

# Phase transition for the interchange and quantum Heisenberg models on the Hamming graph

Radosław Adamczak<sup>a</sup>, Michał Kotowski<sup>b,\*</sup> and Piotr Miłoś<sup>b,†</sup>

<sup>a</sup>*Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland. E-mail: r.adamczak@mimuw.edu.pl*

<sup>b</sup>*Institute of Mathematics of the Polish Academy of Sciences, ul. Śniadeckich 8, 00-656 Warsaw, Poland.*

*E-mail: \*michal.kotowski@mimuw.edu.pl; †pmilos@mimuw.edu.pl*

Received 23 April 2019; revised 25 May 2020; accepted 18 June 2020

**Abstract.** We study a family of random permutation models on the Hamming graph  $H(2, n)$  (i.e., the 2-fold Cartesian product of complete graphs), containing the interchange process and the cycle-weighted interchange process with parameter  $\theta > 0$ . This family contains the random walk representation of the quantum Heisenberg ferromagnet. We show that in these models the cycle structure of permutations undergoes a *phase transition* – when the number of transpositions defining the permutation is  $\leq cn^2$ , for small enough  $c > 0$ , all cycles are microscopic, while for more than  $\geq Cn^2$  transpositions, for large enough  $C > 0$ , macroscopic cycles emerge with high probability.

We provide bounds on values  $C, c$  depending on the parameter  $\theta$  of the model, in particular for the interchange process we pinpoint exactly the critical time of the phase transition. Our results imply also the existence of a phase transition in the quantum Heisenberg ferromagnet on  $H(2, n)$ , namely for low enough temperatures spontaneous magnetization occurs, while it is not the case for high temperatures.

At the core of our approach is a novel application of the cyclic random walk, which might be of independent interest. By analyzing explorations of the cyclic random walk, we show that sufficiently long cycles of a random permutation are uniformly spread on the graph, which makes it possible to compare our models to the mean-field case, i.e., the interchange process on the complete graph, extending the approach used earlier by Schramm.

**Résumé.** Nous étudions une famille de modèles de permutations aléatoires sur le graphe de Hamming  $H(2, n)$  (c'est-à-dire le produit cartésien de deux graphes complets), incluant le processus d'échange et le processus d'échange pondéré par les cycles avec paramètre  $\theta > 0$ . Cette famille comprend la représentation par marches aléatoires du modèle de Heisenberg quantique ferromagnétique. Nous montrons que dans ces modèles, la structure des cycles des permutations satisfait une *transition de phase* – lorsque le nombre de transpositions définissant la permutation est  $\leq cn^2$ , pour  $c > 0$  assez petit, tous les cycles sont microscopiques, tandis que lorsque ce nombre est  $\geq Cn^2$  avec  $C > 0$  assez grand, des cycles macroscopiques apparaissent avec grande probabilité.

Nous déterminons des bornes sur les constantes  $C, c$  dépendant du paramètre  $\theta$  du modèle, en particulier, pour le processus d'échange, nous déterminons exactement le temps critique de la transition de phase. Nos résultats impliquent également l'existence d'une transition de phase pour le modèle de Heisenberg quantique ferromagnétique sur  $H(2, n)$ , stipulant que pour des températures assez basses, une magnétisation spontanée apparaît alors que cela n'est pas le cas aux hautes températures.

Le cœur de notre approche consiste en une nouvelle application de la marche aléatoire cyclique, qui pourrait être intéressante en elle-même. En analysant les explorations de la marche aléatoire cyclique, nous montrons que des cycles suffisamment longs d'une permutation aléatoire sont répartis uniformément sur le graphe, ce qui rend possible une comparaison entre nos modèles au cas du champ moyen, c'est-à-dire le processus d'échange sur le graphe complet, étendant ainsi l'approche antérieure utilisée par Schramm.

*MSC2020 subject classifications:* 60K35; 60J27

*Keywords:* Interchange process; Quantum Heisenberg model; Hamming graph

## Contents

1. Introduction . . . . .	274
2. The cyclic random walk – preliminaries . . . . .	279
3. Isoperimetry of the cyclic random walk . . . . .	280

4. General transposition processes . . . . .	303
5. Macroscopic cycles . . . . .	309
Appendix A: Concentration of point processes . . . . .	316
Appendix B: Estimates on the number of bridges . . . . .	318
Appendix C: Proof of Lemma 3.6 . . . . .	320
Appendix D: Proof of Lemma 5.6 . . . . .	322
Acknowledgements . . . . .	324
References . . . . .	324

## 1. Introduction

In this paper we investigate the cycle structure of random permutations in the interchange process (sometimes called the random stirring process) and its generalizations. For a finite graph  $G = (V, E)$  the interchange process  $\sigma = (\sigma_t)_{t \geq 0}$  on  $G$  is defined as follows: put particles numbered from 1 to  $|V|$  on vertices of the graph and Poisson clocks of rate  $1/|E|$  on each edge. Whenever the clock on an edge  $e \in E$  rings, the particles at the endpoints of  $e$  are swapped. In this way for each  $t \geq 0$  we obtain a permutation  $\sigma_t: V \rightarrow V$ , which is determined by the sequence of transpositions corresponding to swaps occurring up to time  $t$ .

The model has attracted considerable attention, in particular one is interested in how the cycle structure of  $\sigma_t$  changes with  $t$ , especially in the asymptotic case where  $G = G_n$  belongs to a family of graphs with  $|V| = n \rightarrow \infty$ . The starting point of our work is a remarkable result due to Schramm [28], which shows that in the case of the complete graph  $G = K_n$  the model exhibits a phase transition. Suppose the interchange process is run for time  $cn$ , then if  $c > 1/2$ , the resulting permutation will, with high probability, contain a macroscopic cycle (i.e., of size comparable to  $n$ ), while for  $c < 1/2$  all cycles will have size  $o(n)$ . Furthermore, for  $c > 1/2$  after proper rescaling the joint distribution of macroscopic cycle sizes converges to the Poisson-Dirichlet distribution with parameter 1 (which is also the limiting distribution of macroscopic cycles for permutations chosen from the uniform measure on  $S_n$ ). This should be contrasted with the classical result of Diaconis and Shahshahani [14] that the mixing time of the random transposition process on  $K_n$  is  $\frac{1}{2}n \log n$ , in particular Schramm's result shows that long cycles equilibrate long before the distribution of the whole permutation.

Our main interest is twofold – first, to move beyond the complete graph and extend these results to graphs with non-trivial geometry, and second, to obtain similar results for a certain generalization of the interchange process, motivated by studies of models in statistical physics.

Namely, we will be interested in the *cycle-weighted interchange process*, depending on an additional parameter  $\theta > 0$ , in which the probability of a sequence of transpositions is weighted depending on the number of cycles in the resulting permutation (a more precise definition will be given shortly). The physical importance of this model is that for  $\theta = 2$  it corresponds to the random walk representation of the quantum Heisenberg ferromagnet.

We will first state our main result informally and just for the case of the interchange process, with more precise statements given afterwards.

**Main result.** *Let  $G$  be the Hamming graph  $H(2, n) = K_n \times K_n$ , where  $K_n$  is the complete graph on  $n$  vertices. Consider the interchange process on  $G$  run up to time  $t = \beta n^2$ . Then the permutation  $\sigma_t$  obtained at time  $t$  exhibits a phase transition: for  $\beta > 1/2$  with high probability  $\sigma_t$  contains macroscopic cycles, while for  $\beta < 1/2$  all cycles of  $\sigma_t$  are with high probability of size  $o(n)$ .*

We will now introduce the more general setup, which will allow us to formulate our results rigorously.

We consider the Hamming graph  $H = H_n = H(2, n) = (V, E)$ . The vertices  $V = \{0, \dots, n-1\}^2$  are given by a subset of the square lattice and an edge is present between a pair of vertices if they are either in the same row or the same column, where for  $i \in \{0, \dots, n-1\}$  the sets  $L_i = \{0, \dots, n-1\} \times \{i\}$  are called rows and  $D_i = \{i\} \times \{0, \dots, n-1\}$  columns. One can check that  $|V| = n^2$  and  $|E| = n^2(n-1)$ . In the whole paper, we assume implicitly that  $n \geq 2$ .

Let  $\mathfrak{X}$  be the space of finite subsets of  $E \times [0, 1)$ , which we will call *configurations*. Given  $X \in \mathfrak{X}$  we denote by  $X \uparrow$  the sequence  $(e_1, \dots, e_{|X|})$ , where  $(e_1, t_1), \dots, (e_{|X|}, t_{|X|})$  are all points of  $X$  ordered with respect to the second coordinate (and an arbitrary fixed order on  $E$  if  $t_i = t_j$ ). We define  $\sigma(X)$ , a permutation associated with the configuration  $X$ , by

$$\sigma(X) := e_{|X|} \circ e_{|X|-1} \circ \dots \circ e_1, \tag{1}$$

where any edge  $e_i \in E$  is identified with the transposition of its endpoints.

We call a function  $\mathcal{C}: \mathfrak{X} \mapsto \mathbb{R}_+$  *admissible* if for any  $X, Y \in \mathfrak{X}$  we have  $\mathcal{C}(X) = \mathcal{C}(Y)$  whenever  $X^\uparrow = Y^\uparrow$  and the following Lipschitz condition holds

$$|\mathcal{C}(X) - \mathcal{C}(Y)| \leq |X \Delta Y|, \quad (2)$$

where  $\Delta$  is the symmetric difference of sets.

Fix  $\beta > 0$ . Let  $\mathcal{B}$  be the law of a Poisson point process on  $E \times [0, 1)$  with intensity measure given by  $\frac{\beta}{n-1} \#(\cdot) \otimes \text{Leb}$ , where  $\#(\cdot)$  is the counting measure and  $\text{Leb}$  is the Lebesgue measure. Here we treat the Poisson process as a random element of  $\mathfrak{X}$  (we endow this space with a  $\sigma$ -field  $\mathcal{S}$ , being a completion with respect to the measure  $\mathcal{B}$  of the  $\sigma$ -field generated by all functions of the form  $X \mapsto X \cap A$ , where  $A \in 2^E \otimes \text{Bor}([0, 1))$ ).

By rescaling time and changing intensity on the edges to  $1/|E|$  one easily sees that if a configuration  $X$  is drawn at random according to  $\mathcal{B}$ , the resulting sequence of transpositions has the same distribution as the interchange process on  $H$  after time  $\beta n^2$ . In particular the intensity is chosen so that  $X$  has size  $|E| \cdot \frac{\beta}{n-1} = \beta n^2$  on average.

Fix  $\theta > 0$  and an admissible function  $\mathcal{C}$ . We define a probability distribution  $\mu_{\beta, \theta, \mathcal{C}}$  on  $\mathfrak{X}$  which will be the main object of our study

$$\mu_{\beta, \theta, \mathcal{C}}(U) := Z_{\beta, \theta, \mathcal{C}}^{-1} \int_{\mathfrak{X}} \mathbb{1}_U(X) \theta^{\mathcal{C}(X)} \mathcal{B}(dX), \quad (3)$$

where  $Z_{\beta, \theta, \mathcal{C}} = \int_{\mathfrak{X}} \theta^{\mathcal{C}(X)} \mathcal{B}(dX)$  is the partition function normalizing the measure to 1. Note that the condition (2) ensures that the measure is well defined. Throughout the paper we set  $\Theta := \max(\theta^{-1}, \theta)$ .

The process defined by  $\mu_{\beta, \theta, \mathcal{C}}$  will be called the *weighted interchange process* in general and we will use the name *cycle-weighted interchange process* if  $\mathcal{C}(X)$  is the number of cycles in  $\sigma(X)$  (which is easily seen to be an admissible function). Note that  $\theta = 1$  corresponds simply to the interchange process.

The main result of our paper states that for  $\beta$  large enough the random permutation induced by the measure  $\mu_{\beta, \theta, \mathcal{C}}$  has macroscopic cycles with high probability

**Theorem 1.1.** *Let  $\beta, \theta > 0$  be such that  $\beta > \Theta/2$  and let  $\mathcal{C}$  be an admissible function. Let  $X$  be randomly sampled from  $\mu_{\beta, \theta, \mathcal{C}}$ . Then*

$$\lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow +\infty} \mathbb{P}(\text{there exists a cycle of } \sigma(X) \text{ of length at least } \varepsilon n^2) = 1.$$

This contrasts with the situation when  $\beta$  is small.

**Theorem 1.2.** *Let  $\beta, \theta > 0$  be such that  $\beta < \Theta^{-1}/2$  and let  $\mathcal{C}$  be an admissible function. Let  $X$  be randomly sampled from  $\mu_{\beta, \theta, \mathcal{C}}$ . Then for some  $C > 0$*

$$\lim_{n \rightarrow +\infty} \mathbb{P}(\text{all cycles of } \sigma(X) \text{ are shorter than } C \log n) = 1.$$

Together our results imply the existence of a phase transition from microscopic cycles when  $\beta < \Theta^{-1}/2$  to macroscopic ones when  $\beta > \Theta/2$ . We expect that the point of the phase transition is unique, possibly under some mild assumptions on  $\mathcal{C}$ .

The two most important cases in which our results apply are  $\theta = 1$  and  $\theta = 2$  (with  $\mathcal{C}$  being the number of cycles).

*Interchange process.* In the special case of the interchange process, corresponding to  $\theta = 1$ , we have  $\Theta^{-1}/2 = \Theta/2 = 1/2$ , so the above theorems determine precisely the transition point for the occurrence of large cycles: for  $\beta < 1/2$  (and large  $n$ ) with high probability all the cycles are of logarithmic size, while for  $\beta > 1/2$  we get cycles of length comparable to the size of the graph with probability arbitrarily close to one.

*Quantum Heisenberg model.* The case  $\theta = 2$  is particularly interesting from the point of view of statistical physics, since it corresponds to the random walk representation of the quantum Heisenberg ferromagnet. In this representation, introduced by Tóth in [30], the existence of macroscopic cycles translates to nonvanishing spontaneous magnetization in the model. We describe the connection very briefly here, referring the reader to the survey [17] for more details on this model.

The quantum Heisenberg ferromagnet is a model of a spin system whose physical properties depend on a parameter  $\beta > 0$  called the *inverse temperature*. One of the major questions about the model is whether a spontaneous ordering of

spins occurs at low enough temperatures (corresponding to high  $\beta$ ). *Tóth's random walk representation* of the Heisenberg model is given by the measure  $\mu_{\beta,\theta,\mathcal{C}}$  with  $\beta$  equal to the inverse temperature,  $\theta = 2$  and  $\mathcal{C}(X)$  being the number of cycles in the permutation  $\sigma(X)$ . One can then express all physical quantities of interest in terms of the cycle-weighted interchange process given by  $\mu_{\beta,\theta,\mathcal{C}}$ . For example, the correlation between spins at sites  $u$  and  $v$  corresponds to the probability that  $u$  and  $v$  are in the same cycle of  $\sigma(X)$  when  $X$  is sampled from  $\mu_{\beta,\theta,\mathcal{C}}$ .

Crucially, our results in Theorem 1.1 and Theorem 1.2 imply the existence of a phase transition for the model on the Hamming graph – for  $\beta < 1/4$  there is no magnetic ordering as  $n \rightarrow \infty$ , while it emerges for  $\beta > 1$ . Our methods do not imply sharpness of the phase transition, although we conjecture that it is indeed sharp, with the critical value being  $\beta = 1$ . It is a major open problem to determine whether a similar phase transition occurs for  $G_n = [-n, n]^d \cap \mathbb{Z}^d$ ,  $d \geq 2$ , as  $n \rightarrow \infty$ .

For a more precise relation between the existence of macroscopic cycles and the phase transition for spontaneous magnetization see Section 4 of [17].

*Outline of proof strategy.* We now outline the main ideas behind the proof of Theorem 1.1 (the proof of Theorem 1.2 is much less involved). First, however, we would like to stress that the novelty of our paper comes not only from the result itself but also from the methods developed for this purpose. Our new techniques establish a precise geometric picture of cycles which is believed to hold for a much broader family of graphs. We also note that the tools developed in this paper proved useful in the analysis of loop models related to the XXZ model on the complete graph [10].

Broadly speaking, we would like to follow the approach used by Schramm for the complete graph, which consists of showing that after a long enough time cycles of mesoscopic size appear and then quickly merge into macroscopic ones. This in turn relies on analyzing a split-merge process of cycles in the complete graph, with each new transposition either causing two cycles to split or to merge. In the case of the complete graph it is easy to give an upper bound on the rate at which cycles split and a lower bound on the rate at which (long enough) cycles merge.

The key difficulty which appears on any graph with non-trivial geometry, in particular in the case of the Hamming graph  $H_n$ , is that, unlike on the complete graph, the split-merge probabilities depend not only on sizes of the cycles, but also on their spatial structure, more precisely on their isoperimetric properties. We are able to prove that long enough fragments of cycles on  $H_n$  are typically “uniformly spread” on the graph, resembling an exploration of a simple random walk and thus making their isoperimetric properties easy to analyze. In particular, the split-merge probabilities (and thus the behavior of the interchange process) can be approximated by the mean-field (complete graph) case.

This intermediate result is at the core of our arguments and we believe it might be of independent interest when analyzing random transposition processes on other graphs. The crucial tool that we employ is the so-called *cyclic random walk* (abbreviated by CRW), introduced in [30] and later used by Angel in [5] under the name cyclic time random walk. This is an exploration process which, given a configuration  $X$  of transpositions and a starting vertex  $v \in V$ , visits subsequent vertices of the cycle of  $\sigma(X)$  containing  $v$ . The fundamental difficulty in the analysis of the CRW is that it is a non-Markovian process, involving interactions of the random walk with vertices visited in the past. The bulk of our effort is devoted to the analysis of these interactions.

We analyze the behavior of the CRW at a mesoscopic timescale which is:

- short enough so that the interactions with the history are tractable and it is possible to exploit methods similar to excursion theory for random walks,
- long enough so that the trace of the CRW occupies the vertices of the graph in a uniform way, and probabilistic bounds we obtain are strong enough to extend results (by union bound) to longer timescales, including macroscopic (i.e., of the order of  $n^2$ ).

A more detailed outline of this part of the proof is given in Section 3.1. Once we know that the trace of the CRW with high probability occupies the graph in a uniform way, we can extend this property to cycles of  $\sigma(X)$  and carry out the analysis of the corresponding split-merge process described above. We note that the assumption  $\beta > \Theta/2$  is crucial in our approach, as it enables us to show that the explorations of the CRW are sufficiently long.

Schramm's approach requires as a prerequisite the existence of mesoscopic cycles. To prove that indeed they exist with high probability, we employ a natural coupling between the random transposition process and a percolation process, with the quality of the coupling on a fixed timescale depending on the isoperimetric bound.

We use this coupling twice to obtain cycles of mesoscopic length, and then employ the argument by Schramm relying on the mean-field behavior of the split-merge probabilities. An abstract version of Schramm's argument that we use is presented in Section 5.1 and we believe that this part of the paper might be of independent interest, as the results are formulated in a way convenient for application to general transposition processes (e.g., the interchange process on more general graphs).

It is worth noting that an additional difficulty is present in both parts of the proof in the case of models with  $\theta \neq 1$ , as subsequent transpositions appear there in a non-i.i.d. fashion. Our methods are based on the observation that on small

timescales having  $\theta \neq 1$  tilts the measure in a controllable way. Roughly speaking, adding or removing a transposition from a configuration changes the number of cycles only by one, which can change the relative probability of the configuration by at most a factor of  $\Theta^2$ . Thus we can compare this process with an i.i.d process, which makes the analysis of the cyclic random walk and emergence of macroscopic cycles still possible. For  $\theta \neq 1$  this approach allows to analyze the size of cycles for small and large  $\beta$ , however it does not give the critical value of the phase transition.

We end this part with a plan of the rest of the paper. The whole Section 3 is devoted to the analysis of the cyclic random walk. The main results of this part are encapsulated in Proposition 3.1 and Proposition 3.3, providing an upper and a lower bound on the typical isoperimetry of the trace of the CRW. In Section 4 we flesh out the connection between the CRW and the cycles of the transposition process. The crucial property that split-merge probabilities are comparable to the mean-field case is stated in Proposition 4.2. In Section 5.1 we provide an abstract formulation of Schramm's argument regarding macroscopic cycles. It is given in Lemma 5.1 and we believe it might of independent interest. We then use it together with results from Section 4 to prove Theorem 1.1. The (much simpler) proof of Theorem 1.2 is given in Section 5.4.

*Related works.* By now the interchange process and its generalizations have attracted considerable attention, both from the point of view of probability theory and mathematical physics. Here we mention some of the work that is most closely related to the topic of this paper.

- The existence of a phase transition for the appearance of macroscopic cycles in the interchange process on the complete graph, together with convergence of the law of macroscopic cycles to the Poisson-Dirichlet distribution, is due to Schramm [28]. An alternative, simpler proof of the statement that large cycles appear after  $cn$  transpositions, for  $c > 1/2$ , was given by Berestycki [6].

To the best of our knowledge, the only rigorous results concerning finite graphs other than the complete graph, all in the case  $\theta = 1$ , are due to Kotecký, Miłoś and Ueltschi [22] and Miłoś and Şengül [25]. In [22] it is proved that on the hypercube  $\{0, 1\}^n$  with  $N = 2^n$  vertices for any  $\varepsilon > 0$  and large enough times a positive fraction of vertices is contained in cycles of length at least  $N^{\frac{1}{2}-\varepsilon}$ .

In [25] it is proved that on the Hamming graph  $H(2, n)$  for  $\beta > 1/2$  and any  $\varepsilon > 0$  asymptotically almost surely a constant fraction of vertices is contained in cycles of length at least  $n^{2-\varepsilon}$ . In this work we build on the approach of [25] and obtain significantly stronger results, proving the existence of truly macroscopic cycles and also considering the case  $\theta \neq 1$ , which, as mentioned above, poses an additional difficulty. The techniques of [25] enabled the authors to analyze only relatively short explorations of the cyclic random walk (on a timescale of the order of  $n$ ) and thus obtain rather crude bounds on the isoperimetry of the cycles. The main improvement in our work is a detailed analysis of the structure of explorations of the cyclic random walk (in particular its interactions with its history) on a long timescale, enabling us to give tight isoperimetric bounds on the cycles' spatial structure.

- Another approach to the analysis of the cycle structure of random permutations, based on representation theory of the symmetric group, was developed in [2] and [7], with the second paper providing yet another proof of the existence of macroscopic cycles in the interchange process on the complete graph. This approach was recently extended in [3] to prove a sharp phase transition in the  $\theta = 2$  case on the complete graph (mean-field Heisenberg ferromagnet). Another recent result by the same authors [4] implies that in the case  $\theta = 1$  macroscopic cycles emerge with positive probability on the Hamming graph for large enough  $\beta$  (note, however, that their result does not identify the critical value of  $\beta$  and concerns only cycles of length at least  $n^2/2$ , instead of  $\varepsilon n^2$  for any  $\varepsilon > 0$ ).
- In [9] Björnberg, also using representation theory, computed the free energy and the critical temperature in a family of quantum spin models on the complete graph corresponding to the cycle-weighted interchange process with  $\theta = 2, 3, 4, \dots$ . This extends previous results obtained by Penrose [27] and Tóth [29] for  $\theta = 2$ .

The existence of large cycles was established for arbitrary  $\theta > 1$  by a different method in [8], where it is proved that macroscopic cycles appear on the complete graph for  $\theta > 1$  as soon as  $\beta > \theta$  (note, however, that apart from the case  $\theta = 2$  this is strictly larger than the critical value  $\beta_c$  determined in [9] for  $\theta = 2, 3, 4, \dots$ ).

- The interchange process can be defined in a natural way also on infinite bounded-degree graphs. Here one asks whether infinite orbits appear almost surely when time exceeds certain critical value. It is conjectured that a phase transition occurs if the underlying infinite graph is transient. The case of a  $d$ -regular infinite tree was first considered by Angel [5], who proved that infinite orbits exist in an appropriate bounded time interval, and then in subsequent work by Hammond [18,19], where an actual phase transition was established for large enough  $d$ . These results were recently extended to more general random loop models (see below).
- Another generalization of the interchange process are the so-called random loop models, corresponding to a family of quantum spin models containing, among others, the quantum Heisenberg antiferromagnet [17]. Recently it has been proved by Hammond and Hegde [20], building upon earlier work by Björnberg and Ueltschi [11], that there exists a phase transition for the appearance of infinite loops on the infinite  $d$ -regular tree for  $d$  large enough. In the case of

finite graphs, it is proved [10] that on the complete graph the distribution of macroscopic loops for  $\beta > 1$  converges to the Poisson-Dirichlet distribution with parameter  $1/2$ .

*Further research and open questions.* There are a number of open questions closely related to our paper. We believe that techniques we have developed here could be useful in approaching some of them.

- We expect that the same techniques as for the Hamming graph  $H(2, n)$  could be used to analyze other families of Hamming graphs  $\{H(d, n)\}_{n \geq 2}$ , for fixed  $d \in \mathbb{N}$ . Here  $H(d, n)$  has vertex set  $\{0, \dots, n-1\}^d$  and an edge is present between any two vertices which differ in exactly one coordinate. In this paper we decided to focus only on the case  $d = 2$ , so as not to obfuscate already long proofs.

On the other hand, it would be of interest to extend our results to Hamming graphs  $H(d, n)$  which satisfy  $d \rightarrow \infty$  (as well as possibly  $n \rightarrow \infty$ ). An extreme example is the hypercube  $\{H(d, 2)\}_{d \geq 2}$ , which is interesting as the degree of each of its vertices diverges as  $d \rightarrow \infty$ , but only slowly (as it is logarithmic in the number of vertices of the graph). We believe that many ideas from our paper should be applicable to this case, although some new insights will also be required, as the geometry of the hypercube is more complicated than that of  $\{H(2, n)\}_{n \geq 2}$ . We expect that understanding the hypercube would essentially enable one to analyze any Hamming graph.

- The results of this paper do not establish the critical value of  $\beta$  at which the phase transition occurs (apart from the case  $\theta = 1$ ). It is conjectured (and partially proved, see results and discussion in [9]) that on the complete graph the critical value is given by

$$\beta_c(\theta) = \begin{cases} \theta & \text{if } 0 < \theta \leq 2, \\ 2\left(\frac{\theta-1}{\theta-2}\right) \log(\theta-1) & \text{if } \theta > 2, \end{cases}$$

which coincides with the critical parameter of the random-cluster model on the complete graph with  $q = \theta$  [12]. It would be an interesting question to explore the possible connection between the two models further and determine the critical value of  $\beta$  for the Hamming graph.

- We conjecture that the properly normalized list of macroscopic cycle lengths obtained in the weighted interchange process with parameter  $\theta$  should converge to the Poisson-Dirichlet distribution  $\text{PD}(\theta)$ . This would extend the convergence to  $\text{PD}(1)$  in the case of  $\theta = 1$  on the complete graph proved in [28].

## Glossary

To help the reader we include the glossary of notation used in the paper.

$H = H_n$	the Hamming graph on $n^2$ vertices	274
$L_i, D_i$	rows and columns of the Hamming graph $H$	274
$\mathfrak{X}$	the space of configurations (finite subsets of $E \times [0, 1)$ )	274
$X^\uparrow$	sequence of edges corresponding to a configuration $X$	274
$\mu_{\beta, \theta, \mathcal{C}}$	measure defining the weighted interchange process with parameters $\beta, \theta, \mathcal{C}$	275
$\mathcal{X}_s$	the cyclic random walk at time $s$	279
$\mathcal{X}_I, \mathcal{Z}_I$	the path and trace of the cyclic random walk on interval $I$	279
$\mathcal{Z}_s, Z_k$	trace of the cyclic random walk (up to times $s$ , resp. first $k$ vertices)	279
$T_k$	time of entering the $k$ -th new vertex by the CRW	280
$\mathcal{O}, \mathcal{O}_k$	the orbit of $v$ under the permutation $\sigma(X)$ and its first $k$ vertices	280
$\mathcal{F}_s, \mathcal{G}_k$	filtrations of the cyclic random walk	280
$\iota, \chi$	isoperimetry upper and lower bound of a given set	280
$T$	$T \leq n \log^2 n$ , timescale on which we study the trace of the CRW	282
$G_t$	graph induced by the cyclic random walk	282
$\Upsilon_t$	the bad set at time $t$	282
$\tau_{\text{iso}}^\delta$	time until which the trace of the cyclic random walk has small $\iota$	283
$\tau_c$	time when the cyclic random walk closes into a cycle	283
$I, I^b, I^d$	processes counting internal, bad and direct jumps	284
$\lambda, \lambda^b, \lambda^d$	intensities of internal, bad and direct jumps	284
$\mathcal{P}_t$	potential of the path $\mathcal{X}_{[0, t)}$	286
$\mathcal{E}_t(k)$	event that the CRW makes an excursion of length $k$ at time $t$	289
$G_t^\circ$	core of the graph $G_t$	291
$\mathcal{D}_t^\delta$	event that the graph $G_t$ does not contain too many high degree vertices	292
$\mathcal{Q}_t$	event that all straight paths of length $\log^2 n$ in $G_t^\circ$ have good potential	295

$\tau_k$	time of the visit to the $k$ -th new vertex in $L_0$	298
$\text{orb}_s(v), \text{orb}_s^\ell(v)$	orbit of vertex $v$ in the permutation $\sigma_s$ and its $\ell$ first elements	304
$\mathcal{I}$	event that that orbits of $\sigma_t$ have good isoperimetric properties	305
$\mathcal{C}_k(\ell)$	set of vertices belonging to components of $\sigma_k$ of size at least $\ell$	309
$\{G_u^s\}_{u=0, \dots,  X -s}$	random graph process coupled to $\sigma_{s+u}$	312
$\mathcal{G}_{s,u}(\ell)$	vertices belonging to components of size at least $\ell$ in $G_u^s$	313

**2. The cyclic random walk – preliminaries**

Let us now introduce the *cyclic random walk* (abbreviated as CRW), which will be the crucial tool in our analysis of permutations arising from the distribution  $\mu_{\beta, \theta, \mathcal{C}}$ .

Recall that  $\mathfrak{X}$  consists of finite subsets of  $E \times [0, 1)$ . For a configuration  $X = \{(e_1, t_1), \dots, (e_k, t_k)\} \in \mathfrak{X}$  the pairs  $(e_i, t_i)$  will be called *bridges*. For a vertex  $v \in V$  the set  $\{v\} \times [0, 1)$  will be called the *bar* of vertex  $v$ . If  $e_i = \{v, w\}$ , we think of a bridge  $(e_i, t_i)$  as joining the bars of a vertex  $v$  and a vertex  $w$  at time  $t_i \in [0, 1)$ .

We note that in the sequel  $X$  will always be sampled from a distribution for which almost surely all  $t_i$  are pairwise different and there are no bridges at prescribed deterministic times, hence there is no ambiguity in how the process is defined.

Consider now a (possibly random) configuration  $X \in \mathfrak{X}$ . The associated *cyclic random walk*  $\mathcal{X} = (\mathcal{X}_s : s \geq 0)$  is a continuous time process with values in  $V \times [0, 1)$ , exploring the bridges given by  $X$ . It starts at a point  $\mathcal{X}_0 = (v, t) \in V \times [0, 1)$ , then moves upwards on the bar of the vertex  $v$  at unit speed, starting at height  $t$ , until it encounters a bridge  $(\{v, w\}, s) \in X$ . Upon encountering a bridge, the CRW jumps to its other end and continues moving on the new bar. Once it gets to height 1, the CRW moves to the bottom of the bar, at height 0. Note that the CRW can encounter an already traversed bridge, in which case we say that it makes a backtrack.

Notice that the CRW is periodic. Once  $\mathcal{X}$  reaches its starting point again (which will happen in a finite time), then it will repeat itself. See Figure 1 for an example of a configuration  $X$  and the trajectory of the corresponding CRW.

The CRW as well as other jump processes we will consider in the paper will be always càdlàg.

