition (left) and a 'reversal' (right).

In both cases, the stochastic representation of the spin system involves randomly placing these objects in the product of the graph with an interval. In the case of the ferromagnetic model, the relevant measure has transpositions appearing randomly at each edge, in the manner of independent Poisson processes. For the antiferromagnetic model, the structure is the same except that the transpositions are replaced by 'reversals' as on the right in Figure 1. Many quantities of interest for the spin systems, such as correlation functions, may be expressed using expected values of suitable random variables in these processes.

Recently, Ueltschi [25] explained that weighted combinations of the two processes described above also lead to representations of certain quantum spin systems (known as xxz-models). The relevant measure has independent Poisson processes on the edges as before, but the objects are now randomly chosen to be either transpositions or 'reversals', independently over the points of the process and with some fixed probability (see Figure 2). In this paper we study such a process defined on the complete graph. Our main result is that the correlation structure in this model, above a critical point, is described by a probability distribution on random partitions called the Poisson–Dirichlet distribution with parameter $\frac{1}{2}$. To state our results more precisely, let us give the relevant definitions.

1.1 Definitions

We consider the complete graph $K_n = (V_n, E_n)$ on $n > 2$ vertices. The vertex set is $V_n = \{1, 2, \ldots, n\}$ and the edge-set consists of all pairs $\{i, j\}$ of vertices $i \neq j$. To each edge and vertex we attach a circle of circumference 1, which we denote by S^1. We will sometimes identify S^1 with the unit interval $[0, 1)$. A *configuration* ω is a finite subset of $E_n \times S^1 \times \{\bowtie, \between\}$, where $\bowtie, \between$ are two possible marks which we call a *cross* and a *bar*, respectively. The collection of configurations is denoted Ω. An element $(e, \varphi, m) \in \omega$ of the configuration is called a *link* and if (e, φ, m) is a link then we say that ω *has a link at* $(e, \varphi) \in E_n \times S^1$.

We will primarily be interested in configurations obtained as samples of a (marked) Poisson point process defined in the following way. Fix $\nu \in [0, 1)$ and $\beta > 0$. For each edge $e \in E_n$ we consider a Poisson point process with intensity $\frac{\beta}{n-1}$ on $e \times S^1$, these Poisson processes being independent for different edges e. This defines a configuration of unmarked links. The configuration ω is then obtained by assigning to each link a mark, independently of all other links, which is either a cross, $\bowtie$, with probability ν, or a bar, $\between$, with probability $1 - \nu$. The probability measure corresponding to this point process will be denoted by $\mathbb{P}_\beta$ (we consider ν to be fixed and it will be suppressed in the notation), and the corresponding expectation will be denoted by $\mathbb{E}_\beta$. We will refer to this process as the *interchange process with reversals* (the usual interchange process would correspond to taking $\nu = 1$).

Such a configuration ω gives rise to a set of *loops* $\gamma \subseteq V_n \times S^1$. We first give an

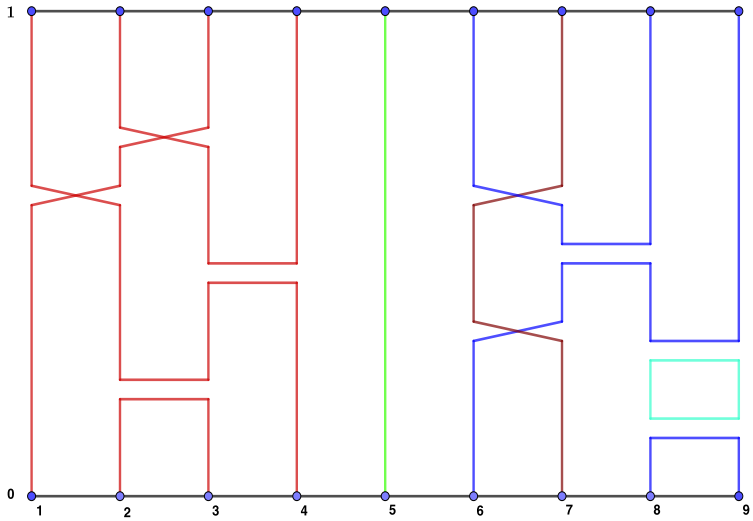

Figure 2: An example of a realisation of links with loops coloured. The circles $\{v\}\times S^1$ are represented as intervals placed vertically, thus there are periodic boundary conditions vertically. The cycles are $(1^\uparrow, 3^\uparrow, 2^\downarrow, 4^\uparrow)$, $(5^\uparrow)$, $(6^\uparrow, 9^\uparrow, 8^\downarrow)$ and $(7^\uparrow)$. There is one small loop which does not give rise to a cycle.

informal description and then a more precise definition. For a fixed point $(v,\varphi) \in V_n \times S^1$ the unique loop, $\gamma(v,\varphi)$, containing it is constructed by the following process. Starting from (v,φ) we move on the associated circle in the positive direction, i.e. after time $\mathrm{d}t$ we are at a point $(v, \varphi + \mathrm{d}t)$. If we encounter a point (v,φ') such that ω has a link at $(\{v,w\},\varphi')$ for some $w \in V_n$, then we traverse this link to $(w,\varphi') \in V_n \times S^1$. We then continue moving in the positive direction if the link was a cross, or in the negative direction if the link was a bar. Each time we encounter a link we follow this rule of traversing the link, reversing our direction if the link was a bar. Continuing until we arrive back at (v,φ) we have traced out a single loop, $\gamma(v,\varphi)$.

More formally, following [10, 25] we may define the loops as follows. A loop of length L is a function $\gamma : [0,L) \to V_n \times S^1$ such that, writing $\gamma(t) = (v(t),\varphi(t))$, the following properties hold:

1. γ is injective and satisfies $\lim_{t\uparrow L} \gamma(t) = \gamma(0)$.
2. γ is piecewise continuous, and if it is continuous on the interval $I \subset [0,L)$ then $v(t)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t)$ are constant on I, with $\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) \in \{-1,1\}$.
3. γ is discontinuous at the point t if and only if ω has a link at $(\{v(t-),y\},t)$ for some $y \neq v(t-)$, in which case $v(t+) = y$.
4. If $I_1 = (t_1,t_2)$ and $I_2 = (t_2,t_3)$ with γ continuous on I_1 and I_2 but discontinuous at t_2 then for any $s_1 \in I_1$ and $s_2 \in I_2$ we have that $\frac{\mathrm{d}}{\mathrm{d}t}\varphi(s_1) = \frac{\mathrm{d}}{\mathrm{d}t}\varphi(s_2)$ if the link at $(\{v(t_2-),v(t_2+)\},t_2)$ is a cross and $\frac{\mathrm{d}}{\mathrm{d}t}\varphi(s_1) = -\frac{\mathrm{d}}{\mathrm{d}t}\varphi(s_2)$ if it is a bar.

Loops with the same support but different parameterisations are identified. This means that the functions $\gamma(t)$, $\gamma(-t)$ and, for $s \in \mathbb{R}$, $\gamma(s \pm t)$ are identified. From this description we can give ω a natural pictorial representation, see Figure 2.

A *cycle* is a sequence of vertices $v \in V_n$ such that the points $(v,0)$ are visited by a single loop. Namely, suppose we start at a point $(v_1,0)$ and follow the loop $\gamma = \gamma(v_1,0)$, in either direction, until we return to the starting point. If we enumerate the successive visits to $V_n \times \{0\}$ as $(v_1,0),\dots,(v_\ell,0),(v_1,0)$ then the corresponding cycle is

$$\mathcal{C} = (v_1^{d_1}, v_2^{d_2}, \dots, v_\ell^{d_\ell}), \qquad v_i \in V_n.$$

Here the cycle $\mathcal{C}$ has length $|\mathcal{C}| = \ell$, and the $d_i \in \{\uparrow, \downarrow\}$ denote the direction in which we pass through the point $(v_i, 0) \in V_n \times S^1$, with $\uparrow$ corresponding to the positive direction ($\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = +1$) and $\downarrow$ corresponding to negative direction ($\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = -1$). Note that the directions d_i in a cycle are defined up to an overall reversal and that we made an arbitrary choice of the first vertex v_1. It is also worth noting that not every loop gives rise to a cycle, see Figure 2. A fixed configuration of links, ω, has an associated set of cycles which we denote by $\mathfrak{C}_\omega$.

1.2 Main result

Let ω be sampled from the measure $\mathbb{P}_\beta$. Consider the random graph where an edge is present between vertices u and v if there is at least one link on $\{u, v\} \times S^1$ in ω. By the Erdős-Rényi theorem, if $\beta > 1$ then the largest connected component of this graph, V_G^β, has size approximately zn where z is the positive solution to $1 - z = e^{-\beta z}$. (If $\beta < 1$ the largest component has size smaller than $(\log n)^2$ and the same holds for the largest cycle.) Let $\mathfrak{X}_\omega$ denote the list $(|\mathcal{C}|/|V_G^\beta| \colon \mathcal{C} \in \mathfrak{C}_\omega)$ of rescaled cycle sizes, ordered by decreasing size (we make it into an infinite list by appending infinitely many 0's). Our main result is the following.

Theorem 1.1. *Let $\beta > 1$ and $\nu \in [0, 1)$. Let ω be sampled from the corresponding measure $\mathbb{P}_\beta$. As $n \to \infty$ the law of $\mathfrak{X}_\omega$ converges weakly to the Poisson–Dirichlet distribution* PD($\frac{1}{2}$).

More precisely, we will show that for given $\beta > 1$ and $\varepsilon > 0$ there exists $n(\beta, \varepsilon)$ such that for $n > n(\beta, \varepsilon)$ there is a coupling of the interchange process with reversals with a PD($\frac{1}{2}$) sample $\mathcal{Y}$ such that

$$\mathbb{P}\big(\|\mathcal{Y} - \mathfrak{X}\|_\infty < \varepsilon\big) > 1 - \varepsilon. \tag{1.1}$$

Note that this result holds for any $\nu < 1$.

The Poisson–Dirichlet distribution with parameter $\theta > 0$, PD(θ), can be defined via the 'stick-breaking' construction as follows. Let $B_1, B_2, \dots$ be independent Beta$(1, \theta)$ random variables, thus $\mathbb{P}(B_i > s) = (1 - s)^\theta$ for $s \in [0, 1]$. We construct a random partition $\{P_i\}_{i \in \mathbb{N}}$ of $[0, 1)$ using the B_i by letting $P_1 = B_1$ and $P_{k+1} = B_{k+1}(1 - P_1 - \cdots - P_k)$. We can think of constructing $\{P_i\}_{i \in \mathbb{N}}$ by progressively breaking off pieces of $[0, 1)$, with the $(k+1)^{\text{th}}$ removed piece being a fraction B_{k+1} of what remained after k pieces had been removed. The law of the partition $\{P_i\}_{i \in \mathbb{N}}$ is called the GEM(θ) distribution. The PD(θ) distribution is obtained by sorting the P_i in order of decreasing size.

Returning to the context of Theorem 1.1, let us comment on the case $\nu = 1$ (only crosses allowed) which is excluded by our result. This is the (usual) interchange process, and was considered on the complete graph in a famous paper by Schramm [21]. To be precise, he considered the closely related process where the configuration ω is obtained by placing the crosses successively one after the other, uniformly and independently at each step. Viewing the crosses as transpositions, as above, the process is a random-transposition random walk on the set of permutations of n objects. In this case Schramm proved that, when the number of transpositions exceeds cn for $c > \frac{1}{2}$, then the rescaled cycle sizes of the resulting random permutation converge in distribution to PD(1).

The main tool in Schramm's argument was a coupling with a split-merge process which has PD(1) as an invariant distribution. Roughly speaking, the important feature is what happens to an existing cycle when a uniformly chosen transposition is applied. If the transposition transposes two points which belonged to different cycles then those cycles merge; if they belonged to the same cycle then the cycle is split. A similar principle applies to the loops, which on the addition of another cross to ω either merge, if the ends are in different loops, or split, if both ends are in the same loop (Figure 3).

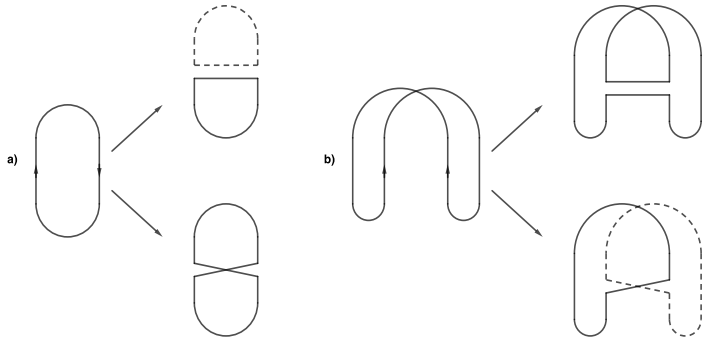

Figure 3: Geometric interpretation of the effect of adding a cross or bar with both endpoints in the same loop: a) in the case when the points have opposite vertical orientation within the loop, a bar splits the loop whereas a cross 'twists'; b) in the case of the same vertical orientation, a cross splits the loop whereas a bar 'twists'.

Now we may explain how the case $\nu < 1$ is different from $\nu = 1$, and why we get PD($\frac{1}{2}$) rather than PD(1). The key point is that the presence of bars (⫤⫣) introduces changes of orientation within the loops. This means that on adding a link (cross or bar) with both endpoints in the same loop, this loop will not always split. Whether or not the loop splits depends on the orientation of the loop at the points where the new link is placed. Specifically, if the link is a cross then a split occurs if and only if the orientation is the same; if it is a bar then the opposite applies. The situation is depicted in Figure 3.

Intuitively, when $\nu < 1$ one would expect large loops to encounter many bars. Hence a uniformly chosen pair of points on a large loop (with the same S^1-coordinate) should have probability close to $\frac{1}{2}$ of having the same orientation, meaning that the probability of splitting is close to $\frac{1}{2}$. The corresponding split-merge dynamics, where proposed splits occur with probability $\frac{1}{2}$, has PD($\frac{1}{2}$) as its invariant distribution.

1.3 Outline and related works

In order to prove Theorem 1.1 we need three ingredients. Firstly, we need that with high probability (converging to 1 as $n \to \infty$) there are cycles of size $\Theta(n)$, and that these large cycles occupy almost the entire giant component V_G^β. This is proved in Section 2 by a straightforward adaptation of arguments in [21]. Secondly, we need that in large cycles roughly half of all vertices are passed through in the positive direction ($\uparrow$) and roughly half in the negative direction ($\downarrow$). In fact, we need the stronger statement that the large cycles are 'well-balanced', namely: one may partition them into much smaller segments such that each segment consists of roughly half $\uparrow$ and half $\downarrow$ (ruling out, for example, a situation in which a cycle of size k consists of a block of $k/2$ vertices passed in direction $\uparrow$ followed by a block of $k/2$ vertices with direction $\downarrow$). This is the main novel contribution of the present paper, and is the content of Section 4. In proving this result we rely on a process which we call the *exploration process*, which we study in Section 3. Thirdly, we show that Schramm's coupling, when combined with the previous two ingredients, can be adapted to couple a PD($\frac{1}{2}$) sample with $\mathfrak{X}$ such that the two samples are close. This appears in Section 5.

We now briefly summarise some other related works apart from Schramm's paper [21]. First note that interchange processes, with reversals ($\nu < 1$) or without ($\nu = 1$), can be defined on more general graphs, by placing independent Poisson processes of links on the edges of the graph. Most papers dealing with graphs other than the complete graph have studied the question of whether there can be large cycles. The case when the

graph is a hypercube, and $\nu = 1$, has been investigated by Kotecký, Miłoś and Ueltschi [15]. The case of Hamming graphs, also for $\nu = 1$, has been investigated by Miłoś and Şengül [17] and by Adamczak, Kotowski and Miłoś [1]. In the case when the graph is an infinite tree one may ask about the occurrence of *infinite* cycles. For $\nu = 1$ this question was investigated by Angel [3] and by Hammond [12, 13]; and for $\nu < 1$ by Björnberg and Ueltschi [8] as well as Hammond and Hegde [14].

As mentioned above, the original interest in the process was due to its connections with quantum spin systems. When the measure $\mathbb{P}_\beta$ defining the process is given an additional weighting of $\vartheta^{\#\text{loops}}$ for $\vartheta \in \mathbb{N}$, the loop-model is essentially equivalent to a spin system on the same graph. This was first proved by Tóth [23] in the case $\nu = 1$ (spin-$\frac{1}{2}$ Heisenberg ferromagnet for $\vartheta = 2$) and Aizenman and Nachtergaele [2] in the case $\nu = 0$ (Heisenberg antiferromagnet, provided the graph is bipartite). This connection was extended to the case $\nu \in [0, 1]$ by Ueltschi [25]. From a probabilistic point of view, any $\vartheta > 0$ makes sense. Such models have been considered on trees [9, 5] and on the Hamming graph [1]. In very recent work there has been some limited progress in the direction of establishing Poisson–Dirichlet structure in these and related loop models [4, 7]. For the Heisenberg model ($\nu = 1$ and $\vartheta = 2$) on the complete graph, the critical point for the appearance of cycles of diverging length was established already in the early 1990's by Tóth and by Penrose [18, 22].

Notation

$V_n = \{1, 2, ..., n\}$	vertex set of the complete graph; typical elements denoted $u, v, w, \dots$
$E = \binom{V}{2}$	edge set of the complete graph
$m \in \{\between, \barwedge\}$	the 'mark' of a link as either a cross or a bar
$\varphi \in S^1$	'phase' or vertical coordinate
Ω	space of configurations, i.e. finite subsets of $E \times S^1 \times \{\between, \barwedge\}$
ω, ω_A	element of Ω, its restriction to $A \subseteq E \times S^1$
$\vec{\omega}$, $\vec{\omega}_k$	ordered sequence of links in ω, its first k elements
$\beta > 0$	intensity parameter
$\mathbb{P}_\beta$	measure of the Poisson links process with intensity $\frac{\beta}{n-1}$ per edge
$\nu \in [0, 1)$	probability of marking a link as a cross $\between$
$\mathfrak{C}_\omega$	set of cycles defined by ω
$\mathfrak{X}_\omega$	list of rescaled cycle sizes
V_G^β	the largest cluster in the random graph whose edges support links
$X_t = X_t(v, \varphi)$	exploration process for loops (started at $(v, \varphi) \in V_n \times S^1$)
$Y_t = Y_t(v, \varphi)$	simple exploration process (started at $(v, \varphi) \in V_n \times S^1$)
$\tau^{X(Y)}(v, \varphi)$	closing time of (simple) exploration started from (v, φ)
$\{\mathcal{F}_t\}_{t \geq 0}$	natural filtration of the exploration process
$\mathcal{I}_t^{X(Y)}$	number of times X_t (resp. Y_t) traverses a link
$\mathcal{J}_t^{X(Y)}$	number of links discovered by X_t (resp. Y_t)
$\mathcal{L}_t^{X(Y)}$	number of windings of X_t (resp. Y_t) around S^1
$\mathcal{K}_t^{X(Y)}$	number of visits to $\varphi = 0 \in S^1$ by X_t (resp. Y_t)
ℓ_k	frontier times of the simple exploration
Δ_k	$\ell_{k+1} - \ell_k$
$c_\omega(v, k)$	set of the next k vertices in a cycle starting from v
$c_\omega^{\uparrow(\downarrow)}(v, k)$	number of vertices in $c_\omega(v, k)$ passed in positive or negative direction
$\mathfrak{B}_\omega(v, k)$	$\big\lvert\lvert c_\omega^{\uparrow}(v, k)\rvert - \lvert c_\omega^{\downarrow}(v, k)\rvert\big\rvert$
PD(θ)	Poisson–Dirichlet distribution with parameter θ
$\mathcal{Y}, \mathcal{Z}$	samples from PD(θ)

2 Large cycles

In this section we show that, for $\beta > 1$, with high probability there are cycles of length of order n. The precise statement appears in Lemma 2.4 at the end of the section. The argument is a minor adaptation of Schramm's [21, Section 2].

As in [21], we will actually work not with configurations ω sampled from the Poisson measure $\mathbb{P}_\beta$, but instead with configurations constructed sequentially one link at a time. Given any configuration $\omega \in \Omega$, note that the set of cycles $\mathfrak{C}_\omega$ only depends on the relative order of the links of ω as well as their position relative to $0 \in S^1$, but not on their precise S^1-coordinates. Given $\omega \in \Omega$, let us order its elements (links) with respect to the S^1-coordinate, namely we write $\omega = \{(e_1, \varphi_1, m_1), \dots, (e_{|\omega|}, \varphi_{|\omega|}, m_{|\omega|})\}$, with $0 < \varphi_1 < \ldots < \varphi_{|\omega|} < 1$ (we can assume that there are no distinct links with $\varphi_i = \varphi_j$, since under $\mathbb{P}_\beta$ this occurs with probability 1). We denote by $\vec{\omega} = ((e_1, m_1), \dots, (e_{|\omega|}, m_{|\omega|}))$ the ordered list of links with S^1-coordinates suppressed. With a slight abuse of terminology we will also refer to the entries (e_i, m_i) of $\vec{\omega}$ as links. As noted above, $\mathfrak{C}_\omega$ is a function of $\vec{\omega}$ only, hence we may write $\mathfrak{C}_{\vec{\omega}}$.

In the rest of this section we will work with a random $\vec{\omega}$ obtained by sequentially laying down a fixed number t of random links. More precisely, first let e_1 be chosen uniformly from the edge-set E_n and let $m_1 \in \{\asymp, \eqsim\}$ be chosen independently of e_1, with probability ν for $\asymp$. Next, given the first s links $(e_1, m_1), \dots, (e_s, m_s)$, we select e_{s+1} uniformly from E_n and the mark $m_{s+1} \in \{\asymp, \eqsim\}$, with probability ν for $\asymp$, independently of each other and of the previous choices. Write $\vec{\omega}_s = ((e_1, m_1), \dots, (e_s, m_s))$ and let $\mathfrak{C}_s := \mathfrak{C}_{\vec{\omega}_s}$ denote the set of cycles after $s \le t$ steps. Note that, if t is taken to be Poisson-distributed with mean $\frac{\beta}{n-1}\binom{n}{2} = \frac{\beta}{2}n$, then $\mathfrak{C}_t$ is equal in distribution to $\mathfrak{C}_\omega$ for ω sampled from $\mathbb{P}_\beta$. Due to concentration properties of the Poisson-distribution, there is very little difference between $\mathfrak{C}_t$ for $t = \lfloor \frac{\beta}{2}n \rfloor$ on the one hand, and $\mathfrak{C}_\omega$ for ω sampled from $\mathbb{P}_\beta$ on the other. We will not make this statement more precise at this point, deferring this to later (see Section 5.3).

We now describe in detail the effect that appending the next link (e_{s+1}, m_{s+1}) to $\vec{\omega}_s$ has on the cycles, that is, the transition $\mathfrak{C}_s \to \mathfrak{C}_{s+1}$. See Figures 4, 5 and 6 for illustrations in the case when $m_{s+1} = \eqsim$.

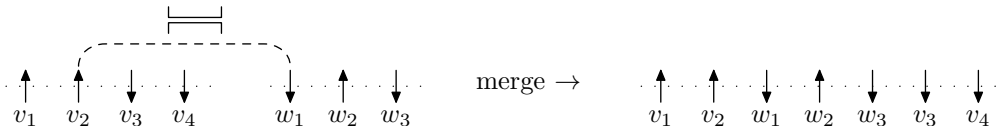

Figure 4: Example of two cycles merging.

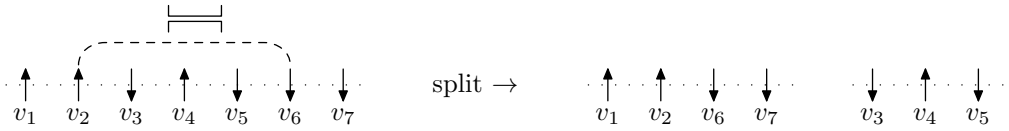

Figure 5: Example of a cycle splitting.