For a time interval  $I$  by  $\mathcal{X}_I$  we denote the path of the CRW during  $I$ , likewise  $\mathcal{Z}_I$  is the set of vertices visited by  $\mathcal{X}_I$ , i.e.,

$$\mathcal{Z}_I := \{w \in V : \mathcal{X}_s = (w, z) \text{ for some } s \in I \text{ and } z \in [0, 1)\}.$$

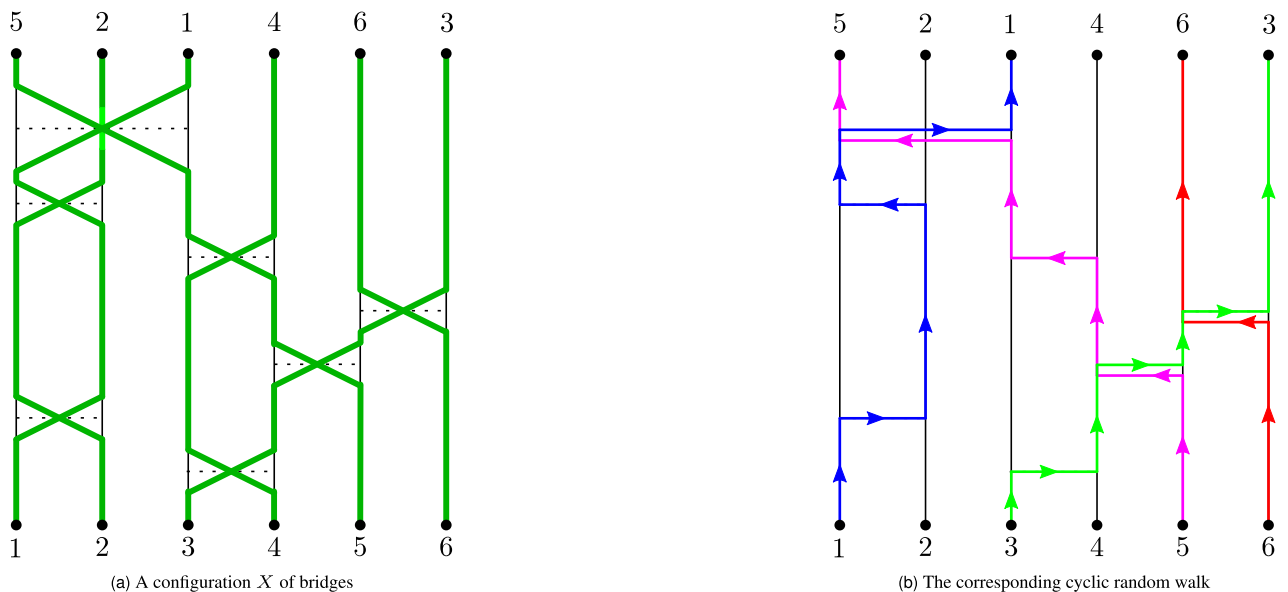


Fig. 1. A configuration of bridges and the corresponding CRW. (left) The dotted lines represent the bridges of  $X$ . The permutation  $\sigma(X)$  is determined in the following way – the labels at the bottom are the labels of the vertices and at the top we have put where they map to under  $\sigma(X)$ . In this example we have  $\sigma(X) = (1365)(2)(4)$ . (right) The path of the CRW using the bridges of  $X$ . The direction in which the CRW travels is indicated by the arrows. Note that some bridges are traversed twice, which corresponds to backtracks.

We will use the abbreviation  $\mathcal{Z}_s := \mathcal{Z}_{[0,s]}$  and simply write  $\mathcal{Z}$  for  $\mathcal{Z}_{[0,\infty)}$ . For  $k \in \mathbb{N}$  we denote by  $T_k := \inf\{s \geq 0 : |\mathcal{Z}_s| \geq k\}$  the time at which the CRW discovers a previously unvisited vertex for the  $k$ -th time (where we use the convention that  $\inf \emptyset = \infty$  and note that  $T_1 = 0$ ). For  $k \in \mathbb{N}$  we set  $Z_k := \mathcal{Z}_{T_k}$ .

The cyclic random walk started at  $(v, 0)$  will be denoted by  $\mathcal{X}(v)$  (and likewise for  $\mathcal{Z}(v)$ ,  $T_k(v)$ ,  $Z_k(v)$  etc.). We will often write simply  $\mathcal{X} = \mathcal{X}(v)$ ,  $\mathcal{Z} = \mathcal{Z}(v)$  etc. if  $v$  is fixed. We will also often abuse notation and write  $\mathcal{X}_s = v$  for  $\mathcal{X}_s = (v, t)$ . Note that since the bars are of height one, the second coordinate  $t$  can be read from the time  $s$ , i.e.,  $t = s \bmod 1$ .

The reason for introducing the cyclic random walk is the following relation between the CRW  $\mathcal{X}$  using the bridges of  $X$  and the permutation  $\sigma(X)$ . Consider the sets

$$\mathcal{O}(v) := \{w \in V : \exists_{t \geq 0} \mathcal{X}_t(v) = (w, 0)\}, \quad (4)$$

$$\mathcal{O}_k(v) := \{w \in V : \exists_{t \in [0,k)} \mathcal{X}_t(v) = (w, 0)\}. \quad (5)$$

It is readily seen that  $\mathcal{O}(v)$  is equal to the orbit of the vertex  $v$  under the permutation  $\sigma(X)$ . Moreover if the orbit has size  $\ell$ , then  $\mathcal{O}_k(v)$  consists of the first  $k \wedge \ell$  elements of the orbit. In other words, vertices visited by  $\mathcal{X}(v)$  at integer times enter the orbit of  $v$  under the permutation  $\sigma(X)$ . Note that we have  $\mathcal{O}(v) \subset \mathcal{Z}(v)$  but not necessarily  $\mathcal{O}_k(v) \subset Z_k(v)$ . For example, the CRW shown in Figure 1, started at 1, visits vertices 1, 2 and 3 up to time 1 (blue path), but  $\mathcal{O}_2(1) = \{1, 3\} \not\subset \{1, 2\} = Z_2(1)$ .

Suppose now that the set of bridges  $X$  is random. Then for  $v \in V$  the cyclic random walk  $\mathcal{X}(v) = (\mathcal{X}_s(v) : s \geq 0)$  is itself a stochastic process. Let  $\mathcal{F} = (\mathcal{F}_s : s \geq 0)$  denote its natural filtration. Set also  $\mathcal{G}_k := \mathcal{F}_{T_k}$ .

In what follows we will be interested in the situation where  $X$  is drawn from the distribution  $\mu_{\beta,\theta,\mathcal{C}}$  defined in (3). Fix  $\beta, \theta > 0$ , an admissible function  $\mathcal{C}$ , and let  $X \in \mathfrak{X}$  be distributed according to  $\mu_{\beta,\theta,\mathcal{C}}$ . For a fixed vertex  $v \in V$ , by  $\mathcal{X}^{\beta,\theta,\mathcal{C}}(v)$  we will denote the corresponding cyclic random walk and call it the cyclic random walk associated to  $\mu_{\beta,\theta,\mathcal{C}}$ , started at  $v$ .

As a final note, we remark that on several occasions we will work with events defined in terms of uncountable intersections over a set of times. Since the processes we are considering almost surely make only countably many jumps, all such events will be in fact measurable.

### 3. Isoperimetry of the cyclic random walk

#### 3.1. The setting and main results

We will now define a notion of isoperimetry for subsets of  $H_n$ . For a set  $A \subset V$  let

$$\iota(A) := \max \left\{ \max_{i \in \{0, \dots, n-1\}} |L_i \cap A|, \max_{i \in \{0, \dots, n-1\}} |D_i \cap A| \right\} \quad (6)$$

and

$$\chi(A) := \min \left\{ \min_{i \in \{0, \dots, n-1\}} |L_i \cap A|, \min_{i \in \{0, \dots, n-1\}} |D_i \cap A| \right\}. \quad (7)$$

Given  $A \subset V$  by  $E(A)$  we will denote the set of edges  $\{v, w\} \in E$  such that  $v, w \in A$ . As each vertex  $v \in A$  has at least  $2(\chi(A) - 1)$  and at most  $2(\iota(A) - 1)$  neighbors in  $A$ , we have the inequalities

$$\chi(A) - 1 \leq \frac{|E(A)|}{|A|} \leq \iota(A) - 1,$$

which justifies the name ‘‘isoperimetry’’. Note also the following subadditivity property of  $\iota$ : for any two subsets  $A, B \subset V$  we have  $\iota(A \cup B) \leq \iota(A) + \iota(B)$ . We also have  $\iota(A) \leq \iota(B)$ ,  $\chi(A) \leq \chi(B)$  whenever  $A \subset B$ , in particular both  $\iota(\mathcal{Z}_t)$  and  $\chi(\mathcal{Z}_t)$  are nondecreasing in  $t$ .

Our main technical result is the following upper bound on  $\iota$  of  $\mathcal{Z}_t$

**Proposition 3.1.** *Fix  $\theta > 0$ , an admissible function  $\mathcal{C}$  and let  $\beta_0, \beta_1$  be such that  $\beta_1 > \beta_0 > \Theta/2$ . Consider  $\beta \in [\beta_0, \beta_1]$  and let  $\mathcal{Z}(v) := \mathcal{Z}^{\beta,\theta,\mathcal{C}}(v)$  be the trace of the cyclic random walk associated to  $\mu_{\beta,\theta,\mathcal{C}}$ , started at  $v \in V$ . Then there exist  $C, c > 0$  (depending only on  $\beta_0, \beta_1, \theta$ , in particular independent of  $\mathcal{C}$ ) such that*

$$\mathbb{P}(\forall v \in V \iota(\mathcal{Z}_{n \log^2 n}(v)) \leq C \log^2 n) \geq 1 - Ce^{-c \log^2 n}.$$



The above proposition will be a key tool in the proof of existence of long cycles. Together with a corresponding (easier) lower bound on the quantity  $\chi$ , given in Proposition 3.3 below, it gives good control of the isoperimetric properties of *mesoscopic* segments of cycles (i.e., of length roughly  $n \log^2 n$ ), which can be then lifted to isoperimetry of full cycles of length at least  $n \log^2 n$ . This will enable us to prove that with high probability the conditional split and merge probabilities in the evolution of cycles behave similarly as in the mean-field case, allowing for an application of a modification of an argument due to Schramm.

The parameters  $\beta_0, \beta_1$  are introduced for technical reasons. When applying the above proposition in the proof of Theorem 1.1 we will need uniformity of constants for  $\beta$  belonging to the interval  $[\beta_0, \beta_1]$ , with appropriately chosen  $\beta_0, \beta_1$ .

The crucial part of the proof of Proposition 3.1 is a bootstrap argument regarding isoperimetry. Informally speaking, we will show that if at some time  $T \leq n \log^2 n$  we have with high probability “good” isoperimetry (of the order of  $n^\alpha$  for some small enough  $\alpha$ ), then actually we have with high probability “very good” isoperimetry (of the order of  $\log^2 n$ ). This is formalized in the following

**Lemma 3.2 (Bootstrap).** *Let  $\theta, \mathcal{C}, \beta_0, \beta_1, \beta, \mathcal{Z}(v)$  be as in Proposition 3.1. Let  $\alpha \in (0, 1/100)$  and  $T \leq n \log^2 n$ . Suppose that for some  $C_1, c_1 > 0$  we have*

$$\mathbb{P}(\forall v \in V \iota(\mathcal{Z}_T(v)) \leq n^\alpha) \geq 1 - C_1 e^{-c_1 \log^2 n}.$$

*Then there exist  $C, C_2, c_2 > 0$  (depending only on  $\theta, \beta_0, \beta_1, \alpha, c_1, C_1$ ) such that we have*

$$\mathbb{P}(\forall v \in V \iota(\mathcal{Z}_T(v)) \leq C \log^2 n) \geq 1 - C_2 e^{-c_2 \log^2 n}.$$

With this lemma the proof of Proposition 3.1 is rather straightforward and is given in Section 3.6. The whole next section will be devoted to the proof of Lemma 3.2.

Later on we will also need a lower bound on  $\chi$ , which is analogous to Proposition 3.1 provided the orbit of the permutation defined by the CRW contains at least  $n \log^2 n$  vertices. Recall the definition (4) of the sets  $\mathcal{O}_k(v)$  and  $\mathcal{O}(v)$ .

**Proposition 3.3.** *Fix  $\theta > 0$ , an admissible function  $\mathcal{C}$  and let  $\beta_0, \beta_1 > 0$ . Consider  $\beta \in [\beta_0, \beta_1]$  and let  $\mathcal{X}(v) := \mathcal{X}^{\beta, \theta, \mathcal{C}}(v)$  be the cyclic random walk associated to  $\mu_{\beta, \theta, \mathcal{C}}$ , started at  $v$ . There exist  $C, c > 0$  (depending only on  $\beta_0, \beta_1, \theta$ , in particular independent of  $\mathcal{C}$ ) such that*

$$\mathbb{P}(\forall v \in V \chi(\mathcal{O}_{n \log^2 n}(v)) \geq c \log^2 n \text{ or } |\mathcal{O}(v)| < n \log^2 n) \geq 1 - C e^{-c \log^2 n}.$$

As the proof of the above proposition is much less involved than for Proposition 3.1, it is given separately in Section 3.7.

*Outline of the proof of Lemma 3.2.* Here we outline the proof strategy of the main technical result of this section – Lemma 3.2. Most of our effort is devoted to the analysis of interactions of the CRW with its history. When the CRW enters a previously visited vertex it may reuse already explored bridges, which can generate a complex behavior depending on the graph  $G_t$  of already visited vertices.

One should keep in mind the following intuitive picture. By the assumption  $\iota(\mathcal{Z}_T) \leq n^\alpha \ll n$  the trace of the CRW is not too concentrated in any row or column. Thus, while moving on the bar of a vertex  $v$ , if the CRW discovers an unexplored bridge, it will typically jump to a yet unexplored vertex. If it fails to discover a new bridge, it backtracks to the vertex visited before  $v$ . Such backtracks are common and may cascade creating some interactions of the CRW with its history. Due to the assumption  $\beta > \Theta/2$ , the rate of discovery of new vertices is fast enough so that the CRW tends to escape its history, making the above mentioned interactions short-ranged and thus fairly easy to analyze. For a very similar reason, unless the CRW closes into a cycle quickly, it makes a fairly long cycle. When  $T$  is at least of the order of  $n$ , the CRW occasionally jumps to a vertex visited a long time before. Analysis of such long-range interactions is the main technical difficulty of the proof.

The main task is to show that the time between two subsequent visits in any fixed row or column, say  $L_0$ , is  $cn$  (for some  $c > 0$ ) with uniformly positive probability. Having done that, by a comparison with a sum of independent random variables it is straightforward to conclude that  $|\mathcal{Z}_{n \log^2 n} \cap L_0| \leq C \log^2 n$  with very high probability and thus also  $\iota(\mathcal{Z}_{n \log^2 n}) \leq C \log^2 n$ .

The CRW can hit  $L_0$  either by a direct jump using a previously unexplored bridge or by entering through its history. As  $T \leq n \log^2 n$  and at each step the CRW has chance roughly  $c/n$  of a direct jump, typically it will make  $c \log^2 n$  visits of the first type.

To analyze entering  $L_0$  through the history we distinguish two cases. The first is when the CRW jumps using a new bridge to a vertex which is close in  $G_t$  to  $L_0$ . The second is when the vertex is far from  $L_0$  and the CRW makes a long backtrack employing already used edges.

To rule out the first possibility we show, using a rather delicate argument, that the dangerous zone (“bad set”) consisting of small balls around  $L_0$  in  $G_t$  is small enough so that the CRW is quite unlikely to jump to it. Thus it is very unlikely to observe  $\log^2 n$  of such jumps.

In the second case it is enough to show that any sufficiently long path in  $G_t$  has what we call “large potential”. Intuitively speaking, the CRW traversing such a path has many chances to escape it by jumping to a new vertex and performing a long excursion avoiding its history. Thus the CRW is unlikely to ever make a long backtrack. This part of the argument is rather technical and again uses crucially the assumption  $\beta > \Theta/2$ .

The roadmap to the proof is as follows. In the rest of this section and Section 3.2 we set up the framework for analyzing the excursions made by the CRW, in particular providing bounds for the intensity of discovering new vertices, number of jumps to the history (Lemma 3.12) and the rate at which the “potential” corresponding to bars is exhausted (Lemma 3.13). This is then used in Section 3.3 to show that the CRW is likely to make short excursions not intersecting its history (Lemma 3.15). In Section 3.4 we show that the bad set described above is typically small (Proposition 3.20) and thus is unlikely to be hit by the CRW (Lemma 3.21). In Section 3.5 we show that long paths typically have large potential (Lemma 3.23) and thus are unlikely to be backtracked (Proposition 3.24). All these pieces are then used in Section 3.6 to show that visits to  $L_0$  are infrequent (Lemma 3.25), which easily implies (Corollary 3.26) good isoperimetry claimed in the conclusion of Lemma 3.2.

*Basic notation and assumptions.* Our goal for the rest of this section is to prove Lemma 3.2. Therefore, from now on we fix

- $\theta > 0$  and an admissible function  $\mathcal{C}$ ,
- $\beta_0, \beta_1$  such that  $\beta_0 > \Theta/2$ , and  $\beta \in [\beta_0, \beta_1]$ ,
- $\alpha \in (0, 1/100)$  and  $T \leq n \log^2 n$ ,
- $\varepsilon \in (0, \frac{1}{20})$  such that  $\alpha < \varepsilon/4$

(the parameter  $\varepsilon$  will play a technical role in intermediate calculations).

We will be considering the cyclic random walk  $\mathcal{X}^{\beta, \theta, \mathcal{C}}(v)$  associated to  $\mu_{\beta, \theta, \mathcal{C}}$ , started at  $(v, 0)$  for a fixed vertex  $v$ . From now on for brevity we write simply  $\mathcal{X}_t = \mathcal{X}_t^{\beta, \theta, \mathcal{C}}(v)$ ,  $\mathcal{Z}_t = \mathcal{Z}_t^{\beta, \theta, \mathcal{C}}(v)$  etc.

In what follows whenever we write about global constants like  $C, c$  we allow them to depend on  $\beta_0, \beta_1, \varepsilon$  (in addition to dependence on  $\theta$  and  $\alpha$ ), so that for  $\beta \in [\beta_0, \beta_1]$  all the statements hold with constants depending uniformly on  $\beta$ . The dependence on  $\theta, \alpha, \beta_0, \beta_1, \varepsilon$  will be suppressed in the notation. Occasionally, if a constant depends additionally on some other parameter  $\delta$ , we will stress it by writing e.g.,  $c = c(\delta)$ . All constants will be independent of the function  $\mathcal{C}$  as long as the Lipschitz condition (2) is satisfied.

In proofs we will often take  $n$  large enough, depending on the other parameters, but unless stated otherwise the propositions being proved will be such that this can be absorbed into constants  $C, c$  appearing in the statements, so the propositions in fact hold for all  $n$ .

We will now introduce several notions which will be useful in analyzing the explorations of the cyclic random walk (see Figure 2 for an illustration of these definitions).

*The graph  $G_t$ .* The vertices  $\mathcal{Z}_t$  and bridges explored up to time  $t$  by the CRW induce a graph denoted by  $G_t$  (we allow multiple edges if there is more than one bridge between two vertices). Let  $d_{G_t}(\cdot, \cdot)$  be its natural graph metric. By  $B_{G_t}(v, r)$  we denote the ball of radius  $r$  in this metric around the vertex  $v$  in  $G_t$ .

*The bad set.* Recall that  $\varepsilon$  is a parameter fixed at the beginning of this section. Let

$$\Upsilon_t := \bigcup_{v \in \mathcal{Z}_t \cap L_0} B_{G_t}(v, n^\varepsilon) \tag{8}$$

be the *bad set* at time  $t$ . Note that  $\Upsilon_t$  is nondecreasing in  $t$ .

The reason for calling this set ‘bad’ is that once the CRW ends up in  $\Upsilon_t$ , it might quickly backtrack its way to  $L_0$  and then make multiple jumps inside the same row, thus ruining good isoperimetric properties of the trajectory.

*Dead vertices.* As the CRW explores the set of vertices, it might happen that at time  $t$  the bar  $\{v\} \times [0, 1)$  corresponding to a vertex  $v$  is completely exhausted, i.e.,  $\{v\} \times [0, 1) \subset \mathcal{X}_{[0, t]}$ . We will call such a vertex  $v$  *dead*.

*Stopping times.* Later on it will be convenient to use the following stopping times related to the CRW. For  $\delta > 0$  we define  $\tau_{\text{iso}}^\delta$  as the time when the CRW loses good isoperimetric properties, i.e.

$$\tau_{\text{iso}}^\delta := \inf\{t \geq 0: \iota(\mathcal{Z}_t) > n^\delta\},$$

and let  $\tau_c$  be the time that the CRW closes into a cycle, i.e.

$$\tau_c := \inf\{t > 0: \mathcal{X}_t = (v, 0)\},$$

where  $v$  is the starting vertex, meaning  $\mathcal{X}_0 = (v, 0)$ .

*Jumps.* It will be important to distinguish several types of jumps that the CRW can make. We will call a jump *fresh* when a previously unexplored bridge is used, otherwise we call it a *backtrack*. Suppose that  $\mathcal{X}$  makes a fresh jump at time  $t$ . We will call it an *internal jump* if  $\mathcal{X}_t \in \mathcal{Z}_{t-}$ , a *bad jump* if  $\mathcal{X}_t \in \Upsilon_{t-}$  and a *direct jump to  $L_0$*  if  $\mathcal{X}_{t-} \notin L_0$  and  $\mathcal{X}_t \in L_0$ .

### 3.2. Intensities of jumps and the potential

Recall that a counting process is a nondecreasing, integer valued càdlàg stochastic process starting at zero and with jumps equal to one.

*Intensity.* Let  $Y$  be an  $\mathcal{F}_t$ -adapted counting process. We will say that a nonnegative process  $\lambda$  is an *intensity* of  $Y$  if  $\lambda$  is  $\mathcal{F}_t$ -progressively measurable,  $\int_0^t \lambda_u du < \infty$  for all  $t$ , and the process  $Y_t - \int_0^t \lambda_s ds$  is an  $\mathcal{F}_t$ -martingale.

In what follows we will often need (conditional) concentration inequalities for counting processes. Two of them, used most frequently, are stated below for convenience and the other two, which will be used only once, are stated in Appendix A. Proofs of all of them can be found in Appendix A as well.

**Lemma 3.4.** *Let  $Y_t$  be a counting process with bounded intensity  $\lambda$  and compensator  $\Lambda_t = \int_0^t \lambda_s ds$ . Assume that  $\sigma, \tau$  are bounded stopping times such that  $\sigma \leq \tau$ . Consider  $\ell > 0$  and let  $X$  be a Poisson variable with parameter  $\ell$ . Then for any  $r \geq 0$  we have almost surely*

$$\mathbb{P}(\{Y_\tau - Y_\sigma \geq r\} \cap \{\Lambda_\tau - \Lambda_\sigma \leq \ell\} | \mathcal{F}_\sigma) \leq \mathbb{P}(X \geq r).$$

If  $r \geq \ell$ , we have in particular

$$\mathbb{P}(\{Y_\tau - Y_\sigma \geq r\} \cap \{\Lambda_\tau - \Lambda_\sigma \leq \ell\} | \mathcal{F}_\sigma) \leq \exp\left(-r \log\left(\frac{r}{e\ell}\right) - \ell\right).$$

We will also need a corresponding lower bound.

**Lemma 3.5.** *Let  $Y_t$  be as in the previous lemma. Let  $\sigma, \tau$  be bounded stopping times such that  $\sigma \leq \tau$  and let  $\delta, \ell > 0$ . Let also  $X$  be a Poisson random variable with parameter  $\ell$ . Then with probability one,*

$$\mathbb{P}(\{Y_\tau - Y_\sigma \leq \ell(1 - \delta)\} \cap \{\Lambda_\tau - \Lambda_\sigma \geq \ell\} | \mathcal{F}_\sigma) \leq \mathbb{P}(X \leq (1 - \delta)\ell) \leq \exp(-\delta^2 \ell / 2),$$

where  $\Lambda_t = \int_0^t \lambda_s ds$ .

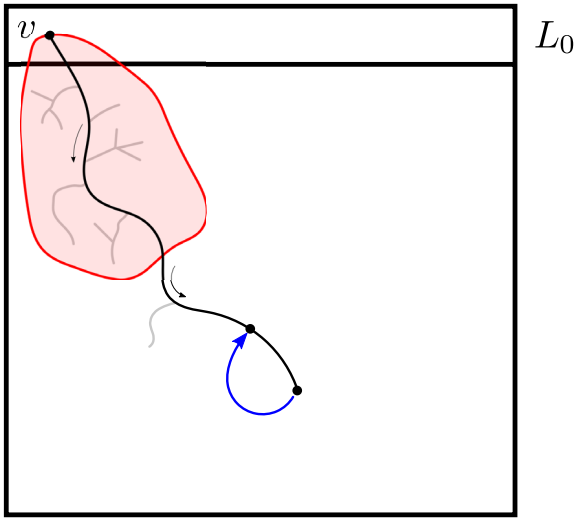
*Set of accessible vertices.* By  $A_t$  we will denote the set of vertices which at time  $t$  are available to the CRW by a fresh jump. Formally,  $A_t = \emptyset$  if the CRW has closed into a cycle before time  $t$ , otherwise let  $(w, z) = \mathcal{X}_t$  and

$$A_t := \{v \in V: \{v, w\} \in E \text{ and } (v, z) \notin \mathcal{X}_{[0, t]}\}.$$

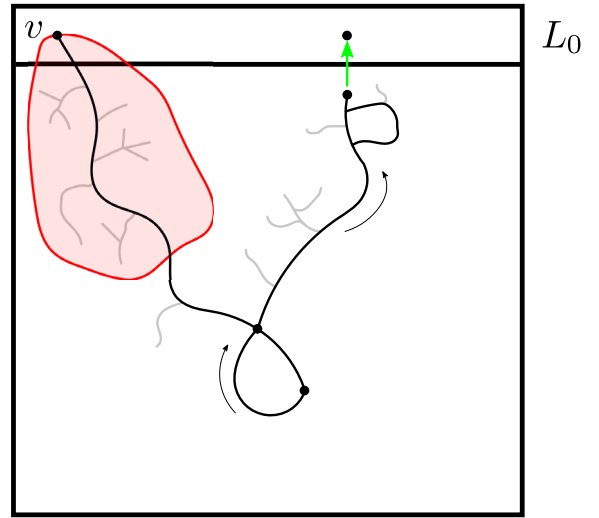
**Lemma 3.6 (Intensity of jumps).** *Let  $Q_t$  be an  $\mathcal{F}_t$ -adapted càdlàg process of subsets of  $V$  such that  $Q_t$  can jump only at times when  $\mathcal{X}_t$  jumps. Let  $J_t = \{|s \leq t: \mathcal{X} \text{ makes a fresh jump at time } s \text{ and } \mathcal{X}_s \in Q_{s-}\}|$ . Then the counting process  $J$  has intensity  $\lambda$  which satisfies*

$$\lambda_t \Theta^{-1} \leq \frac{\beta}{n-1} |A_t \cap Q_t| \leq \lambda_t \Theta.$$

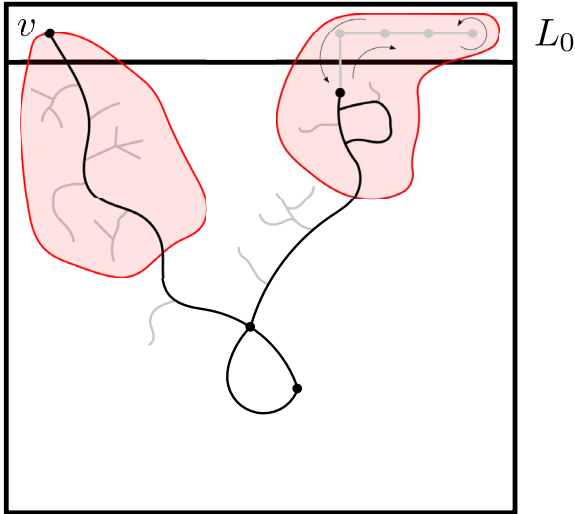
In particular for  $\theta = 1$  we have  $\lambda_t = \frac{\beta}{n-1} |A_t \cap Q_t|$ .



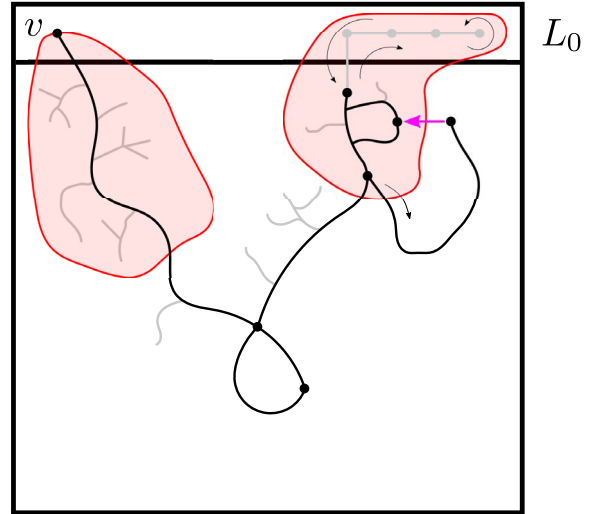
(a) The CRW starts at  $v$  and makes an exploration leaving the bad set. The exploration ends with an *internal jump* (blue edge).



(b) The CRW continues the exploration outside the bad set, making a *direct jump* to  $L_0$  at the end (green edge).



(c) After a jump to  $L_0$  some of the previously visited vertices become a part of the bad set. The CRW then makes an excursion inside  $L_0$  which is completely backtracked (grey edges), ending up outside  $L_0$ .



(d) The CRW makes an excursion leaving the bad set, ending in a *bad jump* (purple edge).

Fig. 2. Sketch of the evolution of the graph  $G_t$ , with the CRW started at  $v$ . Dead vertices are shown in grey. The bad set  $\Upsilon_t$  is marked in red.

The proof of this lemma is presented in Appendix C. We stress that the lemma is not specific to the Hamming graph and its statement holds for any weighted transposition process on a finite graph (with  $\frac{\beta}{n-1}$  replaced by the appropriate edge intensity of the underlying point process).

Let  $I_t$  (resp.  $I_t^b, I_t^d$ ) denote the total number of internal (resp. bad, direct to  $L_0$ ) jumps up to time  $t$ . They are counting processes. We will denote their intensities by  $\lambda$  (resp.  $\lambda^b, \lambda^d$ ). The intensity of the process  $|\mathcal{Z}_t|$  will be denoted by  $\mu$ .

**Corollary 3.7.** For  $t < \tau_c$ ,

$$\mu_t \geq 2\Theta^{-1}\beta \frac{n - \iota(\mathcal{Z}_t)}{n - 1} \tag{9}$$

and

$$\lambda_t \leq \frac{4\Theta\beta}{n} \iota(\mathcal{Z}_t). \tag{10}$$

**Proof.** Let us start with (9). For any  $t \geq 0$  let the current vertex  $w = \mathcal{X}_t$  belong to the column  $D$  and row  $L$ . Using Lemma 3.6 with  $Q_t = V \setminus \mathcal{Z}_t$ , together with the definition of the parameter  $\iota$  and the obvious containment  $(D \cup L) \setminus \mathcal{Z}_t \subset A_t$ , we have

$$\mu_t \geq \Theta^{-1} \frac{\beta}{n-1} |(D \cup L) \setminus \mathcal{Z}_t| = \Theta^{-1} \frac{\beta}{n-1} (|D \setminus \mathcal{Z}_t| + |L \setminus \mathcal{Z}_t|) \geq \Theta^{-1} \frac{2\beta}{n-1} (n - \iota(\mathcal{Z}_t)),$$

as desired. For the proof of (10), we write again by the definition of  $\iota$  and Lemma 3.6 applied with  $Q_t = \mathcal{Z}_t$

$$\lambda_t \leq \Theta \frac{\beta}{n-1} |A_t \cap \mathcal{Z}_t| \leq \Theta \frac{\beta}{n-1} (|D \cap \mathcal{Z}_t| + |L \cap \mathcal{Z}_t|) \leq \Theta \frac{2\beta}{n-1} \iota(\mathcal{Z}_t) \leq \Theta \frac{4\beta}{n} \iota(\mathcal{Z}_t). \quad \square$$

It will be useful to have an estimate of how quickly previously unexplored vertices are discovered by the CRW. We have the following upper bound

**Lemma 3.8 (New vertices are not very frequent).** *There exists an increasing function  $g : (1, \infty) \rightarrow (0, \infty)$  such that for any  $\delta > 2\Theta\beta$ ,  $k > 0$  and  $l \in \mathbb{N}$  we have*

$$\mathbb{P}(|\mathcal{Z}_{T_l+k} \setminus \mathcal{Z}_{T_l}| \leq \lceil \delta k \rceil | \mathcal{G}_l) = \mathbb{P}(T_{l+\lceil \delta k \rceil} \geq T_l + k | \mathcal{G}_l) \geq 1 - e^{-ck},$$

where  $c = 2\Theta\beta g(\frac{\delta}{2\Theta\beta})$ .

**Proof.** By Lemma 3.6 the intensity  $\mu_t$  of  $|\mathcal{Z}_t|$  satisfies

$$\mu_t \leq \Theta \frac{\beta}{n-1} |A_t \setminus \mathcal{Z}_t| \leq \Theta \frac{\beta}{n-1} |A_t| \leq 2\Theta\beta.$$

By Lemma 3.4 (applied with  $Y_t = |\mathcal{Z}_t|$ ,  $\sigma = T_l$ ,  $\tau = T_l + k$ ,  $\lambda = \mu$ ,  $\ell = 2\Theta\beta k$ ,  $r = \lceil \delta k \rceil \geq \ell$ ) we obtain (recall that  $\Delta_t = \int_0^t \lambda_s ds$ ),

$$\begin{aligned} \mathbb{P}(|\mathcal{Z}_{T_l+k} \setminus \mathcal{Z}_{T_l}| > \lceil \delta k \rceil | \mathcal{G}_l) &= \mathbb{P}(\{|\mathcal{Z}_{T_l+k} \setminus \mathcal{Z}_{T_l}| > \lceil \delta k \rceil\} \cap \{\Delta_{T_l+k} - \Delta_{T_l} \leq 2\Theta\beta k\} | \mathcal{G}_l) \\ &\leq \exp\left(-\delta k \log\left(\frac{\delta}{2e\Theta\beta}\right) - 2\Theta\beta k\right). \end{aligned}$$

Using again the inequality  $\delta > 2\Theta\beta$ , it is easy to see that the right hand side is bounded from above by  $e^{-ck}$ , with  $c = 2\Theta\beta g(\frac{\delta}{2\Theta\beta})$  for  $g(x) = x \log x + 1 - x$  as desired.  $\square$

In the remaining part of this section we will work under the following assumption which will not be explicitly stated in the hypotheses of theorems (we recall that we consider the CRW started at  $\mathcal{X}_0 = (v, 0)$  for some fixed vertex  $v$ ).

**Assumption 3.9 (Main assumption).** There exist  $C, c > 0$  such that

$$\mathbb{P}(\iota(\mathcal{Z}_T) > n^\alpha) \leq C e^{-c \log^2 n}.$$

We will often need the following lemma, which is a consequence of Markov's inequality.

**Lemma 3.10.** *For any event  $A$ , any  $\sigma$ -field  $\mathcal{G}$ , and any  $C, r \geq 0$ , if  $\mathbb{P}(A) \geq 1 - C e^{-r}$ , then with probability at least  $1 - C e^{-r/2}$ ,*

$$\mathbb{P}(A | \mathcal{G}) \geq 1 - e^{-r/2}.$$

**Corollary 3.11.** *There exist  $C, c > 0$  such that for any stopping time  $\eta < T$  we have*

$$\mathbb{P}(\iota(\mathcal{Z}_T) \leq n^\alpha | \mathcal{F}_\eta) \geq 1 - e^{-c \log^2 n},$$

with probability at least  $1 - C e^{-c \log^2 n}$ .

In the next lemma we show that  $I_T$ , the total number of internal jumps up to time  $T$ , is small with high probability.

**Lemma 3.12 (Internal jumps are rare).** *There exist  $C, c > 0$  such that*

$$\mathbb{P}\left(I_T \geq \frac{1}{2}n^{2\alpha}\right) \leq Ce^{-c\log^2 n}.$$

**Proof.** We have

$$\mathbb{P}\left(I_T \geq \frac{1}{2}n^{2\alpha}\right) \leq \mathbb{P}\left(I_{T \wedge \tau_{\text{iso}}^\alpha} \geq \frac{1}{2}n^{2\alpha}\right) + \mathbb{P}(\tau_{\text{iso}}^\alpha \leq T).$$

By Assumption 3.9 the second term is small enough. Now we bound, for  $t < \tau_{\text{iso}}^\alpha$ , the intensity  $\lambda_t$  of  $I_t$  by Lemma 3.6

$$\lambda_t \leq \frac{4\Theta\beta}{n} \iota(\mathcal{Z}_t) \leq \frac{4\Theta\beta}{n} n^\alpha = 4\Theta\beta n^{\alpha-1}.$$

Since  $T \leq n\log^2 n$  and  $\frac{1}{2}n^{2\alpha} > 4\Theta\beta n^{\alpha-1}T$  for  $n$  large enough, we can apply Lemma 3.4 to the stopped process  $Y_t = I_{t \wedge \tau_{\text{iso}}^\alpha}$ , obtaining for  $n$  large enough the bound

$$\mathbb{P}\left(I_{T \wedge \tau_{\text{iso}}^\alpha} \geq \frac{1}{2}n^{2\alpha}\right) \leq \exp\left(-\frac{1}{2}n^{2\alpha} \log\left(\frac{\frac{1}{2}n^{2\alpha}}{4\Theta\beta en^{\alpha-1}T}\right)\right) \leq \exp\left(-\frac{1}{2}n^{2\alpha} \log\left(\frac{n^\alpha}{8\Theta\beta e \log^2 n}\right)\right),$$

and the right hand side is at most  $e^{-c\log^2 n}$  for some  $c > 0$ . □

*The potential.* While visiting a vertex the CRW does not necessarily exhaust its whole bar. For  $t > 0$  by  $\mathcal{P}_t$  we denote the Lebesgue measure of unused parts of visited bars and call it the *potential*. Formally,

$$\mathcal{P}_t := |\mathcal{Z}_t| - U_t,$$

where the number of visited vertices  $|\mathcal{Z}_t|$  is equal to the total measure of visited bars and  $U_t$  is the measure of their used parts, which is equal to the Lebesgue measure of the path  $\mathcal{X}_{[0,t]}$ .

Notice that until time  $\tau_c$  the potential  $\mathcal{P}_t$  increases by 1 each time the CRW visits a previously unexplored vertex and otherwise decreases linearly with  $t$ . This means that for  $t \leq \tau_c$  the potential  $\mathcal{P}_t$  follows the equation (with  $\mathcal{P}_0 = 1$ )

$$\mathcal{P}_t = |\mathcal{Z}_t| - t. \tag{11}$$

In the following technical lemma we show that the potential of a path  $\mathcal{X}_{[t,t+s]}$  is typically proportional to  $s$  and cannot drop significantly before time  $T$ . Furthermore, with probability bounded away from 0 it stays strictly positive. These properties will be useful in the forthcoming analysis of excursions and backtracks of the CRW.

**Lemma 3.13 (Controlling the change in potential).** *Fix  $a < \Theta^{-1}\beta_0 - 1/2$ . There exist positive constants  $C = C(a)$ ,  $c = c(a)$ ,  $q$  such that for  $n$  large enough, any stopping time  $\eta$  and any  $s \geq 0$  the following hold with probability at least  $1 - Ce^{-c\log^2 n}$ :*

$$\mathbb{1}_{\{T \geq \eta + s\}} \cdot \mathbb{P}(\{\mathcal{P}_{\eta+s} - \mathcal{P}_\eta \leq as\} \cap \{\tau_c \geq \eta + s\} | \mathcal{F}_\eta) \leq C(e^{-cs} + e^{-c\log^2 n}), \tag{12}$$

$$\mathbb{P}(\forall u \leq T - \eta \mathcal{P}_{\eta+u} - \mathcal{P}_\eta + 1 > 0 | \mathcal{F}_\eta) \geq q, \tag{13}$$

$$\mathbb{P}(\{\exists s \leq u \leq T - \eta \mathcal{P}_{\eta+u} - \mathcal{P}_\eta + 1 \leq as\} \cap \{\tau_c \geq \eta + s\} | \mathcal{F}_\eta) \leq C(e^{-cs} + e^{-c\log^2 n}). \tag{14}$$

**Proof.** Let  $\sigma = \inf\{t > 0: \iota(\mathcal{Z}_t) > n^\alpha\} = \tau_{\text{iso}}^\alpha$ . Set  $\tau = \tau_c \wedge \sigma$ . Let

$$\mathcal{E} = \{\mathcal{P}_{\eta+s} - \mathcal{P}_\eta \leq as\} \cap \{\tau_c \geq \eta + s\} \cap \{T \geq \eta + s\}.$$

By Corollary 3.11 and (11), with probability at least  $1 - Ce^{-c\log^2 n}$ ,

$$\begin{aligned} \mathbb{P}(\mathcal{E} | \mathcal{F}_\eta) &\leq \mathbb{P}(\mathcal{E} \cap \{\sigma > \eta + s\} | \mathcal{F}_\eta) + Ce^{-c\log^2 n} \\ &= \mathbb{P}(\{|\mathcal{Z}_{\eta+s}| - |\mathcal{Z}_\eta| \leq (a+1)s\} \cap \{\tau \geq \eta + s\} | \mathcal{F}_\eta) + Ce^{-c\log^2 n}. \end{aligned}$$

Recall that  $\mu_t$  denotes the intensity of  $|\mathcal{Z}_t|$ . By (9) of Corollary 3.7 and the definition of  $\sigma$ , for  $t < \tau$  we have

$$\mu_t \geq \frac{2\Theta^{-1}\beta}{n-1}(n - \iota(\mathcal{Z}_t)) \geq \frac{2\Theta^{-1}\beta}{n-1}(n - n^\alpha) \geq 1 + 2a =: \beta' > 1, \quad (15)$$

for  $n$  large enough. Thus, setting  $\Lambda_t = \int_0^t \mu_u du$ , we obtain

$$\mathbb{P}(\{|\mathcal{Z}_{\eta+s}| - |\mathcal{Z}_\eta| \leq (a+1)s\} \cap \{\tau \geq \eta + s\} | \mathcal{F}_\eta) \leq \mathbb{P}(\{|\mathcal{Z}_{\eta+s}| - |\mathcal{Z}_\eta| \leq (a+1)s\} \cap \{\Lambda_{\eta+s} - \Lambda_\eta \geq \beta's\} | \mathcal{F}_\eta).$$

Recalling that  $a = \frac{1}{2}(\beta' - 1) > 0$  and using Lemma 3.5 with  $\ell = \beta's$ ,  $\delta = \frac{\beta'-1-a}{\beta'}$  we obtain that the right-hand side above is almost surely bounded by

$$e^{-s\beta'\frac{1}{2}\left(\frac{\beta'-1}{2\beta'}\right)^2} = e^{-cs}.$$

for some  $c > 0$ . This proves (12).

Let us now prove (13). Denoting the event there by  $\mathcal{E}$  and noticing that once the CRW closes into a cycle the potential stays constant, we estimate

$$\begin{aligned} \mathbb{P}(\mathcal{E} | \mathcal{F}_\eta) &= \mathbb{P}(\{\forall u \in [0, T \wedge \tau_c - \eta] \mathcal{P}_{\eta+u} - \mathcal{P}_\eta + 1 \geq 0\} | \mathcal{F}_\eta) \\ &\geq \mathbb{P}(\{\forall u \in [0, T \wedge \tau_c - \eta] |\mathcal{Z}_{\eta+u}| - |\mathcal{Z}_\eta| \geq u - 1\} \cap \{\sigma > T\} | \mathcal{F}_\eta) \\ &\geq \mathbb{P}(\{\forall u \in [0, \tau - \eta] |\mathcal{Z}_{\eta+u}| - |\mathcal{Z}_\eta| \geq u - 1\} \cap \{\sigma > T\} | \mathcal{F}_\eta) \\ &\geq \mathbb{P}(\{\forall u \in [0, \tau - \eta] |\mathcal{Z}_{\eta+u}| - |\mathcal{Z}_\eta| \geq u - 1\} | \mathcal{F}_\eta) - \mathbb{P}(\sigma \leq T | \mathcal{F}_\eta). \end{aligned}$$

By Corollary 3.11 the second term is with probability at least  $1 - Ce^{-c \log^2 n}$  bounded by  $Ce^{-c \log^2 n}$  and thus negligible. As for the first term, note that thanks to (15) the intensity  $\mu_t$  is bounded away from 1 for  $n$  large enough and  $t \in [\eta, \tau]$ , so Lemma A.2 guarantees that the increase of the associated counting process, i.e.,  $|\mathcal{Z}_{\eta+u}|$ , is always at least  $u - 1$ , with probability bounded from below by some  $q > 0$  (observe also that Lemma A.2 does not require any of the involved stopping times to dominate the other one). This concludes the proof of (13).

Now we pass to the proof of (14). Again denoting the event there by  $\mathcal{E}$  and noticing that once the CRW closes into a cycle the potential stays constant, we estimate

$$\begin{aligned} \mathbb{P}(\mathcal{E} | \mathcal{F}_\eta) &\leq \mathbb{P}(\{\exists s \leq u \leq T \wedge \tau_c - \eta \mathcal{P}_{\eta+u} - \mathcal{P}_\eta \leq as - 1\} \cap \{\sigma > T\} | \mathcal{F}_\eta) + \mathbb{P}(\sigma \leq T | \mathcal{F}_\eta) \\ &\leq \mathbb{P}(\exists u \in [s, \tau - \eta] |\mathcal{Z}_{\eta+u}| - |\mathcal{Z}_\eta| \leq u + as | \mathcal{F}_\eta) + Ce^{-c \log^2 n}, \end{aligned}$$

where the second inequality holds with probability at least  $1 - Ce^{-c \log^2 n}$  (for some  $C, c > 0$ ), by Corollary 3.11 applied to the second term. Similarly as in the proof of (13), the intensity of  $|\mathcal{Z}_{\eta+u}|$  is bounded away from 1, so by Lemma A.3 (applied with  $\beta'$  instead of  $\beta$ ), the first term is almost surely bounded by  $e^{-cs}$  for some  $c > 0$  depending only on  $\beta'$ .  $\square$

### 3.3. Excursions

In this section we introduce lemmas which in the final proof will help us show that the CRW with probability bounded away from zero may leave  $L_0$  (or more generally the bad set) and move far away from it (in the metric of the graph  $G_t$ ), thus making a quick return difficult.

The first lemma is of technical nature and asserts that there is a non-negligible probability that the CRW will make a move within its current column. The second lemma will be crucial in proving that the CRW with high probability will not backtrack to  $L_0$ .

**Lemma 3.14 (Jumps within columns are quite likely).** *Let  $k \geq 0$  and let  $D(v)$  denote the column containing vertex  $v$ . Let  $X_k := \mathcal{X}_{T_k}$ . There exist  $C, p > 0$  such that*

$$\mathbb{P}(\{X_{k+1} \in D(X_k)\} \cap \{T_{k+1} < T_k + 1\} | \mathcal{G}_k) \geq p \mathbb{1}_{\{T_k \leq T\}} \quad (16)$$

with probability at least  $1 - Ce^{-c \log^2 n}$  for some  $C, c > 0$ .

**Proof.** Let  $D_t$  (resp.  $L_t$ ) denote the number of jumps of the CRW to a previously unexplored vertex in the same column (resp. row). Then up to time  $\tau_c$  by Lemma 3.6 their intensities (denote them by  $\delta_t, \nu_t$  resp.) satisfy  $\delta_t \Theta^{-1} \leq \frac{\beta}{n-1} |A_t \cap D(\mathcal{X}_t)| \leq \delta_t \Theta$  and  $\nu_t \Theta^{-1} \leq \frac{\beta}{n-1} |A_t \cap L(\mathcal{X}_t)| \leq \nu_t \Theta$ . In particular, on the event  $\{t < \tau_{\text{iso}}^\alpha \wedge \tau_c\}$  we have

$$m := \Theta^{-1} \frac{\beta}{n-1} (n - n^\alpha) \leq \delta_t, \nu_t. \tag{17}$$

We also have trivially

$$\delta_t, \nu_t \leq \Theta \beta. \tag{18}$$

Consider the event  $A = \{\tau_{\text{iso}}^\alpha > T_k\}$  and two stopping times  $\rho = \inf\{t > T_k : D_t > D_{T_k}\}$  and  $\gamma = \inf\{t > T_k : L_t > L_{T_k}\}$ . Define also  $\mathcal{E} = \{\rho < \gamma \wedge (T_k + 1)\}$ . The lemma will follow once we prove that almost surely

$$\mathbb{P}(\mathcal{E} | \mathcal{G}_k) \geq p \mathbb{1}_A. \tag{19}$$

Indeed, the event  $\{\mathbb{1}_A < \mathbb{1}_{\{T_k \leq T\}}\}$  is contained in  $\{\tau_{\text{iso}}^\alpha \leq T\}$  which by Assumption 3.9 has probability at most  $C e^{-c \log^2 n}$ .

Let  $B$  be any element of  $\mathcal{G}_k$  and denote  $P := \mathbb{P}(\mathcal{E} \cap A \cap B)$ . Note that on  $A$  we have  $T_k < \infty$ . Observe also that with probability one  $\rho \neq \gamma \wedge (T_k + 1)$  and so

$$P = \mathbb{E}(D_{\rho \wedge \gamma \wedge (T_k + 1)} - D_{T_k}) \mathbb{1}_{A \cap B} = \mathbb{E} \int_{T_k}^{\rho \wedge \gamma \wedge (T_k + 1)} \delta_s \mathbb{1}_{A \cap B} ds,$$

where we used Doob's theorem and the fact that  $A, B \in \mathcal{G}_k$ .

Since between  $T_k$  and  $\rho \wedge \gamma \wedge (T_k + 1)$  the quantity  $\iota(\mathcal{Z}_s)$  does not change and the CRW does not close into a cycle, we can use (17) to estimate

$$\begin{aligned} P &\geq \mathbb{E}(\rho \wedge \gamma \wedge (T_k + 1) - T_k) m \mathbb{1}_{A \cap B} \\ &= m \mathbb{E}(\rho - T_k) \mathbb{1}_{A \cap B \cap \mathcal{E}} + m \mathbb{E}(\gamma \wedge (T_k + 1) - T_k) \mathbb{1}_{A \cap B} \\ &\quad - m \mathbb{E}(\gamma \wedge (T_k + 1) - T_k) \mathbb{1}_{A \cap B \cap \mathcal{E}} \\ &= m \mathbb{E}(\gamma \wedge (T_k + 1) - T_k) \mathbb{1}_{A \cap B} - m \mathbb{E}(\gamma \wedge (T_k + 1) - \rho) \mathbb{1}_{A \cap B \cap \mathcal{E}}. \end{aligned} \tag{20}$$

Note that

$$\begin{aligned} \mathbb{E}(\gamma \wedge (T_k + 1) - \rho) \mathbb{1}_{A \cap B \cap \mathcal{E}} &\leq \frac{1}{m} \mathbb{E} \int_{\rho}^{\gamma \wedge (T_k + 1)} \nu_s ds \mathbb{1}_{A \cap B \cap \mathcal{E}} \\ &= \frac{1}{m} \mathbb{E}(L_{\gamma \wedge (T_k + 1)} - L_\rho) \mathbb{1}_{A \cap B \cap \mathcal{E}} = \frac{P}{m}, \end{aligned}$$

where in the first equality we used Doob's theorem (note that  $A \cap B \cap \mathcal{E} \in \mathcal{F}_\rho$ ) and in the second one the observation that on  $\mathcal{E}$  we have  $L_{\gamma \wedge (T_k + 1)} - L_\rho \leq 1$ .

Combining the above inequality with (20), we get

$$P \geq \frac{m}{2} \mathbb{E}(\gamma \wedge (T_k + 1) - T_k) \mathbb{1}_{A \cap B}. \tag{21}$$

Integrating by parts we get

$$\begin{aligned} P &\geq \frac{m}{2} \int_0^1 \mathbb{E} \mathbb{1}_{A \cap B} \mathbb{P}(\gamma - T_k > s | \mathcal{G}_k) ds \\ &= \frac{m}{2} \int_0^1 \mathbb{E} \mathbb{1}_{A \cap B} \mathbb{P}(\{L_{T_k+s} - L_{T_k} = 0\} | \mathcal{G}_k) ds. \end{aligned}$$

Using (18) and the concentration estimate from Lemma 3.4 with  $r = 0$ , we get that

$$\mathbb{P}(\{L_{T_k+s} - L_{T_k} = 0\} | \mathcal{G}_k) \geq \mathbb{P}(X = 0),$$



where  $X$  is a Poisson variable with parameter  $\Theta\beta s$ . Thus

$$P \geq \mathbb{P}(A \cap B) \frac{m}{2} \int_0^1 e^{-\Theta\beta s} ds.$$

Setting  $p = \frac{m}{2} \int_0^1 e^{-\Theta\beta s} ds$  and using the definition of  $P$  together with the fact that  $A \in \mathcal{G}_k$ , we get

$$\mathbb{E}(\mathbb{P}(\mathcal{E}|\mathcal{G}_k)\mathbb{1}_A\mathbb{1}_B) = \mathbb{P}(\mathcal{E} \cap A \cap B) \geq \mathbb{E}p\mathbb{1}_A\mathbb{1}_B$$

for all  $B \in \mathcal{G}_k$ , which implies (19) and concludes the proof.  $\square$

*Excursions.* Let  $t$  be a stopping time. We will say that the CRW makes an *excursion* of length  $k$  starting at time  $t$ , the event which we denote by  $\mathcal{E}_t(k)$ , if there exists  $s > 0$  such that the following conditions hold:

- $\mathcal{Z}_{[t, t+s]} \cap (\mathcal{Z}_{t-} \cup L_0) = \emptyset$ ,
- $\mathcal{X}_{t+s}$  is at distance  $k$  from  $\mathcal{X}_t$  in  $G_{t+s}$  or  $t+s = T$ .

Note that in particular if  $t+s < T$  then the CRW has to discover at least  $k$  previously unexplored vertices and it is possible for  $\mathcal{X}_{[t, t+s]}$  to intersect itself. The condition  $t+s = T$  is included as we are interested in the CRW only up to time  $T$ .

**Lemma 3.15 (Excursions are quite likely).** *There exists  $C, c, q > 0$  such that for any  $l \geq 1$  we have*

$$\mathbb{P}(\mathcal{E}_{T_l}(n^\varepsilon)|\mathcal{G}_l) \geq q \cdot \mathbb{1}_{\{\mathcal{X}_{T_l} \notin L_0\} \cap \{T_l \leq T\}}, \quad (22)$$

with probability at least  $1 - Ce^{-c \log^2 n}$ .

**Proof.** Throughout the proof we assume that  $\{\mathcal{X}_{T_l} \notin L_0\} \cap \{T_l \leq T\}$  holds.

Let  $\tau = \inf\{u \in (T_l, \infty) : \mathcal{X}_u \text{ makes an internal jump}\}$  and  $\sigma = \inf\{u \in (T_l, \infty) : \mathcal{X}_u \text{ makes a direct jump to } L_0\}$ . For  $k, s \in \mathbb{N}$  let  $s' = s \wedge (T - T_l)$  and

$$\tilde{\mathcal{E}}(k, s) := \{\tau > T_l + s\} \cap \{\forall_{u \leq s'} \mathcal{P}_{T_l+u} - \mathcal{P}_{T_l-} > 0\} \cap (\{\mathcal{P}_{T_l+s} - \mathcal{P}_{T_l} \geq k\} \cup \{T_l + s > T\}) \cap \{\sigma > T_l + s\}. \quad (23)$$

The lemma will follow once we prove  $\tilde{\mathcal{E}}(k, s) \subset \mathcal{E}_{T_l}(k)$  and that with high probability  $\mathbb{P}(\tilde{\mathcal{E}}(n^\varepsilon, s)|\mathcal{G}_l) \geq q$  for some  $s > 0$  and some constant  $q > 0$ .

The first two conditions of (23) imply that the subgraph  $G_{s'}^l$  of  $G_{T_l+s'}$  induced by the exploration  $\mathcal{X}_{[T_l, T_l+s']}$  is a tree and the CRW does not revisit  $\mathcal{Z}_{T_l-}$ , i.e.,  $\mathcal{Z}_{[T_l, T_l+s']} \cap \mathcal{Z}_{T_l-} = \emptyset$ . Indeed, as there are no internal jumps the CRW can revisit  $\mathcal{Z}_{T_l-}$  only by backtracking the bridge used at time  $T_l$ . This happens only when all vertices  $\mathcal{Z}_{[T_l, T_l+u]}$  are dead at some time  $u \leq s'$  (we again use the fact that there are no internal jumps). This is equivalent to  $\mathcal{P}_{T_l+u} - \mathcal{P}_{T_l-} = 0$  which is impossible. Once we know that  $\mathcal{Z}_{[T_l, T_l+s']} \cap \mathcal{Z}_{T_l-} = \emptyset$  and there are no internal jumps during  $[T_l, T_l + s']$ , the only possibility left is that the exploration is a tree. As a corollary we observe that the first two conditions of (23) imply that the CRW does not close into a cycle, i.e.,  $\tau_c \geq T_l + s'$ .

If  $T_l + s > T$ , then the second condition from the definition of an excursion is trivially satisfied. Assume therefore that  $T_l + s \leq T$ , so that  $s' = s$  in the argument above. As  $G_s^l$  is a tree, one sees that the distance  $d_{G_s^l}$  of the vertex  $\mathcal{X}_{T_l+s}$  from  $\mathcal{X}_{T_l}$  is  $f - b$ , where  $f$  (resp.  $b$ ) is the number of fresh jumps (resp. backtracks) during time  $(T_l, T_l + s]$ .

Notice that  $f = |\mathcal{Z}_{(T_l, T_l+s]}| - 1 = |\mathcal{Z}_{T_l+s}| - |\mathcal{Z}_{T_l}|$ , as there are no internal jumps.

Moreover, we have  $b \leq s$ , since a backtrack occurs only once a whole bar has been exhausted (i.e., the corresponding vertex has become dead) and the CRW moves at unit speed. Altogether this implies

$$d_{G_s^l}(\mathcal{X}_{T_l+s}, \mathcal{X}_{T_l}) \geq |\mathcal{Z}_{T_l+s}| - |\mathcal{Z}_{T_l}| - s = \mathcal{P}_{T_l+s} - \mathcal{P}_{T_l}.$$

To conclude we notice that  $d_{G_{T_l+s}}(\mathcal{X}_{T_l+s}, \mathcal{X}_{T_l}) = d_{G_s^l}(\mathcal{X}_{T_l+s}, \mathcal{X}_{T_l})$  and by the third condition of (23) the assumption  $T_l + s \leq T$  implies that the right hand side above is at least  $k$ .

The final condition of (23) together with  $\{\mathcal{X}_{T_l} \notin L_0\}$  ensures that  $\mathcal{Z}_{[T_l, T_l+s]} \cap L_0 = \emptyset$ .

These arguments proved that  $\tilde{\mathcal{E}}(k, s) \subset \mathcal{E}_{T_l}(k)$ , with the excursion taking total time  $s'$ . Now we are left with showing that with probability at least  $1 - Ce^{-c \log^2 n}$  we have  $\mathbb{P}(\tilde{\mathcal{E}}(n^\varepsilon, s)|\mathcal{G}_l) \geq q$  for some  $s > 0$  and  $q > 0$ . For this part we fix

$k = n^\varepsilon$  and  $s = k/a$ , for some  $a \in (0, \Theta^{-1}\beta - 1/2)$ . We recall that the first two conditions of (23) imply  $\tau_c \geq T_l + s'$ . Thus we have

$$\begin{aligned} \mathbb{P}(\tilde{\mathcal{E}}(k, s)|\mathcal{G}_l) &\geq \mathbb{P}(\{\tau > T_l + s\} \cap \{\forall_{u \leq s'} \mathcal{P}_{T_l+u} - \mathcal{P}_{T_l-} > 0\} \cap \{\sigma > T_l + s\}|\mathcal{G}_l) \\ &\quad - \mathbb{P}(\{\mathcal{P}_{T_l+s} - \mathcal{P}_{T_l} < k\} \cap \{T_l + s \leq T\} \cap \{\tau_c \geq T_l + s'\}|\mathcal{G}_l) \\ &\geq \mathbb{P}(\forall_{u \leq s'} \mathcal{P}_{T_l+u} - \mathcal{P}_{T_l-} > 0|\mathcal{G}_l) - \mathbb{P}(\tau \leq T_l + s|\mathcal{G}_l) - \mathbb{P}(\sigma \leq T_l + s|\mathcal{G}_l) \\ &\quad - \mathbb{P}(\{\mathcal{P}_{T_l+s} - \mathcal{P}_{T_l} < k\} \cap \{T_l + s \leq T\} \cap \{\tau_c \geq T_l + s\}|\mathcal{G}_l), \end{aligned}$$

where on the event  $\{T_l + s \leq T\}$  we replaced  $s'$  with  $s$ .

By Lemma 3.13 with probability at least  $1 - Ce^{-c \log^2 n}$  (for some  $C, c > 0$ ) we have  $\mathbb{P}(\forall_{u \leq s'} \mathcal{P}_{T_l+u} - \mathcal{P}_{T_l-} > 0|\mathcal{G}_l) \geq q$  for some  $q > 0$  (notice that  $\mathcal{P}_{T_l-} + 1 = \mathcal{P}_{T_l}$ ). Thus to conclude the proof it is enough to show that the other terms are  $o(1)$ . Fix  $\delta > 2\Theta\beta$ . We have

$$\mathbb{P}(\tau \leq T_l + s|\mathcal{G}_l) \leq \mathbb{P}(\{\tau \leq T_l + s\} \cap \{T_{l+\lceil \delta s \rceil} \geq T_l + s\}|\mathcal{G}_l) + \mathbb{P}(T_{l+\lceil \delta s \rceil} < T_l + s|\mathcal{G}_l).$$

The second term is  $o(1)$  by Lemma 3.8. For the first term we write

$$\mathbb{P}(\{\tau \leq T_l + s\} \cap \{T_{l+\lceil \delta s \rceil} > T_l + s\}|\mathcal{G}_l) = \mathbb{P}(\{N_s \geq 1\} \cap \{T_{l+\lceil \delta s \rceil} > T_l + s\}|\mathcal{G}_l),$$

where  $N_u = I_{T_l+u} - I_{T_l}$  is the number of internal jumps during time  $[T_l, T_l + u]$ . Recall that by Corollary 3.11 the event  $\{\iota(\mathcal{Z}_{T_l}) \leq n^\varepsilon\}$  has high conditional probability. On this event for  $u \leq s \wedge (T_{l+\lceil \delta s \rceil} - T_l)$  we bound the intensity  $\lambda_u$  of  $N_u$  by (10) from Lemma 3.6

$$\lambda_u \leq \frac{4\Theta\beta}{n} \iota(\mathcal{Z}_{T_l+u}) \leq \frac{4\Theta\beta}{n} (\iota(\mathcal{Z}_{T_l}) + \lceil \delta s \rceil) \leq \frac{4\Theta\beta}{n} (n^\varepsilon + \delta n^\varepsilon/a + 1) =: \bar{\lambda}.$$

Now we have  $\bar{\lambda} \cdot s \leq \bar{\lambda} \cdot n^\varepsilon/a = Cn^{2\varepsilon-1}$ , for some  $C > 0$ . Thus by Markov's inequality, Doob's theorem and monotonicity of  $N$ ,

$$\begin{aligned} \mathbb{P}(\{\tau \leq T_l + s\} \cap \{T_{l+\lceil \delta s \rceil} > T_l + s\}|\mathcal{G}_l) \mathbb{1}_{\{\iota(\mathcal{Z}_{T_l}) \leq n^\varepsilon\}} &\leq \mathbb{E}(N_{s \wedge (T_{l+\lceil \delta s \rceil} - T_l)} \mathbb{1}_{\{\iota(\mathcal{Z}_{T_l}) \leq n^\varepsilon\}}|\mathcal{G}_l) \\ &\leq \mathbb{E}(\bar{\lambda} s \mathbb{1}_{\{\iota(\mathcal{Z}_{T_l}) \leq n^\varepsilon\}}|\mathcal{G}_l) \leq Cn^{2\varepsilon-1} = o(1), \end{aligned}$$

and consequently with probability at least  $1 - Ce^{-c \log^2 n}$  (for some  $C, c > 0$ )

$$\mathbb{P}(\tau \leq T_l + s|\mathcal{G}_l) = o(1).$$

Analogously one can show that

$$\mathbb{P}(\sigma \leq T_l + s|\mathcal{G}_l) = o(1).$$

Indeed, by Lemma 3.6 we can easily estimate the intensity of direct jumps by  $\lambda_u^d \leq \bar{\lambda}$ , as  $|A_u \cap L_0| \leq 1$ .

Finally, by the choice of  $s$  above and (12) in Lemma 3.13 we have

$$\mathbb{P}(\{\mathcal{P}_{T_l+s} - \mathcal{P}_{T_l} < k\} \cap \{T_l + s \leq T\} \cap \{\tau_c \geq T_l + s'\}|\mathcal{G}_l) = o(1), \quad (24)$$

thus the proof is finished.  $\square$

As in the forthcoming proofs we will need to apply Lemma 3.14 and Lemma 3.15 for random  $k$ , we state the following easy corollary

**Corollary 3.16.** *Let  $\tau$  be a stopping time such that with probability one  $\tau \in \{T_i : i \leq n^2\}$  and let  $\tau'$  be such that  $\tau' = T_{i+1}$  on the event  $\{\tau = T_i\}$ . Then with probability at least  $1 - Ce^{-c \log^2 n}$ ,*

$$\mathbb{P}(\{\mathcal{X}_{\tau'} \in D(\mathcal{X}_\tau)\} \cap \{\tau' < \tau + 1\}|\mathcal{F}_\tau) \geq p \mathbb{1}_{\{\tau \leq T\}} \quad (25)$$

and

$$\mathbb{P}(\mathcal{E}_\tau(n^\varepsilon)|\mathcal{F}_\tau) \geq q \cdot \mathbb{1}_{\{\mathcal{X}_\tau \notin L_0\} \cap \{\tau \leq T\}}. \quad (26)$$

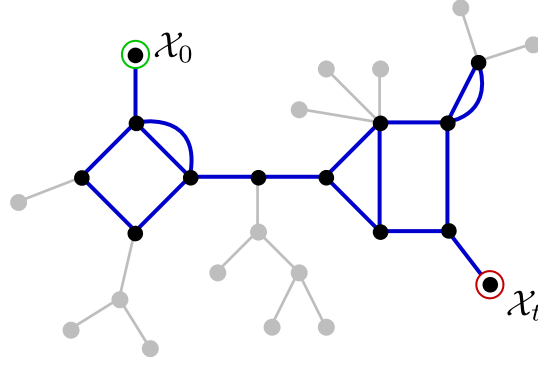


Fig. 3. Graph  $G_t$  and its core  $G_t^{\circ}$  (edges shown as thicker blue lines). Parts of the graph removed in the construction of the core are shown in grey. The starting and end vertices of the CRW,  $\mathcal{X}_0$  and  $\mathcal{X}_t$ , have been marked in green and red, respectively.