For $d \in \{\uparrow, \downarrow\}$ we write $-d$ for the reversed arrow. We have the following:

- If the endpoints of e_{s+1} are in different cycles of $\mathfrak{C}_s$ then those cycles merge.
- If the endpoints are in the same cycle $\mathcal{C}$ then the result depends on the mark m_{s+1} in the following way. Let us assume that $\mathcal{C} = (v_1^{d_1}, v_2^{d_2}, \dots, v_\ell^{d_\ell})$ and that $e_{s+1} = \{v_i, v_j\}$

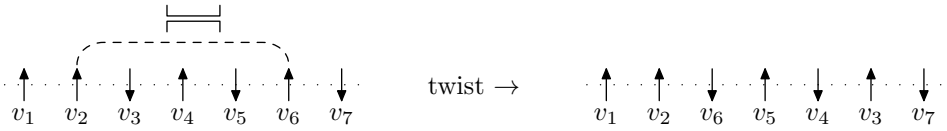

Figure 6: Example of a cycle 'twisting', i.e. it stays intact but is restructured internally.

where $i < j$. Without loss of generality (since directions are defined up to an overall reversal) we may assume that $d_i = \uparrow$.

- If $m_{s+1} = \bowtie$ is a cross then $\mathcal{C}$ splits if and only if $d_j = \uparrow$; in this case the two resulting cycles $\mathcal{C}'$ and $\mathcal{C}''$ are given by:

$$\mathcal{C}' = (v_1^{d_1}, \dots, v_i^{d_i}, v_{j+1}^{d_{j+1}}, \dots, v_\ell^{d_\ell}), \qquad \mathcal{C}'' = (v_{i+1}^{d_{i+1}}, v_{i+2}^{d_{i+2}} \dots, v_j^{d_j}).$$

  On the other hand, if $d_j = \downarrow$ then $\mathcal{C}$ is not split; instead it is modified into $\mathcal{C}'$ where

$$\mathcal{C}' = (v_1^{d_1}, \dots, v_i^{d_i}, v_{j-1}^{-d_{j-1}}, v_{j-2}^{-d_{j-2}}, \dots, v_{i+1}^{-d_{i+1}}, v_j^{d_j}, \dots, v_\ell^{d_\ell}).$$

- If $m_{s+1} = \amalg$ is a bar then $\mathcal{C}$ splits if and only if $d_j = \downarrow$; in this case the two resulting cycles $\mathcal{C}'$ and $\mathcal{C}''$ are given by:

$$\mathcal{C}' = (v_1^{d_1}, \dots, v_i^{d_i}, v_j^{d_j}, \dots, v_\ell^{d_\ell}), \qquad \mathcal{C}'' = (v_{i+1}^{d_{i+1}}, v_{i+2}^{d_{i+2}} \dots, v_{j-1}^{d_{j-1}}).$$

  On the other hand, if $d_j = \uparrow$ then $\mathcal{C}$ is not split but modified into $\mathcal{C}'$ where

$$\mathcal{C}' = (v_1^{d_1}, \dots, v_i^{d_i}, v_j^{-d_j}, v_{j-1}^{-d_{j-1}}, \dots, v_{i+1}^{-d_{i+1}}, v_{j+1}^{d_{j+1}}, \dots, v_\ell^{d_\ell}).$$

Note that the edge e_{s+1} may be selected by first choosing v_i uniformly from E_n and then v_j uniformly from $E_n \setminus \{v_i\}$. In particular we see that, just as in [21, Lemma 2.1], we have:

Lemma 2.1. *In the step from $\vec{\omega}_s$ to $\vec{\omega}_{s+1}$, the probability that some cycle is split into two cycles, with at least one containing at most k vertices, is at most $2k/(n-1)$.*

Building on this, and replicating the arguments of Schramm [21], we obtain the following sequence of lemmas. Lemma 2.2 is proved exactly as [21, Lemma 2.2]. Lemma 2.3 is a version of [21, Lemma 2.3] and is proved in a similar way, see [1, Lemma 5.1] for details. Briefly, the reason that these results hold exactly as in [21] is that, firstly, if the endpoints of e_{s+1} are in different cycles then those cycles always merge, and, secondly, the cycle may or may not split if the endpoints are in the same cycle. This means that both the upper bounds on the probability of splitting, as well as the lower bounds on the probability of merging, are identical to [21]. This is all that we need.

In the following statements we consider a random graph G^s with vertex set V_n obtained by placing an edge between a pair $\{i, j\}$, $i \neq j$, if in $\vec{\omega}_s$ there is at least one link (e_r, m_r), $r \leq s$, such that $e_r = \{i, j\}$. We write $V_G^s(k)$ for the set of vertices in connected components of G^s containing at least k vertices. Similarly, we write $V_{\mathfrak{C}}^s(k)$ for the set of vertices belonging to cycles of $\mathfrak{C}_s$ which are of length at least k.

Lemma 2.2. *For any $s \geq 0$,*

$$\mathbb{E}|V_G^s(k) \setminus V_{\mathfrak{C}}^s(k)| \leq \frac{4sk^2}{n-1} \tag{2.1}$$

Lemma 2.3. *Let $t_0 \in \mathbb{N}$, $\delta \in (0, 1]$, $\varepsilon \in (0, 1/8)$ and $j \in \mathbb{N}$ be such that $2^j \leq \varepsilon\delta n$. Assume that the following conditions hold in the transition from $\vec{\omega}_s$ to $\vec{\omega}_{s+1}$:*

1. there exists $c_1 > 0$ *such that for any* $s, k \in \mathbb{N}$*, we have*

$$\mathbb{P}(\textit{some } \mathcal{C} \in \mathfrak{C}_s \textit{ is split into } \mathcal{C}', \mathcal{C}'' \textit{ s.t. } \min(|\mathcal{C}'|, |\mathcal{C}''|) \le k \mid \mathfrak{C}_s) \le c_1 \frac{k}{n}.$$

2. there exists $c_2 > 0$ *such that for any* $s \in \mathbb{N}$ *and any two cycles* $\mathcal{C}', \mathcal{C}'' \in \mathfrak{C}_s$ *we have*

$$\mathbb{P}(\mathcal{C}', \mathcal{C}'' \textit{ are merged} \mid \mathfrak{C}_s) \ge c_2 \frac{|\mathcal{C}'||\mathcal{C}''|}{n^2}.$$

Then there exist $c_3, c_4 > 0$*, depending only on* c_1, c_2*, such that if*

$$t_1 := t_0 + \lceil \Delta t \rceil, \quad \Delta t = c_3 \delta^{-1} \frac{n}{2^j} \log_2 \left(\frac{n}{2^j} \right), \tag{2.2}$$

then

$$\mathbb{E}\Big(|V_{\mathfrak{C}}^{t_0}(2^j) \setminus V_{\mathfrak{C}}^{t_1}(\varepsilon\delta n)| \;\Big|\; \mathfrak{C}_{t_0} \Big) \mathbb{1}_{\{|V_{\mathfrak{C}}^{t_0}(2^j)| \ge \delta n\}} \le c_4 \delta^{-1} \varepsilon |\log_2(\varepsilon\delta)| n. \tag{2.3}$$

Note that in our case condition (1) is satisfied because of Lemma 2.1 and condition (2) is trivially satisfied. In notation of [1, Lemma 5.1] this corresponds to taking the stopping time $\tau = +\infty$, i.e. conditions (1) and (2) hold for all times $s \in \mathbb{N}$. The lemma states that if, at some time t_0, enough vertices are in reasonably large cycles (size $\ge 2^j$) then at some carefully chosen later time most of these vertices will be in cycles of size of the order n. Here one should think of 2^j as approximately $n^{1/4}$ and of $t_0 \ge c_0 n$ for some $c_0 > \frac{1}{2}$. Then $|V_G^{t_0}(2^j)| \approx zn$ by the Erdős–Rényi theorem, hence by Lemma 2.2 also $|V_{\mathfrak{C}}^{t_0}(2^j)| \approx zn$. Note that if $2^j = n^{1/4}$ then Δt is of the order $n^{3/4} \log n \ll n$, thus for any $c > \frac{1}{2}$ and $t_1 \ge cn$ we may select $c_0 > \frac{1}{2}$ such that $t_0 = t_1 - \lceil \Delta t \rceil \ge c_0 n$.

The final result of this section paraphrases [21, equation (2.4) in Lemma 2.4]. It tells us that most of the vertices in V_G^t (the largest connected component in G^t) belong to large cycles. The proof is precisely as in [21] as the only appeal to the particular structure of the cycles is through invoking the previous lemmas, which all hold as in [21].

Lemma 2.4. *Fix* $c > 1/2$*, take* $t \ge cn$*,* $t \in \mathbb{N}$*. There is some* $C_2 > 0$ *such that for any* $\varepsilon \in (0,1)$*, if* n *is large enough we have*

$$\mathbb{E}\big[|V_G^t \setminus V_{\mathfrak{C}}^t(\varepsilon n)| \big] \le C_2 \varepsilon \log(\tfrac{1}{\varepsilon}) n. \tag{2.4}$$

3 Exploration processes

An important tool in proving Theorem 1.1 is the *exploration process*, which we will define in this section. The exploration process is sometimes also called the *cyclic-time random walk*, see e.g. [13, 14]. It will allow us to uncover the loop containing some specified point $(v_0, \varphi_0) \in V_n \times S^1$ at the same time as we uncover the configuration $\omega \in \Omega$ itself. We will also define a process which we call the *simple exploration process* which is easier to analyse and which may be coupled with the exploration process. In this section we work with a random $\omega \in \Omega$ sampled from the Poisson measure $\mathbb{P}_\beta$ for some fixed $\beta > 1$ (the definitions will make sense for all $\beta > 0$). Recall that $\nu \in [0, 1)$ is fixed throughout.

3.1 Definitions

The exploration process will be denoted $X(\omega) = (X_t(\omega) : t \ge 0)$ and takes values in $V_n \times S^1 \times \{-1, +1\}$. Recall that we identify S^1 with the unit interval $[0, 1)$ using periodic boundary conditions. We will write $X_t = (v_t, \varphi_t, d_t)$. Let $d_0 \in \{-1, +1\}$ be an initial

direction and set $X_0 = (v_0, \varphi_0, d_0)$, where $(v_0, \varphi_0) \in V_n \times S^1$. The process starts by traversing $\{v_0\} \times S^1$ at unit speed in the direction specified by d_0, meaning that $v_t = v_0$, $\varphi_t = \varphi_0 + d_0 t$ and $d_t = d_0$. This continues until it either encounters a link of ω, or it returns to its starting point, i.e. $(v_t, \varphi_t) = (v_0, \varphi_0)$. If a link is encountered first, say at time t and with other endpoint in $\{w\} \times S^1$, then the process jumps to $\{w\} \times S^1$ and proceeds in a direction which depends on whether the link was a cross or a bar. That is, we set $v_t = w$ and $\varphi_{t+s} = \varphi_t + d_0 s$ and $d_{t+s} = d_0$ if the link was a cross or $\varphi_{t+s} = \varphi_t - d_0 s$ and $d_{t+s} = -d_0$ if the link was a bar. We define the process to be right-continuous (càdlàg). The process proceeds in this way, traversing links and adjusting its direction accordingly, until it returns to the starting point (v_0, φ_0, d_0). We let

$$\tau^X = \tau^X(v_0, \varphi_0, d_0) := \inf\{t > 0 : X_t = X_0\} \tag{3.1}$$

be the time when this happens. After this time the process is no longer useful to us, but to be definite we declare that the process continues by repeating itself periodically after time τ^X. Note that at time τ^X, the loop containing (v_0, φ_0) has been fully discovered.

Let us consider those links that, by time τ^X, have been traversed by X at least once. Some of them have been traversed only once, others twice (no link can be traversed more than twice before time τ^X as this would entail visiting a previously visited point (v_t, φ_t, d_t)). We say that a link is *discovered* at the time of its first traversal, and *backtracked* on its second traversal (if traversed twice). Let $\mathcal{J}_t^X(v, \varphi)$ denote the number of times the exploration X, started at (v, φ) and run for time t, has discovered a link ('jumped'). Let $\mathcal{I}_t^X(v, \varphi)$ denote the number of times it has traversed some link, including backtracking. Thus $\mathcal{J}_t^X(v, \varphi) \le \mathcal{I}_t^X(v, \varphi) \le 2\mathcal{J}_t^X(v, \varphi)$ for all $t \le \tau^X$. Next, define the *history* of X as

$$H_t^X := \{(v, \varphi) \in V_n \times S^1 \;\big|\; \exists\, s \le t \text{ s.t. } X_s = (v, \varphi, d) \text{ for some } d \in \{-1, 1\}\}. \tag{3.2}$$

This is the set of points in $V_n \times S^1$ visited by X up to time t. Finally, let $\{\mathcal{F}_t\}_{t \ge 0}$ denote the natural filtration of the exploration process, namely $\mathcal{F}_t := \sigma\big((X_s)_{0 \le s \le t}\big)$, and $\bar{\mathcal{F}}_t := \bigcap_{s > t} \mathcal{F}_s$.

When ω is randomly sampled from the Poisson measure $\mathbb{P}_\beta$ we may, thanks to the memorylessness of Poisson processes, construct (part of) ω itself simultaneously with X. This fact is central to our approach. We formulate the construction as a proposition. In the following result we will be using a Poisson process N on $[0, \infty)$ and we will say that N *rings at time* t if it has an arrival at that time.

Proposition 3.1 (Construction of the exploration process)**.** *Let $v_0 \in V_n$, $\varphi_0 \in S^1$ and $d_0 \in \{-1, 1\}$. Consider the following independent objects:*

- *a Poisson process $N = (N_t : t \ge 0)$ with intensity $n\frac{\beta}{n-1}$,*
- *a sequence $\{v_i\}_{i \in \mathbb{N}}$ of i.i.d. random variables distributed uniformly on V_n,*
- *a sequence $\{\xi_i\}_{i \in \mathbb{N}}$ of i.i.d. random variables taking values ± 1 and satisfying $\mathbb{P}(\xi_i = +1) = \nu$.*

When ω has law $\mathbb{P}_\beta$, then the law of the exploration X, started at $X_0 = (v_0, \varphi_0, d_0)$, may be constructed as follows:

1. *The process starts at $X_0 := (v_0, \varphi_0, d_0)$ and initially only φ_t changes, according to $\varphi_t = \varphi_0 + d_0 t$.*
2. *Whenever N rings, say at time t, we inspect vertex $w = v_{N_t}$. We have two cases:*

(a) If $(w, \varphi_{t-}) \in H^X_{t-}$ or $w = v_{t-}$ then nothing happens and the process continues on, i.e. $X_t = X_{t-}$.

(b) Otherwise we set $X_t = (w, \varphi_{t-}, \xi_{N_t} d_{t-})$ and then φ_t evolves according to $\varphi_{t+s} = \varphi_{t-} + \xi_{N_t} d_{t-} s$.

3. Between successive rings of N the process may backtrack across previously discovered links. More precisely, a backtrack occurs at time t if there exists another time $s < t$ such that $v_s \neq v_{s-}$ and either:

- *$(v_{t-}, \varphi_{t-}) = (v_{s-}, \varphi_{s-})$, in which case we set $X_t = (v_s, \varphi_s, \xi_{N_s} d_{t-})$, or*
- *$(v_{t-}, \varphi_{t-}) = (v_s, \varphi_s)$, in which case we set $X_t = (v_{s-}, \varphi_{s-}, \xi_{N_s} d_{t-})$.*

See Figure 7.

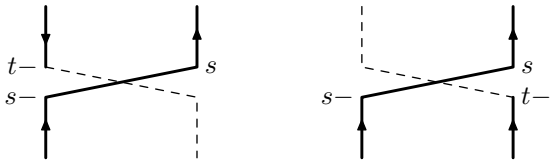

Figure 7: Two possibilities for backtracking a cross. Thick, solid lines belong to the history H^X_{t-}, dashed lines the future trajectory. Left: $(v_{t-}, \varphi_{t-}) = (v_{s-}, \varphi_{s-})$. Right: $(v_{t-}, \varphi_{t-}) = (v_s, \varphi_s)$.

The construction of Proposition 3.1 is fairly standard and has been used previously in for example [1, 13, 14], hence we do not give a proof. Let us however draw attention to the condition that, when $(w, \varphi_{t-}) \in H^X_{t-}$ or $w = v_{t-}$, then the jump proposed by N is canceled. This means that X cannot jump to a previously visited point (w, φ), which effectively amounts to a reduction of the intensity of jumps (see Lemma 3.5 below).

The main difficulty in analysing X is that it may discover a new link which takes it to a previously visited copy of S^1, i.e. it may jump at time t to a point (w, φ) satisfying $(\{w\} \times S^1) \cap H^X_s \neq \varnothing$ for some $s < t$. We refer to this as *jumping to the history*. In this case $\{w\} \times S^1$ has already been partially explored, making X quite difficult to analyse directly.

To get around this problem we introduce what we call the *simple* exploration process $Y = (Y_t : t \geq 0)$, which is easier to analyse and (on time intervals which are not too long) can be coupled with the exploration process X. Roughly speaking, the idea is that for Y we replace the vertex set V_n with an augmented vertex set $\mathbb{N} \times V_n$, where the $\mathbb{N}$-coordinate increases on discovering a new link. The interpretation is that each newly discovered link brings us to a 'fresh' circle $\{(k, v)\} \times S^1$.

Below we give a detailed definition. Notice that the wording is very similar to Proposition 3.1, the main difference being what happens at the jump times of N.

Definition 3.2 (Simple exploration process)**.** *Let $v_0 \in V_n$, $\varphi_0 \in S^1$ and $d_0 \in \{-1, 1\}$. We construct the simple exploration process $Y_t = (k_t, v_t, \varphi_t, d_t)$ as a càdlàg process, using the following independent objects:*

- *a Poisson process $N = (N_t : t \geq 0)$ with intensity $n\frac{\beta}{n-1}$,*
- *a sequence $\{v_i\}_{i \in \mathbb{N}}$ of i.i.d. random variables distributed uniformly on V_n,*
- *a sequence $\{\xi_i\}_{i \in \mathbb{N}}$ of i.i.d. random variables taking values ± 1 and satisfying $\mathbb{P}(\xi_i = 1) = \nu$.*

Using these sources of randomness the process is constructed as follows:

1. *The process starts at* $Y_0 := (0, v_0, \varphi_0, d_0)$ *and initially only* φ_t *changes, according to* $\varphi_t = \varphi_0 + d_0 t$.

2. *Whenever* N *rings, say at time* t, *we inspect vertex* $w = v_{N_t}$; *we have two cases:*

   (a) *If* $w = v_{t-}$ *then nothing happens and the process continues on, i.e.* $Y_t = Y_{t-}$.

   (b) *If* $w \neq v_{t-}$ *we set* $Y_t = (N_t, w, \varphi_{t-}, \xi_{N_t} d_{t-})$ *and then* φ_t *evolves according to* $\varphi_{t+s} = \varphi_{t-} + \xi_{N_t} d_{t-} s$.

3. *Between successive rings of* N *the process may backtrack across previously discovered links. Now there is only one possibility for backtracking: a backtrack occurs at time* t *if there exists another time* $s < t$ *such that* $(k_s, v_s) \neq (k_{s-}, v_{s-})$ *and* $(k_{t-}, v_{t-}, \varphi_{t-}) = (k_s, v_s, \varphi_s)$, *and then we set* $Y_t = (k_{s-}, v_{s-}, \varphi_{s-}, \xi_{N_s} d_{t-})$.

As we did for X we let

$$\tau^Y := \inf\{t > 0 : Y_t = Y_0\} \tag{3.3}$$

be the first time at which the simple exploration process Y arrives back at the starting point. Note that τ^Y, in contrast to τ^X, may take the value $+\infty$, see Proposition 3.4. If $\tau^Y < \infty$ we assume that the process Y stops evolution after τ^Y.

For Y we define the history by

$$H_t^Y := \{(k, v, \varphi) \in \mathbb{N} \times V_n \times S^1 \big| \exists s \leq t \text{ s.t. } Y_s = (k, v, \varphi, d) \text{ for some } d \in \{-1, 1\}\}. \tag{3.4}$$

By slight abuse of terminology, if $Y_0 = (0, v, \varphi, d)$ for some $d \in \{-1, +1\}$ then we say that Y *started at* (v, φ). As for X we denote by $\mathcal{J}_t^Y(v, \varphi)$ (respectively, $\mathcal{I}_t^Y(v, \varphi)$) the number of times the simple exploration process has discovered (respectively, traversed) a link when started at $(0, v, \varphi)$ and run for time t. We denote by Y' the restriction of Y to $V_n \times S^1 \times \{-1, +1\}$, that is if $Y_t = (k, v, \varphi, d)$ then $Y'_t = (v, \varphi, d)$.

The important point which makes Y simpler to analyse than X is that, each time Y discovers a new link (i.e. N rings and $v_{N_t} \neq v_{t-}$), we set k_t to a previously unused value, namely N_t. This means that Y, by construction, can never jump to its history. Crucially, it does still backtrack across previously discovered links.

3.2 Coupling with the simple exploration process

The following lemma shows that when ω is randomly sampled from $\mathbb{P}_\beta$, one can couple the exploration process X and the simple exploration process Y so that they evolve in the same way on a sufficiently short time scale. One should think of $T = o(n^{1/2})$.

Lemma 3.3 (Coupling with the simple exploration process)**.** *Fix* $T > 0$. *Let* X *be the exploration process and let* σ *be a stopping time with respect to the filtration* $\{\bar{\mathcal{F}}_t\}_{t \geq 0}$ *such that* X *jumps to a previously unvisited vertex at time* σ. *Conditionally on* $\bar{\mathcal{F}}_\sigma$, *there exists a coupling* $\mathbb{P}$ *of a process* $\tilde{X}$ *with a process* Y *such that:*

1. *The process* Y *is a simple exploration starting from* $Y_0 = (0, X_\sigma) = (0, \tilde{v}_0, \tilde{\varphi}_0, \tilde{d}_0)$ *where* $X_\sigma = (\tilde{v}_0, \tilde{\varphi}_0, \tilde{d}_0)$.

2. $\mathbb{P}\big(\{\tilde{X}_t\}_{t \geq 0} \in \cdot \mid \bar{\mathcal{F}}_\sigma\big) = \mathbb{P}\big(\{X_{t+\sigma}\}_{t \geq 0} \in \cdot \mid \bar{\mathcal{F}}_\sigma\big)$.

3. $\mathbb{P}\big(\forall_{t < \tau^Y \wedge T}\ \tilde{X}_t = Y'_t \mid \bar{\mathcal{F}}_\sigma\big) \geq 1 - 4\beta T(\mathcal{J}_\sigma^X + \beta T)/n$.

If at some time $t \leq \tau^Y$ we have $\tilde{X}_t \neq Y'_t$ then we consider the coupling as *failed* at time t. The history $H_t^{\tilde{X}}$ of $\tilde{X}$ is defined as in (3.2).

Proof. In the following we write simply $\mathbb{P}(\cdot)$ for $\mathbb{P}(\cdot \mid \bar{\mathcal{F}}_\sigma)$. We will construct $\{\tilde{X}_t\}_{t\geq 0}$ using the same sources of randomness as for Y, namely the same N, $\{v_i\}_{i\in\mathbb{N}}$ and $\{\xi_i\}_{i\in\mathbb{N}}$ as given in Definition 3.2. We write $\tilde{X}_t = (\tilde{v}_t, \tilde{\varphi}_t, \tilde{d}_t)$.

1. The process $\tilde{X}$ starts at $\tilde{X}_0 := X_\sigma$ and initially only $\tilde{\varphi}_t$ changes, by $\tilde{\varphi}_t = \tilde{\varphi}_0 + t\tilde{d}_0$.
2. Whenever N rings, say at time t, we inspect vertex $w = v_{N_t}$; we have two cases:
   (a) If $(w, \tilde{\varphi}_{t-}) \in H^{\tilde{X}}_{t-} \cup H^X_\sigma$ or $w = \tilde{v}_{t-}$ then nothing happens and the process continues on.
   (b) Otherwise the process jumps to $(w, \tilde{\varphi}_{t-}, \xi_{N_t}\tilde{d}_{t-})$.
3. Between successive rings of N the process may backtrack as before, using links of both X and $\tilde{X}$.

It follows from Proposition 3.1 that $\tilde{X}$ and X have the same distribution, giving statements (1) and (2) from the lemma.