**Proof.** It is enough to note that

$$\mathbb{P}(\{\mathcal{X}_{\tau'} \in D(\mathcal{X}_{\tau})\} \cap \{\tau' < \tau + 1\} | \mathcal{F}_{\tau}) = \sum_{k=1}^{n^2} \mathbb{P}(\{X_{k+1} \in D(X_k)\} \cap \{T_{k+1} < T_k + 1\} | \mathcal{G}_k) \mathbb{1}_{\{\tau=T_k\}},$$

so if (25) does not hold then (16) fails for some  $k \leq n^2$ . Now the statement follows from Lemma 3.14 and a union bound over  $k$ .

Analogously, we can write

$$\mathbb{P}(\mathcal{E}_{\tau}(n^{\varepsilon}) | \mathcal{F}_{\tau}) = \sum_{l=1}^{n^2} \mathbb{P}(\mathcal{E}_{\tau}(n^{\varepsilon}) | \mathcal{F}_{\tau}) \mathbb{1}_{\{\tau=T_l\}} = \sum_{l=1}^{n^2} \mathbb{P}(\mathcal{E}_{T_l}(n^{\varepsilon}) | \mathcal{G}_l) \mathbb{1}_{\{\tau=T_l\}},$$

and (26) fails only if (22) fails for some  $l \leq n^2$ . Again we finish by Lemma 3.15 and a union bound.  $\square$

### 3.4. The bad set

Recall that the bad set  $\Upsilon_t$ , defined in (8), consists of vertices which are close in  $G_t$  to  $L_0$ . Our goal is to show that the CRW is unlikely to ever hit the bad set. Since it is difficult to know exactly which part of  $\Upsilon_t$  is accessible to the CRW at time  $t$ , we will in fact prove a stronger statement, namely that with high probability  $\Upsilon_t$  itself is small for all  $t \leq T$ .

*The core of  $G_t$ .* We now introduce a special subgraph of  $G_t$ , which will play an important role in the analysis of bad jumps and backtracks.

For a graph  $G$  and a vertex  $v \in V_G$  denote by  $\deg_G(v)$  the degree of  $v$  in  $G$ , counted with multiplicities. Below, to simplify the notation we will often write  $v \in G$  instead of  $v \in V_G$ .

Let  $G_t^{\circ}$  be the subgraph of  $G_t$  obtained by successively removing dead vertices of degree one (i.e., we remove dead vertices of degree one in  $G_t$ , obtaining the graph  $G_t^{(1)}$ , next we remove dead vertices of degree one in  $G_t^{(1)}$ , etc. until no more vertices can be removed). We will call the graph  $G_t^{\circ}$  the *core* of  $G_t$  (this is similar to what is called the 2-core of  $G_t$  in graph theory, except that we allow possibly two vertices of degree one which are not dead). See Figure 3 for an example of a graph  $G_t$  and its core.

The procedure described above corresponds to removing trees consisting of dead vertices, connected to the core. Note that  $G_t^{\circ}$  can still contain dead vertices and it is not necessarily nondecreasing in  $t$ . The role the graph  $G_t^{\circ}$  will play in subsequent arguments is twofold. In the analysis of the size of the bad set, it will be convenient to handle the intersection of the bad set with the core and the trimmed trees separately. In the subsequent part we will also use the special structure of the core to show that if the CRW is outside the bad set then it is very unlikely to backtrack all the way to  $L_0$ .

The degrees of vertices in  $G_t^{\circ}$  are in a simple relation with the number of internal jumps  $I_t$ .

**Lemma 3.17.** For any  $t \geq 0$ ,

$$\sum_{v \in G_t^{\circ}} (\deg_{G_t^{\circ}}(v) - 2) - 2I_t = -2.$$

**Proof.** We will first prove that for any  $t \geq 0$ ,

$$\sum_{v \in G_t} (\deg_{G_t}(v) - 2) - 2I_t = -2.$$

The expression on the left hand side equals  $-2$  at time  $t = 0$  and changes only when the CRW makes a fresh jump. The increase of the degrees caused by such a jump is compensated either by the summand  $-2$  (if the CRW explores a new vertex) or by the increase of  $I_t$  (in the case of internal jumps). This proves the above formula.

To pass from  $G_t$  to  $G_t^\circ$  note that whenever one removes a vertex of degree one then the sum of degrees decreases by 2 and the number of vertices decreases by one, so the sum in question does not change.  $\square$

For  $\delta > 0$  let us define

$$\mathcal{D}_t^\delta := \left\{ \sum_{v \in G_t^\circ} (\deg_{G_t^\circ}(v) - 2) \leq n^\delta \right\}.$$

Combining the above lemma with Lemma 3.12 we obtain immediately

**Lemma 3.18 (The core has few excess edges).** *There exist  $C, c > 0$  such that*

$$\mathbb{P}(\forall t \leq T \mathcal{D}_t^{2\alpha}) \geq 1 - Ce^{-c \log^2 n}.$$

Moreover, the intersection of the core and the bad set is small, as asserted in the following lemma

**Lemma 3.19.** *There exist  $C, c > 0$  such that*

$$\mathbb{P}\left(\sup_{t \leq T} |V_{G_t^\circ} \cap \Upsilon_t| \geq n^{4\epsilon}\right) \leq Ce^{-c \log^2 n}.$$

**Proof.** We will first estimate the size of the  $k$ -neighborhood of any vertex in  $G_t^\circ$  for arbitrary  $k \geq 1$ .

To this end, fix any vertex  $r \in G_t^\circ$  and consider a spanning tree  $\mathcal{T}$  of  $G_t^\circ$  obtained by a breadth first search starting from  $r$ , so that the distances between  $r$  and any other vertex of  $G_t^\circ$  are the same in  $G_t^\circ$  and in  $\mathcal{T}$ .

Let  $B_i$  denote the ball  $B_{\mathcal{T}}(r, i)$  (we assume  $B_{-1} = \emptyset$ ) and let  $\deg_{\mathcal{T}}(w)$  be the degree of  $w$  in  $\mathcal{T}$ . Obviously we have  $\deg_{\mathcal{T}}(w) \leq \deg_{G_t^\circ}(w)$ . We have  $|B_0| = 1$ ,  $|B_1| - |B_0| = \deg_{\mathcal{T}}(r)$  and for any  $i \geq 1$ ,

$$|B_{i+1}| - |B_i| = \sum_{w \in B_i \setminus B_{i-1}} (\deg_{\mathcal{T}}(w) - 1),$$

since  $\mathcal{T}$  is a tree. For  $j \geq 1$  we can sum these equalities from  $i = 0$  to  $j - 1$ , getting

$$|B_j| - 1 = \sum_{w \in B_{j-1}} (\deg_{\mathcal{T}}(w) - 1) + 1,$$

so

$$|B_j| = \sum_{w \in B_{j-1}} (\deg_{\mathcal{T}}(w) - 2) + |B_{j-1}| + 2.$$

In particular

$$|B_j| \leq \sum_{w \in B_{j-1}} (\deg_{G_t^\circ}(w) - 2) + |B_{j-1}| + 2. \tag{27}$$

Note that  $G_t^\circ$  contains at most two vertices of degree one,  $\mathcal{X}_0$  and  $\mathcal{X}_t$ . Therefore

$$\sum_{w \in B_{j-1}} (\deg_{G_t^\circ}(w) - 2) \leq 2 + \sum_{w \in G_t^\circ} (\deg_{G_t^\circ}(w) - 2).$$

By Lemma 3.18 with probability at least  $1 - Ce^{-c \log^2 n}$  (for some  $C, c > 0$ ) for all  $t \leq T$  the right hand side above is bounded by  $n^{2\alpha} + 2 \leq n^\varepsilon + 2$ . From this and (27) we obtain

$$|B_j| \leq n^\varepsilon + 2 + |B_{j-1}| + 2$$

with high enough probability. This in turn implies that for any  $k \geq 1$  we have  $|B_k| \leq k(n^\varepsilon + 5)$ .

We now apply this estimate to bound the size of the intersection of the core with the bad set. For any  $v \in \mathcal{Z}_t \cap L_0$  let  $G(v)$  be the induced subgraph of  $G_t^\circledast$  with the set of vertices equal to  $B_{G_t}(v, n^\varepsilon) \cap V_{G_t^\circledast}$ . Note that it may happen that  $v \notin G_t^\circledast$ , but if  $G(v)$  is nonempty, it is connected and of diameter at most  $2n^\varepsilon$ . Indeed,  $G_t^\circledast$  is connected and distances in  $G_t^\circledast$  between any two vertices  $w, u \in G_t^\circledast$  are the same as in  $G_t$  (since the shortest paths between any elements of the core are disjoint from the trees which are removed during its construction).

Now choose any  $r \in G(v)$  and let  $B_k$  be defined as in the first part of the proof. Taking  $k = 2n^\varepsilon \geq \text{diam}(G(v))$  we get  $G(v) \subset B_k$ , which by our bound on  $|B_k|$  proves that for  $n$  large enough with probability at least  $1 - Ce^{-c \log^2 n}$ , for any  $t \leq T$  and  $v \in \mathcal{Z}_t \cap L_0$  we have

$$|B_{G_t}(v, n^\varepsilon) \cap V_{G_t^\circledast}| \leq 3n^{2\varepsilon}. \quad (28)$$

Now using the definition of the bad set we get

$$V_{G_t^\circledast} \cap \Upsilon_t = \bigcup_{v \in \mathcal{Z}_t \cap L_0} B_{G_t}(v, n^\varepsilon) \cap V_{G_t^\circledast}.$$

By Assumption 3.9 with probability at least  $1 - C_1 e^{-c_1 \log^2 n}$ , for some  $C_1, c_1 > 0$ , we have  $|\mathcal{Z}_t \cap L_0| \leq n^\varepsilon$ . This estimate together with (28) concludes the proof.  $\square$

To prove that the whole bad set is small up to time  $T$ , we will show that dead trees removed in the construction of  $G_t^\circledast$  cannot be too large and then estimate the total number of such trees.

**Proposition 3.20 (The bad set is small).** *There exist  $C, c > 0$  such that*

$$\mathbb{P}(|\Upsilon_T| \geq n^{7\varepsilon}) \leq Ce^{-c \log^2 n}.$$

**Proof.** Let  $K$  be the graph removed from  $G_T$  in the construction of  $G_T^\circledast$ . We will first prove that

$$\mathbb{P}(|V_K \cap \Upsilon_T| \geq n^{6\varepsilon}) \leq Ce^{-c \log^2 n}. \quad (29)$$

To this end we first estimate the size of the largest connected component of  $K$  (i.e., the largest tree removed in the construction of the core). Let  $\mathcal{A}$  be the event that there is a component  $K'$  of size  $l \geq n^\varepsilon$ . Note that after entering  $K'$  for the first time, say at time  $T_k$ , the CRW traverses the whole tree  $K'$  in a depth first search manner, exhausting all the bars corresponding to visited vertices. Due to the constant speed of the CRW, the time needed for this equals exactly  $l$ . Since the number of vertices and edges in a tree differ by one, during that time the CRW makes  $l - 1$  jumps to previously unexplored vertices and the same number of backtracks. It follows that

$$\mathcal{P}_{T_k+l} - \mathcal{P}_{T_k} = l - 1 - l = -1. \quad (30)$$

Using (14) from Lemma 3.13 with  $\eta = T_k$  and  $s = n^\varepsilon/2$  we get that for fixed  $k, l$  the probability that the condition (30) is fulfilled is smaller than  $Ce^{-c \log^2 n}$  for some  $C, c > 0$ . Observing that  $k, l \leq n^2$  and applying a union bound we get

$$\mathbb{P}(\mathcal{A}) \leq C_1 e^{-c_1 \log^2 n}, \quad (31)$$

for some  $C_1, c_1 > 0$ . To prove (29) it remains to bound the number of trees with nonempty intersection with  $\Upsilon_T$ . Note that such a tree either

- (a) is attached to a vertex from  $\Upsilon_T \cap V_{G_T^\circledast}$
- (b) or contains an element from  $\mathcal{Z}_T \cap L_0$ .

The number of the former trees is at most the sum of degrees of vertices in  $\Upsilon_T \cap V_{G_T^\circ}$ . The number of the latter trees equals at most  $\iota(\mathcal{Z}_T)$ .

Now we show that with probability at least  $1 - Ce^{-cn^\varepsilon}$ , for some  $C, c > 0$ , no vertex has degree greater than  $n^\varepsilon$ . Indeed, for a fixed vertex  $w$  we can estimate the number of bridges incident to  $w$  by using Lemma B.1 (which gives a general bound on the number of bridges in a given subset of  $E \times [0, 1)$  in terms of a Poisson process). More specifically, we apply the second part of the lemma with  $A = E_w \times [0, 1)$ , where  $E_w$  is the set of all edges incident to  $w$ . Note that the Lebesgue measure of  $A$  satisfies  $|A| = 2(n - 1)$ . Thus we obtain that with probability at least  $1 - C'e^{-c'n^\varepsilon}$ , for some  $C', c' > 0$ , the total number of bridges incident to  $w$  is at most  $n^\varepsilon$ . A union bound over  $w$  finishes the argument.

Using Lemma 3.19 with probability at least  $1 - Ce^{-c \log^2 n}$  we have  $|V_{G_T^\circ} \cap \Upsilon_T| \leq n^{4\varepsilon}$ . Recall that the size of a single dead tree is smaller than  $n^\varepsilon$  with high probability, see (31). Thus the total number of vertices belonging to trees from case (a) above is at most  $n^{4\varepsilon} \cdot n^\varepsilon \cdot n^\varepsilon = n^{6\varepsilon}$  with high probability.

On the other hand, by Assumption 3.9 the total number of vertices from trees satisfying case b) is at most  $n^\varepsilon \cdot n^\varepsilon = n^{2\varepsilon}$  with high probability. Combining the two cases yields (29).

Thus with the required probability,  $|\Upsilon_T| = |\Upsilon_T \cap V_K| + |\Upsilon_T \cap V_{G_T^\circ}| \leq n^{4\varepsilon} + n^{6\varepsilon} \leq n^{7\varepsilon}$  for  $n$  large enough.  $\square$

*Bad hits process.* Here we prove that  $I_T^b$ , i.e., the number of bad jumps up to time  $T$  is small with high probability. The argument is similar as in Lemma 3.12, but more subtle.

**Lemma 3.21.** *There exist  $C, c, \delta > 0$  such that for any stopping time  $\eta$  we have*

$$\mathbb{P}(I_T^b - I_\eta^b \geq 1 | \mathcal{F}_\eta) \leq \frac{1}{n^\delta},$$

with probability at least  $1 - Ce^{-c \log^2 n}$ .

**Proof.** We will work on the event  $\{\eta \leq T\}$  (on its complement the probability in question vanishes due to monotonicity of  $I^b$ ). Let us set

$$\tau = \inf\{t \geq 0: \iota(\mathcal{Z}_t) > n^\varepsilon \text{ or } |\Upsilon_t| > n^{7\varepsilon}\}.$$

Using a union bound we write

$$\mathbb{P}(I_T^b - I_\eta^b \geq 1 | \mathcal{F}_\eta) \leq \mathbb{P}(I_{T \wedge \tau}^b - I_\eta^b \geq 1 | \mathcal{F}_\eta) + \mathbb{P}(\tau \leq T | \mathcal{F}_\eta). \tag{32}$$

Observe that  $\mathbb{P}(\tau \leq T | \mathcal{F}_\eta) \leq e^{-c \log^2 n}$  with probability at least  $1 - Ce^{-c \log^2 n}$  (for some  $C, c > 0$ ), since by Proposition 3.20 the bad set is small with high probability (and thus, by Lemma 3.10, also conditionally on  $\mathcal{F}_\eta$ ), and by Corollary 3.11  $\iota(\mathcal{Z}_t)$  is small with high probability conditionally on  $\mathcal{F}_\eta$ .

To deal with the first term we estimate the intensity  $\lambda^b$  of  $I^b$ . For any  $t$  such that  $\eta \leq t < T \wedge \tau$  let  $D(\Upsilon_t)$  (resp.  $L(\Upsilon_t)$ ) denote the set of columns (resp. rows) which have non-empty intersection with  $\Upsilon_t$ . Formally,  $D(\Upsilon_t) = \{D_i : D_i \cap \Upsilon_t \neq \emptyset\}$  (and likewise for  $L(\Upsilon_t)$ ). Let  $K_t = D(\Upsilon_t) \cup L(\Upsilon_t)$ . Recall that by  $A_t$  we denote the set of vertices accessible at time  $t$  by a fresh jump. Observe that if  $w = \mathcal{X}_t$  is a vertex in a row or column belonging to  $K_t$ , then  $|A_t \cap \Upsilon_t| \leq n^{7\varepsilon}$ , otherwise  $|A_t \cap \Upsilon_t| = 0$ . Note also that  $A_t = \emptyset$  if  $t \geq \tau_c$ . By Lemma 3.6 we have

$$\Lambda := \int_\eta^{T \wedge \tau} \lambda_t^b dt \leq \frac{\Theta\beta}{n-1} \int_\eta^{T \wedge \tau} |A_t \cap \Upsilon_t| dt \leq \frac{\Theta\beta}{n-1} n^{7\varepsilon} \int_\eta^{T \wedge \tau \wedge \tau_c} \mathbb{1}_{\{\mathcal{X}_t \in \bigcup K_t\}} dt,$$

where  $\mathcal{X}_t \in \bigcup K_t$  means that  $\mathcal{X}_t$  is at a vertex belonging to a row or column having nonempty intersection with  $\Upsilon_t$ . As  $\Upsilon_t$  is nondecreasing in  $t$ , so is  $K_t$ , thus we can further estimate

$$\Lambda \leq \frac{\Theta\beta}{n-1} n^{7\varepsilon} \int_\eta^{T \wedge \tau \wedge \tau_c} \mathbb{1}_{\{\mathcal{X}_t \in \bigcup K_{T \wedge \tau}\}} dt = \frac{\Theta\beta}{n-1} n^{7\varepsilon} \sum_{F \in K_{T \wedge \tau}} \int_\eta^{T \wedge \tau \wedge \tau_c} \mathbb{1}_{\{\mathcal{X}_t \in F\}} dt.$$

For any row or column  $F \in K_{T \wedge \tau}$  we have

$$\int_\eta^{T \wedge \tau \wedge \tau_c} \mathbb{1}_{\{\mathcal{X}_t \in F\}} dt = \sum_{v \in F} \int_\eta^{T \wedge \tau \wedge \tau_c} \mathbb{1}_{\{\mathcal{X}_t = v\}} dt \leq |F \cap \mathcal{Z}_{T \wedge \tau}| \leq \iota(\mathcal{Z}_{T \wedge \tau}).$$

We used the fact that for a fixed vertex  $v$  the integral is bounded by 1, since  $t \leq \tau_c$  and the bar corresponding to  $v$  has height 1. Combining the above facts we obtain

$$\Lambda \leq \frac{\Theta\beta}{n-1} n^{7\varepsilon} |K_{T \wedge \tau}| \iota(\mathcal{Z}_{T \wedge \tau}).$$

We have  $|\Upsilon_{T \wedge \tau}| \leq n^{7\varepsilon}$ , thus  $|K_{T \wedge \tau}| \leq 2n^{7\varepsilon}$ . Moreover,  $\iota(\mathcal{Z}_{T \wedge \tau}) \leq n^\varepsilon$  and thus  $\Lambda \leq 4\Theta\beta n^{15\varepsilon-1}$ .

As  $I_t^b - \int_0^t \lambda_u^b du$  is a martingale, we have by Doob's theorem

$$\mathbb{E}((I_{T \wedge \tau}^b - I_\eta^b) \mathbb{1}_{\{T \wedge \tau \geq \eta\}} | \mathcal{F}_\eta) = \mathbb{E}\left(\mathbb{1}_{\{T \wedge \tau \geq \eta\}} \int_\eta^{T \wedge \tau} \lambda_t^b dt | \mathcal{F}_\eta\right) \leq 4\Theta\beta n^{15\varepsilon-1}.$$

Now by an application of the conditional Markov inequality we bound the first term of (32) by  $n^{-\delta}$  for some  $\delta > 0$ . This concludes the proof.  $\square$

### 3.5. Backtracks

We will now show that after an internal jump the CRW is unlikely to backtrack its steps back to  $L_0$ .

We start with a deterministic lemma about the structure of the core  $G_t^\circ$ . Recall that a path in a graph is a sequence of pairwise distinct vertices  $v_1, \dots, v_k$  such that for all  $i < k$  vertices  $v_i$  and  $v_{i+1}$  are adjacent.

**Lemma 3.22.** *Suppose that  $G_t^\circ$  contains at most  $n^\delta$  vertices of degree greater than 2, for some  $\delta \in (0, \varepsilon)$ . Then there exists  $n_0 \geq 1$  such that for any  $n \geq n_0$  and any  $v \in L_0 \cap G_t^\circ$  the following holds: every simple path in  $G_t$  which connects  $v$  with  $\Upsilon_t^c \cap G_t^\circ$  must contain a subpath of length at least  $\log^2 n$  consisting only of vertices of degree two in  $G_t^\circ$ .*

**Proof.** By the definition of the bad set any path connecting  $v$  with  $\Upsilon_t^c$  has length  $n^\varepsilon$ . Note that if the end vertex of the path is in  $G_t^\circ$  then the whole path is fully contained in  $G_t^\circ$  (otherwise to return to  $G_t^\circ$  the path would have to repeat one of its vertices). Dividing the path into consecutive subpaths of length  $\log^2 n$  we obtain  $n^\varepsilon / \log^2 n$  subpaths. Since  $n^\delta < n^\varepsilon / \log^2 n$  for  $n$  large enough and we assumed there are at most  $n^\delta$  vertices of degree more than 2 in  $G_t^\circ$ , there must be at least one subpath with only vertices of degree 2.  $\square$

Let  $P$  be a path in  $G_{t-}^\circ$  consisting only of vertices of degree two in  $G_{t-}^\circ$ . We will call such a path *straight*. If  $t < \tau_c$ , we will call the *potential* of  $P$  the total Lebesgue measure of these parts of bars corresponding to vertices of  $P$  which have not been explored up to time  $t$ . For  $t \geq \tau_c$  we set the potential to be zero. Formally, if we denote the potential of  $P$  by  $\mathcal{P}(P)$ , we have

$$\mathcal{P}(P) := \mathbb{1}_{\{t < \tau_c\}} \text{Leb}((V(P) \times [0, 1]) \setminus \mathcal{X}_{[0, t]}),$$

where  $V(P)$  is the set of vertices of  $P$ . We stress that this notion as well as the property of being a straight path depends strongly on  $t$ , which is suppressed in the notation but should not lead to misunderstanding in the sequel.

Suppose that  $t < \tau_c$ . Let  $t_1(P)$  be the time the CRW entered a vertex of  $P$  for the first time, and  $t_2(P) \leq t$  the time when it left a vertex of  $P$  for the last time before  $t$ . Because of the property (11) at time  $t$  the potential of  $P$  equals to  $\mathcal{P}_{t_2(P)-} - \mathcal{P}_{t_1(P)}$ .

Given  $b > 0$ , we will say that  $P$  has *large potential with constant  $b$*  if

$$\mathcal{P}(P) \geq b|P|, \tag{33}$$

where  $|P|$  is the length of  $P$ , i.e., the number of its vertices minus one. Later on the parameter  $b$  will be fixed and we will simply use the term ‘‘large potential’’, with  $b$  being implicit.

We now introduce the event

$$\mathcal{Q}_t(b) := \{\text{all straight paths in } G_t^\circ \text{ of length } \lceil \log^2 n \rceil \text{ have large potential with constant } b\}. \tag{34}$$

**Lemma 3.23 (Straight paths have large potential).** *There exist  $C, c > 0$  and  $b > 0$  such that*

$$\mathbb{P}(\forall_{t < T \wedge \tau_c} \mathcal{Q}_t(b)) \geq 1 - Ce^{-c \log^2 n}.$$

**Proof.** Let  $\gamma = 1/4\Theta\beta_1$ , fix arbitrary  $a < \Theta^{-1}\beta_0 - 1/2$  and let  $b = a\gamma$ . Suppose that there exists  $t < T \wedge \tau_c$  and a straight path  $P$  in  $G_t^\circledast$  of length  $m = \lceil \log^2 n \rceil$  which does not have large potential. In particular this implies that there exist  $l \geq 1$  and  $k \geq m$  such that

$$\mathcal{P}_{T_{l+k}} - \mathcal{P}_{T_l} < bm. \quad (35)$$

Indeed, we can take  $t_1(P) = T_l$  and let  $T_{l+k} \leq t$  be the moment the CRW enters the other end vertex of  $P$ . Note that  $k \geq m$  since the CRW may have traversed some dead trees between  $T_l$  and  $T_{l+k}$ .

The lemma will be shown once we prove that

$$p := \mathbb{P}(\exists_{l \geq 1} \exists_{k \geq m} \mathcal{P}_{T_{l+k}} - \mathcal{P}_{T_l} < bm \text{ and } T_{l+k} \leq T) \leq Ce^{-c \log^2 n}, \quad (36)$$

for some  $C, c > 0$ . Note that writing  $T_{l+k} \leq T$  we implicitly assume that the CRW visits at least  $l+k$  vertices. Now, if the event in the definition of  $p$  holds, then either after some time  $T_l$  we have discovered at least  $k \geq m$  additional vertices very quickly (so that  $T_{l+k} - T_l \leq \gamma m$  and  $T_{l+k} \leq T$ ), or after  $T_l$  the potential failed to increase by  $bm$  in a time interval of length at least  $\gamma m$ . Thus we may estimate

$$p \leq \mathbb{P}(\exists_{l \geq 1} \exists_{u \in [\gamma m, T - T_l]} \mathcal{P}_{T_{l+u}} - \mathcal{P}_{T_l} < bm \text{ and } \tau_c \geq T_l + \gamma m) + \mathbb{P}(\exists_{l \geq 1} \exists_{k \geq m} T_{l+k} - T_l \leq \gamma m \text{ and } T_{l+k} \leq T).$$

Note that we have chosen  $\gamma$  small enough so that by Lemma 3.8 we get

$$\mathbb{P}(\exists_{l \geq 1} \exists_{k \geq m} T_{l+k} - T_l < \gamma m) \leq \sum_{l=1}^{n^2} \sum_{k \geq m} e^{-ck} \leq C_1 e^{-c_1 m}.$$

Thus

$$p \leq \mathbb{P}(\exists_{l \geq 1} \exists_{u \in [\gamma m, T - T_l]} \mathcal{P}_{T_{l+u}} - \mathcal{P}_{T_l} < bm \text{ and } \tau_c \geq T_l + \gamma m) + C_1 e^{-c_1 m},$$

for some  $C_1, c_1 > 0$ . Now using a union bound over  $l$  (observe that  $T_l = \infty$  for  $l > n^2$ ) and (14) from Lemma 3.13 (recall that  $b = a\gamma$ ) we obtain (36), thus concluding the proof.  $\square$

From now on we fix  $b$  to be the constant guaranteed by Lemma 3.23. Let  $S_t$  be the set of endpoints of all straight paths in  $G_{t-}^\circledast$  with potential at time  $t$  at least  $b \log^2 n$  and let  $\rho_k$  denote the  $k$ -th moment  $t$  when the CRW enters a vertex from  $S_t$  by a backtrack. More precisely, set  $\rho_0 = 0$ , and for  $k \geq 1$ ,

$$\rho_k = \inf\{t > \rho_{k-1} : \mathcal{X}_t \in S_t, \mathcal{X}_{t-} \notin S_t \text{ and the bridge } (\{\mathcal{X}_t, \mathcal{X}_{t-}\}, t) \text{ has been traversed before time } t\}.$$

Note that  $\rho_k$  are stopping times and  $\rho_k < \tau_c$  (since if  $t > \tau_c$ , then  $S_t = \emptyset$ ). Moreover, the minimal path in  $G_{\rho_k}^\circledast$  with potential at time  $\rho_k$  at least  $b \log^2 n$ , and starting at  $\mathcal{X}_{\rho_k}$ , is uniquely determined. Let us denote this path by  $P_k$ .

Let us say that the CRW completely covers  $P_k$  if after entering its end vertex by a backtrack it eventually exhausts all the bars corresponding to vertices of  $P_k$ , possibly departing from them at some intermediate time intervals. Denote by  $\mathcal{A}_k$  the event that  $\rho_k < \infty$ , the CRW completely covers  $P_k$  before time  $T$  and while traversing  $P_k$  it does not make an excursion of length  $n^\varepsilon$  (recall Definition 3.3). Formally,

$$\mathcal{A}_k = \bigcup_{t \in (0, T]} (\{t > \rho_k\} \cap \{P_k \times [0, 1) \subset \mathcal{X}_{[0, t]}\} \cap \bigcap_{s \in (0, t]} (\mathcal{E}_s(n^\varepsilon) \cap \{s > \rho_k\})^c).$$

**Proposition 3.24 (No straight paths of large potential are covered without excursions).** *There exist  $C, c > 0$  such that*

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} \mathcal{A}_k\right) \leq Ce^{-c \log^2 n}.$$

*In other words, the probability that before time  $T$  the CRW completely covers some straight path of potential (at the time of entry) at least  $b \log^2 n$  without making an excursion of length  $n^\varepsilon$  in the process, is bounded by  $Ce^{-c \log^2 n}$ .*



**Proof.** Note first that for some  $C > 0$  we have  $\mathbb{P}(\rho_{Cn^2} < \infty) \leq e^{-n^2}$ . Indeed, each  $\rho_k$  involves a backtrack and before the CRW closes into a cycle each bridge can be used in at most one backtrack, so the number of moments  $\rho_k$  is bounded by the total number of bridges in the process. It follows from Lemma B.2 (applied with  $t = 0$ ,  $s = 1$  and  $k = Cn^2$  for suitably chosen  $C > 0$ ) that this number is at most  $Cn^2$ , for some  $C > 0$ , with the required probability.

Thus, as we can perform a union bound over  $k \leq Cn^2$ , it is enough to show that for every  $k \geq 1$  we have  $\mathbb{P}(\mathcal{A}_k) \leq Ce^{-c \log^2 n}$  for some  $C, c > 0$ .

To simplify the notation, in what follows we will drop the subscript  $k$  and write simply  $\rho$ ,  $P$  for  $\rho_k$ ,  $P_k$ . We will also identify  $P$  with  $(V(P) \times [0, 1]) \setminus \mathcal{X}_{[0, \rho]}$ , i.e., with the part of all the bars corresponding to vertices from  $P$  which at time  $\rho$  was unused.

Let us first introduce a change of time to merge into one interval all the random intervals of time during which the CRW stays on  $P$ . Define

$$H_t = \int_0^t \mathbb{1}_P(\mathcal{X}_{\rho+s}) ds,$$

where in the case  $\rho = \infty$  we interpret  $\mathbb{1}_P(\mathcal{X}_{\rho+s})$  as zero (we will use this convention throughout the proof). The process  $H_t$  measures how much of the potential that the path  $P$  had at time  $\rho$  has been used up to time  $t$ . In particular  $P$  gets completely backtracked up to time  $t$  if and only if  $H_{t-\rho}$  equals the potential of the path at time  $\rho$ .

Define now  $\sigma_0 = 0$  and for  $s > 0$

$$\sigma_s = (\inf\{t > 0: H_t \geq s\}) \wedge (T - \rho)_+.$$

Consider also the processes

- $J_s = |\mathcal{Z}_s \setminus L_0|$ ,
- $M_t = \int_0^t \mathbb{1}_P(\mathcal{X}_{(\rho+s)-}) dJ_{\rho+s}$  – the number of jumps to previously unexplored vertices outside  $L_0$  that the CRW makes from  $P$  between times  $\rho$  and  $\rho + t$ .
- $N_s = M_{\sigma_s}$  – the number of jumps to previously unexplored vertices outside  $L_0$  the CRW makes during the first  $s$  time units spent on  $P$  during the backtrack.

Note that  $\sigma_s$  is a stopping time with respect to the filtration  $(\mathcal{F}_{\rho+t})_{t \geq 0}$ . Moreover, if  $\sigma_s < T - \rho$ , then  $\mathcal{X}_{(\rho+\sigma_s)-} \in P$ .

Our strategy for proving that  $\mathbb{P}(\mathcal{A}_k)$  is small is as follows. First we will show that on the set  $\mathcal{A}_k \cap \{\tau_{\text{iso}}^\alpha > T\}$  we have  $N_{b \log^2 n} > c \log^2 n$ , i.e., the CRW makes many jumps to previously unexplored vertices while backtracking  $P$ . Then, to finish the argument, we will use Lemma 3.15 to show that the probability that none of those jumps is a beginning of a forbidden excursion is small.

Let us thus first estimate the intensity of  $N_t$  for  $t \leq b \log^2 n$  with respect to the filtration  $(\mathcal{F}_{\rho+\sigma_s})_{s \geq 0}$  on the event  $\{\sigma_{b \log^2 n} < T - \rho\} \cap \{\tau_{\text{iso}}^\alpha > T\}$ .

Denote by  $\mu_t^J$  is the intensity of  $J_t$  and note that by Doob's theorem and the properties of integrals with respect to counting processes, the intensity of  $M_t$  with respect to  $(\mathcal{F}_{\rho+t})_{t \geq 0}$  equals  $\mathbb{1}_P(\mathcal{X}_{(\rho+t)-}) \mu_{\rho+t}^J$ .

Now, again we employ Doob's theorem together with a change of variables (note that function  $s \mapsto \sigma_s$  is constant on intervals where  $\mathbb{1}_P(\mathcal{X}_{(\rho+\sigma_s)-})$  vanishes and otherwise increases linearly with speed one) to conclude that the intensity of  $N_s$  with respect to  $(\mathcal{F}_{\rho+\sigma_s})_{s \geq 0}$  equals

$$\tilde{\mu}_s = \mathbb{1}_P(\mathcal{X}_{(\rho+\sigma_s)-}) \mu_{\rho+\sigma_s}^J.$$

Therefore we can use the same argument as in the proof of (9) in Lemma 3.7 to conclude that if  $T < \tau_{\text{iso}}^\alpha$ , then for Lebesgue almost all  $s$  such that  $\rho + \sigma_s < T \wedge \tau_c$  we have

$$\tilde{\mu}_s \geq \frac{\Theta^{-1} \beta}{n-1} (n-2-n^\alpha) \geq c' > 0$$

for some  $c' > 0$ . Since  $\mathcal{A}_k \subset \{\rho + \sigma_{b \log^2 n} \leq T \wedge \tau_c\}$  we have

$$\Lambda_{b \log^2 n} := \int_0^{b \log^2 n} \tilde{\mu}_s ds \geq c' b \log^2 n$$

on the event  $\mathcal{A}_k \cap \{\tau_{\text{iso}}^\alpha > T\}$  and thus by Lemma 3.5 and Assumption 3.9, we get for some constants  $C, c > 0$

$$\mathbb{P}(\mathcal{A}_k \cap \{N_{b \log^2 n} \leq c \log^2 n\}) \leq Ce^{-c \log^2 n}. \quad (37)$$

It remains to bound the probability  $\mathbb{P}(\mathcal{A}_k \cap \{N_{b \log^2 n} > c \log^2 n\})$ .

Denote thus by  $\gamma_1, \gamma_2, \dots$  the times of subsequent jumps from  $P$  to previously unexplored vertices outside  $L_0$  made after time  $\rho$ . The condition  $N_{b \log^2 n} > c \log^2 n$  translates into  $\gamma_{\lceil c \log^2 n \rceil} \leq \rho + \sigma_{b \log^2 n}$ . Thus  $\mathcal{A}_k \cap \{N_{b \log^2 n} > c \log^2 n\} \subset \mathcal{A}_k \cap \{\gamma_{\lceil c \log^2 n \rceil} < T\}$ .

Let

$$J := \inf\{r \geq 1 : \mathbb{P}(\mathcal{E}_{\gamma_r}(n^\varepsilon) | \mathcal{F}_{\gamma_r}) < q \cdot \mathbb{1}_{\{\mathcal{X}_{\gamma_r} \notin L_0\}} \mathbb{1}_{\{\gamma_r \leq T\}}\},$$

with  $q$  as in Corollary 3.16.  $J$  is a stopping time with respect to the discrete time filtration  $(\mathcal{F}_{\gamma_r})_{r \geq 1}$ . By Corollary 3.16, for each  $r$  the probability that the inequality in the definition of  $J$  holds is bounded by  $Ce^{-c \log^2 n}$ . Taking the union bound over  $r \leq \lceil c \log^2 n \rceil$  we get  $\mathbb{P}(J \leq \lceil c \log^2 n \rceil) \leq C'e^{-c' \log^2 n}$  for some  $C', c' > 0$ .

We can now estimate

$$\begin{aligned} \mathbb{P}(\mathcal{A}_k \cap \{N_{b \log^2 n} > c \log^2 n\}) &\leq \mathbb{P}(\mathcal{A}_k \cap \{\gamma_{\lceil c \log^2 n \rceil} < T\}) \\ &\leq \mathbb{P}\left(\bigcap_{r=1}^{\lceil c \log^2 n \rceil} \mathcal{E}_{\gamma_r}(n^\varepsilon)^c \cap \{\gamma_{\lceil c \log^2 n \rceil} \leq T\} \cap \{J > \lceil c \log^2 n \rceil\}\right) + C'e^{-c' \log^2 n} \\ &\leq (1-q)^{c \log^2 n} + C'e^{-c' \log^2 n} \leq 2e^{-c'' \log^2 n}, \end{aligned}$$

where the third inequality is obtained by a sequence of conditionings with respect to  $\mathcal{F}_{\gamma_r}$ ,  $r = \lceil c \log^2 n \rceil, \dots, 1$ . Together with (37) this shows that for all  $k$ ,

$$\mathbb{P}(\mathcal{A}_k) \leq Ce^{-c \log^2 n},$$

which ends the proof of the proposition. □

### 3.6. Isoperimetry upper bound

Let  $\tau_0 = 0$  and for  $k \geq 1$  let

$$\tau_k := \inf\{t > \tau_{k-1} : \mathcal{X}_t \in L_0, \mathcal{X}_t \notin \mathcal{Z}_{t-}\}$$

be the  $k$ -th time a new vertex from  $L_0$  is visited by  $\mathcal{X}$ .

First we present the main technical lemma stating that the CRW does not visit  $L_0$  too often. This result will be used to prove the forthcoming Corollary 3.26 claiming good isoperimetry.

**Lemma 3.25.** *There exists  $C, c, p > 0$  such that for any  $k \in \mathbb{N}$*

$$\mathbb{P}(\tau_{k+1} - \tau_k > n \wedge (T - \tau_k) | \mathcal{F}_{\tau_k}) \geq p \cdot \mathbb{1}_{\{\tau_k \leq T-1\}}$$

holds with probability at least  $1 - Ce^{-c \log^2 n}$ .

**Proof.** Throughout the proof we will use “with high probability” as a shorthand for “with probability at least  $1 - Ce^{-c \log^2 n}$ ” for some constants  $C, c > 0$  (whose values may change from line to line).

At time  $\tau_k$  a new vertex  $w \in L_0$  is visited (unless  $k = 0$  and  $\mathcal{X}_0 \notin L_0$ , which is easily dealt with below). Our first aim is to show that the CRW escapes from  $L_0$  and then from the bad set (with conditional probability uniformly bounded away from 0). Define the stopping time  $\sigma = \inf\{t > \tau_k : \mathcal{X}_t \notin \Upsilon_t\}$ . By Lemma 3.14 with high probability the CRW has a chance bounded away from 0 of performing (in time less than 1) a jump from  $w$  to a previously unvisited vertex from the same column, thus leaving  $L_0$ . From the new vertex the CRW can, with probability bounded away from 0, make an excursion described in Corollary 3.16. Note that if this happens, then  $\sigma < \tau_{k+1}$ . Thus with high probability we have  $\mathbb{P}(\sigma < \tau_{k+1} | \mathcal{F}_{\tau_k}) > q$  for some  $q > 0$ , independent of  $n$ .

Note that if  $\tau_{k+1} > \sigma > T$  then,  $\tau_{k+1} - \tau_k > \sigma - \tau_k > T - \tau_k$ , so it is enough to prove that for some  $q' > 0$  we have

$$\mathbb{P}(\mathcal{Z}_{[\sigma, (\sigma+n) \wedge T]} \cap L_0 = \emptyset | \mathcal{F}_\sigma) \geq q' \mathbb{1}_{\{\sigma \leq T, \sigma < \tau_{k+1}\}}, \quad (38)$$

with high probability. This also takes care of the case  $k = 0$ ,  $\mathcal{X}_0 \notin L_0$ , since then  $\sigma = \tau_0$ .

From now on we will implicitly work on the event  $\{\sigma \leq T, \sigma < \tau_{k+1}\}$ . Denote by  $\tau$  the time of the first hit of  $L_0$  after time  $\sigma$ . This hit can happen either by a direct jump or by a backtrack. Denote the event that the former (resp. the latter) situation happens and  $\tau_{k+1} \leq (\sigma + n) \wedge T$  by  $\mathcal{R}$  (resp. by  $\mathcal{K}$ ). Clearly on  $\mathcal{R}$  and  $\mathcal{K}$  we also have  $\tau \leq (\sigma + n) \wedge T$ . Moreover on  $\mathcal{K}$  we have  $\tau < \tau_{k+1} < \infty$ , so in particular  $\tau < \tau_c$ .

Recall that  $I_t^d$  is the number of direct jumps to  $L_0$  in the time interval  $[0, t]$ . By Lemma 3.6 the intensity  $\lambda_t^d$  of  $I_t^d$  is at most  $\frac{\Theta\beta}{n-1}$  as long as the CRW is outside  $L_0$ , in particular on the interval  $[\sigma, \tau)$  (the intensity can be 0 if the only possible vertex in  $L_0$  is dead). Denoting  $\Lambda_t = \int_0^t \lambda_t^d$ , we thus get

$$\mathbb{P}(\mathcal{R}|\mathcal{F}_\sigma) = \mathbb{P}\left(\left\{I_{\tau \wedge (\sigma+n) \wedge T}^d - I_\sigma^d \geq 1\right\} \cap \left\{\Lambda_{\tau \wedge (\sigma+n)} - \Lambda_\sigma \leq \frac{n}{n-1}\Theta\beta\right\} \middle| \mathcal{F}_\sigma\right)$$

and by Lemma 3.4 the right hand side is bounded from above by  $\mathbb{P}(X \geq 1)$ , where  $X$  is a Poisson random variable with parameter  $\frac{n}{n-1}\Theta\beta \leq 2\Theta\beta$ . We conclude that almost surely  $\mathbb{P}(\mathcal{R}|\mathcal{F}_\sigma)$  is uniformly bounded away from 1 or, equivalently, for some  $q > 0$  we have

$$\mathbb{P}(\mathcal{R}^c|\mathcal{F}_\sigma) \geq q \mathbb{1}_{\{\sigma \leq T \wedge \tau_{k+1}\}} \geq q. \quad (39)$$

As  $\mathcal{K} \subset \mathcal{R}^c$ , to obtain (38) it thus suffices to show that  $\mathbb{P}(\mathcal{K}|\mathcal{F}_\sigma) = o(1)$  with high probability. We will first show that we can restrict to an event on which for all  $t \leq T$  the sets  $\Upsilon_t$  have at most  $n^{\varepsilon/2}$  vertices of degree greater than 2 in  $G_t^\circledast$ , there are no bad jumps for  $t \leq \tau$ , and all straight paths in  $G_t^\circledast$  of length at least  $\log^2 n$  have potential greater than  $b \log^2 n$ . To this end recall the notation of Lemmas 3.18, 3.21, 3.23 and define

$$\mathcal{C} := \bigcap_{t \leq T} \mathcal{D}_t^{\varepsilon/2} \cap \{I_T^b = 0\} \cap \bigcap_{t \leq T} \mathcal{Q}_t(b),$$

with  $b$  as guaranteed by Lemma 3.23.

We see that  $\mathcal{C}$  satisfies all the properties mentioned above and moreover by the aforesaid lemmas (and Lemma 3.10), with probability at least  $1 - Ce^{-c \log^2 n}$  we have  $\mathbb{P}(\mathcal{C}|\mathcal{F}_\sigma) \geq 1 - Cn^{-\delta}$  for some  $\delta > 0$ . Thus it is enough to show that  $\mathbb{P}(\mathcal{K} \cap \mathcal{C}|\mathcal{F}_\sigma) = o(1)$  with high probability.

By definition we have  $\mathcal{X}_\sigma \notin \Upsilon_\sigma$ . Let

$$t = \max\{s \in [\sigma, \tau) : \mathcal{X}_s \notin \Upsilon_s, \text{ and } \mathcal{X}_s \notin \Upsilon_s\}$$

i.e.,  $t$  is the time of the last jump of the CRW before  $\tau$  such that  $\mathcal{X}_t \notin \Upsilon_t$  (note that  $t < T$ ). On  $\mathcal{K} \cap \mathcal{C}$  the assumptions of Lemma 3.22 are satisfied, so every simple path in  $G_t$  from  $\mathcal{X}_t$  to  $L_0$  must contain a straight subpath in  $G_t^\circledast$  of length at least  $\log^2 n$ . By the definition of the event  $\mathcal{C}$  at time  $t$  every such path has potential at least  $b \log^2 n$ . Since on  $\mathcal{K} \cap \mathcal{C} \subset \mathcal{R}^c \cap \mathcal{C}$  there are no bad jumps or direct hits to  $L_0$ , to get from  $\mathcal{X}_t$  to  $L_0$  the CRW must completely cover at least one such path (we give a formal proof of this intuitively clear fact below).

Moreover, between  $t$  and  $\tau$  the CRW does not make an excursion of length  $n^\varepsilon$ , since at the end of such an excursion it would be outside the current bad set, which would contradict the definition of  $t$ . Therefore, by Proposition 3.24,  $\mathbb{P}(\mathcal{K} \cap \mathcal{C}) \leq Ce^{-c \log^2 n}$ , which by Lemma 3.10 shows that  $\mathbb{P}(\mathcal{K} \cap \mathcal{C}|\mathcal{F}_\sigma) \leq e^{-\frac{c}{2} \log^2 n}$  with probability at least  $1 - Ce^{-\frac{c}{2} \log^2 n}$ . We have thus proved that  $\mathbb{P}(\mathcal{K}|\mathcal{F}_\sigma) = o(1)$  with probability at least  $1 - Ce^{-c \log^2 n}$  for some  $C, c > 0$ , which together with (39) proves (38).

We finish with a formal proof of the existence of a straight path with large potential which is covered by the CRW between time  $t$  and  $\tau$  (assuming that the event  $\mathcal{C}$  holds). First, it is easy to see that the next vertex visited by the CRW after time  $t$  must belong to  $\Upsilon_t$ . Indeed, assume that the next jump happened at time  $s$ . Then by the definition of  $t$ ,  $\mathcal{X}_s \in \Upsilon_s$ . Moreover the shortest path in  $G_s$  from  $\mathcal{X}_s$  to  $L_0$  avoids  $\mathcal{X}_t$  (otherwise we would have  $\mathcal{X}_t \in \Upsilon_t$ ). Thus this path uses only edges from  $G_t$ , and so  $\mathcal{X}_s \in \Upsilon_t$ .

Now, between  $t$  and  $\tau$  there are finitely many jumps. Denote by  $v_1, \dots, v_M$  consecutive vertices from  $\Upsilon_t$ , visited by the CRW between times  $t$  and  $\tau$ . Let  $t_1, \dots, t_M$  be the times of visits to  $v_1, \dots, v_M$ . Note that in particular  $t_M = \tau$ , since by the definition of  $\tau$  and the event  $\mathcal{K}$ ,  $\mathcal{X}_\tau \in G_\sigma \subset G_t$ .

We will show by induction that each  $v_k$  is connected to  $\Upsilon_t^c \cap G_t^\circledast$  by a path in  $G_t$  consisting only of vertices which have been visited by the CRW between times  $t$  and  $t_k$ .

This is true for  $v_1 = \mathcal{X}_s$  because it belongs to  $\Upsilon_t = \Upsilon_{s-}$  and so (due to absence of bad jumps) the jump from  $\mathcal{X}_t$  to  $\mathcal{X}_s$  is necessarily a backtrack. Assuming that the statement in question holds for  $v_1, \dots, v_k$ , we have three possibilities:

- The vertex  $v_{k+1}$  is visited directly after  $v_k$ , in which case the statement extends to  $v_{k+1}$  since due to absence of bad jumps between times  $t$  and  $t_{k+1}$  the bridge used for this jump corresponds to an edge in  $G_t$ .

- The CRW leaves  $\Upsilon_t$  after visiting  $v_k$  and the re-entry into  $\Upsilon_t$  happens by a bridge corresponding to an edge in  $G_t$ . In this case the statement in question also holds for  $v_{k+1}$ , since then  $\mathcal{X}_{t_{k+1}-} \in G_t^\circ$ .
- The CRW leaves  $\Upsilon_t$  after visiting  $v_k$  and the re-entry into  $\Upsilon_t$  happens via a bridge which up to time  $t$  was unexplored. Since  $v_{k+1} \in \Upsilon_t$  and after time  $t$  there were no bad jumps, in this situation, the vertex  $v_{k+1}$  must equal  $v_l$  for some  $l \leq k$ , since the bridge used for re-entry must have been used for the first time only after time  $t$  and it could have been used only to leave  $\Upsilon_t$ .

Thus in particular  $\mathcal{X}_\tau$  is connected with  $\Upsilon_t^c \cap G_t^\circ$  by a path  $P$  in  $G_t$  consisting only of vertices visited by the CRW between times  $t$  and  $\tau$ . Note that since  $\tau < \tau_c$ , we must have  $\mathcal{X}_\tau \in G_t^\circ$ . By Lemma 3.22 the path  $P$  contains a straight subpath  $P'$  of length at least  $\log^2 n$ . It is now easy to see that this subpath is completely covered between  $t$  and  $\tau$ . Indeed, using again the absence of bad jumps and the fact that on  $\mathcal{K}$  we have  $\tau < \tau_c$ , we see that the order in which the vertices of  $P'$  are visited by the CRW between times  $t$  and  $\tau$  is uniquely determined – the first entry into each consecutive vertex of  $P'$  must be made by a backtrack from the previous one and after each departure from  $P'$  the CRW returns by backtrack using the same bridge through which it has left. This shows that the whole path  $P'$  must be exhausted before time  $\tau$ . Denote by  $t'$  the first time after  $t$  when the CRW enters a vertex of  $P'$ . Then clearly  $P'$  is a straight path in  $G_{t'-}^\circ$  and therefore by the definition of the event  $\mathcal{C}$  it has large potential at time  $t'$ . In particular  $t' = \rho_k$  for some  $k$  (recall the definition of  $\rho_k$  given before Proposition 3.24).  $\square$

**Corollary 3.26.** *There exist  $C, c > 0$  such that*

$$\mathbb{P}(|\mathcal{Z}_T \cap L_0| \geq C \log^2 n) \leq C e^{-c \log^2 n}.$$

**Proof.** Let  $\mathcal{H}_k = \mathcal{F}_{\tau_k}$  and let  $A_k$  denote the event that the estimate from Lemma 3.25 holds. For  $k \geq 0$  consider the events

$$\mathcal{E}_k = (\{\tau_{k+1} - \tau_k > n \wedge (T - \tau_k)\} \cap A_k \cap \{\tau_k \leq T - 1\}) \cup A_k^c \cup \{\tau_k > T - 1\}.$$

We have

$$\mathbb{P}(\mathcal{E}_k | \mathcal{H}_k) = \mathbb{1}_{A_k \cap \{\tau_k \leq T-1\}} \mathbb{P}(\{\tau_{k+1} - \tau_k > n \wedge (T - \tau_k)\} | \mathcal{H}_k) + \mathbb{1}_{A_k^c \cup \{\tau_k > T-1\}},$$

and by applying Lemma 3.25 we can estimate

$$\mathbb{P}(\mathcal{E}_k | \mathcal{H}_k) \geq p \mathbb{1}_{A_k} \mathbb{1}_{\{\tau_k \leq T-1\}} + \mathbb{1}_{A_k^c \cup \{\tau_k > T-1\}} \geq p. \quad (40)$$

Now let  $K = \frac{2}{p} \lceil \log^2 n \rceil$  and consider the sum  $\mathbb{1}_{\mathcal{E}_1} + \dots + \mathbb{1}_{\mathcal{E}_K}$ . Introducing  $\xi_k := \mathbb{1}_{\mathcal{E}_k} - p$ , we observe that by (40) the sum  $M_N = \xi_1 + \dots + \xi_N$  (with  $M_0 = 0$ ) forms a submartingale with increments bounded by 1. As any submartingale can be written as a martingale plus a nonnegative predictable term and here the martingale part has bounded increments, by Azuma's inequality for martingales with bounded increments [24] we get

$$\mathbb{P}(M_N \leq -t) \leq e^{-\frac{t^2}{2N}}$$

for any  $t \geq 0$  and  $N \geq 1$ . Taking  $N = K$ ,  $t = \frac{pK}{2}$  and rewriting the inequality in terms of  $\mathbb{1}_{\mathcal{E}_k}$  we obtain the estimate

$$\mathbb{P}\left(\mathbb{1}_{\mathcal{E}_1} + \dots + \mathbb{1}_{\mathcal{E}_K} \leq \frac{pK}{2}\right) \leq e^{-\frac{p^2 K}{8}}.$$

Therefore with high probability at least  $pK/2$  of the events  $\mathcal{E}_k$  hold. Since  $\mathbb{P}(A_k) \geq 1 - C e^{-c \log^2 n}$  for some  $C, c > 0$ , by doing a union bound over  $k$  we can assume that none of the events  $A_k^c$  hold. This implies that either  $\tau_k > T - 1$  for some  $k = 1, \dots, K$ , or the event  $\{\tau_{k+1} - \tau_k > n\}$  holds at least  $pK/2$  times, which implies  $\tau_K \geq npK/2 \wedge (T - 1)$ . As  $T \leq n \log^2 n$ , in both cases we have  $\tau_K \geq T - 1$ , so with probability at least  $1 - e^{-c \log^2 n}$  (for some  $c > 0$ ) we have at most  $K = \lceil \frac{2}{p} \log^2 n \rceil$  vertices from  $L_0$  visited up to time  $T - 1$ . An easy estimate shows that with high enough probability there are at most  $c \log^2 n$  visits to  $L_0$  between times  $T - 1$  and  $T$ , which ends the proof.  $\square$

*Proofs of Lemma 3.2 and Proposition 3.1.* Using the results of Section 3.6 the proof of Lemma 3.2 is now immediate.

**Proof of Lemma 3.2.** Since in Corollary 3.26 the starting vertex of the CRW was arbitrary, by symmetry of the graph  $H_n$  the claim of the corollary holds with  $L_0$  replaced by any other row or column. Thus by performing a union bound

over all rows and columns, and all starting vertices  $v \in V$ , we obtain that there exist constants  $C, C', c' > 0$  such that

$$\mathbb{P}(\forall v \in V \iota(\mathcal{Z}_T(v)) \leq C \log^2 n) \geq 1 - C' e^{-c' \log^2 n}.$$

Since we worked under Assumption 3.9, this finishes the proof of Lemma 3.2.  $\square$

With this lemma we can finally prove Proposition 3.1.

**Proof of Proposition 3.1.** Fix  $\alpha \in (0, 1/100)$  and  $T = n^{\alpha/2}$ . Consider the cyclic random walk  $\mathcal{X} = \mathcal{X}(v)$  started at a vertex  $v$ . By Lemma 3.8 we have

$$\mathbb{P}(T_{\lceil 4\Theta\beta n^{\alpha/2} \rceil} < \lfloor n^{\alpha/2} \rfloor) \leq C e^{-cn^{\alpha/2}}$$

for some  $C, c > 0$ . Since  $\lceil 4\Theta\beta n^{\alpha/2} \rceil < n^\alpha$  and  $e^{-cn^{\alpha/2}} \leq e^{-c \log^2 n}$  for  $n$  large enough, we obtain that with probability at least  $1 - C_1 e^{-c_1 \log^2 n}$ , for some  $C_1, c_1 > 0$ , until time  $T$  fewer than  $n^\alpha$  vertices have been explored by the CRW. In particular, this implies  $\iota(\mathcal{Z}_T) < n^\alpha$ .

Now by a union bound over starting vertices  $v$  and Lemma 3.2 we obtain that for some  $C, C_2, c_2 > 0$  we have

$$\mathbb{P}(\forall v \in V \iota(\mathcal{Z}_T(v)) \leq C \log^2 n) \geq 1 - C_2 e^{-c_2 \log^2 n}.$$

The rest of the argument is inductive. Suppose that for some  $T \in [n^{\alpha/2}, n^{1-\alpha/2} \log^2 n]$  we have

$$\mathbb{P}(\forall v \in V \iota(\mathcal{Z}_T(v)) \leq C \log^2 n) \geq 1 - C' e^{-c' \log^2 n} \quad (41)$$

for some  $C', c' > 0$ . Consider the CRW started at a fixed vertex  $w$  and run up to time  $T' = \lfloor T \rfloor \lfloor n^{\alpha/2} \rfloor$ . Divide the time interval  $[0, T']$  into  $k = \lfloor n^{\alpha/2} \rfloor$  intervals  $I_i = [S_i, S_{i+1})$  of length  $S = \lfloor T \rfloor$ . Observe now that for any  $v \in V$  by construction of the cyclic random walk we have  $\mathcal{Z}_{[S_i, S_{i+1})}(v) = \mathcal{Z}_{[0, S]}(\mathcal{X}_{S_i}(v))$ . Since  $S \leq T$  and the bound in (41) is uniform over all vertices, we obtain

$$\mathbb{P}(\forall_{i=1, \dots, k} \forall v \in V \iota(\mathcal{Z}_{[S_i, S_{i+1})}(v)) \leq C \log^2 n) \geq 1 - C' e^{-c' \log^2 n}.$$

for some  $C', c' > 0$ . Finally, by subadditivity of  $\iota$  we can bound  $\iota(\mathcal{Z}_{T'}(v))$  by the sum of  $\iota(\mathcal{Z}_{[S_i, S_{i+1})}(v))$  for  $i = 1, \dots, k$ , obtaining

$$\mathbb{P}(\forall v \in V \iota(\mathcal{Z}_{T'}(v)) \leq C \lfloor n^{\alpha/2} \rfloor \log^2 n) \geq 1 - C' e^{-c' \log^2 n}.$$

As  $C \lfloor n^{\alpha/2} \rfloor \log^2 n < n^\alpha$  for  $n$  large enough, we obtain that for some constants  $C_1, c_1 > 0$

$$\mathbb{P}(\forall v \in V \iota(\mathcal{Z}_{T'}(v)) \leq n^\alpha) \geq 1 - C_1 e^{-c_1 \log^2 n}.$$

Now an application of Lemma 3.2 gives for some constants  $C, C_2, c_2 > 0$

$$\mathbb{P}(\forall v \in V \iota(\mathcal{Z}_{T'}(v)) \leq C \log^2 n) \geq 1 - C_2 e^{-c_2 \log^2 n},$$

which finishes the inductive step.

Now, since we started from  $T = n^{\alpha/2}$  and at each step we increase the time by a factor of  $\lfloor n^{\alpha/2} \rfloor$ , after at most  $\lceil \frac{2}{\alpha} \rceil + 1$  steps we obtain

$$\mathbb{P}(\forall v \in V \iota(\mathcal{Z}_{n \log^2 n}(v)) \leq C \log^2 n) \geq 1 - C e^{-c \log^2 n}$$

for some constants  $C, c > 0$  depending on  $\alpha$ , but not on  $n$ . This finishes the proof.  $\square$

### 3.7. Isoperimetry lower bound

In this section we prove the isoperimetry lower bound given by Proposition 3.3. The proof is independent of the previous section (we will only make use of Lemma 3.6 and Lemma 3.8, which do not require Assumption 3.9).

**Proposition 3.27.** Fix  $\theta > 0$ , an admissible function  $\mathcal{C}$  and  $\beta_0, \beta_1 > 0$ . Consider  $\beta \in [\beta_0, \beta_1]$  and let  $\mathcal{X}(v) := \mathcal{X}^{\beta, \theta, \mathcal{C}}(v)$  be the cyclic random walk associated to  $\mu_{\beta, \theta, \mathcal{C}}$ , started at  $v$ . There exist  $C, c > 0$  (depending only on  $\beta_0, \beta_1, \theta$ ) such that

$$\mathbb{P}(\{|Z_{n \log^2 n}(v) \cap L_0| \leq c \log^2 n\} \cap \{T_{n \log^2 n}(v) < \infty\}) \leq C e^{-c \log^2 n}.$$

**Proof.** Let  $\kappa < 1/2\Theta\beta_1$  and  $T \geq \kappa n \log^2 n$ . Let  $Y_t = |Z_t \cap L_0|$  and let  $\lambda_t$  be the corresponding intensity of the process  $Y_t$ . Fix  $\delta < \beta_0 \Theta^{-1} \kappa / 2$ . We will first show that

$$\mathbb{P}(\{Y_T \leq \delta \log^2 n\} \cap \{\tau_c \geq T\}) \leq C e^{-c \log^2 n} \quad (42)$$

for some  $C, c > 0$ .

Consider  $\Lambda_t = \int_0^t \lambda_s ds$ . Observe that

$$\left\{ Y_T \leq \frac{\kappa}{2} \log^2 n \right\} \cap \{\tau_c \geq T\} \subset \left\{ \Lambda_T \geq \frac{\beta_0 \Theta^{-1} \kappa}{2} \log^2 n \right\}. \quad (43)$$

Indeed, up to time  $\tau_c$ , unless the CRW is at a vertex from a column containing an already visited vertex from  $L_0$ , the intensity of making a direct jump to  $L_0$  (and thus necessarily discovering a previously unvisited vertex from  $L_0$ ) is bounded from below by  $\frac{\beta \Theta^{-1}}{n-1} > \frac{\beta_0 \Theta^{-1}}{n}$  by Lemma 3.6. Note that there are at most  $n Y_T$  such bad vertices and until time  $\tau_c$  the CRW can spend time at most 1 at any given vertex. Thus on the event  $\{Y_T \leq \frac{\kappa}{2} \log^2 n\} \cap \{\tau_c \geq T\}$  the total time spent at bad vertices before  $T$  is at most  $n Y_T \leq \frac{\kappa}{2} n \log^2 n$ . We thus have

$$\Lambda_T = \int_0^T \lambda_t dt \geq \frac{\beta_0 \Theta^{-1}}{n} \left( T - \frac{\kappa}{2} n \log^2 n \right) \geq \frac{\beta_0 \Theta^{-1} \kappa}{2} \log^2 n,$$

proving (43). As  $\delta < \beta_0 \Theta^{-1} \kappa / 2$  we thus get

$$\mathbb{P}(\{Y_T \leq \delta \log^2 n\} \cap \{\tau_c \geq T\}) \leq \mathbb{P}\left(\{Y_T \leq \delta \log^2 n\} \cap \left\{ \Lambda_T \geq \frac{\beta_0 \Theta^{-1} \kappa}{2} \log^2 n \right\}\right)$$

and by Lemma 3.5 the right-hand side above is bounded by  $C e^{-c \log^2 n}$  for some  $C, c > 0$ . This proves (42).

Now we prove the statement of the proposition. Let  $M = n \log^2 n$  and  $A = \{|Z_M \cap L_0| \leq \delta \log^2 n\} \cap \{T_M < \infty\}$ , with the same  $\delta$  as above. We have

$$\begin{aligned} \mathbb{P}(A) &\leq \mathbb{P}(A \cap \{T_M \leq \kappa n \log^2 n\}) + \mathbb{P}(A \cap \{T_M > \kappa n \log^2 n\}) \\ &\leq \mathbb{P}(T_M \leq \kappa n \log^2 n) + \mathbb{P}(A \cap \{T_M > \kappa n \log^2 n\}). \end{aligned}$$

As  $1/\kappa > 2\Theta\beta_1 \geq 2\Theta\beta$ , by Lemma 3.8 the first term on the right hand side does not exceed  $e^{-c' \log^2 n}$  for some  $c' > 0$ . For the second term we write

$$\begin{aligned} \mathbb{P}(A \cap \{T_M > \kappa n \log^2 n\}) &= \mathbb{P}(\{|Z_M \cap L_0| \leq \delta \log^2 n\} \cap \{T_M < \infty\} \cap \{T_M > \kappa n \log^2 n\}) \\ &\leq \mathbb{P}(\{|Z_{\kappa n \log^2 n} \cap L_0| \leq \delta \log^2 n\} \cap \{\tau_c \geq \kappa n \log^2 n\}) \end{aligned}$$

and by (42) the right hand side is small enough, which finishes the proof.  $\square$

The proof of Proposition 3.3 is now rather straightforward.

**Proof of Proposition 3.3.** Fix a starting vertex  $v$ . Let  $\tau_0 = 0$  and let

$$\tau_k = \inf\{t > \tau_{k-1} : \mathcal{X}_t \in L_0, \mathcal{X}_t \notin \mathcal{Z}_{t-}\}$$

be the  $k$ -th time a new vertex from  $L_0$  is visited. Let  $J_t$  be the total number of fresh jumps made up to time  $t$ . By Lemma 3.6 the intensity  $\lambda$  of  $J$  is bounded from above by  $2\beta\Theta$ . By Lemma 3.4 we obtain

$$\mathbb{P}(J_{\tau_{k+1}} - J_{\tau_k} \geq 1 | \mathcal{F}_{\tau_k}) \mathbb{1}_{\{\tau_k < \infty\}} \leq \mathbb{P}(X \geq 1),$$

where  $X$  is a Poisson variable with parameter  $2\beta\Theta$ . In particular this implies that for some  $p > 0$  bounded away from 0 the following holds: each time the CRW visits a new vertex  $\mathcal{X}_{\tau_k} = w$  from  $L_0$ , with probability at least  $p$  it makes no fresh jumps for time 1, thus exhausting the whole bar of  $w$ . In particular with probability  $p > 0$  the vertex  $w$  enters the orbit  $\mathcal{O}(v)$ .