For the proof of (3), let

$$V^X_\sigma := \{v \in V_n : (v,\varphi) \in H^X_s \text{ for some } s < \sigma, \varphi \in S^1\},$$
$$V^{\tilde{X}}_t := \{v \in V_n : (v,\varphi) \in H^{\tilde{X}}_s \text{ for some } s < t, \varphi \in S^1\}$$

be the sets of vertices visited by X up to time σ and by $\tilde{X}$ up to time t, respectively. We define

$$\rho := \inf\big\{t \geq 0 : N \text{ rings at time } t \text{ and } v_{N_t} \in V^{\tilde{X}}_t \cup V^X_\sigma\big\}.$$

Until time $\rho \wedge \tau^Y$ the processes $\tilde{X}_t$ and Y'_t are equal, thus

$$\mathbb{P}\big(\forall_{t<\tau^Y\wedge T}\tilde{X}_t = Y'_t\big) \geq \mathbb{P}(\rho \geq T).$$

Each time t when N rings there is a chance that $v_{N_t} \in V^{\tilde{X}}_t \cup V^X_\sigma$. This has probability at most the number of previously visited vertices divided by n. Since the number of visited vertices cannot exceed the number of discovered links by more than 1, we get

$$\begin{aligned}\mathbb{P}(\rho < T) &\leq \mathbb{E}\left[\sum_{i=1}^{N_T}\frac{(\mathcal{J}^X_\sigma + i)}{n}\right] = \frac{\mathcal{J}^X_\sigma}{n}\mathbb{E}[N_T] + \frac{1}{2n}\mathbb{E}\big[N_T(N_T+1)\big] \\ &= \frac{\mathcal{J}^X_\sigma \beta T}{n-1} + \frac{\beta T}{n-1} + (\beta T)^2\frac{n}{2(n-1)^2},\end{aligned}$$

where in the last equality we used that N_T has Poisson distribution with mean $n\frac{\beta}{n-1}T$. Since $\mathcal{J}^X_\sigma \geq 1$ this gives the claimed bound on $\mathbb{P}\big(\forall_{t\leq\tau^Y\wedge T}\tilde{X}_t = Y'_t\big)$. □

3.3 Properties of the exploration processes

Next we present some basic properties of the processes X and Y, starting with the simple exploration Y.

First note that $\mathcal{J}^Y_t$, the number of links discovered by time t, is a Poisson process with rate β, stopped at time τ^Y (at which time Y itself terminates). It will be convenient to extend this process beyond time τ^Y. For this purpose we let N'_t denote a Poisson process of rate β which agrees with $\mathcal{J}^Y_t$ up to time τ^Y.

Many relevant properties of Y can be understood in terms of the process Z given by $Z_t := N'_t - t$. For example, if Z_t hits -1 then this corresponds to Y returning to its starting point, that is to say we have that $\tau^Y = \inf\{t \geq 0 : Z_t = -1\}$. To see this, note

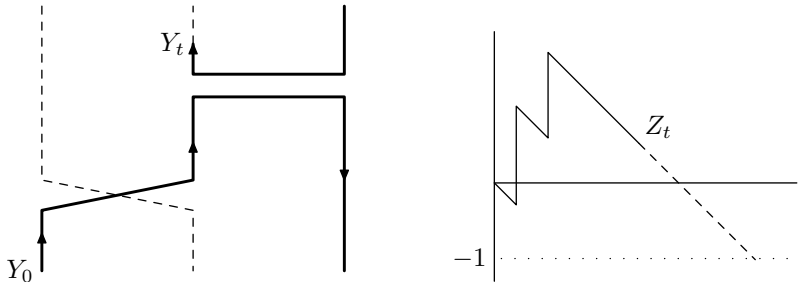

Figure 8: Left: in bold, simple exploration Y up to time t; dashed, the future trajectory of Y assuming no more links are discovered. Right: corresponding plot of Z, dashed line giving extrapolation until time τ^Y assuming no more links are discovered.

that $N'_t + 1$ counts the number of copies of S^1 that Y has visited by time t. If $Z_t = -1$ then $N'_t + 1 = t$, which means that the total time spent equals the number of S^1's visited. Hence at this time Y has explored the entirety of each copy of S^1 it has visited, meaning that it must have returned to its starting point. See Figure 8

Note that $\beta > 1$ implies that $Z_t \to +\infty$ almost surely. We define a sequence of random times which we call *frontier times* ℓ_k, as well as processes $Z^{(k)} = (Z_{\ell_k + t} - Z_{\ell_k})_{t \geq 0}$, as follows. First, we let $\ell_0 := 0$ and $Z^{(0)} := Z$. Next, we let ℓ_1 be the time when Z_{t-} attains its *global* minimum (note that as $Z_t \to +\infty$ almost surely, this time is almost surely finite). This is necessarily a jump time of Z (equivalently, of N') but it is not a stopping time. Inductively, ℓ_{k+1} is the time when $Z^{(k)}_{t-}$ attains its global minimum. We also write $\Delta_k = \ell_{k+1} - \ell_k$ for the time spent between successive frontier times. See Figure 9 for a sample trajectory of the process Z_t with frontier times marked.

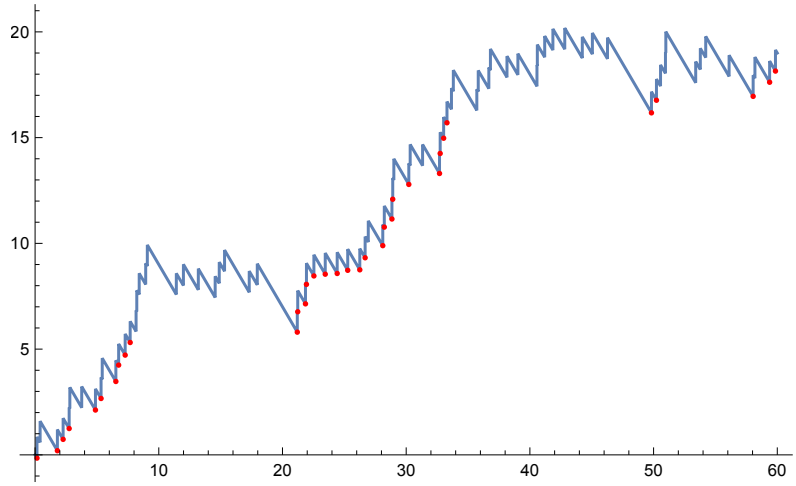

Figure 9: A simulation of the process Z_t with $\beta = 1.3$. The frontier times are marked red.

In terms of the simple exploration Y, the frontier times ℓ_k play the following role. Recall that the jump times of Z are exactly the times when Y discovers a new link. The frontier times are the times when Y discovers a new link *which is never backtracked.*

Proposition 3.4 (Survival and increments of the simple exploration)**.** *Let $\beta > 1$ and write $\mathcal{S} = \{\tau^Y = \infty\}$.*

1. We have that $\mathbb{P}(\mathcal{S}) = z$, where z is the unique positive solution of $1 - z = e^{-\beta z}$.

2. *There exists $C, c > 0$ such that*

$$\mathbb{P}(\mathcal{S}^c \cap \{\tau^Y \geq t\}) \leq Ce^{-ct}. \tag{3.5}$$

3. *Conditionally on $\mathcal{S}$, the sequence $\big\{\big(\Delta_k, (Z_t^{(k)})_{0\leq t<\Delta_k}\big)\big\}_{k=0}^{+\infty}$ is i.i.d.*

4. *There exists $C, c > 0$ such that for any $k \geq 0$*

$$\mathbb{P}(\Delta_k \geq t \mid \mathcal{S}) \leq Ce^{-ct}, \quad t \geq 0. \tag{3.6}$$

The proof is based on well-known properties of Poisson processes, for completeness we provide details in Appendix A. The first two parts of the proposition tell us that the simple exploration either continues indefinitely, or it closes 'quickly'. Intuitively, the former scenario parallels the situation when the (true) exploration process X explores a large cycle. The other two parts tell us that, conditionally on Y 'surviving', the frontier times ℓ_k are renewal times, and the renewal intervals Δ_k are typically short.

We now turn to discussing some properties of the exploration process X. Recall that $\mathcal{J}_t^X := \#\{s \leq t \colon X \text{ discovers a new link at time } s\}$. Let

$$\mathcal{N}_t^X := \#\{v \in V_n : (v, \varphi) \in H_t^X \text{ for some } \varphi \in S^1\}$$

denote the total number of vertices visited by X up to time t, and let $\mathcal{A}_t^{\mathrm{his}} := \{\mathcal{N}_t^X \leq \frac{n}{2}\}$ denote the event that no more than $n/2$ vertices have been visited up to time t. Note that $\mathcal{N}_t^X \leq \mathcal{J}_t^X + 1$ and that $\mathcal{J}_t^X \leq N_t$ where N is the Poisson process of rate $n\frac{\beta}{n-1}$ in Proposition 3.1. From this and a simple argument using Laplace transform we see that

$$\mathbb{P}((\mathcal{A}_t^{\mathrm{his}})^c) = \mathbb{P}(\mathcal{N}_t^X > \tfrac{n}{2}) \leq \exp\big(-\tfrac{n}{3}(\log(\tfrac{n-1}{3t\beta}) - 1)\big). \tag{3.7}$$

In particular, if $t = o(n)$ then $\mathbb{P}(\mathcal{A}_t^{\mathrm{his}}) \geq 1 - e^{-cn}$ for some $c > 0$.

By A_t^X we will denote the set of vertices available to the exploration X at time t by means of a new jump, i.e. $A_t^X = \emptyset$ if $t \geq \tau^X$, otherwise if $X_t = (v, \varphi, d)$ then

$$A_t^X := \{w \in V_n \setminus \{v\} \colon (w, \varphi) \notin H_t^X\}.$$

Recall that a *counting process* is a nondecreasing, integer valued càdlàg stochastic process starting at zero and with jumps equal to one. Let J be an $\mathcal{F}_t$-adapted counting process. We will say that a nonnegative process λ is an *intensity* of J if λ is $\mathcal{F}_t$-progressively measurable, $\int_0^t \lambda_u du < \infty$ a.s. for all t, and the process $J_t - \int_0^t \lambda_s \, ds$ is an $\mathcal{F}_t$-martingale.

Lemma 3.5 (Intensity of jumps)**.** *The processes $\mathcal{J}^X$ and $\mathcal{N}^X$ are counting processes with intensities λ, μ given respectively by*

$$\lambda_t = \frac{\beta}{n-1}|A_t^X| \qquad \text{and} \qquad \mu_t = \frac{\beta}{n-1}(n - \mathcal{N}_t^X).$$

In particular, on the event $\mathcal{A}_t^{\mathrm{his}}$ we have $\mu_t \geq \frac{\beta}{2}$.

A proof of this rather intuitive statement may be found (in a more general setting) in [1, Lemma 3.7]. The following lemma also appears in a more general form in [1, Lemma A.2], we include its proof here for the sake of completeness.

Lemma 3.6. *Suppose M is a counting process with intensity λ and let $\Lambda_t = \int_0^t \lambda_s \, ds$. Let σ, τ be stopping times such that $\sigma \leq \tau$. Let $\ell > 0$. Then we have*

$$\mathbb{P}\left(\{M_\tau - M_\sigma \leq \ell/2\} \cap \{\Lambda_\tau - \Lambda_\sigma \geq \ell\} | \mathcal{F}_\sigma\right) \leq e^{-\ell/8}.$$

Proof. Consider any $A \in \mathcal{F}_\sigma$ with positive probability and the process $\widetilde{M}_t = M_{\sigma+t} - M_\sigma$, which is a counting process with intensity $\widetilde{\lambda}_t = \lambda_{\sigma+t}$ with respect to the filtration $\widetilde{\mathcal{F}}_t = \mathcal{F}_{\sigma+t}$. Let $\widetilde{\mathbb{P}}(\cdot) = \mathbb{P}(\cdot|A)$. We have $\widetilde{\Lambda}_t = \int_0^t \widetilde{\lambda}_s ds = \Lambda_{\sigma+t} - \Lambda_\sigma$. Let N be a Poisson process with intensity 1 such that $N_t = M_{\Lambda_t}$ almost surely (see [1, Theorem A.1] and references there). We get

$$\begin{aligned}\widetilde{\mathbb{P}}(\{M_\tau - M_\sigma \le \ell/2\} \cap \{\Lambda_\tau - \Lambda_\sigma \ge \ell\}) &= \widetilde{\mathbb{P}}(\{\widetilde{M}_{\tau-\sigma} \le \ell/2\} \cap \{\widetilde{\Lambda}_{\tau-\sigma} \ge \ell\}) \\ &\le \widetilde{\mathbb{P}}(N_\ell \le \ell/2).\end{aligned}$$

Using the form of the Laplace transform of N_ℓ and Chebyshev's inequality we obtain

$$\begin{aligned}\widetilde{\mathbb{P}}(N_\ell \le \ell/2) &\le \inf_{a\ge 0} \exp\left((e^{-a}-1)\ell + a\ell/2\right) \\ &\le \inf_{a\ge 0} \exp\left(\frac{1}{2}a^2\ell - \frac{a\ell}{2}\right) = e^{-\ell/8},\end{aligned}$$

where in the second step we have used the elementary inequality $e^{-a} - 1 + a \le \frac{1}{2}a^2$ valid for $a \ge 0$. Thus we get

$$\mathbb{P}(\{M_\tau - M_\sigma \le \ell/2\} \cap \{\Lambda_\tau - \Lambda_\sigma \ge \ell\} \,|A) \le e^{-\ell/8},$$

for arbitrary $A \in \mathcal{F}_\sigma$ of positive probability, which implies the lemma. □

Corollary 3.7 (Visits to previously unvisited vertices)**.** *Let σ be a stopping time with respect to the filtration of the exploration process X and let $\eta^X(\sigma)$ be the first time after σ when X makes a jump to a previously unvisited vertex. For any $t > 0$ on the event $\mathcal{A}^{\mathrm{his}}_\sigma$ we have*

$$\mathbb{P}\left(\eta^X(\sigma) \wedge \tau^X - \sigma \ge t \mid \mathcal{F}_\sigma\right) \le e^{-t/16}$$

Proof. By definition of $\eta^X(\sigma)$, between times σ and $\eta^X(\sigma) \wedge \tau^X$ there are no jumps to previously unvisited vertices. In particular $\eta^X(\sigma) \wedge \tau^X - \sigma \ge t$ implies that $\mathcal{N}^X_{\sigma+t} - \mathcal{N}^X_\sigma = 0$ and that $\mathcal{A}^{\mathrm{his}}_{\sigma+t}$ holds. Thus Lemma 3.5 implies, with μ_t being the intensity of $\mathcal{N}^X_t$ and $\Lambda_t = \int_0^t \mu_s\, ds$, that then $\Lambda_{\sigma+t} - \Lambda_t \ge t/2$. Applying Lemma 3.6 with $M_t = \mathcal{N}^X_t$, $\tau = \sigma + t$ and $\ell = t/2$ easily gives the desired estimate. □

4 Balance

This section contains the main work of the paper. The goal of the section is to prove that large cycles are 'balanced' in the sense that they contain roughly equal numbers of vertices passed in the directions ↑ and ↓. In fact we show that, with high probability, in a cycle which is at least $\lfloor n^{1/2} \rfloor$ long each segment of $\lfloor n^{1/2} \rfloor$ consecutive vertices is balanced in this sense. Throughout the section we work with a random $\omega \in \Omega$ sampled from the Poisson measure $\mathbb{P}_\beta$ for some fixed $\beta > 1$ (recall that $\nu \in [0, 1)$ is fixed).

We start by introducing some notation. Given $\omega \in \Omega$, $v \in V_n$ and $k \in \mathbb{N}$, let us write $\mathcal{C}_\omega(v) = (v_1^{d_1}, v_2^{d_2}, \dots, v_\ell^{d_\ell}) \in \mathfrak{C}_\omega$ for the cycle containing v. Without loss of generality we may assume that $v = v_1$ and that $d_1 = \uparrow$. Under these assumptions, we let

$$c_\omega(v, k) := \{v_i \in \mathcal{C}_\omega(v) : 1 \le i \le k \wedge \ell\}$$

denote the set of the first k vertices of $\mathcal{C}$ following v. Note that if the cycle containing v has length smaller than k then $|c_\omega(v, k)| = |\mathcal{C}_\omega(v)| < k$. We further let

$$c^{\uparrow}_\omega(v, k) := \{v_i \in c_\omega(v, k) : d_i = \uparrow\}, \qquad c^{\downarrow}_\omega(v, k) := \{v_i \in c_\omega(v, k) : d_i = \downarrow\}$$

denote those vertices in $c_\omega(v,k)$ which are passed in the same, respectively opposite, direction as v. Finally we define the *balance* $\mathfrak{B}_\omega(v,k)$ of the segment of length k after v in the loop, as

$$\mathfrak{B}_\omega(v,k) := |c_\omega^{\uparrow}(v,k)| - |c_\omega^{\downarrow}(v,k)|.$$

The main result of this section is the following proposition, which tells us that $\mathfrak{B}_\omega(v,\lfloor n^{1/2}\rfloor)$ is typically of much smaller order than $n^{1/2}$.

Proposition 4.1 (Segments of cycles are balanced)**.** *Let $\beta_1 > \beta_0 > 1$. There exist $C, c > 0$ such that for all $\beta \in [\beta_0, \beta_1]$ and for any $v \in V_n$, we have*

$$\mathbb{P}_\beta\left(\left\{\left|\mathfrak{B}_\omega(v,\lfloor n^{1/2}\rfloor)\right| \geq n^{5/12}\log^3 n\right\} \cap \left\{|\mathcal{C}_\omega(v)| \geq \lfloor n^{1/2}\rfloor\right\}\right) \leq Ce^{-c\log^2 n}.$$

This says that cycles of length at least $\lfloor n^{1/2}\rfloor$ are very likely to have balance $|\mathfrak{B}| < n^{5/12}\log^3 n \ll n^{1/2}$. Cycles containing fewer than $\lfloor n^{1/2}\rfloor$ vertices may possibly be unbalanced, but this does not concern us. A key feature of this result is that the upper bound kills any polynomial in n, making it possible to use quite crude union bounds later in the proof of Theorem 1.1. Also note that the bound is claimed to be uniform in $\beta \in [\beta_0, \beta_1]$ for any $\beta_1 > \beta_0 > 1$. This will allow us to derive a version of the proposition stated above where we 'remove' a deterministic number of links from ω, which will be important for the coupling with PD($\frac{1}{2}$) in Section 5.

To formulate the last claim precisely, recall the notation $\vec{\omega}$ for the ordered list of links of ω, and note that $\mathcal{C}_\omega(v)$, $c_\omega(v,k)$, $c_\omega^{\uparrow}(v,k)$, $c_\omega^{\downarrow}(v,k)$ and $\mathfrak{B}_\omega(v,k)$ all depend on $\vec{\omega}$ only. Also recall that, for $s \leq |\omega|$, we write $\vec{\omega}_s = \{(e_1,m_1),\dots,(e_s,m_s)\}$ for the sequence of the first s links. If X_ω is a random variable which only depends on the relative order of links in ω we write $X_{\vec{\omega}}$ for its value on any link configuration with the same relative order.

Proposition 4.2. *Let $\beta > 1$ and $\rho \in [0,1)$. For $1 \leq s \leq |\omega|$ write*

$$\mathcal{A}(v,s) := \left\{\mathfrak{B}_{\vec{\omega}_s}(v,\lfloor n^{1/2}\rfloor) \geq n^{5/12}\log^3 n\right\} \cap \left\{|\mathcal{C}_{\vec{\omega}_s}(v)| \geq \lfloor n^{1/2}\rfloor\right\}.$$

There exist $C, c > 0$ such that

$$\mathbb{P}_\beta\left(\bigcup_{v\in V_n}\bigcup_{|\omega|-n^\rho\leq s\leq|\omega|}\mathcal{A}(v,s)\right) \leq Ce^{-c\log^2 n}.$$

The proofs are given at the end of the section, after several preparatory results.

4.1 Winding processes

Recall the notation $X_t = (v_t, \varphi_t, d_t)$ and $Y_t = (k_t, v_t, \varphi_t, d_t)$ for the exploration and simple exploration, in particular that $d_t \in \{-1,+1\}$ indicates the direction of motion. We will use superscripts X and Y on v_t, φ_t, d_t to distinguish between the two processes. Define the *winding processes* $\{\mathcal{L}_t^X\}_{t\geq 0}$ and $\{\mathcal{L}_t^Y\}_{t\geq 0}$ by

$$\mathcal{L}_t^X := \int_0^t d_s^X\,ds, \qquad \mathcal{L}_t^Y := \int_0^t d_s^Y\,ds.$$

Thus $\mathcal{L}^X$ increases at rate 1 when the process X travels in the positive direction, otherwise it decreases at rate 1, and the same is true for $\mathcal{L}^Y$.

To prove Proposition 4.1 we will first estimate $\mathcal{L}^X$, and then transfer these estimates to $\mathfrak{B}$. In order to estimate $\mathcal{L}^X$ we will use the coupling of X and Y introduced in Lemma 3.3, together with the following estimate on $\mathcal{L}^Y$:

Proposition 4.3 (Winding of the simple exploration process)**.** *Let $\beta_1 > \beta_0 > 1$. There exist $C, c > 0$ such that for any $\beta \in [\beta_0, \beta_1]$, $T > 0$ and $s \in [1, T^{1/2}]$ we have*

$$\mathbb{P}_\beta\Big(\sup_{t \le T} |\mathcal{L}^Y_t| \ge sT^{1/2} \,\Big|\, \mathcal{S} \Big) \le C \exp(-cs^2),$$

where $\mathcal{S} = \{\tau^Y = +\infty\}$.

Proof. Fix $\beta \in [\beta_0, \beta_1]$. We use Proposition 3.4 and the notation therein, we also write $\overline{\mathbb{P}}(\cdot) = \mathbb{P}_\beta(\cdot \mid \mathcal{S})$. For lighter notation, within this proof let $\mathcal{L}_t = \mathcal{L}^Y_t$.

At each frontier time ℓ_k the process Z jumps, meaning that Y discovers a new link. We let $\ell^*_0 = 0$ and let $\{\ell^*_k\}_{k \ge 1}$ be the subsequence consisting of the times ℓ_k at which the link is marked as a bar (i.e. $\xi_i = -1$ in the notation of Definition 3.2). As the choice of markings is independent of Z, using Proposition 3.4 we conclude that $\Delta^*_k := \ell^*_{k+1} - \ell^*_k$ form an i.i.d. sequence under $\overline{\mathbb{P}}$, satisfying

$$\overline{\mathbb{P}}(\Delta^*_k \ge s) \le \tilde{C} e^{-\tilde{c}s}, \quad s \ge 0,$$

for some $\tilde{C}, \tilde{c} > 0$. Also, the increments $\mathcal{L}_{\ell^*_{k+1}} - \mathcal{L}_{\ell^*_k}$ are independent under $\overline{\mathbb{P}}$.

Now, the key observation is that we have the equality in distribution

$$\mathcal{L}_{\ell^*_{k+2}} - \mathcal{L}_{\ell^*_{k+1}} \overset{(\mathrm{d})}{=} -\big(\mathcal{L}_{\ell^*_{k+1}} - \mathcal{L}_{\ell^*_k}\big),$$

because upon crossing a bar the winding processes changes its orientation. Using these facts we infer that for any $k \in \{0, 1, \dots\}$

$$\mathcal{Q}_k := \mathcal{L}_{\ell^*_{k+2}} - \mathcal{L}_{\ell^*_k} = \big(\mathcal{L}_{\ell^*_{k+2}} - \mathcal{L}_{\ell^*_{k+1}}\big) + \big(\mathcal{L}_{\ell^*_{k+1}} - \mathcal{L}_{\ell^*_k}\big)$$

are symmetric random variables with $\{\mathcal{Q}_{2k}\}_{k \ge 0}$ being independent. Moreover, by $|\mathcal{L}_{\ell^*_{k+1}} - \mathcal{L}_{\ell^*_k}| \le \Delta^*_k$ and (4.1) they have exponential tails.