The rest of the proof is a rather standard concentration estimate. Let  $K = \lceil c \log^2 n \rceil$ , with  $c$  as in Proposition 3.27. Let  $\mathcal{H}_k = \mathcal{F}_{\tau_k}$  and

$$A_k = (\{\tau_k < \infty\} \cap \{\mathcal{X}_{\tau_k} \in \mathcal{O}(v)\}) \cup \{\tau_k = \infty\}.$$

By the argument above we have  $\mathbb{P}(A_k | \mathcal{H}_k) \geq p$ , which implies that for  $\xi_k := \mathbb{1}_{A_k} - p$  the sum  $M_N = \xi_1 + \dots + \xi_N$  (with  $M_0 = 0$ ) forms a submartingale with increments bounded by 1. As any submartingale can be written as a martingale plus a nonnegative predictable term and here the martingale part has bounded increments, by Azuma's inequality for martingales with bounded increments [24] we get

$$\mathbb{P}(M_N \leq -t) \leq e^{-\frac{t^2}{2N}}$$

for any  $t \geq 0$  and  $N \geq 1$ . Taking  $N = K$  and  $t = \frac{pK}{2}$  we obtain the estimate

$$\mathbb{P}\left(\mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_K} \leq \frac{pK}{2}\right) \leq e^{-\frac{p^2 K}{8}}.$$

Together with Proposition 3.27 this implies that with probability at least  $1 - C'e^{-c' \log^2 n}$  (for some  $C', c' > 0$ ) we have either

- $T_{n \log^2 n}(v) = \infty$ , which implies  $|\mathcal{O}(v)| < n \log^2 n$ ,
- or  $\tau_K \leq T_{n \log^2 n} < \infty$  and  $\sum_{k=1}^K \mathbb{1}_{A_k} \geq \frac{pK}{2}$ .

In the latter case we get that the event  $\{\tau_k < \infty\} \cap \{\mathcal{X}_{\tau_k} \in \mathcal{O}(v)\}$  holds at least  $\frac{cp}{2} \log^2 n$  times for  $k \leq K$ , implying in particular  $|\mathcal{O}_{n \log^2 n}(v) \cap L_0| \geq c'' \log^2 n$  for some  $c'' > 0$  (note that for  $k \geq 1$  if  $T_k < \infty$ , then  $T_k \leq k - 1$ ).

To finish the proof we note that by symmetry of  $H_n$  the above argument is valid with  $L_0$  replaced by any other row or column (as Proposition 3.27 has the same symmetry) and the starting vertex  $v$  was arbitrary. Thus by performing a union bound over starting vertices  $v \in V$  and all rows and columns we obtain the desired bound on  $\chi(\mathcal{O}_{n \log^2 n}(v))$ .  $\square$

#### 4. General transposition processes

In this section we introduce the notion of a general transposition process, which is another perspective on the permutation model  $\mu_{\beta, \theta, \mathcal{C}}$  defined in (1) and (3). In this formalism we will state Lemma 4.1 and Proposition 4.2. These are all prerequisites needed in the next, final section.

For  $X \in \mathfrak{X}$  let  $(e_1, t_1), (e_2, t_2), \dots, (e_{|X|}, t_{|X|})$  be the points of  $X$  sorted by the second coordinate. We define  $\{\sigma_t\}_{t \in \{0, 1, \dots, |X|\}}$  by  $\sigma_0 := \text{id}$  and for  $t \in \{1, \dots, |X|\}$

$$\sigma_t := e_t \circ \dots \circ e_1,$$

where any edge is identified with the transposition of its endpoints. This becomes a stochastic process when  $X$  is sampled according to  $\mu_{\beta, \theta, \mathcal{C}}$ . Note that  $\sigma_{|X|} = \sigma(X)$ , with the latter permutation defined in (1).

In what follows we will consider  $(\sigma_i)_{i=0}^k$  conditionally on  $\Omega_k := \{|X| = k\}$ . We use  $(\mathcal{F}_i)_{i=0}^k$  to denote the filtration on  $\Omega_k$  associated with the process.

Obviously this transposition model depends on the parameters of  $\mu_{\beta, \theta, \mathcal{C}}$ . What might be surprising is that its evolution is not far from the i.i.d. transposition process on the edges of  $H_n$ . This observation will play a crucial role in the proof of forthcoming Proposition 4.2.

**Lemma 4.1.** *Let  $\beta, \theta > 0$  and  $\mathcal{C}$  be an admissible function. Let  $X$  be sampled from  $\mu_{\beta, \theta, \mathcal{C}}$  and let  $\{\sigma_t\}_{t \geq 0}$  be the associated transposition process. For any  $i, k \in \mathbb{N}$ ,  $i < k$ ,  $e \in E$  we have*

$$\mathbb{P}(\sigma_{i+1} \circ \sigma_i^{-1} = e | \mathcal{F}_i, |X| = k) \in \left[ \frac{\Theta^{-2}}{|E|}, \frac{\Theta^2}{|E|} \right]. \quad (44)$$

**Proof.** Let us fix  $e \in E$  and a sequence of permutations  $\sigma = (\sigma_1, \dots, \sigma_i)$  such that for any  $j \in \{1, \dots, i-1\}$  the composition  $\sigma_{j+1} \circ \sigma_j^{-1}$  is a transposition. Let  $U_{\sigma,e}^{i+1}$  be the set of all  $X \in \mathfrak{X}$  such that  $|X| = k$ ,  $e_{i+1} = e$  and the transposition process associated to  $X$  agrees with  $\sigma$  up to time  $i$ . For any  $\tilde{e} \in E$  let  $T : U_{\sigma,e}^{i+1} \mapsto U_{\sigma,\tilde{e}}^{i+1}$  be the mapping which swaps the  $(i+1)$ -th point of  $X$  from  $e$  to  $\tilde{e}$ , i.e.,  $(e_{i+1}, t_{i+1}) = (e, t_{i+1})$  in  $X$  is replaced by  $(\tilde{e}, t_{i+1})$ . Note that  $T$  is a bijection which preserves the Poisson point process  $\mathcal{B}$ . Thus we have

$$\frac{\mu_{\beta,\theta,\mathcal{C}}(U_{\sigma,\tilde{e}}^{i+1})}{\mu_{\beta,\theta,\mathcal{C}}(U_{\sigma,e}^{i+1})} = \frac{\int_{\mathfrak{X}} \mathbb{1}_{U_{\sigma,\tilde{e}}^{i+1}}(X) \theta^{\mathcal{C}(X)} \mathcal{B}(dX)}{\int_{\mathfrak{X}} \mathbb{1}_{U_{\sigma,e}^{i+1}}(X) \theta^{\mathcal{C}(X)} \mathcal{B}(dX)} = \frac{\int_{\mathfrak{X}} \mathbb{1}_{U_{\sigma,e}^{i+1}}(X) \theta^{\mathcal{C}(T(X))} \mathcal{B}(dX)}{\int_{\mathfrak{X}} \mathbb{1}_{U_{\sigma,e}^{i+1}}(X) \theta^{\mathcal{C}(X)} \mathcal{B}(dX)}.$$

By the Lipschitz property (2) of  $\mathcal{C}$  clearly we have  $|\mathcal{C}(T(X)) - \mathcal{C}(X)| \leq 2$ . Thus the integrands in the numerator and the denominator above can differ by a factor of at most  $\theta^2$ , which leads to the estimate

$$\mu_{\beta,\theta,\mathcal{C}}(U_{\sigma,\tilde{e}}^{i+1}) / \mu_{\beta,\theta,\mathcal{C}}(U_{\sigma,e}^{i+1}) \leq \Theta^2.$$

From this and an analogous argument for the lower bound it is straightforward to obtain (44).  $\square$

In the final arguments we will need statements holding uniformly in a large enough time window before time  $|X|$ . Thus, conditionally on  $|X| = k$ , we let  $k' = \max\{0, k - 2\lceil n^{11/6} \rceil\}$  and define the time interval  $I := \{k', \dots, k-1\}$ . We will now prove that on  $I$  the transposition process corresponding to the measure  $\mu_{\beta,\theta,\mathcal{C}}$  behaves in a certain sense similarly to the mean-field case (corresponding to  $\theta = 1$  on the complete graph), i.e., for most values of  $k$  the process, when conditioned on  $|X| = k$ , with high probability has splitting and merging probabilities comparable to the mean-field case.

To formalize this intuition let us introduce a stopping time  $\tau$ , corresponding to the moment when the cycles lose good isoperimetric properties. Denote by  $\text{orb}_\sigma(v)$  the cycle of the permutation  $\sigma$  containing  $v \in V$  and by  $\text{orb}_\sigma^\ell(v)$  its first  $\ell \wedge |\sigma|$  elements. We will write  $\text{orb}_i^\ell(v)$  as a shorthand for  $\text{orb}_{\sigma_i}^\ell(v)$ .

For constants  $c_1, c_2 > 0$  which will be fixed later we define

$$\begin{aligned} \tau^\ell &:= \inf\{s \in I : \exists v \in V \ell(\text{orb}_s^{n \log^2 n}(v)) > c_1 \log^2 n\}, \\ \tau^\chi &:= \inf\{s \in I : \exists v \in V |\text{orb}_s(v)| \geq n \log^2 n \text{ and } \chi(\text{orb}_s^{n \log^2 n}(v)) < c_2 \log^2 n\}, \\ \tau &:= \tau^\ell \wedge \tau^\chi. \end{aligned}$$

Note that for each  $k$ ,  $\tau$  is a stopping time with respect to the natural filtration of the process  $(\sigma_t)_{t=0}^k$  on  $\{|X| = k\}$ .

**Proposition 4.2.** *Let  $\beta, \theta > 0$  be such that  $\beta > \Theta/2$  and let  $\mathcal{C}$  be an admissible function. Let  $X$  be sampled from  $\mu_{\beta,\theta,\mathcal{C}}$  and let  $\{\sigma_t\}_{t \in \{0, \dots, |X|\}}$  be the associated transposition process.*

*Then there exists  $K_n \subset \mathbb{N}$ ,  $C, c > 0$  and constants  $c_1, c_2$  in the definition of  $\tau$ , depending only on  $\beta, \theta$  and  $\mathcal{C}$ , such that the following properties hold.*

- (i)  $\mathbb{P}(|X| \in K_n) \geq 1 - Ce^{-c \log^2 n}$
- (ii) For  $k \in K_n$ ,  $\mathbb{P}(\tau = \infty | |X| = k) \geq 1 - Ce^{-c \log^2 n}$ .
- (iii) Consider  $\ell \geq n \log^2 n$ . Denote by  $\mathcal{D}_i$  the event that in the transition from step  $i$  to  $i+1$  a cycle of  $\sigma_i$  is split into two cycles, one of which has size smaller than  $\ell$ . Then for every  $i \in I$ ,

$$\mathbb{P}(\mathcal{D}_i | \mathcal{F}_i, |X| = k) \mathbb{1}_{\{k \in K_n\}} \mathbb{1}_{\{\tau > i\}} \leq C \frac{\ell}{n^2}.$$

- (iv) Let  $\mathcal{C}_1, \mathcal{C}_2$  be two cycles of  $\sigma_i$  such that  $|\mathcal{C}_j| \geq n \log^2 n$ , for  $j \in \{1, 2\}$ . Denote by  $\mathcal{M}_i$  the event that they are merged in the transition from step  $i$  to  $i+1$ . Then for every  $i \in I$ ,

$$\mathbb{P}(\mathcal{M}_i | \mathcal{F}_i, |X| = k) \geq c \frac{|\mathcal{C}_1| |\mathcal{C}_2|}{n^4} \mathbb{1}_{\{k \in K_n\}} \mathbb{1}_{\{\tau > i\}}. \quad (45)$$

The proof is deferred to the end of the next subsection.



#### 4.1. Isoperimetry and its consequences

Here we use the notation from the previous section and assume that  $\{\sigma_t\}_{t \in \{0, \dots, |X|\}}$  is a transposition process associated to the distribution  $\mu_{\beta, \theta, \mathcal{C}}$ . The proof of Proposition 4.2 is rather easy once we know that any (long enough) fragment of  $\sigma_t$  is “spread evenly on the graph”. This is formalised in events  $\mathcal{I}^t$  and  $\mathcal{I}^X$  defined below.

Recall that  $I = \{k', \dots, k-1\}$ , where  $k' = \max\{0, k - 2\lceil n^{11/6} \rceil\}$ , conditionally on  $|X| = k$ . Let  $c_1, c_2 > 0$  and consider the events

$$\mathcal{I}^t := \{\tau^t \geq |X|\} = \{\forall s \in I \forall v \in V \iota(\text{orb}_s^{n \log^2 n}(v)) \leq c_1 \log^2 n\}, \quad (46)$$

$$\mathcal{I}^X := \{\tau^X \geq |X|\} = \{\forall s \in I \forall v \in V \chi(\text{orb}_s^{n \log^2 n}(v)) \geq c_2 \log^2 n \mathbb{1}_{\{|\text{orb}_s(v)| \geq n \log^2 n\}}\}. \quad (47)$$

Finally, let  $\mathcal{I} := \mathcal{I}^t \cap \mathcal{I}^X$ .

We will show that the event  $\mathcal{I}$  holds (for appropriate choice of  $c_1, c_2$ ) with high probability. As a first step we prove that the conclusions of Proposition 3.1 and Proposition 3.3 hold uniformly for cyclic random walks using their bars only up to a certain level.

More precisely, for  $s, t \in (0, 1)$  let  $\mathfrak{X}^{s,t}$  be the restriction of  $\mathfrak{X}$  to  $[s, t]$ , i.e., the space of finite subsets of  $E \times [s, t]$ . We can define measures  $\mu_{\beta, \theta, \mathcal{C}}^{s,t}$  on  $\mathfrak{X}^{s,t}$  by a formula analogous to (3), i.e.,

$$\mu_{\beta, \theta, \mathcal{C}}^{s,t}(U) := \frac{1}{Z_{\beta, \theta, \mathcal{C}}^{s,t}} \int_{\mathfrak{X}^{s,t}} \mathbb{1}_U(Y) \theta^{\mathcal{C}(Y)} \mathcal{B}(dY). \quad (48)$$

If  $X \in \mathfrak{X}$  and  $X = \{(e_1, t_1), \dots, (e_k, t_k)\}$ , we define the restriction of  $X$  to  $[s, t]$ , denoted by  $X^{s,t}$ , by including only these pairs  $(e_i, t_i)$  for which  $t_i \in [s, t]$ . We have  $X^{s,t} \in \mathfrak{X}^{s,t}$ . For simplicity we will write  $\mathfrak{X}^t$  instead of  $\mathfrak{X}^{0,t}$ , etc.

For  $Y \in \mathfrak{X}^t$  let  $\mathcal{Z}^Y(v)$  denote the trace of the cyclic random walk (started at  $(v, 0)$ ) using the bridges of  $Y$  and running on bars of height  $t$  instead of height 1. Note that if  $Y$  is distributed according to  $\mu_{\beta, \theta, \mathcal{C}}^t$  (for some  $\beta, \theta, \mathcal{C}$ ), then the process  $(\mathcal{Z}_T^Y(v), T \geq 0)$  has the same distribution as  $(\mathcal{Z}_{T/t}^X(v), T \geq 0)$ , where  $X \in \mathfrak{X}$  is distributed according to  $\mu_{\beta, \theta, \mathcal{C}}$  (this follows directly by properties of the Poisson point process  $\mathcal{B}$ ).

Furthermore, if  $X$  is distributed according to  $\mu_{\beta, \theta, \mathcal{C}}$ , then the law of  $X^t$  under  $\mu_{\beta, \theta, \mathcal{C}}$  is given by  $\mu_{\beta, \theta, \mathcal{C}^t}^t$ , where  $\mathcal{C}^t : \mathfrak{X}^t \mapsto \mathbb{R}$  is a function defined by

$$\theta^{\mathcal{C}^t(Y)} := \int_{\mathfrak{X}} \theta^{\mathcal{C}(Y \cup Z|_{E \times [t, 1]})} \mathcal{B}(dZ),$$

where  $Z|_{E \times [t, 1]} = Z \cap (E \times [t, 1])$ . It is easy to check that if  $\mathcal{C}$  is admissible, then so is  $\mathcal{C}^t$ .

Let us define the analogues of (46) and (47) for the CRW. Recall the sets  $\mathcal{O}_k, \mathcal{O}$ , defined in (4), and denote by  $\mathcal{O}_k^Y, \mathcal{O}^Y$  the analogous sets for the process  $Y$ . For fixed  $t_0 \in (0, 1)$  and  $c_1, c_2 > 0$  we set

$$\mathcal{A}^t := \{\forall t \in [t_0, 1) \forall v \in V \iota(\mathcal{Z}_{tn \log^2 n}^{X^t}(v)) \leq c_1 \log^2 n\},$$

$$\mathcal{A}^X := \{\forall t \in [t_0, 1) \forall v \in V \chi(\mathcal{O}_{n \log^2 n}^{X^t}(v)) \geq c_2 \log^2 n \mathbb{1}_{\{|\mathcal{O}^{X^t}(v)| \geq n \log^2 n\}}\}.$$

We also set  $\mathcal{A} := \mathcal{A}^t \cap \mathcal{A}^X$ . We will now prove

**Proposition 4.3.** *Fix  $\beta, \theta, \mathcal{C}$  such that  $\beta > \Theta/2$  and let  $t_0 \in (0, 1)$  be such that  $t_0 \beta > \Theta/2$ . Let  $X$  be distributed according to  $\mu_{\beta, \theta, \mathcal{C}}$ . Then there exist  $C, c > 0$  and  $c_1, c_2 > 0$  in the definition of  $\mathcal{A}^t$  and  $\mathcal{A}^X$  above such that*

$$\mathbb{P}(\mathcal{A}) \geq 1 - C e^{-c \log^2 n}.$$

**Proof.** It will be convenient to divide the time interval  $[0, 1)$  into subintervals small enough so that each of them contains at most one bridge, as then it will be enough to control isoperimetry at the endpoints and use a union bound.

Let  $\kappa > 0$  (to be specified later in the proof). Let  $t_0 = s_0 < s_1 < \dots < s_k = 1$  be such that  $k = \lceil \beta n^2 e^{\kappa \log^2 n} \rceil$  and  $|s_{i+1} - s_i| \leq \beta^{-1} n^{-2} e^{-\kappa \log^2 n}$  for  $i \in \{0, 1, \dots, k-1\}$ . For simplicity we will write  $X^i := X^{s_i}$  and  $X^{i,i+1} := X^{s_i, s_{i+1}}$  (and likewise for  $\mathfrak{X}$ ).

Consider the event that there is at most one bridge in each interval  $[s_i, s_{i+1})$

$$\mathcal{E} := \{\forall i \in \{0, \dots, k-1\} |X^{i,i+1}| \leq 1\}.$$

First note that

$$\mathbb{P}(\mathcal{E}^c) \leq C e^{-c \log^2 n} \quad (49)$$

for some  $C, c > 0$ . Indeed, for any fixed  $i = 0, \dots, k-1$  by Lemma B.2 applied with  $t = s_i, s = |s_{i+1} - s_i|$  we have

$$\mathbb{P}(|X^{i,i+1}| > 1) \leq e^{C\lambda_i} \mathbb{P}(Y_{\lambda_i} \geq 2),$$

for some  $C > 0$ , where  $\lambda_i \leq \Theta e^{-\kappa \log^2 n}$  and  $Y_{\lambda_i}$  is a Poisson variable with parameter  $\lambda_i$ . Using the simple estimate  $\mathbb{P}(Y_{\lambda_i} \geq 2) \leq \lambda_i^2$  (valid for  $\lambda_i$  small enough) we obtain

$$\mathbb{P}(|X^{i,i+1}| > 1) \leq e^{C\Theta e^{-\kappa \log^2 n}} \Theta^2 e^{-2\kappa \log^2 n} \leq C' e^{-2\kappa \log^2 n}$$

for some  $C' > 0$ . Applying union bound over all  $i = 0, \dots, k-1$ , gives  $\mathbb{P}(\mathcal{E}) \leq k \cdot C' e^{-2\kappa \log^2 n} \leq 2C' \beta n^2 e^{-\kappa \log^2 n} \leq C e^{-c \log^2 n}$  for some  $C, c > 0$ .

Now we prove that each of the events defining  $\mathcal{A}$  holds with high enough probability. We start with the event  $\mathcal{A}^t$ . Recall that  $\mathcal{Z}^{X^t}(v)$  is the trace of the CRW using bridges of  $X^t$ . Consider the event

$$\mathcal{J} := \left\{ \forall_{i \in \{0, 1, \dots, k\}} \forall_{v \in V} \iota(\mathcal{Z}_{s_i n \log^2 n}^{X^{s_i}}(v)) \leq C \log^2 n \right\}$$

with  $C$  as in Proposition 3.1.

As remarked before, for any fixed  $i \in \{0, 1, \dots, k\}$  the law of  $X^{s_i}$  under  $\mu_{\beta, \theta, \mathcal{C}}$  is given by  $\mu_{\beta_i, \theta, \mathcal{C}^{s_i}}^{s_i}$  and the trace  $\mathcal{Z}_T^{X^{s_i}}(v)$  has the same distribution as the trace  $\mathcal{Z}_{T/s_i}^{\tilde{X}^i}(v)$ , where  $\tilde{X}^i$  has distribution  $\mu_{\beta_i, \theta, \mathcal{C}^{s_i}}$  with  $\beta_i = s_i \beta$  (note that the latter CRW uses bars of height 1). Note that  $\beta_i > \Theta/2$ , as by assumption  $t_0 \beta > \Theta/2$ , and we have  $\beta_i \in [t_0 \beta, \beta]$ .

Therefore, by Proposition 3.1 we obtain for any fixed  $i$

$$\mathbb{P}(\forall_{v \in V} \iota(\mathcal{Z}_{s_i n \log^2 n}^{X^{s_i}}(v)) \leq C \log^2 n) \geq 1 - C e^{-c \log^2 n},$$

with constants  $C, c$  depending only on  $\Theta, \beta$  and  $t_0$  (but not on  $\mathcal{C}^{s_i}$ ).

Applying a union bound over  $i = 0, \dots, k$  we get

$$\mathbb{P}(\mathcal{J}^c) \leq (k+1) C e^{-c \log^2 n} \leq 4C \beta n^2 e^{\kappa \log^2 n} e^{-c \log^2 n}.$$

Now we can fix  $\kappa < c/2$  to obtain that  $\mathbb{P}(\mathcal{J}) \geq 1 - C' e^{-c' \log^2 n}$  for some  $C', c' > 0$ .

On the event  $\mathcal{E}$  there is at most one bridge in each interval  $[s_i, s_{i+1})$ , which implies for any  $v \in V$

$$\sup_{t \in [s_i, s_{i+1})} \iota(\mathcal{Z}_{t n \log^2 n}^{X^t}(v)) \leq \max\{\iota(\mathcal{Z}_{s_i n \log^2 n}^{X^{s_i}}(v)), \iota(\mathcal{Z}_{s_{i+1} n \log^2 n}^{X^{s_{i+1}}}(v))\}.$$

Thus on the event  $\mathcal{E} \cap \mathcal{J}$  we obtain

$$\sup_{t \in [t_0, 1)} \max_{v \in V} \iota(\mathcal{Z}_{t n \log^2 n}^{X^t}(v)) \leq C \log^2 n.$$

Since each event  $\mathcal{E}$  and  $\mathcal{J}$  occurs with high enough probability, we can take  $c_1 = C$  to obtain that  $\mathcal{A}^t$  holds with the required probability.

The proof for the event  $\mathcal{A}^X$  is analogous. Using the same notation as above, let

$$\mathcal{K} := \left\{ \forall_{i \in \{0, 1, \dots, k\}} \forall_{v \in V} \chi(\mathcal{O}_{n \log^2 n}^{X^{s_i}}(v)) \geq C \log^2 n \text{ or } |\mathcal{O}^{X^{s_i}}(v)| < n \log^2 n \right\}$$

with  $C$  as in Proposition 3.3.

As before,  $(\mathcal{Z}_T^{X^{s_i}}(v), T \geq 0)$  has the same distribution as  $(\mathcal{Z}_{T/s_i}^{\tilde{X}^i}(v), T \geq 0)$ , where  $\tilde{X}^i$  is distributed according to  $\mu_{\beta_i, \theta, \mathcal{C}^{s_i}}$  with  $\beta_i = s_i \beta$ . Applying Proposition 3.3 and performing a union bound over  $i$  gives us that  $\mathbb{P}(\mathcal{K}) \geq 1 - C' e^{-c' \log^2 n}$  for some  $C', c' > 0$  depending on  $\theta, \beta, t_0$ .

On the event  $\mathcal{E}$  for each  $t \in [s_i, s_{i+1})$  we have either  $\mathcal{O}_{t n \log^2 n}^{X^t}(v) = \mathcal{O}_{s_i n \log^2 n}^{X^{s_i}}(v)$  and  $\mathcal{O}^{X^t}(v) = \mathcal{O}^{X^{s_i}}(v)$ , or  $\mathcal{O}_{t n \log^2 n}^{X^t}(v) = \mathcal{O}_{s_{i+1} n \log^2 n}^{X^{s_{i+1}}}(v)$  and  $\mathcal{O}^{X^t}(v) = \mathcal{O}^{X^{s_{i+1}}}(v)$ . Thus  $\mathcal{E} \cap \mathcal{K} \subseteq \mathcal{A}^X$ . We finish by observing that both  $\mathcal{E}$  and  $\mathcal{K}$  hold with high enough probability, so in the definition of  $\mathcal{A}^X$  we can take  $c_2 = C$  with  $C$  as in the definition of  $\mathcal{K}$ .  $\square$

Now we can prove that the good isoperimetric properties  $\mathcal{I}^t$  and  $\mathcal{I}^X$  defined in (46) and (47) hold in the discrete time setting (for some  $c_1, c_2$ ) with high probability. Recall that  $\mathcal{I} = \mathcal{I}^t \cap \mathcal{I}^X$ .

**Lemma 4.4.** *Let  $\beta, \theta > 0$  be such that  $\beta > \Theta/2$  and  $C$  be an admissible function. There exist  $c_1, c_2 > 0$  in the definitions (46) and (47),  $C, c, c' > 0$  and  $K_n \subset \mathbb{N}$  such that*

$$\mathbb{P}(\mathcal{I}||X| = k) \geq (1 - Ce^{-c \log^2 n}) \mathbb{1}_{\{k \in K_n\}} \quad \text{and} \quad \mathbb{P}(|X| \in K_n) \geq 1 - e^{-c' \log^2 n}.$$

**Proof.** Let  $X$  be sampled from  $\mu_{\beta, \theta, C}$  and fix  $t_0 \in (0, 1)$  such that  $t_0 \beta > \Theta/2$ .

We first prove that the event  $\mathcal{I} = \mathcal{I}^t \cap \mathcal{I}^X$  holds with probability at least  $1 - Ce^{-c \log^2 n}$ , for some  $C, c > 0$ , when  $c_1, c_2$  are chosen appropriately.

Consider the bridges  $\sigma_s = (e_s, t_s)$  for  $s = |X| - 2\lceil n^{11/6} \rceil, \dots, |X|$ . The event

$$\mathcal{S} = \{|X| \geq 2\lceil n^{11/6} \rceil\} \cap \{\forall_{s \in \{|X| - 2\lceil n^{11/6} \rceil, \dots, |X|\}} t_s > t_0\}$$

holds with high probability. Indeed, if  $|X| < 2\lceil n^{11/6} \rceil$  or  $t_s \leq t_0$  for some  $s$ , then necessarily  $|X \cap (E \times [t_0, 1])| \leq 2\lceil n^{11/6} \rceil$ . An application of the second part of Lemma B.2 with  $k_n = 2\lceil n^{11/6} \rceil$  shows that with probability at least  $1 - Ce^{-cn^2}$ , for some  $C, c > 0$ , this does not happen.

From now on we work on the event  $\mathcal{S}$ . As in the proof of Proposition 4.3 let  $X^{t_s}$  denote the restriction of  $X$  to the interval  $[0, t_s)$  and let  $\mathcal{Z}^{X^{t_s}}(v)$  be the trace of the corresponding cyclic random walk started at  $v$ .

By the construction of the cyclic random walk we have  $\text{orb}_s(v) \subseteq \mathcal{Z}^{X^{t_s}}(v)$  and  $\text{orb}_s^{n \log^2 n}(v) \subseteq \mathcal{Z}_{t_s n \log^2 n}^{X^{t_s}}(v)$ . Recalling the definition of the event  $\mathcal{A}^t$ , by Proposition 4.3 there exist  $C, c > 0$  such that with probability at least  $1 - Ce^{-c \log^2 n}$  we have

$$\sup_{t \in [t_0, 1)} \max_{v \in V} \iota(\mathcal{Z}_{t n \log^2 n}^{X^t}(v)) \leq C \log^2 n.$$

As on  $\mathcal{S}$  we have  $t_s > t_0$  for all  $s = |X| - 2\lceil n^{11/6} \rceil, \dots, |X|$ , together with the observation about the orbits this shows that  $\mathcal{I}^t$ , with  $c_1 = C$  in the definition (46), holds with probability at least  $1 - C'e^{-c' \log^2 n}$  for some  $C', c' > 0$ .

For the proof that the event  $\mathcal{I}^X$  holds with high probability, note that  $\mathcal{O}_{t_s n \log^2 n}^{X^{t_s}}(v) = \text{orb}_s^{n \log^2 n}(v)$ , in particular  $\text{orb}_s(v) = \mathcal{O}^{X^{t_s}}(v)$ . Therefore, recalling the definition of  $\mathcal{A}^X$ , we can use Proposition 4.3 to conclude that with probability at least  $1 - C'e^{-c' \log^2 n}$ , for some  $C', c' > 0$ , we have

$$\forall s \in I \quad \forall v \in V \quad \chi(\text{orb}_s^{n \log^2 n}(v)) \geq C \log^2 n \mathbb{1}_{\{|\text{orb}_s(v)| \geq n \log^2 n\}}.$$

As before on the event  $\mathcal{S}$  we have  $t_s > t_0$ , so if we take  $c_2 = C$  in (47) we obtain that the event  $\mathcal{I}^X$  holds with probability at least  $1 - C'e^{-c' \log^2 n}$  for some  $C', c' > 0$ .

Now suppose that  $C, c > 0$  are such that  $\mathbb{P}(\mathcal{I}) \geq 1 - Ce^{-c \log^2 n}$ . Let

$$K_n := \{k \in \mathbb{N} : \mathbb{P}(\mathcal{I}||X| = k) \geq 1 - Ce^{-\frac{c}{2} \log^2 n}\}.$$

We write by definition of  $K_n$

$$\begin{aligned} \mathbb{P}(\mathcal{I}) &= \sum_{k \in K_n} \mathbb{P}(\mathcal{I}^c||X| = k) \mathbb{P}(|X| = k) + \sum_{k \notin K_n} \mathbb{P}(\mathcal{I}^c||X| = k) \mathbb{P}(|X| = k) \\ &\geq \mathbb{P}(|X| \notin K_n) \cdot Ce^{-\frac{c}{2} \log^2 n} \end{aligned}$$

and now the lower bound on  $\mathbb{P}(\mathcal{I})$  together with a simple calculation gives us

$$\mathbb{P}(|X| \in K_n) \geq 1 - \frac{Ce^{-c \log^2 n}}{Ce^{-\frac{c}{2} \log^2 n}} = 1 - e^{-\frac{c}{2} \log^2 n}$$

as desired, which proves the second assertion of the lemma with  $c' = c/2$ .  $\square$

We will now relate good isoperimetric properties of cycles to the probabilities of splits and merges in the corresponding transposition process.