Let us set $K = \lfloor c_1 T \rfloor$, for some $c_1 > 0$ to be chosen later. We consider the cases $\sum_{i=0}^{2K} \Delta^*_i > T$ and $\sum_{i=0}^{2K} \Delta^*_i \le T$ separately. If $\sum_{i=0}^{2K} \Delta^*_i > T$ then we can cover $[0, T]$ with the intervals $[\ell^*_{2k}, \ell^*_{2k+2})$ for $k \le K$. As $\mathcal{L}_t$ is continuous in t, we can replace the supremum of $|\mathcal{L}_t|$ by maximum. The maximum $\max_{0 \le t \le T} \mathcal{L}_t$ can then be bounded by the maximum at endpoints ℓ^*_{2k} plus the maximum increment over all the intervals $[\ell^*_{2k}, \ell^*_{2k+2})$. The maximum increment on $[\ell^*_{2k}, \ell^*_{2k+2})$ is in turn bounded by the length $\Delta^*_{2k} + \Delta^*_{2k+1}$. We thus get

$$\begin{aligned}\overline{\mathbb{P}}\left(\max_{t \le T} |\mathcal{L}_t| \ge sT^{1/2} \right) \le \overline{\mathbb{P}}\left(\max_{k \le K} \Big| \sum_{i=0}^{k} \mathcal{Q}_{2i} \Big| + \max_{k \le K} \big(\Delta^*_{2k} + \Delta^*_{2k+1} \big) \ge sT^{1/2} \right) \\ + \overline{\mathbb{P}}\left(\sum_{i=0}^{2K} \Delta^*_i \le T \right).\end{aligned}$$

The first of these two terms can be bounded by applying a union bound and Etemadi's inequality [6, Thm. 22.5], giving

$$\begin{aligned}\overline{\mathbb{P}}\left(\max_{k \le K} \Big| \sum_{i=0}^{k} \mathcal{Q}_{2i} \Big| + \max_{k \le K} \big(\Delta^*_{2k} + \Delta^*_{2k+1} \big) \ge sT^{1/2} \right) \\ \le 3 \max_{k \le K} \overline{\mathbb{P}}\left(\Big| \sum_{i=0}^{k} \mathcal{Q}_{2i} \Big| \ge \tfrac{1}{6} sT^{1/2} \right) + K \overline{\mathbb{P}}\left(\Delta^*_0 + \Delta^*_1 \ge \tfrac{1}{2} sT^{1/2} \right), \quad (4.1)\end{aligned}$$

where we have used that Δ_k^* are i.i.d. By symmetry of $\mathcal{Q}_{2i}$ and Markov's inequality we have for any $\theta > 0$ that

$$\overline{\mathbb{P}}\left(\Big|\sum_{i=0}^{k}\mathcal{Q}_{2i}\Big| \geq \tfrac{1}{6}sT^{1/2}\right) = 2\overline{\mathbb{P}}\left(\sum_{i=0}^{k}\mathcal{Q}_{2i} \geq \tfrac{1}{6}sT^{1/2}\right) \leq 2\frac{\left(\overline{\mathbb{E}}[e^{\theta\mathcal{Q}_0}]\right)^k}{\exp(\frac{1}{6}\theta sT^{1/2})}.$$

For small enough θ the Laplace transform $\overline{\mathbb{E}}[e^{\theta\mathcal{Q}_0}]$ is finite, since the $\mathcal{Q}_i$'s have exponential tails. Using that $\overline{\mathbb{E}}[\mathcal{Q}_0] = 0$ we see that there exist $\theta_0, c_2 > 0$ such that if $\theta \leq \theta_0$ then $\overline{\mathbb{E}}[e^{\theta\mathcal{Q}_0}] \leq e^{c_2\theta^2}$. Setting $\theta = sT^{-1/2}/(12c_1c_2)$ with c_1 chosen large enough so that $\theta < \theta_0$ we get (using $k \leq K \leq c_1T$ and $s \leq T^{1/2}$)

$$\overline{\mathbb{P}}\left(\Big|\sum_{i=0}^{k}\mathcal{Q}_{2i}\Big| \geq \tfrac{1}{6}sT^{1/2}\right) \leq 2\exp\left(c_1c_2\theta^2T - \tfrac{1}{6}\theta sT^{1/2}\right) \leq 2\exp(-c_3s^2),$$

with $c_3 = (72c_1c_2)^{-1}$.

For the second term on the right in (4.1) we recall that the Δ_k^*'s have exponential tails. As $s \in [1, T^{1/2}]$, we thus obtain for some $C_4, c_4, C_5, c_5 > 0$

$$K\overline{\mathbb{P}}\left(\Delta_0^* + \Delta_1^* \geq \tfrac{1}{2}sT^{1/2}\right) \leq KC_4e^{-c_4sT^{1/2}} \leq C_5e^{-c_5s^2}.$$

Finally, by standard large deviation considerations for i.i.d. variables we have

$$\overline{\mathbb{P}}\left(\sum_{i=0}^{2K}\Delta_i^* \leq T\right) \leq C_6e^{-c_6K},$$

for some $C_6, c_6 > 0$, provided we choose c_1 large enough so that the mean of the sum above is larger than T.

This proves the claim for any fixed $\beta \in [\beta_0, \beta_1]$. The uniformness over such β follows since the upper bound can be chosen as a continuous function of β. □

In order to compare $\mathcal{L}$ with $\mathfrak{B}$ we need to keep track of how many times the exploration passes level $0 \in S^1$. To this end we make the following definitions. Denote by $\mathcal{K}_t^X(v,\varphi)$ (respectively, $\mathcal{K}_t^Y(v,\varphi)$) the number of times X (respectively, Y) passes through $0 \in S^1$ when started at (v,φ) moving in the positive direction ($d_0 = +1$) and run for time t. We will write $\mathcal{K}_t^X$ (respectively, $\mathcal{K}_t^Y$) when the starting point is not ambiguous. Recall the definition of $\mathcal{I}_t^X(v,\varphi)$ given right after (3.4).

Proposition 4.4 (Winding and the number of visited vertices)**.** *For any $t > 0$ and any $(v,\varphi) \in V_n \times S^1$ we have*

$$\mathcal{K}_t^X(v,\varphi) \leq t + \mathcal{I}_t^X(v,\varphi) + 1, \quad \mathcal{K}_t^Y(v,\varphi) \leq t + \mathcal{I}_t^Y(v,\varphi) + 1. \tag{4.2}$$

Moreover, there exists $c = c(\beta) > 0$ and $t_0 = t_0(\beta)$ such that

$$\mathbb{E}_\beta[\mathcal{K}_t^Y \mid \mathcal{S}] \geq ct, \quad \text{for all } t \geq t_0, \tag{4.3}$$

where as usual $\mathcal{S} = \{\tau^Y = +\infty\}$.

Proof. The proofs of (4.2) for X and Y are the same. We write $\mathcal{K}_t$ and $\mathcal{I}_t$, omitting X, Y, v and φ in order to simplify the notation.

Define a sequence of times $\tau_0 := 0$ and, for $i \geq 1$,

$$\tau_i := \inf\{t > \tau_{i-1} : \varphi_t = 0\}.$$

We first claim that for $i \geq 1$ we have

$$\text{if} \quad \mathcal{I}_{\tau_{i+1}} - \mathcal{I}_{\tau_i} = 0 \quad \text{then} \quad \tau_{i+1} - \tau_i = 1. \tag{4.4}$$

Indeed, at time τ_i the exploration passed $0 \in S^1$ and if it passes 0 again without traversing a link this means that it has completed a full lap on one copy of S^1. From (4.4) we deduce that

$$(\tau_{i+1} - \tau_i) + (\mathcal{I}_{\tau_{i+1}} - \mathcal{I}_{\tau_i}) \geq 1, \text{ for all } i \geq 1.$$

Let k be such that $t \in [\tau_k, \tau_{k+1})$. Then

$$\mathcal{K}_t - 1 = \mathcal{K}_{\tau_k} - 1 = k - 1 \leq \sum_{i=0}^{k-1} \left[(\tau_{i+1} - \tau_i) + (\mathcal{I}_{\tau_{i+1}} - \mathcal{I}_{\tau_i}) \right] = \tau_k + \mathcal{I}_{\tau_k} \leq t + \mathcal{I}_t,$$

as claimed.

Now we turn to (4.3). Recall from Proposition 3.4, and the discussion preceding it, the notation Δ_k and $Z^{(k)}$ as well as the notion of frontier times. Let us use the term *return times* for the jump times of Z which are not frontier times, and *return links* for the corresponding links traversed by Y. Observe that $\mathcal{K}_t^Y \geq \mathcal{R}_t^Y$, where $\mathcal{R}_t^Y$ denotes the number of return links which have been backtracked by Y up to time t. This is because Y does not visit its own history other than by backtracking, hence between discovering a return link and backtracking it Y must complete at least one circle.

Let R_k denote the total number of return times of $(Z_t)_{\ell_k \leq t \leq \ell_{k+1}}$. By Proposition 3.4, conditionally on $\mathcal{S}$ the sequence $\{(\Delta_k, R_k)\}_{k \geq 1}$ is a renewal-reward process, and by the basic renewal-reward theorem [11, Theorem 10.5] it thus follows that

$$\frac{\mathbb{E}[\mathcal{K}_t^Y \mid \mathcal{S}]}{t} \geq \frac{\mathbb{E}[\mathcal{R}_t^Y \mid \mathcal{S}]}{t} \to \frac{\mathbb{E}[R_1 \mid \mathcal{S}]}{\mathbb{E}[\Delta_1 \mid \mathcal{S}]}, \quad \text{as } t \to \infty.$$

The result (4.3) follows from $\mathbb{E}[R_1 \mid \mathcal{S}] > 0$, which is easily checked. For example, it suffices to check that $\mathbb{P}(R_1 = 1 \mid \mathcal{S}) > 0$. Letting $\sigma_1, \sigma_2, \ldots$ denote the jump times of Z and using that $\mathbb{P}(R_1 = 1 \mid \mathcal{S}) = \mathbb{P}(\ell_1 = \sigma_2 \mid \mathcal{S})$ we get

$$\begin{aligned} \mathbb{P}(R_1 = 1 \mid \mathcal{S}) &= \frac{\mathbb{P}(\sigma_1 < 1, \sigma_2 - \sigma_1 \in (1, 2 - \sigma_1), (Y_{\sigma_2 + t} - Y_{\sigma_2})_{t \geq 0} \in \mathcal{S})}{\mathbb{P}(\mathcal{S})} \\ &= \mathbb{P}(\sigma_1 < 1, \sigma_2 - \sigma_1 \in (1, 2 - \sigma_1)) = e^{-2\beta}(e^\beta - 1 - \beta) > 0. \end{aligned}$$ □

We now come to the key technical result of the paper, an upper bound on $\mathcal{L}^X$ when X explores part of a large cycle. At the same time we also provide a lower bound on $\mathcal{K}^X$ since the proof follows a similar structure.

Proposition 4.5 (Winding for the exploration process is small)**.** *Let $\beta_1 > \beta_0 > 1$, and consider the exploration X started at an arbitrary point (v, φ, d). There exist $C_1, c_1 > 0$ such that for any $\beta \in [\beta_0, \beta_1]$ we have*

$$\mathbb{P}_\beta \left(|\mathcal{L}^X_{n^{1/3}}| \mathbb{1}_{\{\tau^X \geq n^{1/3}\}} \geq 3n^{1/4} \log n \right) \leq C_1 e^{-c_1 \log^2 n}.$$

Moreover, we have for some $c_2 > 0$

$$\mathbb{P}_\beta \left(\left\{ \mathcal{K}^X_{n^{1/3}} < c_2 n^{1/3} \right\} \cap \left\{ \tau^X \geq n^{1/3} \right\} \right) \leq C_1 e^{-c_1 \log^2 n}.$$

Before giving the proof we outline the main ideas. We want to use the coupling of the exploration process X to the simple exploration process Y from Lemma 3.3, as well as the concentration result for the latter process, Proposition 4.3. To get good concentration we will decompose $[0, n^{1/3}]$ into many shorter time intervals $[t_i, t_{i+1})$ of

length approximately $n^{1/6}$ each. On each $[t_i, t_{i+1})$ we will wait for a 'good' coupling with a simple exploration: first we wait until the exploration X jumps to a new vertex so we can start a coupling, then we check if the simple exploration survives indefinitely, which it does with probability $z > 0$. If so, we can apply Proposition 4.3 in this interval. If not, then we repeat the procedure, waiting for a jump of X to a new vertex and looking at the coupled simple exploration. Typically we only need to perform this a small number of times until we get a coupling with a simple exploration which survives.

Let us make these ideas formal and introduce the setup that will be used in the proof. Set $a_n := n^{1/3}$. We define $t_i := i \cdot b_n$, where $i \in \{0, 1, \dots, \lfloor n^{1/6} \rfloor\}$, $b_n := a_n / \lfloor n^{1/6} \rfloor$. Writing $m = \lfloor n^{1/6} \rfloor - 1$, we decompose

$$\mathcal{L}^X_{a_n} = \sum_{i=0}^{m} (\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{t_i}).$$

Fix $i \in \{0, 1, \dots, m\}$. Let us first analyse the change of the winding process on one interval $[t_i, t_{i+1})$. To this end we will define two sequences of times $\{\sigma^i_k\}_{k=0}^{+\infty}$ and $\{\tau^i_k\}_{k=0}^{+\infty}$ as well as a sequence of simple explorations $\{Y^i_k\}_{k=1}^{\infty}$. The σ^i_k will form a non-decreasing sequence, taking values in $[t_i, t_{i+1}]$, and will be defined so that, for $k \geq 1$ and as long as $\sigma^i_k < t_{i+1}$, the process X jumps to a new vertex at time σ^i_k. For such k, the process Y^i_k is defined to be an independent copy of a simple exploration, coupled with X as in Lemma 3.3, starting at time σ^i_k. The possibility $\sigma^i_k = t_{i+1}$ signifies that we have finished with the interval $[t_i, t_{i+1})$ and must move on to the next one.

We now define the times σ^i_k and τ^i_k. First we set $\sigma^i_0 := t_i, \tau^i_0 := 0$. Next, for $k = 1, 2, \dots$ we set

$$\sigma^i_k := t_{i+1} \wedge \inf\left\{s \geq \sigma^i_{k-1} + \tau^i_{k-1} : X \text{ jumps to a new vertex at time } s\right\}$$

where $\tau^i_k := \tau^{Y^i_k}$ is the time when Y^i_k terminates (returns to its starting point). Note that if $\sigma^i_{k-1} = t_{i+1}$ then $\sigma^i_k = \sigma^i_{k+1} = \ldots = t_{i+1}$. In particular, this will occur if $\tau^i_{k-1} = \infty$. In this case we do not need to define $Y^i_k, Y^i_{k+1}, \dots$ Also note that, since the coupling of Y^i_k with X entails constructing both processes using the same sources of randomness, we may work with the τ^i_k as if they are adapted to the filtration of X, even though they are defined in terms of Y.

In words, these definitions mean that, firstly, Y^i_1 is a simple exploration coupled with X, started at time σ^i_1, the first time in $[t_i, t_{i+1})$ that X jumps to a new vertex. This coupling is then run either for the remaining time in $[t_i, t_{i+1})$, or until Y^i_1 returns to its starting point (after time τ^i_1). For $k \geq 2$, if the simple exploration Y^i_{k-1} has returned to its starting point, at time $\sigma^i_{k-1} + \tau^i_{k-1} \in [t_i, t_{i+1})$, then we wait until X jumps to a new vertex again. We call the time when this occurs σ^i_k and we begin a new coupling with a simple exploration, Y^i_k, from the location of X at this time.

Let

$$k_0 = k^i_0 := \min\left\{k \in \mathbb{N} : \tau^i_k = +\infty \text{ or } \sigma^i_{k+1} = t_{i+1}\right\}.$$

The first possibility, $\tau^i_{k_0} = +\infty$, means that at attempt number k_0 the coupled simple exploration $Y^i_{k_0}$ survives (and is the first one with this property). The other possibility, that $\tau^i_{k_0} < +\infty$ but $\sigma^i_{k_0+1} = t_{i+1}$, means that after time σ_{k_0} the exploration X never jumps to a new vertex until the end of the interval $[t_i, t_{i+1})$. Included in this possibility is the case when X closes the loop before jumping again. Intuitively, k_0 is the number of attempts at coupling X with a simple exploration which survives, until we either succeed or run out of time.

We now turn to the proof of the proposition.

Proof of Proposition 4.5. Fix $\beta \in [\beta_0, \beta_1]$. We first show (4.5) for this β. Recall the definition of the event $\mathcal{A}_t^{\mathrm{his}}$ given above (3.7). First note that it suffices to show that

$$\mathbb{P}_\beta \left(\mathcal{A}_{a_n}^{\mathrm{his}} \cap \left\{ |\mathcal{L}_{a_n}^X| \mathbb{1}_{\{\tau^X \geq a_n\}} \geq 3n^{1/4} \log n \right\} \right) \tag{4.5}$$

satisfies the claimed bound, due to (3.7). Also note that $\mathcal{A}_{a_n}^{\mathrm{his}} \subseteq \mathcal{A}_s^{\mathrm{his}}$ for $s \leq a_n$.

Consider the interval $[t_i, t_{i+1})$, the stopping times σ_k^i, τ_k^i and the variable k_0^i from the preceding discussion. Keeping i fixed for now, we will drop it from the superscript on σ_k, τ_k and k_0. We claim that, under $\mathbb{P}(\cdot|\mathcal{F}_{t_i})$, the random variable k_0 is stochastically dominated by a geometric distribution with parameter z, that is to say,

$$\text{for all } k \geq 1, \quad \mathbb{P}(k_0 \geq k \mid \mathcal{F}_{t_i}) \leq (1-z)^{k-1}. \tag{4.6}$$

Here $z > 0$ is the survival probability of a simple exploration, see Proposition 3.4. The claim is easily established by induction, using

$$\mathbb{P}(k_0 \geq k+1 \mid \mathcal{F}_{t_i}) = \mathbb{P}(k_0 \geq k+1 \mid \mathcal{F}_{t_i}, \{k_0 \geq k\}) \mathbb{P}(k_0 \geq k \mid \mathcal{F}_{t_i})$$

and

$$\mathbb{P}(k_0 \geq k+1 \mid \mathcal{F}_{t_i}, \{k_0 \geq k\}) \leq \mathbb{P}(\tau_k < \infty \mid \mathcal{F}_{t_i}, \{k_0 \geq k\}) = 1 - z.$$

We now show that there exist constants $C_2, c_2 > 0$, uniform in n and in i, such that for all $t > 0$, on the event $\mathcal{A}_{a_n}^{\mathrm{his}} \cap \{\tau^X \geq a_n\}$ we have

$$\mathbb{P}(\sigma_{k_0} - t_i \geq t \mid \mathcal{F}_{t_i}) \leq C_2 e^{-c_2 t}. \tag{4.7}$$

First we establish that there are $C_1, c_1 > 0$ such that for any $k \geq 0$ and any $t > 0$, on $\mathcal{A}_{a_n}^{\mathrm{his}} \cap \{\tau^X \geq a_n\}$ we have

$$\mathbb{P}\left((\sigma_{k+1} - \sigma_k) \mathbb{1}_{\{k_0 > k\}} \geq t \mid \mathcal{F}_{\sigma_k} \right) = \mathbb{P}(k_0 > k, \sigma_{k+1} - \sigma_k \geq t \mid \mathcal{F}_{\sigma_k}) \leq C_1 e^{-c_1 t}. \tag{4.8}$$

Indeed, for any $k < k_0$ we have $\tau_k < \infty$, so

$$\begin{aligned} &\mathbb{P}\left(k_0 > k, \sigma_{k+1} - \sigma_k \geq t \mid \mathcal{F}_{\sigma_k}\right) \\ &\leq \mathbb{P}\left(\{\sigma_{k+1} - \sigma_k \geq t\} \cap \{\tau_k \leq t/2\} \mid \mathcal{F}_{\sigma_k}\right) + \mathbb{P}\left(\tau_k > t/2,\, \tau_k < +\infty \mid \mathcal{F}_{\sigma_k}\right). \end{aligned}$$

The second term is at most Ce^{-ct}, for some $C, c > 0$, by (3.5) from Proposition 3.4. To estimate the first term, note that $\tau_k \leq t/2$ together with $\sigma_{k+1} - \sigma_k \geq t$ implies that $\sigma_{k+1} - (\sigma_k + \tau_k) \geq t/2$, in particular X does not visit previously unexplored vertices for time at least $t/2$ after $\sigma_k + \tau_k$. Thus

$$\begin{aligned} &\mathbb{P}\left(\{\sigma_{k+1} - \sigma_k \geq t\} \cap \{\tau_k \leq t/2\} \mid \mathcal{F}_{\sigma_k}\right) \mathbb{1}_{\mathcal{A}_{a_n}^{\mathrm{his}}} \mathbb{1}_{\{\tau^X \geq a_n\}} \\ &\qquad \leq \mathbb{E}\big[\mathbb{P}\big(\{\sigma_{k+1} \wedge \tau^X - (\sigma_k + \tau_k) \geq t/2\} \cap \mathcal{A}_{\sigma_k + \tau_k}^{\mathrm{his}} \mid \mathcal{F}_{\sigma_k + \tau_k}\big) \mid \mathcal{F}_{\sigma_k}\big]. \end{aligned} \tag{4.9}$$

By Corollary 3.7 the probability is at most $e^{-c't}$ for some $c' > 0$, which together with the previous estimate proves (4.8).

Now note that

$$\sigma_{k_0} - t_i = \sigma_{k_0} - \sigma_0 = \sum_{j=0}^{k_0 - 1} \left(\sigma_{j+1} - \sigma_j\right),$$

where by (4.8) each summand, conditionally on all previous terms, has exponential tails. Since k_0, the number of summands, is by (4.6) itself dominated by a geometric random

variable, one may conclude that the sum itself has exponential tails, as claimed in (4.7). In more detail, we have for any $k > 0$ that

$$\mathbb{P}(\sigma_{k_0} - \sigma_0 \geq t \mid \mathcal{F}_{t_i}) \leq \mathbb{P}\Big(\sum_{j=0}^{k} (\sigma_{j+1} - \sigma_j) \mathbb{1}_{\{k_0 > j\}} \geq t \,\Big|\, \mathcal{F}_{t_i} \Big) + \mathbb{P}(k_0 > k \mid \mathcal{F}_{t_i}).$$

Now for any $\theta > 0$ we have

$$\mathbb{P}\Big(\sum_{j=0}^{k} (\sigma_{j+1} - \sigma_j) \mathbb{1}_{\{k_0 > j\}} \geq t \,\Big|\, \mathcal{F}_{t_i} \Big) \leq e^{-\theta t} \mathbb{E}\Big[\exp\Big(\theta \sum_{j=0}^{k} (\sigma_{j+1} - \sigma_j) \mathbb{1}_{\{k_0 > j\}} \Big) \,\Big|\, \mathcal{F}_{t_i} \Big],$$

where (using that $\{k_0 > k-1\} \in \mathcal{F}_{\sigma_k}$)

$$\begin{aligned} &\mathbb{E}\Big[\exp\Big(\theta \sum_{j=0}^{k} (\sigma_{j+1} - \sigma_j) \mathbb{1}_{\{k_0 > j\}} \Big) \,\Big|\, \mathcal{F}_{t_i} \Big] \\ &\qquad = \mathbb{E}\Big[\exp\Big(\theta \sum_{j=0}^{k-1} (\sigma_{j+1} - \sigma_j) \mathbb{1}_{\{k_0 > j\}} \Big) \mathbb{E}\big[e^{\theta(\sigma_{k+1} - \sigma_k) \mathbb{1}_{\{k_0 > k\}}} \mid \mathcal{F}_{\sigma_k} \big] \,\Big|\, \mathcal{F}_{t_i} \Big]. \end{aligned}$$

Here the inner factor may be written as

$$\mathbb{E}\big[e^{\theta(\sigma_{k+1} - \sigma_k) \mathbb{1}_{\{k_0 > k\}}} \mid \mathcal{F}_{\sigma_k} \big] = \int_0^\infty \mathbb{P}\big(e^{\theta(\sigma_{k+1} - \sigma_k) \mathbb{1}_{\{k_0 > k\}}} > s \mid \mathcal{F}_{\sigma_k} \big) ds.$$

Using (4.8) we conclude that we may choose $\theta > 0$ (depending on constants c_1, C_1 in (4.8)) such that, on $\mathcal{A}^{\mathrm{his}}_{a_n} \cap \{\tau^X \geq a_n\}$,

$$\mathbb{E}\big[e^{\theta(\sigma_{k+1} - \sigma_k) \mathbb{1}_{\{k_0 > k\}}} \mid \mathcal{F}_{\sigma_k} \big] \leq e,$$

say, for all $k \geq 0$. It follows by induction that

$$\mathbb{E}\Big[\exp\Big(\theta \sum_{j=0}^{k} (\sigma_{j+1} - \sigma_j) \mathbb{1}_{\{k_0 > j\}} \Big) \,\Big|\, \mathcal{F}_{t_i} \Big] \mathbb{1}_{\mathcal{A}^{\mathrm{his}}_{a_n}} \mathbb{1}_{\{\tau^X \geq a_n\}} \leq e^k,$$

and hence

$$\mathbb{P}(\sigma_{k_0} - \sigma_0 \geq t \mid \mathcal{F}_{t_i}) \mathbb{1}_{\mathcal{A}^{\mathrm{his}}_{a_n}} \mathbb{1}_{\{\tau^X \geq a_n\}} \leq e^{k - \theta t} + \mathbb{P}(k_0 > k \mid \mathcal{F}_{t_i}).$$

Setting $k = \lfloor \frac{\theta}{2} t \rfloor$ and using (4.6), this gives (4.7).