**Lemma 4.5.** *Let  $\sigma$  be a permutation and let  $e = \{u, w\}$  be an edge chosen at random according to a distribution  $\{p_e\}_{e \in E}$  satisfying  $c/|E| \leq p_e \leq C/|E|$ , for some  $C, c > 0$ . Let  $(u, w)$  denote the transposition of endpoints of  $e$ .*

(i) *Suppose that for some  $k, m \in \mathbb{N}$  and each  $v$  we have*

$$\iota(\text{orb}_\sigma^k(v)) \leq m. \quad (50)$$

*Then for any  $\ell \geq k$  the probability that a cycle of  $\sigma$  is split in  $(u, w) \circ \sigma$  into two cycles, one of which has size smaller than  $\ell$ , is at most*

$$\frac{4C\ell}{kn}m.$$

(ii) *Suppose that for some  $k, m \in \mathbb{N}$  and each  $v$  satisfying  $|\text{orb}_\sigma(v)| \geq k$  we have*

$$\chi(\text{orb}_\sigma^k(v)) \geq m. \quad (51)$$

*Then given two cycles  $\mathcal{C}_1, \mathcal{C}_2$  of  $\sigma$  of length at least  $k$ , the probability that they are merged in  $(u, w) \circ \sigma$  into one cycle is at least*

$$\frac{c}{2} \frac{|\mathcal{C}_1||\mathcal{C}_2|m^2}{n^2k^2}.$$

**Proof.** We start with the proof of (i). Fix a vertex  $v \in V$ , let  $D, L$  denote respectively the column and the row containing  $v$ , and let  $\text{orb}_\sigma^{-\ell}(v) = \text{orb}_{\sigma^{-1}}^\ell(v)$ . The number of  $w \in V$  for which  $(v, w)$  is an edge such that a cycle of  $\sigma$  is split in  $(v, w) \circ \sigma$  into two cycles, one of which has size smaller than  $\ell$ , is equal to

$$|(\text{orb}_\sigma^\ell(v) \cup \text{orb}_\sigma^{-\ell}(v)) \cap (D \cup L \setminus \{v\})| \leq 2\iota(\text{orb}_\sigma^\ell(v) \cup \text{orb}_\sigma^{-\ell}(v)).$$

By dividing  $\text{orb}_\sigma^\ell(v) \cup \text{orb}_\sigma^{-\ell}(v)$  into pieces of length  $k$  and exploiting subadditivity of  $\iota$  we obtain

$$\iota(\text{orb}_\sigma^\ell(v) \cup \text{orb}_\sigma^{-\ell}(v)) \leq \left\lceil \frac{2\ell}{k} \right\rceil \max_{v' \in V} \iota(\text{orb}_\sigma^k(v')) \leq \left\lceil \frac{2\ell}{k} \right\rceil m.$$

Thus for fixed  $v$  there are at most  $\lceil \frac{2\ell}{k} \rceil m$  edges with one endpoint equal to  $v$  which would cause a cycle of  $\sigma$  to split with one of the resulting pieces smaller than  $\ell$ . By our assumptions each such edge  $e$  is chosen with probability  $p_e \leq \frac{C}{|E|}$ . As each vertex has degree  $2(n-1)$  in  $H_n$ , we obtain that the total probability of such a split is at most

$$\frac{C}{n-1} \left\lceil \frac{2\ell}{k} \right\rceil m \leq \frac{4C\ell}{kn}m,$$

as desired.

For the proof of (ii), consider two cycles  $\mathcal{C}_1 = \text{orb}_\sigma(v_1), \mathcal{C}_2 = \text{orb}_\sigma(v_2)$  of length at least  $k$ . By the assumption of the lemma we have  $\chi(\text{orb}_\sigma^k(v_j)) \geq m$  for  $j = 1, 2$ . By dividing  $\mathcal{C}_j$  into segments of length  $k$  and recalling the definition of  $\chi$ , we obtain that each  $\mathcal{C}_j$  has at least  $\lfloor \frac{|\mathcal{C}_j|}{k} \rfloor m \geq \frac{1}{2} \frac{|\mathcal{C}_j|m}{k}$  vertices in each row and each column of  $H_n$ . This implies that there are at least  $2n(\frac{1}{2} \frac{|\mathcal{C}_1|m}{k})(\frac{1}{2} \frac{|\mathcal{C}_2|m}{k}) = \frac{n}{2} \frac{|\mathcal{C}_1||\mathcal{C}_2|m^2}{k^2}$  edges joining a vertex from  $\mathcal{C}_1$  with a vertex from  $\mathcal{C}_2$ .

For each such an edge  $e$  we have  $p_e \geq \frac{c}{|E|}$ . As choosing such an edge results in a merge between  $\mathcal{C}_1$  and  $\mathcal{C}_2$ , we obtain

$$\mathbb{P}(\mathcal{C}_1 \text{ and } \mathcal{C}_2 \text{ are merged in } (u, w) \circ \sigma) \geq \frac{c}{n^2(n-1)} \cdot \frac{n}{2} \frac{|\mathcal{C}_1||\mathcal{C}_2|m^2}{k^2} \geq \frac{c}{2} \frac{|\mathcal{C}_1||\mathcal{C}_2|m^2}{n^2k^2},$$

as desired. □

Now we can finally employ Lemmas 4.1, 4.4 and 4.5 to prove Proposition 4.2.

**Proof of Proposition 4.2.** Let  $K_n \subset \mathbb{N}$  and  $C, c, c' > 0$  be as in Lemma 4.4 (in particular (i) of Proposition 4.2 is satisfied). We have

$$\mathbb{P}(\mathcal{I}||X| = k) \geq (1 - Ce^{-c \log^2 n}) \mathbb{1}_{\{k \in K_n\}},$$

which in particular implies (ii), i.e.,

$$\mathbb{P}(\tau = \infty | |X| = k) \geq (1 - Ce^{-c \log^2 n}) \mathbb{1}_{\{k \in K_n\}}. \quad (52)$$

Fix  $i \in I$  and let  $\mathcal{C}_1 = \text{orb}_i(v_1)$ ,  $\mathcal{C}_2 = \text{orb}_i(v_2)$  for some  $v_1, v_2 \in V$ . Let  $\ell \geq n \log^2 n$ . As on the event  $\{\tau > i\}$  we have  $\iota(\text{orb}_i^{n \log^2 n}(v)) \leq c_1 \log^2 n$  for all  $v$ , by Lemma 4.1 and part (i) of Lemma 4.5 we obtain

$$\mathbb{P}(\mathcal{D}_i | \mathcal{F}_i, |X| = k) \mathbb{1}_{\{k \in K_n\}} \mathbb{1}_{\{\tau > i\}} \leq \frac{4\Theta^{-1} \ell}{n \cdot n \log^2 n} c_1 \log^2 n = C' \frac{\ell}{n^2}$$

for some  $C' > 0$ , giving (iii).

Let us now pass to the proof of (iv). Fix  $i \in I$  and let  $\mathcal{C}_1 = \text{orb}_i(v_1)$ ,  $\mathcal{C}_2 = \text{orb}_i(v_2)$  be two cycles of length at least  $n \log^2 n$ . On the event  $\{\tau > i\}$  we have  $\chi(\text{orb}_i^{n \log^2 n}(v)) \geq c_2 \log^2 n$  for  $j = 1, 2$ . Thus by Lemma 4.1 and part (ii) of Lemma 4.5 we obtain

$$\mathbb{P}(\mathcal{M}_i | \mathcal{F}_i, |X| = k) \geq \frac{\Theta^{-1}}{2} \frac{|\mathcal{C}_1| |\mathcal{C}_2| (c_2 \log^2 n)^2}{n^2 (n \log^2 n)^2} \mathbb{1}_{\{k \in K_n\}} \mathbb{1}_{\{\tau > i\}} \geq C'' \frac{|\mathcal{C}_1| |\mathcal{C}_2|}{n^4} \mathbb{1}_{\{k \in K_n\}} \mathbb{1}_{\{\tau > i\}}$$

for some  $C'' > 0$  as desired.  $\square$

## 5. Macroscopic cycles

In this section we will prove our main results, namely Theorem 1.1 and Theorem 1.2. We will need one more ingredient, which is a variant of an argument due to Schramm, described in the next subsection.

In most of this section we will work with the general transposition process introduced in Section 4.

### 5.1. Schramm's argument

In this standalone part we develop ideas of [28]. Our contribution is to rephrase them in terms of an abstract split-merge process. We will then apply the results of this subsection to transposition processes coming from measures  $\mu_{\beta, \theta, \mathcal{C}}$  (see Section 5.3) and we believe that they might be useful in more general cases.

Let  $V$  be a finite set and  $h \in \mathbb{N}$ . We say that  $\{\sigma_k\}_{k \in \{0, 1, \dots, h\}}$  is a *random split-merge process* over  $V$  if any  $\sigma_k$  is a (random) partition of  $V$ . We refer to the sets of this partition as components. We call a random split-merge process *simple* if in transition from  $\sigma_k$  to  $\sigma_{k+1}$  (which we will call step  $k$ ) we allow only

- a component to be split into two,
- two components to be merged.

We denote by  $\{\mathcal{F}_k\}_{k \in \{0, 1, \dots, h\}}$  the natural filtration of the process and write  $\mathcal{C}_k(\ell)$  for the set of  $v \in V$  which belong to components of  $\sigma_k$  of size at least  $\ell \in \mathbb{N}$ . We set  $I_h := \{0, 1, \dots, h-1\}$ .

Now we state the main result of this section. Its proof is essentially the same as proofs in Section 2 of [28], with only very slight modifications.

**Lemma 5.1.** *Let  $h \in \mathbb{N}$  and  $\{\sigma_k\}_{k \in \{0, 1, \dots, h\}}$  be a simple split-merge process over a finite set  $V$ . Let  $t_0 \in I_h$ ,  $\delta \in (0, 1]$ ,  $\varepsilon \in (0, 1/8)$  and  $j \in \mathbb{N}$  be such that  $2^j \leq \varepsilon \delta |V|$ . Assume moreover that for a stopping time  $\tau$  the following conditions hold*

- (i) *there exists  $c_1 > 0$  such that for any  $k \in I_h$ ,  $\ell \geq 2^j$  we have*

$$\mathbb{P}(\text{some component } C \text{ of } \sigma_k \text{ is split in transition from step } k \text{ to } k+1 \text{ into } C_1, C_2$$

$$\text{and } \min(|C_1|, |C_2|) \leq \ell | \mathcal{F}_k) \mathbb{1}_{\{\tau > k\}} \leq c_1 \frac{\ell}{|V|}.$$

- (ii) *there exists  $c_2 > 0$  such that for any  $k \in I_h$  and any two components  $C_1, C_2 \in \sigma_k$  such that  $|C_1|, |C_2| \geq 2^j$  we have*

$$\mathbb{P}(\text{components } C_1, C_2 \text{ are merged in transition from step } k \text{ to } k+1 | \mathcal{F}_k) \geq c_2 \frac{|C_1| |C_2|}{|V|^2} \mathbb{1}_{\{\tau > k\}}.$$

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

$$t_1 := t_0 + \lceil \Delta t \rceil, \quad \Delta t = c_3 \delta^{-1} \frac{|V|}{2^j} \log_2 \left( \frac{|V|}{2^j} \right), \quad (53)$$

satisfies  $t_1 \leq h$ , then

$$\mathbb{E}(|\mathcal{C}_{t_0}(2^j) \setminus \mathcal{C}_{t_1}(\varepsilon \delta |V|)| | \mathcal{F}_{t_0}) \mathbb{1}_{\{|\mathcal{C}_{t_0}(2^j)| \geq \delta |V|\}} \leq c_4 \delta^{-1} \varepsilon |\log_2(\varepsilon \delta)| |V| + c_4 \delta^{-1} |V| \mathbb{P}(\tau \leq t_1 | \mathcal{F}_{t_0}). \quad (54)$$

In other words, if sufficiently many vertices are in components of size at least  $2^j$  at time  $t_0$ , most of them will be in components of size  $\varepsilon \delta |V|$  at time  $t_1$  (unless the split-merge properties (i) and (ii) fail, which is reflected in the second term).

The proof is an implementation of the following simple idea. Fix  $K := \lceil \log_2(\varepsilon \delta |V|) \rceil$  and set milestones  $t_0 = T_j < T_{j+1} < \dots < T_K = t_1$ . In each epoch  $\{T_j, T_j + 1, \dots, T_{j+1}\}$  we expect the size of moderately large components to grow by a factor of two. The subtlety lies in formalising this statement and finding the correct lengths of the epochs. They need to be long enough so that most components have a chance to merge into bigger ones and at the same time to be short enough so that not too many splits occur. It turns out that the proper choice is

$$T_{i+1} - T_i = m_i := \lceil a_i \rceil, \quad a_i := \frac{4}{c_2} \delta^{-1} \frac{|V|}{2^i} \log_2 \left( \frac{|V|}{2^i} \right). \quad (55)$$

Now the form of  $\Delta t$  in (53) is rather natural.

In proofs below we will use  $N$  as shorthand for  $|V|$  and  $\log$  for  $\log_2$ . All unspecified constants are assumed to be independent of  $\delta$  and  $\varepsilon$ . We implicitly assume that the condition  $t_1 \leq h$  is met. We can also assume that  $\mathbb{P}(\{|\mathcal{C}_{t_0}(2^j)| \geq \delta |V|\}) > 0$ . For notational simplicity fix an arbitrary  $A \in \mathcal{F}_{t_0}$  such that  $\mathbb{P}(A \cap \{|\mathcal{C}_{t_0}(2^j)| \geq \delta |V|\}) > 0$  and set  $\mathbf{P}(\cdot) = \mathbb{P}(\cdot | A \cap \{|\mathcal{C}_{t_0}(2^j)| \geq \delta |V|\})$ . We denote by  $\mathbf{E}(\cdot)$  the corresponding expectation. The inequality (54) is then equivalent to

$$\mathbf{E}|\mathcal{C}_{t_0}(2^j) \setminus \mathcal{C}_{t_1}(\varepsilon \delta N)| \leq c_4 \delta^{-1} \varepsilon |\log(\varepsilon \delta)| N + c_4 \delta^{-1} N \mathbf{P}(\tau \leq t_1)$$

for all admissible sets  $A$ .

The idea behind the proof of Lemma 5.1 consists of showing that for any epoch  $i \in \{j, \dots, K-1\}$  the number of vertices in  $\mathcal{C}_{T_i}(2^i) \setminus \mathcal{C}_{T_{i+1}}(2^{i+1})$  must be small.

The first reason for a vertex to fall in this set is splitting. Namely, by  $\bar{S}_i$  we denote the set of vertices  $v \in V$  which at some time  $t \in \{T_i, \dots, T_{i+1} - 1\}$  belong to a component which in transition to time  $t + 1$  is split and  $v$  ends up in a component of size smaller than  $2^{i+1}$ . In Lemma 5.2 we show that  $\bar{S}_i$  is small.

The second reason is failure of the components to merge. We define  $\bar{M}_i := \mathcal{C}_{T_i}(2^i) \setminus (\mathcal{C}_{T_{i+1}}(2^{i+1}) \cup \bar{S}_i)$ , the set containing vertices whose components did not split, but failed to merge into a bigger one. In Lemma 5.3 we analyze  $\bar{M}_i$  in more detail and prove that it is small as well.

We also denote

$$S_k := \bigcup_{i=j}^k \bar{S}_i, \quad M_k := \bigcup_{i=j}^k \bar{M}_i.$$

The following inclusion reveals the rationale behind the above definitions:

$$\mathcal{C}_{T_j}(2^j) \setminus \mathcal{C}_{T_{k+1}}(2^{k+1}) \subset S_k \cup M_k. \quad (56)$$

We first deal with splits

**Lemma 5.2.** *Under assumptions of Lemma 5.1 there exists  $c > 0$  such that*

$$\mathbf{E}(|S_{K-1}| \mathbb{1}_{\{\tau > T_K\}}) \leq c \varepsilon |\log_2(\varepsilon \delta)| N. \quad (57)$$

**Proof.** Fix  $i \in \{j, \dots, K-1\}$  and let

$$\mathcal{D}_i := \{\text{a split occurred in transition from time } t \text{ to } t + 1 \text{ creating a component of size smaller than } 2^{i+1}\}.$$

Whenever a split creates one or possibly two components of size smaller than  $2^{i+1}$ , at most  $2 \cdot 2^{i+1}$  vertices are added to the set  $\bar{S}_i$ . Thus

$$|\bar{S}_i| \mathbb{1}_{\{\tau > T_i\}} \leq 2 \cdot 2^{i+1} \sum_{t=T_i}^{T_{i+1}-1} \mathbb{1}_{\mathcal{D}_t} \mathbb{1}_{\{\tau > T_i\}}.$$

Applying assumption (i) of Lemma 5.1 we obtain  $\mathbb{1}_{\{\tau > T_i\}} \mathbf{P}(\mathcal{D}_t) \leq C2^{i+1}/N$ , for some  $C > 0$ . As  $\{\tau > T_K\} \subset \{\tau > T_i\}$ , we have

$$\mathbf{E}(|S_{K-1}| \mathbb{1}_{\{\tau > T_K\}}) \leq 8C \sum_{i=j}^{K-1} m_i 2^{2i} / N.$$

The rest of the proof follows by calculations employing the form of  $m_i$ .

$$\begin{aligned} \frac{1}{N} \sum_{i=j}^{K-1} m_i 2^{2i} &\leq \frac{2}{N} \sum_{i=j}^{K-1} a_i 2^{2i} = \frac{8}{c_2} \delta^{-1} \sum_{i=j}^{K-1} 2^i \log\left(\frac{N}{2^i}\right) \leq \frac{8}{c_2} \delta^{-1} \sum_{i=0}^{K-1} 2^i \log\left(\frac{N}{2^i}\right) \\ &= \frac{8}{c_2} \delta^{-1} ((2^K - 1) \log N - 2 + 2^{K+1} - 2^K K) \\ &\leq \frac{8}{c_2} \delta^{-1} ((\log N - K) 2^K + 2^{K+1}). \end{aligned}$$

Recalling  $K = \lceil \log_2(\varepsilon \delta N) \rceil$ , we check easily that  $2^{K+1} \leq 4\varepsilon \delta N$  and  $(\log N - K) 2^K \leq 2|\log(\varepsilon \delta)| \varepsilon \delta N$ , therefore (57) follows.  $\square$

Analysing  $M_i$  is somewhat tricky. We introduce an additional index

$$\sigma := \min\{i \in \{j, \dots, K-1\} : |C_t(2^i)| < \delta N/2 \text{ for some } t \in \{T_i + 1, \dots, T_{i+1}\}\}, \quad (58)$$

with the convention  $\sigma = +\infty$  when the set is empty. Now we can state

**Lemma 5.3.** *Under assumptions of Lemma 5.1 we have*

$$\mathbf{E}(|M_{(\sigma-1) \wedge (K-1)}| \mathbb{1}_{\{\tau > T_K\}}) \leq 2\varepsilon \delta N. \quad (59)$$

**Proof.** Consider an epoch  $i \in \{j, \dots, K-1\}$  and let  $v$  be any vertex. Recalling the definition of  $\bar{M}_i$ , if  $v \in \bar{M}_i$ , then at time  $T_i$  the vertex  $v$  is in a component of size at least  $2^i$ , at time  $T_{i+1}$  the component of  $v$  is smaller than  $2^{i+1}$ , and  $v \notin \bar{S}_i$ , so there is no splitting between these two times which would put  $v$  in a component smaller than  $2^{i+1}$ . Therefore, we have  $v \in C_t(2^i) \setminus C_t(2^{i+1})$  for all  $t \in \{T_i, \dots, T_{i+1}-1\}$ . For any step  $t$  of the epoch consider the event  $\mathcal{A}_t$  that the component of  $v$  merges with another component of size at least  $2^i$ . Notice that  $\{v \in M_i\} \subset \bigcap_t \mathcal{A}_t^c$ , where the intersection is over  $t \in \{T_i, \dots, T_{i+1}-1\}$ . Indeed, had any  $\mathcal{A}_t$  happened the component of  $v$  would have been of size at least  $2^{i+1}$  and, as  $v \notin \bar{S}_i$ , it would have survived until the end of the epoch. Denote also the event  $\mathcal{E}_t := \{|C_t(2^i)| \geq \delta N/2\} \cap \{v \in C_t(2^i) \setminus C_t(2^{i+1})\}$ . By assumption (ii) of Lemma 5.1 we have

$$\mathbf{P}(\mathcal{A}_t | \mathcal{F}_t) \geq c_2 2^i \frac{\delta N/2 - 2^{i+1}}{N^2} \mathbb{1}_{\mathcal{E}_t} \mathbb{1}_{\{\tau > t\}} \geq c_2 2^i \delta \frac{1/2 - 2\varepsilon}{N} \mathbb{1}_{\mathcal{E}_t} \mathbb{1}_{\{\tau > t\}} \geq \frac{c_2 2^i \delta}{4} \frac{1}{N} \mathbb{1}_{\mathcal{E}_t} \mathbb{1}_{\{\tau > t\}}.$$

Using the facts above we conclude

$$\begin{aligned} \mathbf{P}(\{v \in \bar{M}_i\} \cap \{\sigma > i\} \cap \{\tau \geq T_{i+1}\}) &\leq \mathbf{E}(\mathbb{1}_{\bigcap_{t \in \{T_i, \dots, T_{i+1}-1\}} (\mathcal{A}_t^c \cap \mathcal{E}_t)} \cap \{\tau > t\}) \\ &\leq \mathbf{E}(\mathbb{1}_{\bigcap_{t \in \{T_i, \dots, T_{i+1}-2\}} (\mathcal{A}_t^c \cap \mathcal{E}_t)} \cap \{\tau > t\} \mathbb{1}_{\mathcal{E}_{T_{i+1}-1}} \mathbb{1}_{\{\tau > T_{i+1}-1\}} \mathbf{P}(\mathcal{A}_{T_{i+1}-1}^c | \mathcal{F}_{T_{i+1}-1})) \\ &\leq \left(1 - \frac{c_2 2^i \delta}{4} \frac{1}{N}\right) \cdot \mathbf{E}(\mathbb{1}_{\bigcap_{t \in \{T_i, \dots, T_{i+1}-2\}} (\mathcal{A}_t^c \cap \mathcal{E}_t)} \cap \{\tau > t\}) \leq \dots \leq \left(1 - \frac{c_2 2^i \delta}{4} \frac{1}{N}\right)^{m_i}. \end{aligned}$$

The choice of  $m_i$  in (55) is such that

$$\mathbf{P}(\{v \in \bar{M}_i\} \cap \{\sigma > i\} \cap \{\tau \geq T_{i+1}\} | \mathcal{F}_{T_i}) \leq 2^i / N.$$

The rest of the proof follows by the estimate

$$\mathbf{E}(|M_{(\sigma-1) \wedge (K-1)}| \mathbb{1}_{\{\tau > T_K\}}) \leq \sum_{i=j}^{K-1} \mathbf{E}(|\bar{M}_i| \mathbb{1}_{\{\sigma > i\}} \mathbb{1}_{\{\tau \geq T_{i+1}\}}) \leq N \frac{2^K}{N} \leq 2\varepsilon \delta N. \quad \square$$

We are now ready for

**Proof of Lemma 5.1.** By (56) on the event  $\{\sigma < +\infty\}$  we have  $\mathcal{C}_{T_j}(2^j) \setminus \mathcal{C}_{T_\sigma}(2^\sigma) \subset S_{\sigma-1} \cup M_{\sigma-1}$ . Consider now  $t \in \{T_\sigma + 1, \dots, T_{\sigma+1}\}$  such that  $|\mathcal{C}_t(2^\sigma)| < \delta N/2$  and any  $v \in \mathcal{C}_{T_j}(2^j) \setminus \mathcal{C}_t(2^\sigma)$ . Recall that we work on the event  $\{|\mathcal{C}_{T_0}(2^j)| \geq \delta N\}$ . If  $v \notin \mathcal{C}_{T_\sigma}(2^\sigma)$  then  $v \in S_{\sigma-1} \cup M_{\sigma-1}$ , otherwise  $v \in \mathcal{C}_{T_\sigma}(2^\sigma) \setminus \mathcal{C}_t(2^\sigma)$  and so  $v \in \bar{S}_\sigma \subset S_{K-1}$ . Thus  $\mathcal{C}_{T_j}(2^j) \setminus \mathcal{C}_t(2^\sigma) \subset M_{(\sigma-1) \wedge (K-1)} \cup S_{K-1}$ . Since  $|\mathcal{C}_{T_j}(2^j)| \geq \delta N$ , we conclude that  $\{\sigma < +\infty\} \subset \{|M_{(\sigma-1) \wedge (K-1)} \cup S_{K-1}| \geq \delta N/2\}$ .

Now let  $\Delta := \mathbf{E}|\mathcal{C}_{T_0}(2^j) \setminus \mathcal{C}_{t_1}(\varepsilon \delta N)|$ . Using (56) we have  $\Delta \leq \mathbf{E}|M_{K-1} \cup S_{K-1}|$ . Furthermore,

$$\begin{aligned} \Delta &\leq \mathbf{E}(|M_{K-1} \cup S_{K-1}| \mathbb{1}_{\{\sigma = +\infty\}}) + N\mathbf{P}(\sigma < +\infty) \\ &\leq \mathbf{E}|M_{(\sigma-1) \wedge (K-1)} \cup S_{K-1}| + N\mathbf{P}(|M_{(\sigma-1) \wedge (K-1)} \cup S_{K-1}| \geq \delta N/2) \\ &\leq (1 + 2/\delta)\mathbf{E}|M_{(\sigma-1) \wedge (K-1)} \cup S_{K-1}| \\ &\leq (1 + 2/\delta)(\mathbf{E}|M_{(\sigma-1) \wedge (K-1)}| + \mathbf{E}|S_{K-1}|), \end{aligned}$$

where in the second line we used Markov's inequality. Furthermore, we write

$$\mathbf{E}|M_{(\sigma-1) \wedge (K-1)}| + \mathbf{E}|S_{K-1}| \leq \mathbf{E}(|M_{(\sigma-1) \wedge (K-1)}| \mathbb{1}_{\{\tau > T_K\}}) + (\mathbf{E}|S_{K-1}| \mathbb{1}_{\{\tau > T_K\}}) + 2N\mathbf{P}(\tau \leq T_K)$$

Now applying Lemma 5.2 and Lemma 5.3 we get (recalling that  $T_K = t_1$ )

$$\Delta \leq (1 + 2/\delta)(2\varepsilon \delta N + c\varepsilon |\log(\varepsilon \delta)|N) + (2 + 4/\delta)N\mathbf{P}(\tau \leq t_1).$$

As  $\delta < 1$  and  $\varepsilon < 1/8$ , one readily checks that the right hand side is bounded from above by

$$c'\delta^{-1}\varepsilon |\log(\varepsilon \delta)|N + c'\delta^{-1}N\mathbf{P}(\tau \leq t_1)$$

for some  $c' > 0$ , which concludes the proof. □

## 5.2. Mesoscopic cycles

In this section we will work in the setting of general transposition process, introduced in Section 4, for  $X$  sampled from  $\mu_{\beta, \theta, \mathcal{C}}$ . All quantities like  $\mathcal{C}_t(\cdot)$  are implicitly related to this process. We focus on the supercritical phase  $\beta > \Theta/2$ . Our aim, formalized in Proposition 5.4, is to show that a substantial fraction of vertices belongs to mesoscopic cycles (of size at least  $n \log^2 n$ ). This result will be used in the next section as an input to Lemma 5.1 to prove the existence of macroscopic cycles. The crucial ingredient that we use are the isoperimetric properties of the cycles stated in Proposition 4.2.

**Proposition 5.4.** *Let  $\beta > \Theta/2$ . There exist  $\delta > 0$  and sets  $K_n \subset \{\lceil n^{11/6} \rceil, \lceil n^{11/6} \rceil + 1, \dots\}$  such that  $\lim_{n \rightarrow \infty} \mathbb{P}(|X| \in K_n) = 1$  and*

$$\lim_{n \rightarrow \infty} \inf_{k \in K_n} \min_{t \in \{k - \lceil n^{11/6} \rceil, \dots, k\}} \mathbb{P}(|\mathcal{C}_t(n \log^2 n)| \geq \delta n^2 | |X| = k) = 1.$$

The proof of the proposition hinges on a coupling between the generalized transposition process and a random graph process. Let  $s \in \{0, \dots, |X|\}$  and consider a process  $G^s = \{G_u^s\}_{u \in \{0, \dots, |X| - s\}}$  of random graphs on the vertex set  $V$  defined as follows. Initially,  $G_0^s$  is a graph whose connected components are the cycles of  $\sigma_s$ . There might be many graphs



satisfying this property and for our purposes it will not matter which one is chosen. Next, for any  $u \in \{1, \dots, |X| - s\}$  the edge  $e$  corresponding to transposition  $e_{s+u}$  (i.e.,  $\sigma_{s+u} = e_{s+u} \circ \sigma_{s+u-1}$ ) is added to the edge set of  $G_u^s$ .

Recall that  $\mathcal{C}_t(\ell)$  is the set of vertices which belong to cycles of length at least  $\ell$  in  $\sigma_t$ . Correspondingly, let  $\mathcal{G}_{s,u}(\ell)$  be the set of vertices of  $G_u^s$  which belong to connected components of size at least  $\ell$ . Importantly, any cycle of  $\sigma_{s+u}$  is contained in a connected component  $G_u^s$ . Hence, it follows that  $\mathcal{C}_{s+u}(\ell) \subset \mathcal{G}_{s,u}(\ell)$  for any  $s, u, \ell \in \mathbb{N}$ .

There are two key ingredients in the proof Proposition 5.4. First,  $G^s$  is monotonic and thus can be studied using standard random graph techniques. In particular Lemma 5.6 below shows that macroscopic clusters emerge quickly in  $G^s$ . Second, on sufficiently short time intervals the difference  $\mathcal{G}_{s,u}(\ell) \setminus \mathcal{C}_{s+u}(\ell)$  is small, which is formalized in Lemma 5.5.

Consider an interval  $I \subset \{0, \dots, |X|\}$ ,  $k \in \mathbb{N}$  and let  $c_1 > 0$  be the constant from Proposition 4.2. Let

$$\mathcal{I}_k(I) := \left\{ \sup_{t \in I} \max_{v \in V} \iota(\text{orb}_t^k(v)) \leq c_1 \log^2 n \right\}$$

denote the event that fragments of permutation orbits for  $t \in I$  have good isoperimetric properties. Our first lemma quantifies the quality of the coupling between  $\mathcal{C}_{s+u}$  and  $G_u^s$ .

**Lemma 5.5.** *Let  $s \in \mathbb{N}$  and let  $\Delta$  be an  $|X|$ -measurable  $\mathbb{N}$ -valued random variable. Suppose that  $k, \ell \in \mathbb{N}$  satisfy  $k \leq \ell$ . Then for some  $C > 0$  on the event  $s + \Delta \leq |X|$  we have*

$$\mathbb{E} \left[ \max_{u \in \{0, \dots, \Delta\}} |\mathcal{G}_{s,u}(\ell) \setminus \mathcal{C}_{s+u}(\ell)| \mid |X| \right] \leq C \frac{\ell^2 \Delta}{kn} \log^2 n + 2\ell \Delta \mathbb{P}(\mathcal{I}_k(\{s, \dots, s + \Delta\})^c \mid |X|).$$

**Proof.** The proof is an adaptation of [25, Lemma 4.2] to the discrete time setting and the case  $\Theta \neq 1$ . Let  $I$  be the set of  $u \in [0, \Delta - 1]$  such that  $\sigma$  experiences a fragmentation at time  $s + u$  which splits a cycle and at least one of the resulting cycles has length less than  $\ell$ .

From Lemma 4.1 and point (i) of Lemma 4.5 we obtain that at any time  $u$  the (conditional) probability of a fragmentation in which one piece is smaller than  $\ell$  is at most

$$\frac{4\Theta^2 \ell}{kn} c_1 \log^2 n + \mathbb{1}_{\{\mathcal{I}_k(\{u\})^c\}}.$$

Hence we see that

$$\mathbb{E}[|I| \mid |X|] \leq \frac{4\Theta^2 \ell \Delta}{kn} c_1^2 \log^2 n + \Delta \mathbb{P}(\mathcal{I}_k(\{s, \dots, s + \Delta\})^c \mid |X|). \quad (60)$$

Let  $u \in [1, \Delta]$  and consider any cycle  $\gamma$  of  $\sigma_{s+u}$  such that  $\gamma \subset \mathcal{G}_{s,u}(\ell) \setminus \mathcal{C}_{s+u}(\ell)$ , that is,  $\gamma$  is contained in a component of  $G_u^s$  of size at least  $\ell$  and  $|\gamma| < \ell$ . Then it follows that there must have been a vertex  $v \in \gamma$  such that the cycle containing  $v$  must have fragmented at some time in  $\{s, \dots, s + u\}$  producing a cycle of size smaller than  $\ell$ .

For  $t \in \{s, \dots, s + u\}$ , let  $\gamma_t^{(v)}$  be the cycle of  $\sigma_t$  containing  $v$ . Let  $t' \in \{s, \dots, s + u - 1\}$  be the maximal time such that the size of  $\gamma_{t'}^{(v)}$  jumps downwards, that is, the cycle containing  $v$  experiences a fragmentation. Then at this time  $t'$ ,  $\sigma$  experiences a fragmentation which splits a cycle into two and at least one of the resulting cycles has length less than  $\ell$ . Note that the cycle thus obtained is a part of  $\gamma$ . It follows that  $t' \in I$  and consequently  $|\mathcal{G}_{s,u}(\ell) \setminus \mathcal{C}_{s+u}(\ell)| \leq 2\ell |I|$ . Taking suprema and using (60) we obtain the desired result.  $\square$

The second lemma quantifies how quickly big clusters emerge in the random graph process  $G_u^s$ .

**Lemma 5.6.** *For any  $\delta \in (0, 1/8)$  there exists a sequence  $\{a_n\}_{n \in \mathbb{N}}$  such that the following holds:  $a_n \nearrow 1$  and for any  $s, u, \ell, h \in \mathbb{N}$  satisfying  $\log^2 n \leq \ell \leq n^2$  and  $u \geq (n^2/\sqrt{\ell}) \log n$ ,  $s + u \leq h$  and  $\mathbb{P}(|\mathcal{G}_{s,0}(\ell)| \geq \delta n^2, |X| = h) > 0$ , we have*

$$\mathbb{P}(|\mathcal{G}_{s,u}(\delta n^2/8)| \geq \delta n^2/8 \mid |\mathcal{G}_{s,0}(\ell)| \geq \delta n^2, |X| = h) \geq a_n.$$

As the proof is an adaptation of [25, Lemma 4.3] and is of graph-theoretical nature, it is moved to Appendix C.