Recall the notion of a failed coupling from Lemma 3.3. The bound (4.7) tells us that typically we don't wait too long for a coupling with a simple exploration process that survives. If the coupling doesn't fail, we will be able to transfer estimates of the winding process from the simple exploration to the process X.

To this end we distinguish three possible scenarios of what can happen during a given time interval. We say that the interval $[t_i, t_{i+1})$ is *good*, denoting this event by $\mathcal{G}_i$, if the following hold:

- $\tau^i_{k_0} = +\infty$, and
- none of the k_0 attempted couplings failed until time $T = t_{i+1} - t_i \leq n^{1/6}$.

On the event $\mathcal{G}_i$ the coupling started at time $\sigma^i_{k_0}$ survives and it lasts until time t_{i+1}, in particular X cannot close its loop before time t_{i+1} (as this would entail X returning to some vertex visited before time $\sigma^i_{k_0}$ and hence Y^{k_0} returning to its starting point, i.e.

$\tau^{Y^{k_0}} < \infty$). Thus $\mathcal{G}_i \subseteq \{\tau^X \geq t_{i+1}\}$. Next, we say that the interval $[t_i, t_{i+1})$ is *terminal* if $\tau^i_{k_0} < +\infty$ and $\sigma^i_{k_0+1} = t_{i+1}$, and we denote this event by $\mathcal{T}_i$. Note that $\{\tau^X < t_{i+1}\} \subseteq \mathcal{T}_i$. Finally we let $\mathcal{B}_i = (\mathcal{G}_i \cup \mathcal{T}_i)^c$, and if this event occurs we say that the interval $[t_i, t_{i+1})$ is *bad*. On this event one of the attempted couplings failed.

Let us now estimate the winding process on each of the above events. We have

$$\begin{aligned}
|\mathcal{L}^X_{a_n}| \mathbb{1}_{\{\tau^X \geq a_n\}} &\leq \sum_{i=0}^{m} |\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{t_i}| \mathbb{1}_{\{\tau^X \geq a_n\}} \\
&= \sum_{i=0}^{m} |\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{t_i}| \big(\mathbb{1}_{\mathcal{G}_i} + \mathbb{1}_{\mathcal{B}_i} + \mathbb{1}_{\mathcal{T}_i}\big) \mathbb{1}_{\{\tau^X \geq a_n\}} \\
&\leq \lfloor n^{1/6} \rfloor \max_{0 \leq i \leq m} |\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{t_i}| \mathbb{1}_{\mathcal{G}_i} + b_n \sum_{i=0}^{m} \mathbb{1}_{\mathcal{B}_i} + b_n \sum_{i=0}^{m} \mathbb{1}_{\mathcal{T}_i} \mathbb{1}_{\{\tau^X \geq a_n\}}, \quad (4.10)
\end{aligned}$$

where for the second and third term we used the trivial estimate that $|\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{t_i}| \leq t_{i+1} - t_i = b_n$.

We will now estimate each of the three terms in (4.10) separately. Let us start with the first one. We will estimate $|\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{t_i}|$ on the event $\mathcal{G}_i \cap \mathcal{A}^{\mathrm{his}}_{a_n}$. Since $\mathcal{L}^X$ increases at rate at most 1, we have the estimate

$$|\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{t_i}| \leq \sigma^i_{k_0} - t_i + |\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{\sigma^i_{k_0}}| \mathbb{1}_{\{\sigma^i_{k_0} < t_{i+1}\}}. \quad (4.11)$$

By (4.7), on $\mathcal{A}^{\mathrm{his}}_{a_n}$ the first term on the right hand side in (4.11) is at most $\log^2 n$ with probability at least $1 - C_2 e^{-c_2 \log^2 n}$. In the second term we may, on $\mathcal{G}_i$, replace $\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{\sigma^i_{k_0}}$ by $\mathcal{L}^Y_{t_{i+1} - \sigma^i_{k_0}}$, where $Y = Y^i_{k_0}$ is the simple exploration started at time $\sigma^i_{k_0}$. Now we apply Proposition 4.3 with $s = \frac{1}{2} \log n$ and $T = b_n$. As $t_{i+1} - \sigma^i_{k_0} \leq t_{i+1} - t_i = b_n = n^{1/3}/\lfloor n^{1/6} \rfloor$, we obtain in particular

$$\mathbb{P}\big(\mathcal{G}_i \cap \big\{|\mathcal{L}^Y_{t_{i+1} - \sigma^i_{k_0}}| \geq \tfrac{1}{2} b_n^{1/2} \log n\big\} \mid \mathcal{F}_{\sigma_{k_0}}\big) \leq C e^{-c \log^2 n}.$$

Thus

$$\mathbb{P}\big(\mathcal{A}^{\mathrm{his}}_{a_n} \cap \big\{|\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{t_i}| \mathbb{1}_{\mathcal{G}_i} \geq b_n^{1/2} \log n\big\}\big) \leq C_3 e^{-c_3 \log^2 n},$$

for some $C_3, c_3 > 0$. Furthermore, applying a union bound we obtain that (recalling $m = \lfloor n^{1/6} \rfloor - 1$)

$$\mathbb{P}\Big(\mathcal{A}^{\mathrm{his}}_{a_n} \cap \Big\{ \max_{0 \leq i \leq m} |\mathcal{L}^X_{t_{i+1}} - \mathcal{L}^X_{t_i}| \mathbb{1}_{\mathcal{G}_i} \geq b_n^{1/2} \log n \Big\}\Big) \leq \lfloor n^{1/6} \rfloor C_3 e^{-c_3 \log^2 n}.$$

We now move to the second term of (4.10). We need to estimate $\mathbb{P}(\mathcal{B}_i | \mathcal{F}_{t_i})$. To this end notice that by Lemma 3.3 the probability for any given coupling to fail is bounded above by

$$4\beta b_n (\mathcal{J}^X_{a_n} + \beta b_n)/n \leq 4\beta a_n (\mathcal{J}^X_{a_n} + \beta a_n)/n.$$

Defining $\mathcal{D}_0 := \big\{\mathcal{J}^X_{a_n} \leq 4\beta a_n\big\}$, we have, for any $k > 0$, that

$$\begin{aligned}
\mathbb{P}(\mathcal{B}_i \mid \mathcal{F}_{t_i}) &\leq \mathbb{P}(k_0 \geq k \mid \mathcal{F}_{t_i}) + \mathbb{P}(\{k_0 < k\} \cap \mathcal{B}_i \cap \mathcal{D}_0 \mid \mathcal{F}_{t_i}) + \mathbb{P}(\mathcal{D}_0^c \mid \mathcal{F}_{t_i}) \\
&\leq (1 - z)^k + k \frac{20 \beta^2 a_n^2}{n} + \mathbb{P}(\mathcal{D}_0^c \mid \mathcal{F}_{t_i}).
\end{aligned}$$

Recalling that $a_n = n^{1/3}$ and choosing $k = \lfloor n^{1/12} \rfloor$ it follows that, for some $C > 0$,

$$\mathbb{P}(\mathcal{B}_i \mid \mathcal{F}_{t_i}) \leq p_{n,i} := C n^{-1/4} + \mathbb{P}(\mathcal{D}_0^c \mid \mathcal{F}_{t_i}). \quad (4.12)$$

We claim that, on the event $\mathcal{D}_1 := \{\mathcal{J}^X_{a_n} \le 2\beta a_n\}$, we have for large enough n that $p_{n,i} \le 2Cn^{-1/4}$ with C from (4.12). Indeed,

$$\begin{aligned}\mathbb{P}(\mathcal{D}_0^c \mid \mathcal{F}_{t_i}) &\le \mathbb{P}(\mathcal{J}^X_{t_i} > 2\beta a_n \mid \mathcal{F}_{t_i}) + \mathbb{P}(\mathcal{J}^X_{a_n} - \mathcal{J}^X_{t_i} > 2\beta a_n \mid \mathcal{F}_{t_i}) \\ &= \mathbb{1}\{\mathcal{J}^X_{t_i} > 2\beta a_n\} + \mathbb{P}(\mathcal{J}^X_{a_n} - \mathcal{J}^X_{t_i} > 2\beta a_n \mid \mathcal{F}_{t_i}).\end{aligned}$$

On the right-hand-side, the indicator vanishes on $\mathcal{D}_1$, and the probability is at most $C_1 e^{-c_1 a_n}$ for some $C_1, c_1 > 0$, since $\mathcal{J}^X_t$ is a counting process with intensity bounded above by β, see Lemma 3.5 and the argument for (3.7). The claim follows.

Thus, employing (4.12), we obtain for large enough n that

$$\begin{aligned}\mathbb{P}\left(\textstyle\sum_{i=0}^m \mathbb{1}_{\mathcal{B}_i} \ge n^{1/12}\log n\right) &\le \mathbb{P}\left(\left\{\textstyle\sum_{i=0}^m \mathbb{1}_{\mathcal{B}_i} \ge n^{1/12}\log n\right\} \cap \mathcal{D}_1\right) + \mathbb{P}(\mathcal{D}_1^c) \\ &\le \mathbb{P}\left(\left\{\textstyle\sum_{i=0}^m (\mathbb{1}_{\mathcal{B}_i} - \mathbb{P}(\mathcal{B}_i|\mathcal{F}_{t_i})) \ge n^{1/12}\log n - \sum_{i=0}^m p_{n,i}\right\} \cap \mathcal{D}_1\right) + \mathbb{P}(\mathcal{D}_1^c) \\ &\le \mathbb{P}\left(\textstyle\sum_{i=0}^m (\mathbb{1}_{\mathcal{B}_i} - \mathbb{P}(\mathcal{B}_i|\mathcal{F}_{t_i})) \ge \frac{1}{2} n^{1/12}\log n\right) + \mathbb{P}(\mathcal{D}_1^c).\end{aligned}$$

The sum inside the first probability is a martingale, with increments bounded by 1. Thus by the Azuma inequality (see e.g. [16, Theorem A.10]) we get

$$\mathbb{P}\left(\textstyle\sum_{i=0}^m (\mathbb{1}_{\mathcal{B}_i} - \mathbb{P}(\mathcal{B}_i|\mathcal{F}_{t_i})) \ge \frac{1}{2} n^{1/12}\log n\right) \le 2\exp\left(-\frac{n^{1/6}\log^2 n}{8\lfloor n^{1/6}\rfloor}\right). \tag{4.13}$$

As before we have that $\mathbb{P}(\mathcal{D}_1^c) \le e^{-cn}$ for some $c > 0$. Taken together, these facts give

$$\mathbb{P}\left(\textstyle\sum_{i=0}^m \mathbb{1}_{\mathcal{B}_i} \ge n^{1/12}\log n\right) \le C_4 e^{-c_4 \log^2 n}, \tag{4.14}$$

for some $C_4, c_4 > 0$.

Finally, let us now consider $\sum_{i=0}^m \mathbb{1}_{\mathcal{T}_i}\mathbb{1}_{\{\tau^X \ge a_n\}}$. Observe that for $i \le m-1$ the event $\mathcal{T}_i \cap \{\tau^X \ge a_n\}$ requires that the exploration neither jumps to an unvisited vertex nor closes the loop for a time period of at least $n^{1/6}$. By a similar application of Corollary 3.7 as for (4.9) the latter event has probability smaller than $C_4 e^{-c_4 n^{1/6}}$, for some $C_4, c_4 > 0$. We thus have

$$\begin{aligned}\mathbb{P}\left(\textstyle\sum_{i=0}^m \mathbb{1}_{\mathcal{T}_i}\mathbb{1}_{\{\tau^X \ge a_n\}} \ge 2\right) &\le \mathbb{P}\left(\textstyle\sum_{i=0}^{m-1} \mathbb{1}_{\mathcal{T}_i}\mathbb{1}_{\{\tau^X \ge a_n\}} \ge 1\right) \\ &\le \sum_{i=0}^{m-1} \mathbb{P}\big(\mathcal{T}_i \cap \{\tau^X \ge a_n\}\big) \le C_5 e^{-c_5 n^{1/6}},\end{aligned}$$

for some $C_5, c_5 > 0$.

From (4.10), (4.1), (4.14) and (4.1) we conclude that for some $C_6, c_6 > 0$

$$\mathbb{P}\big(\mathcal{A}^{\mathrm{his}}_{a_n} \cap \{|\mathcal{L}^X_{a_n}|\mathbb{1}_{\{\tau^X \ge a_n\}} \ge \lfloor n^{1/6}\rfloor b_n^{1/2}\log n + b_n n^{1/12}\log n + b_n\}\big) \le C_6 e^{-c_6 \log^2 n}.$$

Since $\lfloor n^{1/6}\rfloor b_n^{1/2}\log n + b_n n^{1/12}\log n + b_n \le 3n^{1/4}\log n$, this concludes the proof of (4.5) for a fixed $\beta \in [\beta_0, \beta_1]$. The uniformness over such β follows since the upper bound can be chosen as a continuous function of β.

Now we turn to (4.5). We aim to do a similar decomposition as above, and as before it suffices to work on the event $\mathcal{A}^{\mathrm{his}}_{a_n}$. For $i \in \{0, 1, \dots, m\}$, let Y_i be the coupled simple exploration started at time $\sigma^i_{k_0}$, and let $\mathcal{S}_i$ be the event that Y_i survives. On the event $\mathcal{G}_i$ we have in particular that $\mathcal{S}_i$ occurs, and we can use $\mathcal{K}^X_{t_{i+1}} - \mathcal{K}^X_{t_i} \ge \mathcal{K}^{Y_i}_{t_{i+1}-\sigma^i_{k_0}}$. On $\mathcal{G}_i^c$ we simply use $\mathcal{K}^X_{t_{i+1}} - \mathcal{K}^X_{t_i} \ge 0$. Therefore

$$\mathcal{K}^X_{a_n} = \sum_{i=0}^m (\mathcal{K}^X_{t_{i+1}} - \mathcal{K}^X_{t_i}) \ge \sum_{i=0}^m (\mathcal{K}^X_{t_{i+1}} - \mathcal{K}^X_{t_i})\mathbb{1}_{\mathcal{G}_i} \ge \sum_{i=0}^m \mathcal{K}^{Y_i}_{t_{i+1}-\sigma^i_{k_0}}\mathbb{1}_{\mathcal{G}_i}.$$

Hence

$$\begin{aligned}
&\mathbb{P}(\mathcal{A}^{\mathrm{his}}_{a_n} \cap \{\mathcal{K}^X_{a_n} < c_2 a_n\} \cap \{\tau^X \geq a_n\}) \\
&\quad \leq \mathbb{P}\Big(\mathcal{A}^{\mathrm{his}}_{a_n} \cap \Big\{ \sum_{i=0}^{m} \mathcal{K}^{Y_i}_{t_{i+1}-\sigma^i_{k_0}} \mathbb{1}_{\mathcal{G}_i} < c_2 a_n \Big\} \cap \{\tau^X \geq a_n\}\Big) \\
&\quad \leq \mathbb{P}\Big(\sum_{i=0}^{\lfloor n^{1/6} \rfloor - 1} \mathcal{K}^{Y_i}_{b_n - \log^2 n} \mathbb{1}_{\mathcal{G}_i} < c_2 a_n \Big) \\
&\qquad + \mathbb{P}(\{\exists i \leq m : \sigma^i_{k_0} - t_i > \log^2 n\} \cap \mathcal{A}^{\mathrm{his}}_{a_n} \cap \{\tau^X \geq a_n\}).
\end{aligned}$$

Using (4.7) we get for some $C_8, c_8 > 0$

$$\mathbb{P}(\{\exists i \leq m : \sigma^i_{k_0} - t_i > \log^2 n\} \cap \mathcal{A}^{\mathrm{his}}_{a_n} \cap \{\tau^X \geq a_n\}) \leq C_8 e^{-c_8 \log^2 n}.$$

To bound the first probability, we use that

$$\sum_{i=0}^{m} \mathcal{K}^{Y_i}_{b_n - \log^2 n} \mathbb{1}_{\mathcal{G}_i} \geq \sum_{i=0}^{m} \mathcal{K}^{Y_i}_{b_n - \log^2 n} \mathbb{1}_{\mathcal{S}_i} - \sum_{i=0}^{m} \mathcal{K}^{Y_i}_{b_n - \log^2 n} (\mathbb{1}_{\mathcal{B}_i} + \mathbb{1}_{\mathcal{T}_i})$$

The processes Y_i are i.i.d. and by (4.3) from Proposition 4.4 we get

$$\mathbb{E}[\mathcal{K}^{Y_i}_{b_n - \log^2 n} \mathbb{1}_{\mathcal{S}_i}] \geq cz(b_n - \log^2 n).$$

Hence by standard large deviations estimates we get for some $C_7, c_7 > 0$

$$\mathbb{P}\Big(\sum_{i=0}^{m} \mathcal{K}^{Y_i}_{b_n - \log^2 n} \mathbb{1}_{\mathcal{S}_i} < c_2 a_n \Big) \leq C_7 e^{-c_7 \log^2 n}$$

provided we pick c_2 small enough.

It remains to bound the contributions involving $\mathcal{B}_i$ and $\mathcal{T}_i$. Recall from Proposition 4.4 that $\mathcal{K}^{Y_i}_t \leq t + 1 + \mathcal{I}^{Y_i}_t \leq t + 1 + 2\mathcal{J}^{Y_i}_t$, and from the observations preceding (3.7) that the number of 'jumps' $\mathcal{J}^{Y_i}_t$ is dominated by a Poisson process with rate $\beta \frac{n}{n-1}$. It follows that, with probability at least $1 - C_8 e^{-c_8 \log^2 n}$, we have that $\mathcal{K}^{Y_i}_{b_n - \log^2 n} \leq 5\beta b_n$ for all $i \leq m$. Consequently, using also (4.14)

$$\begin{aligned}
\mathbb{P}\Big(\sum_{i=0}^{m} \mathcal{K}^{Y_i}_{b_n - \log^2 n} \mathbb{1}_{\mathcal{B}_i} \geq 5\beta n^{1/4} \log n \Big) &\leq \mathbb{P}\Big(\sum_{i=0}^{m} \mathbb{1}_{\mathcal{B}_i} \geq n^{1/12} \log n \Big) + C_8 e^{-c_8 \log^2 n} \\
&\leq C_9 e^{-c_9 \log^2 n} + C_8 e^{-c_8 \log^2 n}.
\end{aligned}$$

Finally, for the terms involving $\mathcal{T}_i$ we again use that $\mathcal{K}^{Y_i}_{b_n - \log^2 n} \leq 5\beta b_n$ for all $i \leq m$, with probability at least $1 - C_8 e^{-c_8 \log^2 n}$, combined with (4.1) to get

$$\begin{aligned}
\mathbb{P}\Big(\sum_{i=0}^{m} \mathcal{K}^{Y_i}_{b_n - \log^2 n} \mathbb{1}_{\mathcal{T}_i} \mathbb{1}_{\{\tau^X \geq a_n\}} \geq 10\beta n^{1/6} \Big) &\leq \mathbb{P}\Big(\sum_{i=0}^{m} \mathbb{1}_{\mathcal{T}_i} \mathbb{1}_{\{\tau^X \geq a_n\}} \geq 2 \Big) + C_8 e^{-c_8 \log^2 n} \\
&\leq C_{10} e^{-c_{10} \log^2 n} + C_8 e^{-c_8 \log^2 n}.
\end{aligned}$$

This establishes (4.5). □

4.2 Proofs of Propositions 4.1 and 4.2

Now we turn to the proofs of the main results of Section 4, namely Propositions 4.1 and 4.2, which concern the balance $\mathfrak{B}(v, \lfloor n^{1/2} \rfloor)$ in cycles of length at least $\lfloor n^{1/2} \rfloor$. We start with the following corollary of Proposition 4.5, which states that the bounds in that proposition hold uniformly over all possible starting points $(v, \varphi) \in V_n \times S^1$ for the exploration process X. We use the notation $\mathcal{L}^X_t(v, \varphi)$, $\mathcal{K}^X_t(v, \varphi)$ and $\tau^X(v, \varphi)$ when X starts at (v, φ).

Corollary 4.6. *Let $\beta_1 > \beta_0 > 1$. There exist $C, c > 0$ such that for all $\beta \in [\beta_0, \beta_1]$ we have*

$$\mathbb{P}_\beta \left(\exists_{(v,\varphi)\in V_n \times S^1} : |\mathcal{L}^X_{n^{1/3}}(v,\varphi)| \mathbb{1}_{\{\tau^X(v,\varphi) \geq n^{1/3}\}} \geq 3n^{1/4} \log^2 n \right) \leq Ce^{-c \log^2 n}, \tag{4.15}$$

and, for some $c_2 > 0$,

$$\mathbb{P}_\beta \left(\exists_{(v,\varphi)\in V_n \times S^1} : \left\{ \mathcal{K}^X_{n^{1/3}}(v,\varphi) < c_2 n^{1/3} \right\} \cap \left\{ \tau^X(v,\varphi) \geq n^{1/3} \right\} \right) \leq Ce^{-c \log^2 n}. \tag{4.16}$$

In the proof we will use the following notation. For a measurable subset $A \subseteq E \times S^1$ and $\omega \in \Omega$ we denote the restriction

$$\omega_A := \{(e,\varphi,m) \in \omega : (e,\varphi) \in A\}.$$

Also recall that we will often identify S^1 with the interval $[0,1)$.

Proof. We give details for (4.15), the argument for (4.16) is very similar. Write

$$\mathcal{B}(v,\varphi) = \left\{ |\mathcal{L}^X_{n^{1/3}}(v,\varphi)| \mathbb{1}_{\{\tau^X(v,\varphi) \geq n^{1/3}\}} \geq 3n^{1/4} \log^2 n \right\}.$$

We fix $\varepsilon > 0$ to be a small enough positive constant (to be specific, ε needs to be smaller than the constant c_1 in the exponent on the right-hand-side of (4.5)). Let $m = \lfloor e^{\varepsilon \log^2 n} \rfloor$ and define the growing sequence of sets $A_i := E \times [0, \beta_0/\beta_1 + i\delta]$, where $\delta = \delta_n := \frac{1}{m}\left(1 - \frac{\beta_0}{\beta_1}\right)$ and $i \in \{0, 1, \ldots, m\}$. We will consider the sequence ω_{A_i} which we think of as revealing the configuration ω in increments of size δ. Consider the event

$$\mathcal{D} := \left\{ \omega : |\omega_{A_i \setminus A_{i-1}}| \leq 1 \text{ for each } i \in \{1, \ldots, m\} \right\}$$

that each step in the sequence reveals at most one more link. Since $|\omega_{A_i \setminus A_{i-1}}|$ is Poisson distributed with mean $\binom{n}{2}\frac{\beta}{n-1}\delta_n$ we have for some $\tilde{C}, C_2, c_2 > 0$ that

$$\mathbb{P}(\mathcal{D}^c) \leq \sum_{i=1}^{m} \mathbb{P}\left(|\omega_{A_i \setminus A_{i-1}}| \geq 2 \right) \leq C_2 e^{-c_2 \log^2 n}. \tag{4.17}$$

Thus it suffices to show that $\mathbb{P}(\cup_{(v,\varphi)} \mathcal{B}(v,\varphi) \cap \mathcal{D})$ satisfies the bound (4.15).

Now on $\mathcal{D}$, to determine if there is some $(v,\varphi) \in V_n \times S^1$ for which $\mathcal{B}(v,\varphi)$ holds it suffices to consider φ of the form $\varphi_i = \beta_0/\beta_1 + i\delta$ for $0 \leq i \leq m$. Indeed, if φ is arbitrary, let i be such that $\varphi_{i-1} \leq \varphi \leq \varphi_i$. Then (on $\mathcal{D}$) the exploration started at (v,φ) agrees either with that started at (v,φ_{i-1}) or that started at (v,φ_i) (up to a small time-shift of size at most δ which we will ignore). Hence, using Proposition 4.5,

$$\mathbb{P}(\cup_{(v,\varphi)} \mathcal{B}(v,\varphi) \cap \mathcal{D}) \leq \sum_{v \in V_n} \sum_{i=0}^{m} \mathbb{P}(\mathcal{B}(v,\varphi_i)) \leq n e^{\varepsilon \log^2 n} C_1 e^{-c_1 \log^2 n}.$$

For $\varepsilon > 0$ small enough, this satisfies the claimed bound. □

To proceed we will need some notations and observations which will allow us to relate the winding process, $\mathcal{L}$, to the balance of cycles, $\mathfrak{B}$. Let $X_s = X_s(v_0, \varphi_0, d_0)$ denote the exploration started at (v_0, φ_0) in the direction $d_0 \in \{-1, +1\}$, viewed at time s. Let us write $X_s = (v_s, \varphi_s, d_s)$ and define

$$\mathfrak{B}^X_t(v_0, \varphi_0, d_0) = \sum_{0 \leq s \leq t} d_s \mathbb{1}_{\{\varphi_s = 0\}}.$$

(Although formally the summation is over an uncountable set, almost surely there is only a finite number of nonzero terms.) In words, $\mathfrak{B}^X_s$ totals the number of visits of X to level

$\varphi = 0$, counted with the sign given by the direction of travel. Note that our previously defined balance-quantity may be written as

$$\mathfrak{B}(v,k) = \mathfrak{B}^X_{\tau_k}(v,0,+1)$$

where τ_k is the first time at which X has made k visits to level $\varphi = 0$.

It is easy to see the following: for any starting point (v_0, φ_0, d_0) and any $t \geq 0$ we have that

$$\big||\mathfrak{B}^X_t(v_0,\varphi_0,d_0)| - |\mathcal{L}^X_t(v_0,\varphi_0,d_0)|\big| \leq 3. \tag{4.18}$$

Indeed, for $t = 0$ the two terms are either 1 and 0 (if $\varphi_0 = 0$) or 0 and 0 (if $\varphi_0 \neq 0$). As t increases, $|\mathfrak{B}^X_t|$ stays constant until X passes level $\varphi = 0$, at which time it changes by 1. Until this time $|\mathcal{L}^X_t|$ can change by at most 1, since if it changes more then this necessarily means that X passes level $\varphi = 0$; hence the difference in (4.18) is certainly bounded by 2. Between successive visits to $\varphi = 0$ it remains bounded by 2 for the same reason. Finally, after the last visit to $\varphi = 0$ we may have that $\mathcal{L}^X_t$ changes by up to 1 while $|\mathfrak{B}^X_t|$ remains constant. Thus the difference is at most 3.

Proof of Proposition 4.1. Let

$$\mathcal{A}_n := \bigcap_{(v,\varphi)\in V_n\times S^1} \Big\{ |\mathcal{L}^X_{n^{1/3}}(v,\varphi)| \mathbb{1}_{\{\tau^X(v,\varphi)\geq n^{1/3}\}} \leq 3n^{1/4}\log^2 n \Big\}$$

and

$$\mathcal{B}_n := \bigcap_{(v,\varphi)\in V_n\times S^1} \Big(\Big\{ \mathcal{K}^X_{n^{1/3}}(v,\varphi) \geq c_2 n^{1/3} \Big\} \cup \Big\{ \tau^X(v,\varphi) \leq n^{1/3} \Big\} \Big),$$

where c_2 is as in (4.5). (To see that $\mathcal{A}_n$ and $\mathcal{B}_n$ are measurable, note that one gets the same events if φ is restricted to rationals.) By Corollary 4.6 we have $\mathbb{P}(\mathcal{A}_n \cap \mathcal{B}_n) \geq 1 - Ce^{-c\log^2 n}$ for some $C, c > 0$, so it suffices to consider $\omega \in \mathcal{A}_n \cap \mathcal{B}_n$.

Suppose v is such that $|\mathcal{C}_\omega(v)| \geq \lfloor n^{1/2} \rfloor$ (otherwise there is nothing to prove). For $i \geq 1$, let $t_i = in^{1/3}$ and let $i_0 := \min\{i \geq 1 : \mathcal{K}^X_{t_i}(v,0) \geq n^{1/2}\}$. Thus by time t_{i_0} the exploration (started at $(v,0)$) has visited the first $\lfloor n^{1/2} \rfloor$ vertices in $\mathcal{C}_\omega(v)$ following v. Since $\omega \in \mathcal{B}_n$, the contributions to $\mathcal{K}^X_t$ between successive t_i are all at least $c_2 n^{1/3}$; using the additivity of $\mathcal{K}^X_t$ we conclude that $i_0 \leq c_2^{-1} n^{1/6}$.

Let us write $X_{t_i} = (v_i, \varphi_i, d_i)$. Note that (using (4.18))

$$\begin{aligned} |\mathfrak{B}(v, \lfloor n^{1/2} \rfloor)| &\leq \sum_{i=1}^{i_0-1} |\mathfrak{B}^X_{t_i}(v_{i-1},\varphi_{i-1},d_{i-1})| + t_{i_0} - t_{i_0-1} \\ &\leq \sum_{i=1}^{i_0-1} \big(|\mathcal{L}^X_{t_i}(v_{i-1},\varphi_{i-1},d_{i-1})| + 3 \big) + n^{1/3}. \end{aligned}$$

As $\omega \in \mathcal{A}_n$ we get

$$|\mathfrak{B}(v, \lfloor n^{1/2} \rfloor)| \leq (i_0 - 1)(3n^{1/4}\log^2 n + 3) + n^{1/3} \leq n^{5/12}\log^3 n,$$

for n large enough, as required. □

Proof of Proposition 4.2. This will follow from Proposition 4.1 using a similar argument as for Corollary 4.6. As in that argument, we fix some small enough $\varepsilon > 0$ and we use the same notation m, δ, A_i and $\mathcal{D}$.

Note that, on $\mathcal{D}$, for each $s \le |\omega|$ there is some (random) i such that $\vec{\omega}_s = \vec{\omega}_{A_i}$. Hence the probability in (4.2) is at most

$$\sum_{v\in V_n}\sum_{i=0}^{m} \mathbb{P}_\beta\left(\left\{\mathfrak{B}_{\vec{\omega}_{A_i}}(v, \lfloor n^{1/2}\rfloor) \ge n^{5/12}\log^3 n\right\} \cap \left\{|\mathcal{C}_{\vec{\omega}_{A_i}}(v)| \ge \lfloor n^{1/2}\rfloor\right\}\right) + \mathbb{P}(\mathcal{D}^c \cup \{|\omega| < n^\rho\}).$$

We observe that under $\mathbb{P}_\beta$ the distribution of $\vec{\omega}_{A_i}$ is the same as the distribution of $\vec{\omega}$ under $\mathbb{P}_{\bar{\beta}_i}$, with $\bar{\beta}_i = \beta\,(\beta_0/\beta_1 + i\delta) \in [\beta_0, \beta_1]$. Using Proposition 4.1 and a straightforward bound on $\mathbb{P}(\mathcal{D}^c \cup \{|\omega| < n^\rho\})$ we deduce that the probability in (4.2) is at most

$$ne^{\varepsilon \log^2 n} Ce^{-c\log^2 n} + C_2 e^{-c_2 \log^2 n}.$$

Choosing $\varepsilon > 0$ small enough concludes the proof. □

5 Poisson–Dirichlet coupling

In this section we prove our main result, Theorem 1.1. From the previous sections, Lemma 2.4 tells us that there are cycles of size of the order n and Propositions 4.1 and 4.2 tell us that these cycles are 'balanced'. The former lemma is stated in terms of a sequentially constructed $\vec{\omega}$ with a fixed number of links, whereas the latter are formulated in terms of ω sampled from the Poisson law $\mathbb{P}_\beta$, so one of our tasks is to combine these two descriptions. Another task is to convert the balance-property of Propositions 4.1 and 4.2 into a quantitative result about the probability of splitting cycles when a uniformly placed link is added, see Lemma 5.1. Following that, the main task will be to provide a coupling of a PD($\frac{1}{2}$) sample with the rescaled cycle sizes. We begin by introducing some relevant notation and facts, as well as an outline of the proof. Throughout this section, $\beta > 1$ and $\nu \in [0, 1)$ are fixed.

5.1 Preparation and outline

The coupling with PD($\frac{1}{2}$) will involve sequentially appending a small number of uniformly, independently placed links to a random configuration ω. We will do this as follows. First, let ω have distribution $\mathbb{P}_\beta$. Next, let $q \ge 0$ be an integer-valued random variable which is independent of ω and bounded (as $n \to \infty$). Recall that $\vec{\omega}$ denotes the ordered sequence of links in ω and that $\vec{\omega}_s$ denotes the first s links of $\vec{\omega}$.

To start the coupling, we will consider $\vec{\omega}_s$ for $s = |\omega| - q$. We construct a sequence $\vec{\omega}'_t$ for $t \in \{s, s+1, \dots, s+q\}$, where $\vec{\omega}'_s = \vec{\omega}_s$ and the following $\vec{\omega}'_t$ are obtained by sequentially and independently appending in total q uniformly placed links one at a time. Obviously the final configuration $\vec{\omega}'_{s+q}$ then has $|\omega| = s + q$ links, and it agrees in distribution with $\vec{\omega}$. Letting $\mathfrak{C}_t = \mathfrak{C}_{\vec{\omega}'_t}$ denote the cycle structure of $\vec{\omega}'_t$, it thus suffices to prove that Theorem 1.1 holds for $\mathfrak{C}_q$.

Before proceeding, let us recall the key features of $\vec{\omega}'_s = \vec{\omega}_s$ which follow from our work in the previous sections. Precise statements are deferred to Section 5.3. First, it is clear that (since $\beta > 1$) we can find a constant $c > \frac{1}{2}$ such that the number of links s satisifes $s \ge cn$ with high probability (converging to 1 as $n \to \infty$). On this event Lemma 2.4 applies to $\vec{\omega}'_s$, meaning (roughly speaking) that there are cycles of size of the order n which together occupy a fraction $\approx zn$ of all vertices. (Here z is the same as in Proposition 3.4.) Second, since q is bounded, Proposition 4.2 certainly applies to $\vec{\omega}'_s$. Thus (with high probability), in any of the large cycles of $\vec{\omega}'_s$, any segment of $\lfloor n^{1/2}\rfloor$ consecutive vertices in that cycle has balance $|\mathfrak{B}| < n^{5/12}\log^3 n$.

Next, let us describe the evolution of $\mathfrak{C}_t$, $0 \le t \le q$, in a way which is suitable for the coupling with PD($\frac{1}{2}$). Since PD($\frac{1}{2}$) is a probability distribution on 'continuous' partitions

of the interval $[0,1)$ it is convenient to represent $\mathfrak{C}_t$ also as a (labelled) partition of $[0,1)$ (in the actual proof we will use a different interval but the idea is the same). The mapping is fairly intuitive so we do not give a completely detailed description. Each vertex $v \in V_n$ is represented as a subinterval $I(v)$ of the form $[\frac{i}{n}, \frac{i+1}{n})$ for $0 \le i = i(v) \le n-1$, and this mapping is chosen so that the cycles of $\mathfrak{C}_t$ become disjoint intervals of the form $[\frac{i}{n}, \frac{j}{n})$ for $0 \le i < j \le n-1$, where if u, v are consecutive in a cycle then $I(u)$ and $I(v)$ are consecutive subintervals of $[\frac{i}{n}, \frac{j}{n})$ (interpreted cyclically). The subintervals $I(v)$ are labelled using the labels $\uparrow, \downarrow$ consistently with the orientations of the vertices within the cycles. See Figure 10.

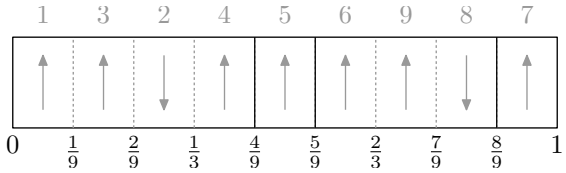

Figure 10: Representation of cycles as subintervals of $[0,1)$. Solid vetical lines delimit the cycles and dashed lines the vertices. The cycles here are the same as in Figure 2, i.e. $(1^{\uparrow}, 3^{\uparrow}, 2^{\downarrow}, 4^{\uparrow})$, $(5^{\uparrow})$, $(6^{\uparrow}, 9^{\uparrow}, 8^{\downarrow})$ and $(7^{\uparrow})$.

Naturally, this mapping is defined up to (i) cyclic rotations within each cycle, (ii) overall reversal of all the labels (arrows) in cycles, and (iii) the relative placement of the intervals $[\frac{i}{n}, \frac{j}{n})$ representing the cycles within $[0,1)$. Regarding the last item, the canonical way to order the intervals would be by decreasing length, but we wish to keep the flexibility of reordering them for the time being.

In this setting the dynamics of uniformly placing links may be constructed using two independent uniform random variables U, U' in $[0,1)$:

- We first sample the mark $m \in \{\rightleftharpoons, \rightleftharpoons\}$ of the link with probability ν for $\rightleftharpoons$.
- We then sample U and set the first endpoint of the link to be u if U falls in the interval $I(u)$.
- Before selecting the other endpoint we (i) move the (interval $[\frac{i}{n}, \frac{j}{n})$ representing the) cycle containing u to the front of $[0,1)$, then (ii) cyclically reorder this cycle so that $I(u) = [0, \frac{1}{n})$.
- Now we select the second endpoint by setting it to be v if $U' \in I(v)$. It may happen that $I(v) = I(u)$; since this has probability $\frac{1}{n}$ we will in practice be able to disregard this possibility, but to be definite let us say that nothing happens to the cycles in this case.

Having selected the endpoints of the link as well as its mark, we apply the rules given in Section 2 for splitting, merging or twisting cycles.

Using this construction, the sequence $\mathfrak{C}_1, \dots, \mathfrak{C}_q$ may be obtained starting with $\mathfrak{C}_0$ and using a sequence $\{(U_t, U'_t, m_t))\}_{t=1}^{q}$ of independent random variables with the above distributions.

We now turn to the task of showing that the probability of splitting a large cycle is close to $\frac{1}{2}$, in a sense which we will make precise. Let us assume that $\vec{\omega}'_s$ belongs to the event

$$\bigcap_{v \in V_n} \left(\{ |\mathfrak{B}(v, \lfloor n^{1/2} \rfloor)| < n^{5/12} \log^3 n \} \cup \{ |\mathcal{C}(v)| < \lfloor n^{1/2} \rfloor \} \right) \tag{5.1}$$

that any cycle of size at least $\lfloor n^{1/2} \rfloor$ is 'balanced'. This event holds with high probability due to Proposition 4.2. Recalling that the cycles $\mathfrak{C}_0$ form a partition of the vertex set V_n,

we define a refinement $\mathfrak{S}$ of this partition into 'segments' as follows. For each cycle $\mathcal{C} \in \mathfrak{C}_0$ satisfying $|\mathcal{C}| \geq \lfloor n^{1/2} \rfloor$ we fix a division of $\mathcal{C}$ into non-intersecting sets of consecutive vertices, each of size between $\lfloor n^{1/2} \rfloor$ and $2\lfloor n^{1/2} \rfloor$. If $|\mathcal{C}| < \lfloor n^{1/2} \rfloor$ then we declare $\mathcal{C}$ to be a segment on its own. On the event in (5.1) we see using the triangle-inequality that each segment $S \in \mathfrak{S}$ satisfying $|S| \geq \lfloor n^{1/2} \rfloor$ has balance $|\mathfrak{B}(S)| < 2n^{5/12} \log^3 n$, where $\mathfrak{B}(S)$ is the difference between the number of $\uparrow$ and number of $\downarrow$ in S.

As we proceed by adding links, and thereby modify the cycle structure, we keep the partition $\mathfrak{S}$ into segments fixed. That is, at all later steps we will 'remember' for each vertex $v \in V_n$ which segment S it belonged to at $t = 0$. After some steps a segment S need no longer be a consecutive set of vertices within a cycle, for example if a cycle is split in the middle of S. We say that a segment S is *untouched* at step $t \in \{1, \dots, q\}$ if none of the links placed in steps $1, 2, \dots, t-1$ had an endpoint in S, otherwise the segment is *touched*. If S is untouched then it is also 'intact' in the sense that it is still consists of consecutive vertices in some cycle, and $|\mathfrak{B}(S)|$ is unchanged from $t = 0$.

In the representation of $\mathfrak{C}_0$ as a collection of marked subintervals of $[0, 1)$, the segments S become subintervals (of length $\leq 2n^{-1/2}$) of the intervals representing the cycles (possibly we may have to interpret these subintervals cyclically). Recall that we used a uniform random variable $U' \in [0, 1)$ to select the second endpoint of a uniformly placed link. We now modify this construction slightly, and will instead use *two* uniform independent $U', U'' \in [0, 1)$. We begin by sampling U', and we note which segment S it falls in (more precisely, which subinterval representing such a segment). If this segment S is *touched* then we let v be the vertex selected by U' as before and we do not use U''. However, if S is untouched then we do not record the precise location of U' within S; instead we use U'' to independently select a uniform location within S and we select the second endpoint of the link to be v if $U'' \in I(v)$.

The following result is now straightforward. Intuitively, it tells us that the probability of splitting a long cycle is very close to $\frac{1}{2}$, moreover the choice of whether or not to split is almost independent of the location where we propose to split.

Lemma 5.1. *Assume that the event in* (5.1) *holds at $t = 0$. At step $t \geq 1$ (i.e. in the transition $\mathfrak{C}_{t-1} \to \mathfrak{C}_t$), let u be the vertex selected by U_t and let $\mathcal{C}(u) \in \mathfrak{C}_{t-1}$ be the cycle containing u. Fix the orientation of $\mathcal{C}(u)$ so that u has label $\uparrow$. Suppose that U'_t selects a segment S which: (i) is untouched, (ii) is in the cycle $\mathcal{C}(u)$, i.e. $S \subseteq \mathcal{C}(u)$, and (iii) has size $|S| \geq \lfloor n^{1/2} \rfloor$. Let $\alpha \in \{\uparrow, \downarrow\}$ be the label of the vertex v selected by U''_t. Then (on the event described) the conditional probability $p_t = \mathbb{P}(\alpha = \uparrow \mid U'_t)$ that v has the same orientation as u satisfies*

$$|p_t - \tfrac{1}{2}| \leq n^{-1/12} \log^3 n.$$

Proof. We have that $p_t = \#(\uparrow \text{ in } S)/|S|$ so

$$|p_t - \tfrac{1}{2}| = \tfrac{1}{2}|\mathfrak{B}(S)|/|S| \leq n^{-1/12} \log^3 n. \qquad \square$$

We now give a brief outline of the rest of this section. First, in Section 5.2, we describe a slight modification of a coupling due to Schramm [21]. The coupling evolves a pair of partitions of the interval $[0, 1)$ such that, firstly, the marginal dynamics have PD(θ) as an invariant distribution, and, secondly, the two partitions become 'close'. Moreover, for $\theta = \frac{1}{2}$ these dynamics are very similar to the dynamics of $\mathfrak{C}_t$ above (Schramm defined the coupling for $\theta = 1$ but as we will see and as has been noted before [10], the extension to $\theta \in (0, 1]$ is completely straightforward). Then, in Section 5.3, we focus on the case $\theta = \frac{1}{2}$ and show how an adaptation of Schramm's coupling allows us to couple a PD($\frac{1}{2}$)-sample to the 'discrete' partition coming from the cycles $\mathfrak{C}_t$. This will allow us to prove Theorem 1.1. Lemma 5.1 comes in here and, intuitively speaking, by using the pair (U', U'') as described above we "trade accuracy for independence": U' will tell us the exact location

for splitting in the PD($\frac{1}{2}$)-distributed partition, whereas in $\mathfrak{C}_t$ this is decided by U''. As we will see, the locations in the segment S defined by U' and U'' are close enough to each other, and p_t is close enough to $\frac{1}{2}$, that the two partitions become more and more similar.

5.2 Schramm's coupling

Fix any $\theta \in (0,1]$, later we will take $\theta = \frac{1}{2}$. We will define a sequence $\big((\mathcal{Y}^t, \mathcal{Z}^t) : t = 0, 1, \dots\big)$ of pairs of random partitions of $[0,1)$ into countably many intervals $[a,b)$, in such a way that (i) the marginal dynamics are stationary for PD(θ), and (ii) regardless of starting configuration, $\mathcal{Y}^t$ and $\mathcal{Z}^t$ become 'closer' in a sense to be defined later.

The subintervals $[a,b)$ of $[0,1)$ constituting the partitions $\mathcal{Y}^t$ and $\mathcal{Z}^t$ will be called *blocks*. We will think of the blocks of $\mathcal{Y}^t$ and $\mathcal{Z}^t$ as distinguishable, and as before leave some flexibility about the relative placement of the blocks within $[0,1)$. By a slight abuse of notation we will identify a block $\mathcal{Y}_i^t \subseteq [0,1)$ with its length $|\mathcal{Y}_i^t| \in [0,1]$.

Some of the blocks of $\mathcal{Y}^t$ will be *matched* with blocks of $\mathcal{Z}^t$, and this relation is symmetric (if $\mathcal{Y}_i^t$ is matched with $\mathcal{Z}_j^t$ then $\mathcal{Z}_j^t$ is matched with $\mathcal{Y}_i^t$). Other blocks are unmatched. Matched pairs of blocks have the same size, and such pairs will be created in some instances of the process we are about to describe. The total length of all unmatched blocks will be denoted by $R = R^t$ and the total length of matched blocks $Q = Q^t$. We place the matched blocks at the end of $[0,1)$ and the unmatched blocks at the beginning, and within the matched and unmatched parts we order the blocks by decreasing size. See Figure 11.

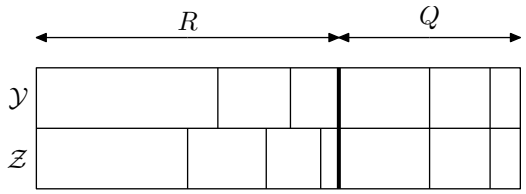

Figure 11: Example of a pair $(\mathcal{Y}, \mathcal{Z})$. The matched blocks account for a total Q of the length, and the unmatched R, where $Q + R = 1$. The thick vertical line indicates the border between the matched and unmatched parts.

A step of the coupling is completed with the help of three independent random variables U, U' and W, all uniformly distributed in $[0,1)$. First U is sampled, and if U falls in the blocks $\mathcal{Y}_i$ and $\mathcal{Z}_j$ of $\mathcal{Y}$ and $\mathcal{Z}$, respectively, then we say that these two blocks of $\mathcal{Y}$ and $\mathcal{Z}$ are *highlighted*. Moreover, the highlighted blocks are moved to the front of $[0,1)$, see Figures 12 and 13. Then U' is sampled and we do the following:

- if (in either $\mathcal{Y}$ or $\mathcal{Z}$) we have that U' falls in a block different from the highlighted one, then this block is merged with the highlighted block;
- if U' falls in a highlighted block then we propose a split of the highlighted block(s) at the position U';
- in the case of proposing a split, the split is carried out if we have that $W \le \theta$.

Thus it is possible to merge blocks in both $\mathcal{Y}$ and $\mathcal{Z}$, to merge blocks in one but (propose a) split in the other, or to (propose a) split in both. In the case when we propose a split in both $\mathcal{Y}$ and $\mathcal{Z}$, note that the same W is used for both, thus either both split or neither. In this case, if they split then at least two of the newly created blocks are of the exact same size (see Figure 14), and those blocks are then declared *matched* and moved to

the matched part. Before the next step the blocks are sorted into the matched and unmatched parts and ordered by size within those parts, as before. Figures 14, 15 and 16 show some of the possible scenarios.

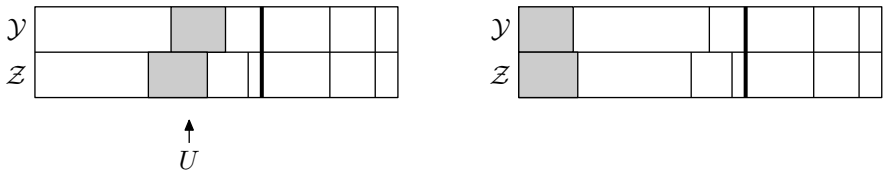

Figure 12: U highlights a block in $\mathcal{Y}$ and a block in $\mathcal{Z}$ and they are moved to the front. In this case the highlighted blocks are unmatched.

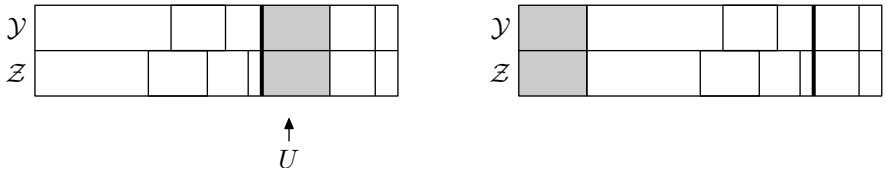

Figure 13: Example when the highlighted blocks are matched.

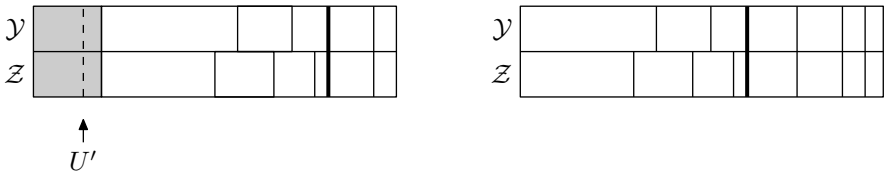

Figure 14: Example when a split is carried out in both $\mathcal{Y}$ and $\mathcal{Z}$. In this case the highlighted blocks are already matched and consequently all the formed blocks are matched.

The following result about the marginal dynamics is due to Tsilevich [24] (for $\theta = 1$) and Pitman [19] (general θ). Another proof can be found in [10, Theorem 7.1].

Lemma 5.2. *If $\mathcal{Y}^0$ (respectively, $\mathcal{Z}^0$) has distribution PD(θ) then $\mathcal{Y}^t$ (respectively, $\mathcal{Z}^t$) has distribution PD(θ) for all $t \geq 0$.*

We will need quantitative results about how the sizes of the largest unmatched blocks evolve under these dynamics. We will present a sequence of results, Lemmas 5.4 to 5.6, which culminate in Corollary 5.7. As the proofs of these lemmas are identical or nearly identical to the corresponding proofs in [21] we omit the details, but give comments where there are differences in the case $\theta < 1$.

Fix $\varepsilon > 0$ and introduce the following notation. Let $N_\varepsilon(\mathcal{Y}^t)$ and $N_\varepsilon(\mathcal{Z}^t)$ denote the number of unmatched blocks of size $\geq \varepsilon$ in $\mathcal{Y}^t$ and $\mathcal{Z}^t$, respectively, and let $N_\varepsilon^t = N_\varepsilon(\mathcal{Y}^t) + N_\varepsilon(\mathcal{Z}^t)$ be the total number of unmatched blocks of size $\geq \varepsilon$ after t steps. Let $\sigma(\varepsilon, \mathcal{Y}^t) = \sum_i \mathcal{Y}_i^t \mathbb{1}_{\{\mathcal{Y}_i^t < \varepsilon\}}$ be the total length of blocks smaller than ε in $\mathcal{Y}^t$, and similarly define $\sigma(\varepsilon, \mathcal{Z}^t)$. Also let $\overline{\varepsilon} = \varepsilon + \sigma(\varepsilon, \mathcal{Y}^0) + \sigma(\varepsilon, \mathcal{Z}^0)$.

Before presenting the lemmas about the coupling, we note the following a-priori estimates:

Proposition 5.3.

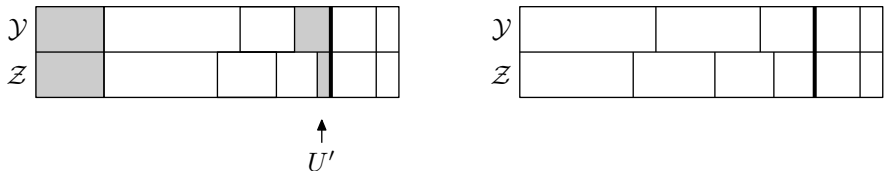

Figure 15: Example when the highlighted blocks are matched but are then merged with some unmatched blocks. In this case the formed blocks are unmatched.

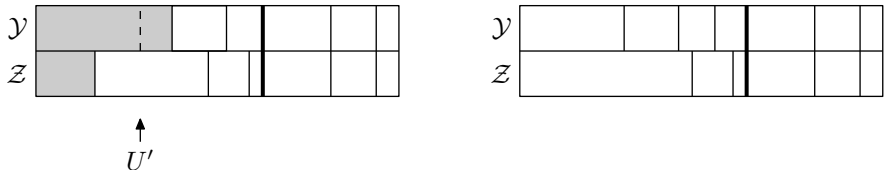

Figure 16: Example when one block is split and one is merged.

1. *If for some $C > 0$ and all $\varepsilon \in (0,1)$ we have*

$$\mathbb{E}[\sigma(\varepsilon, \mathcal{Y}^0)] \le C\varepsilon \log(\tfrac{1}{\varepsilon}) \tag{5.2}$$

   then for some $C' > 0$ we have $\mathbb{E}[N_\varepsilon(\mathcal{Y}^0)] \le C' \log^2(\frac{1}{\varepsilon})$.

2. *If $\mathcal{Y}^0$ has distribution PD(θ) with $\theta \in (0,1]$ then $\mathbb{E}[\sigma(\varepsilon, \mathcal{Y}^0)] \le \varepsilon$.*

The proof is sketched (for $\theta = 1$) in [21]. For completeness we give details in Appendix B. In the proof of Theorem 1.1, $\mathcal{Y}^0$ will have the PD($\frac{1}{2}$)-distribution while $\mathcal{Z}^0$ will satisfy a bound of the form (5.2). Thus, in the following sequence of lemmas, one should think of $\overline{\varepsilon}$ as being of the order $\le \varepsilon \log(\frac{1}{\varepsilon})$ as $\varepsilon \to 0$, and N^0_ε as of the order $\le \log^2(\frac{1}{\varepsilon})$.

In the next few results we will be working conditionally on $(\mathcal{Y}^0, \mathcal{Z}^0)$, hence $\overline{\varepsilon}$ and N^0_ε will be treated as constants. We let q be a random time, independent of the chain $((\mathcal{Y}^t, \mathcal{Z}^t) : t \ge 0)$, and write $\eta = \max\{\mathbb{P}(q = t) : t \ge 0\}$.

Let $y^t_{(1)}$ and $z^t_{(1)}$ denote the largest *unmatched* blocks in $\mathcal{Y}^t$ and $\mathcal{Z}^t$, respectively. In the following result, note that $R^t - y^t_{(1)} \vee z^t_{(1)}$ is small if most of the unmatched length R^t is covered by the largest unmatched block in either $\mathcal{Y}^t$ or in $\mathcal{Z}^t$. Hence the product $R^t(R^t - y^t_{(1)} \vee z^t_{(1)})$ is small if either R^t is small (which is what we want), or the unmatched part contains a large block (which can be handled because such a situation is 'unstable').

Lemma 5.4. *Conditionally on $\mathcal{Y}^0, \mathcal{Z}^0$,*

$$\mathbb{E}\big[R^q(R^q - y^q_{(1)} \vee z^q_{(1)})\big] \le \frac{\eta}{2} N^0_\varepsilon + 5\overline{\varepsilon}\, \mathbb{E}[q].$$

When applying this and the following estimates, the main case will be when q is uniformly distributed on $\{0, 1, \ldots, \lceil \varepsilon^{-1/2} \rceil - 1\}$. Then η is approximately $\varepsilon^{1/2}$ and $\mathbb{E}[q]$ is of the order $\varepsilon^{-1/2}$. If $\overline{\varepsilon}$ and N^0_ε are of the order indicated above then the right-hand-side is small (of the order $\le \varepsilon^{1/2} \log^2(\frac{1}{\varepsilon})$).

Proof. The proof is essentially identical to the proof in [21, Lemma 3.1] and uses that on the event that up to time q no blocks of size $\le \varepsilon$ are created or merged, N^t_ε is non-increasing for $t < q$. The only extra case which arises for $\theta < 1$ is when, going from t to $t+1$, two blocks are merged in $\mathcal{Y}$ (respectively, $\mathcal{Z}$) but a split is proposed for $\mathcal{Z}$ (respectively, $\mathcal{Y}$) and not accepted. In this case we see that $N^{t+1}_\varepsilon = N^t_\varepsilon - 1$, hence N^t_ε is still non-increasing. □

Write $y^t_{(2)}$ for the second-largest unmatched block in $\mathcal{Y}^t$.

Lemma 5.5. *For $\rho \in (0,1)$ we have (conditionally on $\mathcal{Y}^0, \mathcal{Z}^0$)*

$$\mathbb{P}\big(R^q - (y^q_{(1)} + y^q_{(2)}) > \rho\big) \le 2^6\theta^{-1}\rho^{-4}\big(\tfrac{\eta}{2}N^0_\varepsilon + 5\bar{\varepsilon}\mathbb{E}[q+1]\big).$$

The lemma says that the two largest unmatched entries together dominate the unmatched part (if η, q, $\bar{\varepsilon}$ and N^0_ε are of the order indicated above, then the right-hand-side is small as long as, say, $\rho \ge \varepsilon^{1/10}$).

Proof. This proof is virtually identical to the proof of [21, Lemma 3.2]. We consider whether the event $\mathcal{R} = \{R^q - z^q_{(1)} < \rho/4\}$ occurs or not. In the case when $\mathcal{R}$ does not occur we can apply Lemma 5.4 exactly as in [21]. In the case when $\mathcal{R}$ does occur, the key observation in [21] is that there is good probability that $z^q_{(1)}$ splits into two blocks of size $\ge \rho/4$ while two unmatched blocks of $\mathcal{Y}^q$ merge, allowing us to apply Lemma 5.4 in the next step instead. The only extra consideration for $\theta < 1$ is that the split must be accepted, which happens with probability θ, resulting in the factor θ^{-1} in the statement of the lemma. □

We next bound the 'average' probability of having a large unmatched block. Its corollary, Corollary 5.7, is especially important for us.

Lemma 5.6. *Let $\varepsilon, \rho \in (0,1)$ and let $t > 0$ and k be such that $t \ge 2^k/\rho$. Then (conditionally on $\mathcal{Y}^0, \mathcal{Z}^0$)*

$$t^{-1}\sum_{s=0}^{t-1}\mathbb{P}(y^s_{(1)} > \rho) \le C[k^{-1}\rho^{-1} + 2^{4k}\rho^{-5}(N^0_\varepsilon/t + \bar{\varepsilon}t)], \tag{5.3}$$

for some constant C.

Proof. The proof is identical to the proof of [21, Lemma 3.3] where we insert the bounds from Lemmas 5.4 and 5.5 when the bounds from [21, Lemma 3.1] and [21, Lemma 3.2] are used. □

We now make some additional assumptions on $(\mathcal{Y}^0, \mathcal{Z}^0)$ and q, which allow us to obtain a more explicit bound on $\mathbb{P}(y^q_{(1)} \ge \rho)$. As usual we work conditionally on $(\mathcal{Y}^0, \mathcal{Z}^0)$.

Corollary 5.7. *Assume that $\bar{\varepsilon} < 1$ and that*

$$(\bar{\varepsilon})^{1-\gamma} \le \eta \le (\bar{\varepsilon})^\gamma/(N^0_\varepsilon \vee 1), \quad \text{for some } \gamma \in (0, \tfrac{1}{2}). \tag{5.4}$$

Then for some $C = C(\gamma)$ we have that for all $\rho \in (0,1)$,

$$\mathbb{P}(y^q_{(1)} \ge \rho) \le \mathbb{P}(q \ge 1/\eta) + \frac{C}{\rho\log(1/\bar{\varepsilon})}. \tag{5.5}$$

Proof. The proof is identical to the proof of [21, Corollary 3.4] (in [21] there is an additional parameter λ which we have set to 1). □

Note that if q is uniformly distributed on $\{0, 1, \dots, \lceil\varepsilon^{-1/2}\rceil - 1\}$ and N^0_ε is of the order at most $\log^2(\frac{1}{\varepsilon})$, as discussed above, then (5.4) holds. Moreover, in this case $\mathbb{P}(q \ge 1/\eta) = 0$. If $\bar{\varepsilon}$ is at most of the order $\varepsilon\log(\frac{1}{\varepsilon})$ as discussed above then the right-hand-side of (5.5) can be made arbitrarily small by choosing ε small.

5.3 Proof of Theorem 1.1

We now turn to the proof of our main result. Let $\varepsilon > 0$, to be fixed later. Recall the set-up of Section 5.1: ω is sampled from $\mathbb{P}_\beta$ where $\beta > 1$, and we consider $\vec{\omega}'_s = \vec{\omega}_s$ for $s = |\omega| - q$. We take q to be uniformly distributed on $\{0, 1, \dots, \lceil \varepsilon^{-1/2} \rceil - 1\}$. Let $\mathcal{A}_0$ be the event that the following conditions hold:

- for some $c > \frac{1}{2}$ we have $s \geq cn$;
- the event in (5.1) holds for $\vec{\omega}'_s = \vec{\omega}_s$; and
- the random graph $G(n, s)$, which has an edge wherever $\vec{\omega}_s$ has at least one link, has a unique giant connected component V_G containing between $0.99zn$ and $1.01zn$ vertices, and any other connected component has size at most $\log^2 n$. (Here z is the same as in Proposition 3.4.)

In the following discussion we will assume that $\mathcal{A}_0$ holds, as $\mathbb{P}(\mathcal{A}_0^c) = o(1)$ as $n \to \infty$. Note that the cycles $\mathfrak{C}_0$ refine the components of $G(n, s)$, hence (on $\mathcal{A}_0$) a cycle is either contained in V_G or it has size $\leq \log^2 n$.

We take $\mathcal{Y}^0$ to have distribution PD($\frac{1}{2}$). Roughly speaking, $\mathcal{Z}^0$ will be obtained from the cycles $\mathfrak{C}_0$ and we want to use the coupling from Section 5.2 to obtain $(\mathcal{Y}^q, \mathcal{Z}^q)$. The main modification of the coupling is that we use the construction in Lemma 5.1 for splitting in $\mathcal{Z}$. There are also several minor modifications to take into account. In what follows we work conditionally on $\vec{\omega}'_0$.

We subdivide the cycles of $\mathfrak{C}_0$ into segments $S \in \mathfrak{S}$ as in Section 5.1. Write $m = |V_G|$. We let $\mathcal{Z}^0$ be a representation of the cycles $\mathfrak{C}_0$ as intervals as in Section 5.1, but now as subintervals of $[0, \frac{n}{m})$ rather than $[0, 1)$. Thus each vertex v is represented by a subinterval $I(v)$ of the form $[\frac{i}{m}, \frac{i+1}{m})$ where $0 \leq i = i(v) \leq n - 1$. Note that $[0, \frac{n}{m})$ has length roughly $\frac{1}{z}$. In keeping with the terminology of the previous subsection, we refer to the intervals which represent the cycles as *blocks*. The subintervals $I(v)$ representing the vertices are labelled using $\uparrow, \downarrow$, as before. Clearly, a cycle of size $\geq \lfloor n^{1/2} \rfloor$ in $\mathfrak{C}_0$ is represented as a block of size $\geq \lfloor n^{1/2} \rfloor / m$ in $\mathcal{Z}^0$.

We place the blocks representing the cycles of $\mathfrak{C}_0$ which lie in the giant component V_G at the start of $[0, \frac{n}{m})$, i.e. in $[0, 1)$. The blocks representing the remaining cycles are placed in $[1, \frac{n}{m})$. Within $[0, 1)$ we will later have matched and unmatched blocks, as in Section 5.2, and again we place the unmatched blocks first and within the matched and unmatched parts we order the blocks by decreasing size. See Figure 17.

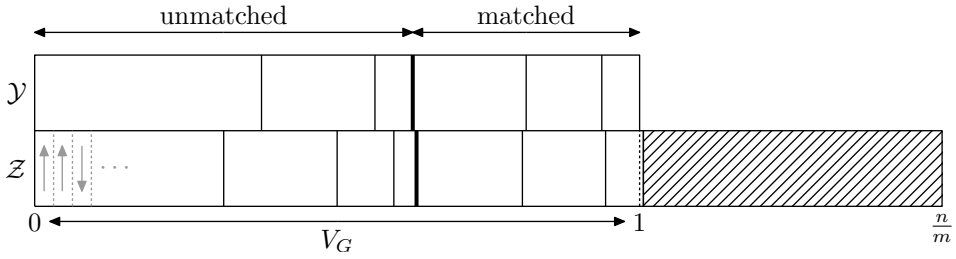

Figure 17: $\mathcal{Z}^t$ consists of blocks representing the cycles $\mathfrak{C}_t$, placed in the interval $[0, \frac{n}{m})$. Cycles belonging to V_G are placed first, roughly in the interval $[0, 1)$, and are sorted into those matched with a block of $\mathcal{Y}^t$ and those not. Matched blocks can differ slightly in size. Also $\mathcal{Z}^t$ will after a few steps have an 'overhang' since the giant component grows. The hatched part consists of blocks representing cycles with vertices that are not in V_G.

We will define dynamics for $(\mathcal{Y}^t, \mathcal{Z}^t)$ such that the marginal dynamics for $\mathcal{Y}^t$ are as in Section 5.2 with $\theta = \frac{1}{2}$, and the marginal dynamics for $\mathcal{Z}^t$ are as in Section 5.1. Thus $\mathcal{Y}^t$

will have distribution PD($\frac{1}{2}$) for all t, due to Lemma 5.2, and $\mathcal{Z}^t$ will be a representation of the cycles $\mathfrak{C}_t$ as intervals.

In order to be able to define a *successful coupling* later on, we will need the notion of a *forbidden set* $F^t \subseteq [0, \frac{n}{m})$ (for any given time t). This set will arise due to small errors which accumulate during the process. Initially, for $t = 0$, we set $F^0 = \varnothing$. In later steps, we define F^t as consisting of the following parts:

- First, all segments which have been *touched* up to time t (or rather, the union of the $I(v)$ for v belonging to touched segments) in $\mathcal{Z}$ are forbidden.

- Second, F^t contains an *overhang* (defined as $\{s \in (1, \frac{n}{m}] : s \in I(u) \text{ for some } u \in V_G\}$) which arises because in $\mathcal{Z}$ cycles outside the giant component may merge with cycles inside the giant component, meaning that the giant component grows with time.

- Third, it will be necessary to allow matched blocks, defined shortly, to have slightly different sizes, rather than the exact same size as in Section 5.2. When the blocks of $\mathcal{Y}^t$ and $\mathcal{Z}^t$ are lined up as in Figure 17, the subset in $[0, \frac{n}{m})$ where part of a matched block does not overlap with its partner belongs to the forbidden set F^t.

- Also, $[0, \frac{1}{m})$ is forbidden. To understand the meaning of this, recall that once U has been sampled, the block it highlighted is moved and rotated so that the corresponding $I(u)$ is moved to $[0, \frac{1}{m})$. Putting this interval in the forbidden set will simply be a way to enforce that all links have two different endpoints.

Once we have described precisely the transitions in our process we will easily be able to bound the size $|F^t|$ by a very small number, see (5.6).

Let us now define a step of the process. Steps will again be accomplished using independent uniform random variables U, U', U'' and W, but now U and U' will be uniformly distributed in $[0, \frac{n}{m})$ while U'' and W are still uniform in $[0, 1)$. For $\mathcal{Z}$ we will also sample the mark $m \in \{\rightleftharpoons, \rightleftharpoons\}$ of the corresponding link in each step. In what follows we will assume that U and U' never fall in the forbidden set. For concreteness, if U or U' fell in the forbidden set we would declare the process failed and stop.

Firstly, if U or U' falls outside $[0, 1)$ we perform the corresponding transition in $\mathcal{Z}$, using the rules from Section 5.1, but do nothing to $\mathcal{Y}$. Let us now assume that U, U' both fall in $[0, 1)$. When U and U' highlight different blocks these blocks are merged, as before. For $\mathcal{Z}^t$ the labels $\uparrow, \downarrow$ must be handled appropriately, taking into account also the mark $m \in \{\rightleftharpoons, \rightleftharpoons\}$ of the link, as in Section 2. In the case when U' falls in a block highlighted by U (in either $\mathcal{Y}$, $\mathcal{Z}$ or both) it has to be decided if a split should be carried out. There are three cases for how this is decided. First, if a split is proposed in $\mathcal{Y}$ only (Figure 16) then it is carried out if $W \leq \frac{1}{2}$, so this case is the same as in Section 5.2. Second, if a split is proposed only in $\mathcal{Z}$ then we decide whether to carry it out by looking at the labels $\uparrow, \downarrow$ in the intervals $I(u)$ and $I(v)$ selected by U and U', as well as the mark $m \in \{\rightleftharpoons, \rightleftharpoons\}$ of the link, and applying the rules of Section 2. In this case we do not need to use U''. However, the third and most important case is when a split is proposed in both $\mathcal{Y}$ and $\mathcal{Z}$. In this case we do the following:

1. In $\mathcal{Y}$ we record the exact location of U'. If we decide to carry out the split in $\mathcal{Y}$, then it will be done at the location of U'.

2. In $\mathcal{Z}$ we only record the segment S in which U' falls;

3. Then we use U'' to independently sample a uniform point within S and make the splitting decision for $\mathcal{Z}$ as in Section 5.1.

It only remains to specify how we decide whether or not to split in $\mathcal{Y}$. Let v be such that $U'' \in I(v)$. Assume that the block of $\mathcal{Z}$ highlighted by U has size $\geq \lfloor n^{1/2} \rfloor / m$, and that U, U' did not fall in the forbidden set F^t. We are then able to apply Lemma 5.1. Thus the conditional probability p_t that $I(u)$ and $I(v)$ have the same label $\uparrow$ is within $n^{-1/12} \log^3 n$ of $\frac{1}{2}$. Depending on whether p_t is bigger or smaller than $\frac{1}{2}$, and also on the mark $m \in \{\rightarrowtail\!\!\!\!\!\!\succ, \rightarrowtail\!\!\!\!\!\!\prec\}$ of the link, the conditional probability of splitting in $\mathcal{Z}^t$ is thus either slightly above or slightly below $\frac{1}{2}$. We wish to 'maximally couple' the decision whether or not to split in $\mathcal{Z}$ with the decision in $\mathcal{Y}$, but keeping the splitting probability for $\mathcal{Y}$ at exactly $\frac{1}{2}$. Let us describe this assuming $p_t \leq \frac{1}{2}$, the other case is similar.

- If the mark $m = \rightarrowtail\!\!\!\!\!\!\succ$, recall that this means that we split in $\mathcal{Z}$ if $I(v)$ has label $\uparrow$, i.e. with probability $p_t \leq \frac{1}{2}$. Our rule for $\mathcal{Y}$ is then: split in $\mathcal{Y}$ if $I(v)$ has label $\uparrow$, but if $I(v)$ has label $\downarrow$ split in $\mathcal{Y}$ anyway with probability $(\frac{1}{2} - p_t)/(1 - p_t)$ (independently of all other choices).
- If the mark $m = \rightarrowtail\!\!\!\!\!\!\prec$, this means that we split in $\mathcal{Z}$ if $I(v)$ has label $\downarrow$, i.e. with probability $1 - p_t \geq \frac{1}{2}$. Our rule for $\mathcal{Y}$ is then: do nothing (no split) in $\mathcal{Y}$ if $I(v)$ has label $\uparrow$, but if $I(v)$ has label $\downarrow$ also do nothing in $\mathcal{Y}$ with probability $(\frac{1}{2} - p_t)/(1 - p_t)$ (independently of all other choices).

It is not hard to check that these rules ensure that the probability of splitting in $\mathcal{Y}$ is exactly $\frac{1}{2}$, independently of the location U' of the proposed split. We can also see from this that the probability of $\mathcal{Y}$ and $\mathcal{Z}$ making different choices (i.e. one splits and the other one twists) is $|p_t - \frac{1}{2}| \leq n^{-1/12} \log^3 n$.

If the decision is to split in *both* $\mathcal{Y}$ and $\mathcal{Z}$, the blocks created are declared *matched* as in Section 5.2 (if the blocks which split were already matched we get two pairs of matched blocks, otherwise one pair). Note that U' and U'' differ by at most $2\lfloor n^{1/2} \rfloor / m$ due to the upper bound on the size of segments S; this will give us a bound on how much matched blocks can differ in size.

There is one final case in which we need to specify the rules for deciding to split, which is when a split is proposed in both $\mathcal{Y}$ and $\mathcal{Z}$ but the block of $\mathcal{Z}$ has size $< \lfloor n^{1/2} \rfloor / m$. This is unlikely and we will see that we can assume that this does not occur, but to be definite let us say that in this case we split in $\mathcal{Y}$ if $W \leq \frac{1}{2}$.

Now we turn to bounding the size of the forbidden set F^t. We claim that, for any $t \in \{0, 1, \ldots, q\}$ we have

$$|F^t| \leq \frac{7\varepsilon^{-1/2} n^{1/2}}{m}. \tag{5.6}$$

Indeed, after t steps we have added at most $4t\lfloor n^{1/2} \rfloor / m$ due to touched segments, at most $t \log^2 n/m$ due to overhang, and at most $2t\lfloor n^{1/2} \rfloor / m$ due to the size-difference of matched blocks. Adding to this $1/m$ for $[0, \frac{1}{m})$ and recalling that $t \leq q \leq \varepsilon^{-1/2}$ we arrive at (5.6).

Let us say that the coupling $(\mathcal{Y}^q, \mathcal{Z}^q)$ was *successful*, denoting this event by $\mathcal{G}$, if the following occur for all $t \in \{1, \ldots, q\}$:

- U_t, U'_t do not fall in the forbidden set F^t;
- in step t we do not propose to split a block of $\mathcal{Z}$ which has size $< \lfloor n^{1/2} \rfloor / m$; and
- if at step t it is proposed to split a block in both $\mathcal{Y}$ and $\mathcal{Z}$ then we either split in both or in neither.

Using (5.6) and Lemma 5.1 and recalling that $q \leq \varepsilon^{-1/2}$ we get:

$$\begin{aligned} \mathbb{P}(\mathcal{G}^c) &\leq 2\varepsilon^{-1/2} \frac{7\varepsilon^{-1/2} n^{1/2}/m}{n/m} + \varepsilon^{-1/2} \frac{n^{1/2}/m}{n/m} + \varepsilon^{-1/2} n^{-1/13} \\ &\leq 16\varepsilon^{-1} n^{-1/13}. \end{aligned} \tag{5.7}$$

Here we bounded the probability that we make different decisions for splitting a large block in $\mathcal{Y}$ and in $\mathcal{Z}$ by $n^{-1/13}$, which is valid for large enough n.

We can now put the different pieces together and wrap up the proof of Theorem 1.1. Having defined $(\mathcal{Y}^q, \mathcal{Z}^q)$, let us now order the blocks in both of them by decreasing size, and let us think of them as two infinite sequences by appending infinitely many 0's at the end. Let $\delta > 0$ be arbitrary and write

$$\mathcal{D} = \big\{\|\mathcal{Y}^q - \mathcal{Z}^q\|_\infty > \delta\big\}.$$

It suffices to show that $\mathbb{P}(\mathcal{D})$ can be made arbitrarily small by choosing $\varepsilon > 0$ small and n large.

Recall that we have been working on the event $\mathcal{A}_0$ defined in the beginning in this proof. Also recall the quantities $\sigma(\varepsilon, \mathcal{Y})$ and $N_\varepsilon(\mathcal{Y})$ defined in Section 5.2. We now define $\sigma(\varepsilon, \mathcal{Z})$ and $N_\varepsilon(\mathcal{Z})$ similarly but counting only those blocks which intersect $[0, 1)$. Given this, $\bar{\varepsilon}$ and N^0_ε are given as in Section 5.2. By (2.4) in Lemma 2.4 we have for some $C > 0$ and any $\varepsilon > 0$ that $\mathbb{E}[\sigma(\varepsilon, \mathcal{Z}^0)] \le C\varepsilon \log(\frac{1}{\varepsilon})$. Hence by Proposition 5.3 we also have $\mathbb{E}[N_\varepsilon(\mathcal{Z}^0)] \le C' \log^2(\frac{1}{\varepsilon})$. By the same Proposition, the same bounds also apply to $\mathcal{Y}^0$. Hence $\mathbb{E}[\bar{\varepsilon}] \le 3C\varepsilon \log(\frac{1}{\varepsilon})$ and $\mathbb{E}[N^0_\varepsilon] \le 2C' \log^2(\frac{1}{\varepsilon})$. Defining the events $\mathcal{A}_1 = \{\bar{\varepsilon} \le \varepsilon^{3/4}\}$ and $\mathcal{A}_2 = \{N^0_\varepsilon \le \varepsilon^{-1/4}\}$ and using Markov's inequality, we get

$$\mathbb{P}(\mathcal{A}_0 \cap \mathcal{A}_1^c) \le 3C\varepsilon^{1/4} \log(\tfrac{1}{\varepsilon}), \qquad \mathbb{P}(\mathcal{A}_0 \cap \mathcal{A}_2^c) \le 2C'\varepsilon^{1/4} \log^2(\tfrac{1}{\varepsilon}).$$

On $\mathcal{G} \cap \mathcal{A}_0 \cap \mathcal{A}_1 \cap \mathcal{A}_2$, we can apply Corollary 5.7, with $\gamma = \frac{1}{5}$ say, and $\rho = \delta/2$. (We use $\rho = \delta/2$ rather than δ to account for such things as the size-difference of matched blocks; thus n should be taken sufficiently large.) This gives, for some $C'' > 0$ and n large

$$\begin{aligned}\mathbb{P}(\mathcal{D}) &\le \mathbb{P}(\mathcal{A}_0^c) + \mathbb{P}(\mathcal{A}_0 \cap (\mathcal{A}_1^c \cup \mathcal{A}_2^c \cup \mathcal{G}^c)) + \mathbb{P}(\{y^q_{(1)} \ge \delta/2\} \cap \mathcal{A}_0 \cap \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{G}) \\ &\le o(1) + C''\varepsilon^{1/4} \log^2(\tfrac{1}{\varepsilon}) + \frac{2C}{\delta \log(\varepsilon^{-3/4})}.\end{aligned}$$

The right-hand-side can be made arbitrarily small by picking $\varepsilon > 0$ small and n large. (Since $\mathcal{G}$ involves the entire process, to be completely rigorous Lemmas 5.4–5.6 and Corollary 5.7 should be proved on the event $\mathcal{G}$. This can be done by working with the time $\min\{q, \tau\}$ where τ is the first time at which $\mathcal{G}$ fails. From (5.7) we see that, with high probability, $\tau > q$ and so the only change is the addition of an $o(1)$ term.) □

A Proof of Proposition 3.4

We use the notation from Proposition 3.4, noting that $\mathcal{S} = \{Z_t > -1 \text{ for all } t \ge 0\}$. Also note that $z = \mathbb{P}(\mathcal{S}) > 0$ when $\beta > 1$, since if $z = 0$ then $\mathbb{P}(\exists t : Z_t = -1) = 1$ which by the Markov property would imply $\mathbb{P}(\liminf_{t\to\infty} Z_t = -\infty) = 1$, contradicting the fact that $Z_t \to +\infty$ almost surely.

It will be useful to consider the *times of record minima* m_k, which are defined as follows. Let $\tau_1, \tau_2, \ldots$ denote the jump times of Z (equivalently, of N'). First we define $m_1 := \tau_1$, and then inductively

$$m_{k+1} := \min\{\tau_j > m_k : Z_{\tau_j -} < Z_{m_k -}\}, \qquad \text{where } \min \varnothing = \infty.$$

Importantly, the m_k are stopping times and they characterize the frontier time ℓ_1 by $\ell_1 = \max\{m_k : m_k < \infty\}$, i.e. ℓ_1 equals the last record time. The later frontier times ℓ_k may be expressed similarly using the record minima of $Z^{(k)}$. Using that the m_k are stopping times we have for all $k \ge 1$ that

$$\mathbb{P}(m_k < \infty) = (1 - z)^{k-1}. \tag{A.1}$$

Proof that $z = 1 - e^{-\beta z}$: Write $\delta = 1 - z = \mathbb{P}(\exists t : Z_t = -1)$. By conditioning on the first time when Z_t hits -1 we see that $\mathbb{P}(\exists t : Z_t = -j) = \delta^j$ for all $j \geq 1$. Since N'_t only takes integer values it follows $\delta^j = \mathbb{P}(\exists k \geq 0 : N'_{k+j} = k)$. From this and the Markov property at time 1 we find that

$$\delta = \sum_{j \geq 0} e^{-\beta} \frac{\beta^j}{j!} \delta^j = e^{-\beta(1-\delta)},$$

as claimed. □

Proof that, given $\mathcal{S}$, the sequence $\big\{\big(\Delta_k, (Z_t^{(k)})_{0 \leq t < \Delta_k}\big)\big\}_{k=0}^{+\infty}$ is i.i.d.: We start by establishing that

$$Z^{(k)} \stackrel{(\mathrm{d})}{=} \big(Z \mid Z \in \mathcal{S}\big), \qquad \text{for all } k \geq 1. \tag{A.2}$$

This is reasonable since, for example, after ℓ_1 we know that Z does not set a new record minimum, which is the same as saying that $Z^{(1)}$ does not hit -1. Using induction, (A.2) follows from these two equalities in law:

$$\begin{aligned} &Z^{(1)} \stackrel{(\mathrm{d})}{=} \big(Z \mid Z \in \mathcal{S}\big), \quad \text{and} \\ &\big(Z^{(1)} \mid Z \in \mathcal{S}\big) \stackrel{(\mathrm{d})}{=} \big(Z \mid Z \in \mathcal{S}\big). \end{aligned} \tag{A.3}$$

To prove (A.3), let $\mathcal{B}$ be some event. Using the description of ℓ_1 in terms of the stopping times m_k we see that

$$\begin{aligned} \mathbb{P}(Z^{(1)} \in \mathcal{B}) &= \sum_{k \geq 1} \mathbb{P}\big(m_k < \infty, m_{k+1} = \infty, (Z_{m_k+t} - Z_{m_k})_{t \geq 0} \in \mathcal{B}\big) \\ &= \sum_{k \geq 1} \mathbb{P}\big(m_k < \infty, (Z_{m_k+t} - Z_{m_k})_{t \geq 0} \in \mathcal{B} \cap \mathcal{S}\big) \\ &= \sum_{k \geq 1} \mathbb{E}\big[\mathbb{1}_{\{m_k < \infty\}} \mathbb{P}\big((Z_{m_k+t} - Z_{m_k})_{t \geq 0} \in \mathcal{B} \cap \mathcal{S} \mid m_k\big)\big] \\ &= \mathbb{P}(Z \in \mathcal{B} \cap \mathcal{S}) \sum_{k \geq 1} \mathbb{P}(m_k < \infty) = \mathbb{P}(Z \in \mathcal{B} \cap \mathcal{S}) \sum_{k \geq 1} (1-z)^{k-1} \\ &= \mathbb{P}(Z \in \mathcal{B} \cap \mathcal{S}) / \mathbb{P}(Z \in \mathcal{S}), \end{aligned}$$

thus $Z^{(1)} \stackrel{(\mathrm{d})}{=} (Z^{(0)} \mid Z^{(0)} \in \mathcal{S})$.

It will be useful to note the following:

$$\begin{aligned} \mathbb{P}(\ell_1 \leq a \mid Z \in \mathcal{S}) &= \sum_{k \geq 1} \mathbb{P}\big(m_k \leq a, m_{k+1} = \infty, Z \in \mathcal{S}\big) / \mathbb{P}(\mathcal{S}) \\ &= \sum_{k \geq 1} \mathbb{P}\big(m_k \leq a, Z_{m_k -} > -1, (Z_{m_k+t} - Z_{m_k}) \in \mathcal{S}\big) / \mathbb{P}(\mathcal{S}) \\ &= \sum_{k \geq 1} \mathbb{P}\big(m_k \leq a, Z_{m_k} > 0\big). \end{aligned}$$

In particular, letting $a \to \infty$,

$$1 = \sum_{k \geq 1} \mathbb{P}\big(m_k < \infty, Z_{m_k} > 0\big).$$

Using this we find that

$$\begin{aligned} \mathbb{P}(Z^{(1)} \in \mathcal{B} \mid Z \in \mathcal{S}) &= \mathbb{P}(\mathcal{S})^{-1} \sum_{k \geq 1} \mathbb{P}(m_k < \infty, Z_{m_k} > 0, (Z_{m_k+t} - Z_{m_k})_{t \geq 0} \in \mathcal{B} \cap \mathcal{S}) \\ &= \frac{\mathbb{P}(Z \in \mathcal{B} \cap \mathcal{S})}{\mathbb{P}(\mathcal{S})} \sum_{k \geq 1} \mathbb{P}\big(m_k < \infty, Z_{m_k} > 0\big) \\ &= \mathbb{P}(Z \in \mathcal{B} \mid Z \in \mathcal{S}). \end{aligned}$$

Thus $\big(Z^{(1)} \mid Z \in \mathcal{S}\big) \stackrel{(\mathrm{d})}{=} \big(Z \mid Z \in \mathcal{S}\big)$.

It remains to prove independence. For this it suffices to show that, for any $r \geq 2$ and events $\mathcal{B}_1, \ldots, \mathcal{B}_r$ we have

$$\mathbb{P}((Z_t^{(i)})_{0 \leq t < \Delta_i} \in \mathcal{B}_i \; \forall i = 1, \ldots, r) = \prod_{i=1}^{r} \mathbb{P}((Z_t^{(i)})_{0 \leq t < \Delta_i} \in \mathcal{B}_i).$$

We give details for the case $r = 2$, the other cases are similar. Using (A.2) we have

$$\begin{aligned} &\mathbb{P}((Z_t^{(1)})_{0 \leq t < \Delta_1} \in \mathcal{B}_1, (Z_t^{(2)})_{0 \leq t < \Delta_2} \in \mathcal{B}_2) \\ &\qquad = \mathbb{P}((Z_t^{(0)})_{0 \leq t < \Delta_0} \in \mathcal{B}_1, (Z_t^{(1)})_{0 \leq t < \Delta_1} \in \mathcal{B}_2 \mid Z^{(0)} \in \mathcal{S}). \end{aligned} \tag{A.4}$$

Note that

$$\begin{aligned} &\mathbb{P}((Z_t^{(0)})_{0 \leq t < \Delta_0} \in \mathcal{B} \mid Z^{(0)} \in \mathcal{S}) \\ &= \mathbb{P}(\mathcal{S})^{-1} \sum_{k \geq 1} \mathbb{P}\big(m_k < \infty, Z_{m_k} > 0, (Z_t)_{0 \leq t < m_k} \in \mathcal{B}, (Z_{m_k + t} - Z_{m_k})_{t \geq 0} \in \mathcal{S}\big) \\ &= \sum_{k \geq 1} \mathbb{P}\big(m_k < \infty, Z_{m_k} > 0, (Z_t)_{0 \leq t < m_k} \in \mathcal{B}\big) \end{aligned} \tag{A.5}$$

The right-hand-side of (A.4) equals

$$\begin{aligned} \mathbb{P}(\mathcal{S})^{-1} \sum_{k \geq 1} \mathbb{P}\Big(m_k < \infty, Z_{m_k} > 0, &(Z_t)_{0 \leq t < m_k} \in \mathcal{B}_1, (Z_{m_k + t} - Z_{m_k})_{t \geq 0} \in \mathcal{S}; \\ &(Z_{m_k + t} - Z_{m_k})_{0 \leq t < \ell_2 - \ell_1} \in \mathcal{B}_2\Big). \end{aligned}$$

By conditioning on m_k and $(Z_t)_{0 \leq t < m_k}$ we find that this equals

$$\mathbb{P}((Z_t^{(0)})_{0 \leq t < \Delta_0} \in \mathcal{B}_2 \mid Z \in \mathcal{S}) \sum_{k \geq 1} \mathbb{P}(m_k < \infty, Z_{m_k} > 0, (Z_t)_{0 \leq t < m_k} \in \mathcal{B}_1),$$

which by (A.5) and (A.2) equals $\mathbb{P}((Z_t^{(1)})_{0 \leq t < \Delta_1} \in \mathcal{B}_1) \mathbb{P}((Z_t^{(2)})_{0 \leq t < \Delta_2} \in \mathcal{B}_2)$. □

Proof of (3.5) *and* (3.6). Both will follow from $\mathbb{P}(\ell_1 \geq t) \leq Ce^{-ct}$. Indeed, for (3.5) we have that

$$\mathbb{P}(\tau^Y < \infty, \tau^Y \geq t) \leq \mathbb{P}(Z_{\ell_1 -} \leq -1, \ell_1 \geq t) \leq \mathbb{P}(\ell_1 \geq t),$$

and for (3.6) we have by what was shown above

$$\mathbb{P}(\Delta_k \geq t \mid \mathcal{S}) = \mathbb{P}(\ell_1 \geq t \mid \mathcal{S}) \leq \frac{1}{z} \mathbb{P}(\ell_1 \geq t).$$

We have

$$\mathbb{P}(\ell_1 \geq t) \leq \sum_{k \geq \lfloor t \rfloor} \mathbb{P}(\ell_1 \in [k, k+1)) \leq \sum_{k \geq \lfloor t \rfloor} \mathbb{P}(Z_k < 1) = \sum_{k \geq \lfloor t \rfloor} \mathbb{P}(N'_k < k+1),$$

since if $\ell_1 \in [k, k+1)$ then in particular Z_t must be < 0 for some $t \in [k, k+1)$ which requires that $Z_k < 1$. Now N'_k is Po(βk)-distributed, so a simple computation with Laplace-transforms gives

$$\mathbb{P}(N'_k < k+1) \leq e\beta \exp[-k(\beta - 1 - \log \beta)].$$

Since $\beta > 1$ we have that $\beta - 1 - \log \beta > 0$, and the bound follows. □

B Proof of Proposition 5.3

Let us write simply $\mathcal{Y}$ for $\mathcal{Y}^0$. For the first part, let $K = \lceil \log_2(1/\varepsilon) \rceil$. We have

$$\begin{aligned}\sum_{k=0}^{K} 2^k \sigma(2^{-k}, \mathcal{Y}) &\geq \sum_i \mathcal{Y}_i \sum_{k=0}^{K} 2^k \mathbb{1}_{\{2^{-K} \leq \mathcal{Y}_i < 2^{-k}\}} \\ &= \sum_i \mathcal{Y}_i \mathbb{1}_{\{\mathcal{Y}_i \geq \varepsilon\}} \sum_{k=0}^{\lfloor \log_2(1/\mathcal{Y}_i) \rfloor} 2^k \\ &\geq \sum_{i \geq 1} \mathcal{Y}_i \mathbb{1}_{\{\mathcal{Y}_i \geq \varepsilon\}} \frac{1}{\mathcal{Y}_i} = N_\varepsilon(\mathcal{Y}).\end{aligned}$$

Hence using (5.2),

$$\mathbb{E}[N_\varepsilon(\mathcal{Y})] \leq \sum_{k=0}^{K} 2^k \mathbb{E}[\sigma(2^{-k}, \mathcal{Y})] \leq C \sum_{k=0}^{K} k \leq C' K^2,$$

which gives the claim.

Next, if $\mathcal{Y}$ is PD(θ)-distributed then a size-biased sample from $\mathcal{Y}$ is Beta$(1, \theta)$-distributed. This means that if we select a random index I in such a way that $\mathbb{P}(I = i \mid \mathcal{Y}) = |\mathcal{Y}_i|$, then

$$\mathbb{E}[\sigma(\varepsilon, \mathcal{Y})] = \mathbb{E}\big[\sum_i \mathbb{1}_{\{|\mathcal{Y}_i| < \varepsilon\}} |\mathcal{Y}_i|\big] = \mathbb{P}(\mathcal{Y}_I < \varepsilon) = 1 - (1-\varepsilon)^\theta \leq \varepsilon,$$

as claimed. □

References

[1] R. Adamczak, M. Kotowski and P. Miłoś, *Phase transition for the interchange and quantum Heisenberg models on the Hamming graph*, preprint arXiv:1808.08902

[2] M. Aizenman and B. Nachtergaele, *Geometric aspects of quantum spin states*, Comm. Math. Phys., 164, 17–63 (1994). MR-1288152

[3] O. Angel, *Random infinite permutations and the cyclic time random walk*, Discrete Math. Theor. Comput. Sci. Proc., 9–16 (2003). MR-2042369

[4] C. Benassi and D. Ueltschi, *Loop correlations in random wire models*, arXiv:1807.06564

[5] V. Betz, J. Ehlert and B. Lees, *Phase transition for loop representations of Quantum spin systems on trees*, J. Math. Phys., 59(11), 113302 (2018). MR-3877248

[6] P. Billingsley, *Probability and Measure*, Third Edition, Wiley (1995). MR-1324786

[7] J. E. Björnberg, J. Fröhlich and D. Ueltschi, *Quantum spins and random loops on the complete graph*, arXiv:1811.12834

[8] J. E. Björnberg and D. Ueltschi, *Critical parameter of random loop model on trees*, Ann. Appl. Probab., 28(4), 2063–2082 (2018). MR-3843823

[9] J. E. Björnberg and D. Ueltschi, *Critical temperature of Heisenberg models on regular trees, via random loops*, J. Statist. Phys. 173, 1369–1385 (2018). MR-3878347

[10] C. Goldschmidt, D. Ueltschi, and P. Windridge, *Quantum Heisenberg models and their probabilistic representations*, Entropy and the Quantum II, Contemp. Math., 552, 177–224 (2011). MR-2868048

[11] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, Third Edition, Oxford (2006). MR-0667520

[12] A. Hammond, *Infinite cycles in the random stirring model on trees*, Bull. Inst. Math. Acad. Sin., 8, 85–104 (2013). MR-3097418

[13] A. Hammond, *Sharp phase transition in the random stirring model on trees*, Probab. Theory Rel. Fields, 161, 429–448 (2015). MR-3334273

[14] A. Hammond and M. Hegde, *Critical point for infinite cycles in a random loop model on trees*, arXiv:1805.11772. MR-3983335

[15] R. Kotecký, P. Miłoś and D. Ueltschi, *The random interchange process on the hypercube*, Electron. Commun. Probab., 21, 9 pp. (2016). MR-3485373

[16] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times*, Second Edition, American Mathematical Society (2017). MR-3726904

[17] P. Miłoś and B. Şengül, *Existence of a phase transition of the interchange process on the Hamming graph*, Electron. J. Probab., 24, 21 pp. (2019). MR-3978214

[18] O. Penrose, *Bose-Einstein condensation in an exactly soluble system of interacting particles*, J. Stat. Phys. 63(3–4), 761–781 (1991). MR-1115812

[19] J. Pitman, *Poisson-Dirichlet and GEM invariant distributions for split-and-merge transformation of an interval partition*, Combin. Probab. Comput., 11(5), 501–514 (2002). MR-1930355

[20] R. T. Powers, *Heisenberg model and a random walk on the permutation group*, Letters in Mathematical Physics 1(2), 125–130 (1976). MR-0418794

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

[22] B. Tóth, *Phase transition in an interacting Bose system. An application of the theory of Ventsel' and Friedlin*, J. Stat. Phys 61(2–4), 749–764 (1990). MR-1086297

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

[24] N. V. Tsilevich, *Stationary random partitions of a natural series*, Teor. Veroyatnost. i Primenen., 44(1), 55–73 (1999). MR-1751188

[25] D. Ueltschi, *Random loop representations for quantum spin systems*, J. Math. Phys., 54, 083301 (2013). MR-3135479

Acknowledgments. The research of JEB is supported by Vetenskapsrådet grant 2015-05195. BL gratefully acknowledges support from the Alexander von Humboldt Foundation. JEB and BL also gratefully acknowledge support from *Stiftelsen Olle Engkvist Byggmästare*. MK acknowledges support from the National Science Centre, Poland, grant no. 2015/18/E/ST1/00214. PM is supported by the National Science Centre, Poland, grant no. 2014/15/B/ST1/02165. We thank Radosław Adamczak for useful discussions about this project.

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`