In the two subsequent lemmas we show that in the supercritical phase the random graph process has macroscopic clusters for times close to  $|X|$ .

**Lemma 5.7.** *Let  $\beta > \Theta/2$ . There exists  $\delta > 0$  such that*

$$\lim_{n \rightarrow +\infty} \mathbb{P}(|X| \geq 2 \lceil n^2 / \log n \rceil \text{ and } |\mathcal{G}_{0,|X|-2 \lceil n^2 / \log n \rceil}(\delta n^2)| \geq \delta n^2) = 1.$$

**Proof.** Let  $I = [0, a)$ , for  $a < 1$ , be an interval such that  $\beta' = \beta|I|^{\Theta-1} > 1/2$ . We also set  $J := [0, 1) \setminus I$ . Recall the notation used in Lemma B.1. By the monotonicity of the graph process it is enough to show that for some  $c > 0$

$$\lim_{n \rightarrow +\infty} \mathbb{P}(|\mathcal{G}_{0,|X_{E \times I}|}(cn^2)| \geq cn^2) = 1, \quad \lim_{n \rightarrow +\infty} \mathbb{P}(|X_{E \times J}| \geq 2 \lceil n^2 / \log n \rceil) = 1. \quad (61)$$

For  $X \in \mathfrak{X}$  we set  $\hat{X} := \{e : \exists t \in (0, 1) (e, t) \in X\}$ . We intend to compare  $\hat{X}_{E \times I}$  and  $\hat{X}_{E \times J}$  with the Bernoulli percolation on  $E$ . To this end we use the Holley theorem [16, Theorem 4.6] with  $\mathcal{L} = 1$  and  $S = \{0, 1\}$ , where 1 indicates that  $e$  is an open edge (i.e., belongs to a given set). For any  $e \in E$  by the first part of Lemma B.1 we get

$$\begin{aligned} \mathbb{P}(e \in \hat{X}_{\{e\} \times I} | \hat{X}_{(E \setminus \{e\}) \times I}) &= \mathbb{E}(\mathbb{P}(|X_{\{e\} \times I}| \geq 1 | X_{(E \times [0, 1]) \setminus (\{e\} \times I)}) | \hat{X}_{(E \setminus \{e\}) \times I}) \\ &\geq 1 - e^{-\frac{\beta|I|^{\Theta-1}}{n-1}} = 1 - e^{-\frac{\beta'}{n-1}} \geq \frac{\beta'}{n-1} - o(1/n) =: p_n. \end{aligned}$$

This yields that  $\hat{X}_{\{e\} \times I}$  is stochastically bounded from below by the Bernoulli percolation process with the probability of opening  $p_n$ . As  $\beta' > 1/2$ , for  $n$  large enough this process is in the supercritical phase. As a consequence, to get the first convergence in (61) we can apply known results on the emergence of the giant component in supercritical percolation (see Theorem 1.1 in [31] and the discussion therein; we note that the results of [31] are formulated only for  $p_n$  in the critical window, but the proof techniques carry over to the strictly supercritical case, see e.g., discussion in Section 3 of the cited paper).

Similarly,  $\hat{X}_{E \times J}$  is bounded by a percolation process with the probability of opening  $\geq 1 - e^{-\frac{\beta|J|^{\Theta-1}}{n-1}} \geq \frac{\beta|J|^{\Theta-1}}{2(n-1)}$ . From this bound and the fact that  $|E| = O(n^3)$  we infer that  $|X_{E \times J}| \geq |\hat{X}_{E \times J}| \geq 2 \lceil n^2 / \log n \rceil$  with probability converging to 1.  $\square$

**Lemma 5.8.** *Let  $\beta > \Theta/2$ . There exist  $\delta > 0$ , sets  $K_n \subset I_n := \{\lfloor n^2 / \log n \rfloor, \dots, \lfloor n^2 \log n \rfloor\}$  and a sequence  $\{a_n\}_{n \geq 1}$  such that the following holds:  $\lim_{n \rightarrow +\infty} \mathbb{P}(|X| \in K_n) = 1$ ,  $a_n \nearrow 1$  and*

$$\mathbb{P}(|\mathcal{G}_{0, k - \lfloor n^2 / \log n \rfloor}(\delta n^2)| \geq \delta n^2 | |X| = k) \geq a_n \mathbb{1}_{\{k \in K_n\}}. \quad (62)$$

**Proof.** Denote  $\mathcal{A}_n := \{|\mathcal{G}_{0, |X| - 2 \lceil n^2 / \log n \rceil}(\delta n^2)| \geq \delta n^2\}$ , with  $\delta$  as in Lemma 5.7, and set  $L_n := \{k \in \mathbb{N} : \mathbb{P}(\mathcal{A}_n | |X| = k) \leq c_n\}$ , for  $c_n \in (0, 1)$  to be fixed later. As  $\mathbb{P}(\mathcal{A}_n) \leq \mathbb{P}(|X| \in L_n)c_n + \mathbb{P}(|X| \notin L_n)$ , by a simple calculation we get

$$\frac{1 - \mathbb{P}(\mathcal{A}_n)}{1 - c_n} \geq \mathbb{P}(|X| \in L_n).$$

By Lemma 5.7 we have  $\mathbb{P}(\mathcal{A}_n) \rightarrow 1$ , so we can find  $c_n, \delta$  such that  $c_n \nearrow 1$  and the left-hand side converges to 0. Consequently, we have  $\mathbb{P}(|X| \in K'_n) \rightarrow 1$  for  $K'_n := \mathbb{N} \setminus L_n$ . We set  $K_n := K'_n \cap I_n$ . Using Lemma B.2 we see that  $\mathbb{P}(|X| \notin I_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Thus we get  $\lim_{n \rightarrow +\infty} \mathbb{P}(|X| \in K_n) = 1$  and  $\mathbb{P}(|\mathcal{G}_{0, k - 2 \lceil n^2 / \log n \rceil}(\delta n^2)| \geq \delta n^2 | |X| = k) \geq c_n \mathbb{1}_{\{k \in K_n\}}$  as desired.  $\square$

The proof of Proposition 5.4 follows by making comparisons of the random graph process and the generalized interchange process on appropriate time intervals, as made possible by Lemma 5.5. Below we make only two such comparisons. It is possible to iterate Lemma 5.5 more times on shorter and shorter time intervals, thus getting a tighter control on the difference  $\mathcal{G}_{s,u}(\ell) \setminus \mathcal{C}_{s+u}(\ell)$ . This method was used in [25] to prove the existence of cycles of size  $n^{2-\varepsilon}$  for any  $\varepsilon > 0$ . To the best of our knowledge this method alone cannot be pushed to obtain macroscopic cycles. Instead, in the next section we use modified Schramm's argument presented in Lemma 5.1, together with Proposition 5.4 as a prerequisite.

**Proof of Proposition 5.4.** Let  $\delta_1, K_n^1, a_n^1$  be  $\delta, K_n, a_n$  asserted by Lemma 5.8. We first apply Lemma 5.5 with  $k_1 = c_1 \log^2 n/2, \ell_1 = n^{1/3} \log n, s_1 = 0$  and  $\Delta_1 = h \leq n^2 \log n$ , conditionally on  $|X| = h$ . Clearly,  $\mathcal{I}_{k_1}(I) = \Omega$  for any interval  $I$ . Thus for  $h \in K_n^1$  using Markov's inequality we get for some  $C > 0$

$$\mathbb{P} \left[ \max_{t \in \{0, \dots, h\}} |\mathcal{G}_{0,t}(\ell_1) \setminus \mathcal{C}_t(\ell_1)| \geq \delta_1 n^2 / 2 | |X| = h \right] \leq \frac{C(n^{1/3} \log n)^2 n^2 \log n}{(\log^2 n/2)n(\delta_1 n^2/2)} \log^2 n = \frac{4C}{\delta_1} n^{-1/3} \log^3 n \rightarrow 0.$$

Combining this with Lemma 5.8 we get

$$\lim_{n \rightarrow +\infty} \inf_{h \in K_n^1} \mathbb{P} \left( \min_{t \in \{h-2\lceil n^{11/6} \rceil, \dots, h\}} |\mathcal{C}_t(\ell_1)| \geq \delta_1 n^2 / 2 \mid |X| = h \right) = 1. \quad (63)$$

Using this result we will be able to repeat the argument above on a short time interval contained in the supercritical phase. Crucially, on this interval we can use Lemma 4.4, which will let us obtain a much better estimate in Lemma 5.5.

Fix  $\tau \in \{h - \lceil n^{11/6} \rceil, \dots, h\}$ , conditionally on  $|X| = h$ , and set  $\Delta_2 := \lceil n^{11/6} \rceil$ . Let  $s_2 := \tau - \Delta_2$ . Starting from (63) we first apply Lemma 5.6 with  $\ell_1 = n^{1/3} \log n$  (note that  $\Delta_2 \geq \frac{n^2}{\sqrt{t_1}} \log n$ ), getting

$$\lim_{n \rightarrow +\infty} \inf_{h \in K_n^1} \mathbb{P} \left( |\mathcal{G}_{s_2, \Delta_2}(\delta_2 n^2)| \geq \delta_2 n^2 \mid |X| = h \right) = 1, \quad (64)$$

for some  $\delta_2 > 0$ . Fix  $\ell_2 = k_2 = n \log^2 n$ . Let  $K_n^2 = K_n^1 \cap K_n$ , where  $K_n$  is given by Lemma 4.4. Let  $h \in K_n^2$ , using Lemma 5.5 we estimate

$$\begin{aligned} \sup_{h \in K_n^2} \mathbb{E} \left[ |\mathcal{G}_{s_2, \Delta_2}(\ell_2) \setminus \mathcal{C}_\tau(\ell_2)| \mid |X| = h \right] &\leq \frac{C_1 (n \log^2 n)^2 n^{11/6}}{(n \log^2 n) n} \log^2 n + 2(n \log^2 n) n^{11/6} e^{-c \log^2 n} \\ &= C_2 n^{11/6} \log^4 n. \end{aligned}$$

for some  $C_1, C_2, c > 0$ . Markov's inequality implies

$$\sup_{h \in K_n^2} \mathbb{P} \left[ |\mathcal{G}_{s_2, \Delta_2}(\ell_2) \setminus \mathcal{C}_\tau(\ell_2)| \geq n^{11/6} \log^6 n \mid |X| = h \right] \leq C_2 / \log^2 n.$$

This combined with (64) yields the statement of the proposition.  $\square$

### 5.3. Macroscopic cycles in the supercritical phase $\beta > \Theta/2$ . Proof of Theorem 1.1

Now we are ready to show our main result. Recall the general transposition process  $\{\sigma_t\}$  introduced in Section 4 and, importantly, that  $\sigma_{|X|} = \sigma(X)$ , where  $\sigma(X)$  defined in (1) is the main object of our study.

**Proposition 5.9.** *Let  $\beta > \Theta/2$ . There exist sets  $K_n \subset \mathbb{N}$  such that  $\lim_{n \rightarrow \infty} \mathbb{P}(|X| \in K_n) = 1$  and*

$$\lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{k \in K_n} \mathbb{P}(\text{there exists a cycle of } \sigma(X) \text{ of length at least } \varepsilon n^2 \mid |X| = k) = 1.$$

**Proof.** Let  $\delta$  and  $K_n^1$  be respectively  $\delta$  and  $K_n$  asserted by Proposition 5.4. This proposition shows that cycles of size at least  $n \log^2 n$  are common. We will use this information to show the existence of macroscopic cycles. The key role in this proof is played by Schramm's argument, encapsulated in Lemma 5.1, and isoperimetric properties of cycles. The latter imply that the split-merge process behaves similarly to the mean-field (the complete graph) case, which is stated conveniently in Proposition 4.2. We denote sets  $K_n$  from that proposition by  $K_n^2$ .

In the proof we work conditionally on  $|X| = k$ , where  $k \in K_n^1 \cap K_n^2$ . Fix  $\varepsilon \in (0, 1/8)$  and consider the largest  $j \in \mathbb{N}$  such that  $2^j \leq n \log^2 n$ . Employing the notation from Lemma 5.1, we set  $t_0 = k - \Delta t$ , where  $\Delta t$  is given by (53). With this choice we have  $t_1 = k$ . Observe that for  $j$  as above we have  $\Delta t = o(n)$ , in particular  $t_0 \geq k - \lceil n^{11/6} \rceil$ . Let  $\tau$  be the stopping time from Proposition 4.2. One easily checks that conditions (i) and (ii) of Lemma 5.1 are fulfilled by assertions (iii) and (iv) of Proposition 4.2. Consequently we get

$$\mathbb{E} \left( |\mathcal{C}_{t_0}(2^j) \setminus \mathcal{C}_k(\varepsilon \delta n^2)| \mid \mathcal{F}_{t_0}, |X| = k \right) \mathbb{1}_{\{|\mathcal{C}_{t_0}(2^j)| \geq \delta n^2\}} \leq c \delta^{-1} \varepsilon |\log_2(\varepsilon \delta)| n^2 + c \delta^{-1} n^2 \mathbb{P}(\tau \leq k \mid \mathcal{F}_{t_0}, |X| = k),$$

for some  $c > 0$ . Using Markov's inequality we get

$$\mathbb{P} \left( |\mathcal{C}_{t_0}(2^j) \setminus \mathcal{C}_k(\varepsilon \delta n^2)| \geq \delta n^2 / 2 \mid \mathcal{F}_{t_0}, |X| = k \right) \mathbb{1}_{\{|\mathcal{C}_{t_0}(2^j)| \geq \delta n^2\}} \leq 2c \delta^{-2} (\varepsilon |\log_2(\varepsilon \delta)|) + \mathbb{P}(\tau \leq k \mid \mathcal{F}_{t_0}, |X| = k).$$

Consequently,

$$\mathbb{P} \left( |\mathcal{C}_k(\varepsilon \delta n^2)| \geq \delta n^2 / 2 \mid |X| = k \right)$$

$$\begin{aligned}
&\geq \mathbb{E}(\mathbb{P}(|\mathcal{C}_{t_0}(2^j) \setminus \mathcal{C}_k(\varepsilon \delta n^2)| < \delta n^2/2 | \mathcal{F}_{t_0}, |X| = k) \mathbb{1}_{\{|\mathcal{C}_{t_0}(2^j)| \geq \delta n^2\}} | |X| = k) \\
&\geq \mathbb{E}(\mathbb{1}_{\{|\mathcal{C}_{t_0}(2^j)| \geq \delta n^2\}} (1 - 2c\delta^{-2}(\varepsilon |\log_2(\varepsilon \delta)| + \mathbb{P}(\tau \leq k | \mathcal{F}_{t_0}, |X| = k)))) \\
&\geq \mathbb{P}(|\mathcal{C}_{t_0}(2^j)| \geq \delta n^2 | |X| = k) (1 - 2c\delta^{-2}\varepsilon |\log_2(\varepsilon \delta)|) - 2c\delta^{-2}\mathbb{P}(\tau \leq k | |X| = k).
\end{aligned}$$

By Proposition 5.4, with our choice of  $j$  and  $t_0$  the first probability on the right hand side approaches 1 as  $n \rightarrow \infty$  uniformly over  $k \in K_n^1 \cap K_n^2$ . The second probability goes to 0 uniformly over  $k \in K_n^1 \cap K_n^2$  by (ii) of Proposition 4.2. Noticing that  $\lim_{n \rightarrow +\infty} \mathbb{P}(|X| \in K_n) = 1$  and taking the limit  $\varepsilon \rightarrow 0$  we obtain our result.  $\square$

**Proof of Theorem 1.1.** Let  $K_n$  be the sets claimed in Proposition 5.10. We write

$$\begin{aligned}
&\mathbb{P}(\text{there exists a cycle of } \sigma(X) \text{ of length at least } \varepsilon n^2) \\
&\geq \sum_{k \in K_n} \mathbb{P}(\text{there exists a cycle of } \sigma(X) \text{ of length at least } \varepsilon n^2 | |X| = k) \mathbb{P}(|X| = k) \\
&\geq \mathbb{P}(|X| \in K_n) \inf_{k \in K_n} \mathbb{P}(\text{there exists a cycle of } \sigma(X) \text{ of length at least } \varepsilon n^2 | |X| = k).
\end{aligned}$$

Since  $\mathbb{P}(|X| \in K_n) \rightarrow 1$  as  $n \rightarrow \infty$ , by taking  $\liminf$  over  $n \rightarrow \infty$  and then the limit  $\varepsilon \rightarrow 0$  using Proposition 5.10 we obtain the statement of the theorem.  $\square$

#### 5.4. Microscopic cycles in the subcritical phase $\beta < \Theta^{-1}/2$ . Proof of Theorem 1.2

We will now sketch a proof of the statement about the behavior of cycle lengths in subcritical phase. This is a much easier task than in the supercritical phase. Our main result follows directly from the following

**Proposition 5.10.** *Let  $\beta < \Theta^{-1}/2$ . There exist  $C > 0$  and sets  $K_n \subset \mathbb{N}$  such that  $\lim_{n \rightarrow \infty} \mathbb{P}(|X| \in K_n) = 1$  and*

$$\lim_{n \rightarrow \infty} \inf_{k \in K_n} \mathbb{P}(|\mathcal{C}_k(C \log n)| = 0 | |X| = k) = 1.$$

**Proof.** Let  $C > 0$  and recall that  $\mathcal{C}_t(C \log n) \subset \mathcal{G}_{0,t}(C \log n)$ . For  $X \in \mathfrak{X}$  we consider  $\bar{X} := \{e : (e, t) \in X\}$ . Similarly as in Lemma 5.8 we can prove that  $\bar{X}$  is stochastically bounded from above by the Bernoulli percolation process with the probability of opening an edge being  $\frac{\beta'}{n-1}$  for some  $\beta' < 1/2$ . Now the result follows by a rather standard argument using coupling with branching processes or a random walk (see e.g., [15, Theorem 2.3.1]).  $\square$

## Appendix A: Concentration of point processes

The proofs of our auxiliary lemmas concerning counting processes will be all based on the following well known result (see e.g., [23] or [13, Chapter II.6]).

**Theorem A.1.** *Let  $Y$  be a counting process with intensity  $\lambda$ . Let  $\Lambda_t = \int_0^t \lambda_s ds$  be the compensator of  $Y$ . Then (on an enlarged probability space) there exists a Poisson process  $N$  with intensity one such that almost surely for all  $t \geq 0$ ,  $X_t = N_{\Lambda_t}$ .*

**Proof of Lemma 3.4.** Let  $A \in \mathcal{F}_\sigma$  be any event of nonzero probability. The process  $\tilde{Y}_t = Y_{\sigma+t} - Y_\sigma$  is a counting process with intensity  $\tilde{\lambda}_t = \lambda_{\sigma+t}$  with respect to the filtration  $\tilde{\mathcal{F}}_t = \mathcal{F}_{\sigma+t}$  and the conditional probability  $\tilde{\mathbb{P}} = \mathbb{P}(\cdot | A)$ .

Set  $\tilde{\Lambda}_t = \int_0^t \tilde{\lambda}_s ds$  and note that  $\tau - \sigma$  is a stopping time with respect to the filtration  $\tilde{\mathcal{F}}_t$ . Let  $N$  be the Poisson process of intensity one, given for  $\tilde{Y}$  by Theorem A.1. We have

$$\tilde{\mathbb{P}}(\{Y_\tau - Y_\sigma \geq r\} \cap \{\Lambda_\tau - \Lambda_\sigma \leq \ell\}) = \tilde{\mathbb{P}}(\{\tilde{Y}_{\tau-\sigma} \geq r\} \cap \{\tilde{\Lambda}_{\tau-\sigma} \leq \ell\}) \leq \tilde{\mathbb{P}}(N_\ell \geq r).$$

If  $r \geq \ell$ , by using the form of the Laplace transform for the Poisson distribution we get

$$\tilde{\mathbb{P}}(N_\ell \geq r) \leq \inf_{u \geq 0} \exp(\ell(e^u - 1) - ur) \leq \exp\left(-r \log\left(\frac{r}{e\ell}\right) - \ell\right).$$

Going back to the original probability measure, we conclude that for any  $A \in \mathcal{F}_\sigma$ ,

$$\mathbb{P}(\{Y_\tau - Y_\sigma \geq r\} \cap \{\Lambda_\tau - \Lambda_\sigma \leq \ell\} \cap A) \leq \mathbb{P}(A)\mathbb{P}(X \geq r),$$

and for  $r \geq \ell$

$$\mathbb{P}(X \geq r) \leq \exp\left(-r \log\left(\frac{r}{e\ell}\right) - \ell\right),$$

which implies the lemma.  $\square$

**Proof of Lemma 3.5.** Again, consider any  $A \in \mathcal{F}_\sigma$  with positive probability and the process  $\tilde{Y}_t = Y_{\sigma+t} - Y_\sigma$ , which is a counting process with intensity  $\tilde{\lambda}_t = \lambda_{\sigma+t}$  with respect to the filtration  $\tilde{\mathcal{F}}_t = \mathcal{F}_{\sigma+t}$ . Let  $\tilde{\mathbb{P}} = \mathbb{P}(\cdot|A)$ . We have  $\tilde{\Lambda}_t = \int_0^t \tilde{\lambda}_s ds = \Lambda_{\sigma+t} - \Lambda_\sigma$ . In particular if  $N$  is the Poisson process given for  $\tilde{Y}$  by Theorem A.1, we get

$$\begin{aligned} \tilde{\mathbb{P}}(\{Y_\tau - Y_\sigma \leq \ell(1 - \delta)\} \cap \{\Lambda_\tau - \Lambda_\sigma \geq \ell\}) &= \tilde{\mathbb{P}}(\{\tilde{Y}_{\tau-\sigma} \leq \ell(1 - \delta)\} \cap \{\tilde{\Lambda}_{\tau-\sigma} \geq \ell\}) \\ &\leq \tilde{\mathbb{P}}(N_\ell \leq \ell(1 - \delta)). \end{aligned}$$

Using the form of the Laplace transform of  $N_\ell$  and Chebyshev's inequality we obtain

$$\begin{aligned} \tilde{\mathbb{P}}(N_\ell \leq \ell(1 - \delta)) &\leq \inf_{a \geq 0} \exp((e^{-a} - 1)\ell + a\ell(1 - \delta)) \\ &\leq \inf_{a \geq 0} \exp\left(\frac{1}{2}a^2\ell - a\ell\delta\right) = \exp\left(-\frac{1}{2}\delta^2\ell\right), \end{aligned}$$

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

$$\mathbb{P}(\{Y_\tau - Y_\sigma \leq \ell(1 - \delta)\} \cap \{\Lambda_\tau - \Lambda_\sigma \geq \ell\} | A) \leq \mathbb{P}(X \leq (1 - \delta)\ell) \leq \exp\left(-\frac{1}{2}\delta^2\ell\right),$$

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

**Lemma A.2.** Let  $Y$  be a counting process with bounded intensity  $\lambda$ . Consider two bounded stopping times  $\sigma, \tau$ . Then for any  $\beta > 1$ , with probability one

$$\mathbb{P}(\{\exists u \in [0, \tau - \sigma] Y_{\sigma+u} - Y_\sigma < u - 1\} \cap \{\forall u \in [0, \tau - \sigma] \lambda_{\sigma+u} \geq \beta\} | \mathcal{F}_\sigma) \leq 1 - q,$$

for some  $q > 0$  depending only on  $\beta$ .

**Proof.** Fixing  $A \in \mathcal{F}_\sigma$  with  $\mathbb{P}(A) > 0$  and using notation from the proof of Lemma 3.4, we have

$$\begin{aligned} &\tilde{\mathbb{P}}(\{\exists u \in [0, \tau - \sigma] Y_{\sigma+u} - Y_\sigma < u - 1\} \cap \{\forall u \in [0, \tau - \sigma] \lambda_{\sigma+u} \geq \beta\}) \\ &= \tilde{\mathbb{P}}(\{\exists u \in [0, \tau - \sigma] N_{\tilde{\Lambda}_u} < u - 1\} \cap \{\forall u \in [0, \tau - \sigma] \lambda_{\sigma+u} \geq \beta\}) \\ &\leq \tilde{\mathbb{P}}(\exists u \geq 0 N_{\beta u} < u - 1). \end{aligned}$$

The law of large numbers and the Markov property for the Poisson process implies that for  $\beta > 1$  the last probability is bounded by  $1 - q$  for some  $q > 0$  depending only on  $\beta$ . Since  $A$  in the above argument is arbitrary, we obtain the lemma.  $\square$

**Lemma A.3.** Let  $Y$  be a counting process with bounded intensity  $\lambda$ . Consider two bounded stopping times  $\sigma, \tau$ . Then for any  $\beta > 1, s \geq 0$ , with probability one,

$$\mathbb{P}(\{\exists u \in [s, \tau - \sigma] Y_{\sigma+u} - Y_\sigma \leq u + as\} \cap \{\forall u \in [0, \tau - \sigma] \lambda_{\sigma+u} \geq \beta\} | \mathcal{F}_\sigma) \leq e^{-cs},$$

where  $a = \frac{\beta-1}{2} > 0$  and  $c = \frac{1}{2}(1 - \frac{1+a}{\beta})^2\beta$ .

**Proof.** Denote the event in question by  $\mathcal{E}$  and fix  $A \in \mathcal{F}_\sigma$  with positive probability. Using the notation from the proof of Lemma 3.4 and arguments from the proof of Lemma A.2 we get

$$\tilde{\mathbb{P}}(\mathcal{E}) \leq \mathbb{P}(\exists_{u \geq s} N_{\beta u} \leq u + as) \leq \mathbb{P}\left(\exists_{u \geq \beta s} \frac{N_u}{u} \leq 1 - \rho\right),$$

where  $N$  is a Poisson process with intensity one and  $\rho = 1 - \frac{1+a}{\beta} \in (0, 1)$ . Let  $\mathcal{G}_u$  be the  $\sigma$ -field generated by  $\{N_t : t \geq u\}$  and note that for any  $0 < u < t$  we have  $\mathbb{E}(N_u/u | \mathcal{G}_t) = N_t/t$ . Thus for any  $a \in \mathbb{R}$  the process  $(\exp(aN_u/u))_{u>0}$  is a reversed submartingale. Using Doob's maximal inequality for reversed submartingales we get

$$\begin{aligned} \mathbb{P}\left(\exists_{u \geq \beta s} \frac{N_u}{u} \leq (1 - \rho)\beta\right) &= \inf_{b>0} \mathbb{P}\left(\sup_{u \geq \beta s} \exp(-bN_u/u) \geq \exp(-b(1 - \rho))\right) \\ &\leq \inf_{b>0} \mathbb{E} \exp\left(-b \frac{N_{\beta s}}{\beta s} + b(1 - \rho)\right) = \inf_{b>0} \exp(\beta s(e^{-b/(\beta s)} - 1) + b(1 - \rho)) \\ &= \inf_{b>0} \exp(s\beta(e^{-b} - 1) + b\beta s(1 - \rho)) \leq \exp\left(-\frac{1}{2}\rho^2\beta s\right). \end{aligned}$$

As in previous lemmas, since  $A$  is arbitrary, this implies the assertion.  $\square$

## Appendix B: Estimates on the number of bridges

Here we collect useful estimates enabling us to compare the distribution of bridges for general  $\theta > 0$  with the i.i.d. case, i.e.,  $\theta = 1$ .

In the first lemma we show that for  $\theta$  not necessarily equal to 1 the number of bridges using any subset of edges can still be approximated by a Poisson distribution (with parameter depending on  $\theta$ ).

For any measurable  $A \subset E \times [0, 1)$  and a configuration  $X \in \mathfrak{X}$  we denote  $X_A := \{x \in X : x \in A\}$ .

**Lemma B.1.** Let  $\lambda_n := \frac{\beta\Theta^{-1}}{n-1}$ ,  $\bar{\lambda}_n := \frac{\beta\Theta}{n-1}$ . Then for any measurable  $A \subset E \times [0, 1)$  we have

$$\mathbb{P}(|X_A| \geq 1 | X_{E \times [0, 1) \setminus A}) \in [1 - e^{-|A|\lambda_n}, 1 - e^{-|A|\bar{\lambda}_n}], \quad (\text{B.1})$$

and

$$\mathbb{P}(|X_A| \geq k | X_{E \times [0, 1) \setminus A}) \leq e^{|A|(\bar{\lambda}_n - \lambda_n)} \mathbb{P}(Y \geq k), \quad (\text{B.2})$$

where  $Y$  has Poisson distribution with parameter  $\bar{\lambda}_n|A|$ .

**Proof.** For  $X \in \mathfrak{X}$  and  $A \subset E \times [0, 1)$  let  $U_A^=0 = \{X \in \mathfrak{X} : |X_A| = 0\}$  and  $U_A^{\geq k} = \{X \in \mathfrak{X} : |X_A| \geq k\}$ . Furthermore, let  $V$  be any event measurable with respect to  $X_{E \times [0, 1) \setminus A}$  such that  $\mathcal{B}(V) > 0$ . We have

$$\frac{\mathbb{P}(U_A^{\geq k} | V)}{\mathbb{P}(U_A^=0 | V)} = \frac{\mu_{\beta, \theta, \mathcal{C}}(U_A^{\geq k} \cap V)}{\mu_{\beta, \theta, \mathcal{C}}(U_A^=0 \cap V)} = \frac{\int_{\mathfrak{X}} \mathbb{1}_{U_A^{\geq k}}(X_A) \mathbb{1}_V \theta^{\mathcal{C}(X_A \cup X_{E \times [0, 1) \setminus A})} \mathcal{B}(dX)}{\int_{\mathfrak{X}} \mathbb{1}_{U_A^=0}(X_A) \mathbb{1}_V \theta^{\mathcal{C}(X_{E \times [0, 1) \setminus A})} \mathcal{B}(dX)}.$$

By the Lipschitz property of  $\mathcal{C}$  we have  $|\mathcal{C}(X_A \cup X_{E \times [0, 1) \setminus A}) - \mathcal{C}(X_{E \times [0, 1) \setminus A})| \leq |X_A|$ . Furthermore, using independence of  $X_A$  and  $X_{E \times [0, 1) \setminus A}$  under  $\mathcal{B}$  we get

$$\frac{\mathbb{P}(U_A^{\geq k} | V)}{\mathbb{P}(U_A^=0 | V)} \leq \frac{\int_{\mathfrak{X}} \mathbb{1}_{U_A^{\geq k}}(X_A) \Theta^{|X_A|} \mathcal{B}(dX)}{\int_{\mathfrak{X}} \mathbb{1}_{U_A^=0}(X_A) \mathcal{B}(dX)} = \sum_{\ell=k}^{\infty} \frac{(\bar{\lambda}_n|A|)^\ell}{\ell!}, \quad (\text{B.3})$$

where in the last equality we used the fact that under  $\mathcal{B}$  the random variable  $|X_A|$  is Poisson with parameter  $\beta|A|/(n-1)$ . As  $\mathbb{P}(U_A^{\geq 1} | V) + \mathbb{P}(U_A^=0 | V) = 1$ , by elementary calculations we get

$$\mathbb{P}(U_A^{\geq 1} | V) \leq 1 - e^{-|A|\bar{\lambda}_n}.$$

Thus we obtain the upper bound in (B.1). The lower bound can be proven analogously. Now from (B.1) and (B.3) we infer (B.2).  $\square$

The second lemma gives tail bounds on the number of bridges in terms of tails of Poisson variables with parameters depending on  $\theta$  and  $\beta$ .

**Lemma B.2.** *Let  $X \in \mathfrak{X}$  be distributed according to  $\mu_{\beta, \theta, c}$ . For any  $k \in \mathbb{N}$ ,  $t \in [0, 1)$  and  $s \in (0, 1]$  such that  $t + s \leq 1$  we have*

$$\begin{aligned} \mathbb{P}(|X \cap (E \times [t, t+s])| \geq k) &\leq e^{s\beta(\Theta - \frac{1}{\Theta})n^2} \mathbb{P}(Y \geq k), \\ \mathbb{P}(|X \cap (E \times [t, t+s])| \leq k) &\leq e^{s\beta(\Theta - \frac{1}{\Theta})n^2} \mathbb{P}(Y \leq k), \end{aligned}$$

where  $Y$  is a Poisson variable with parameter  $s\Theta\beta n^2$ .

In particular for  $k_n = o(n^2)$  we have for sufficiently large  $n$

$$\mathbb{P}(|X \cap (E \times [t, t+s])| \leq k_n) \leq Ce^{-cn^2}$$

for some  $C, c > 0$  depending on  $\beta, \theta$  and  $s$ .

**Proof.** Let  $X_{t,t+s} = X \cap (E \times [t, t+s])$ . Observe that by the Lipschitz condition (2) we have

$$\theta^{\mathcal{C}(X \setminus X_{t,t+s})} \Theta^{-|X_{t,t+s}|} \leq \theta^{\mathcal{C}(X)} \leq \theta^{\mathcal{C}(X \setminus X_{t,t+s})} \Theta^{|X_{t,t+s}|}. \quad (\text{B.4})$$

To prove the first estimate we write

$$\mathbb{P}(|X_{t,t+s}| \geq k) = \frac{\int_{\mathfrak{X}} \mathbb{1}_{\{|X_{t,t+s}| \geq k\}} \theta^{\mathcal{C}(X)} \mathcal{B}(dX)}{\int_{\mathfrak{X}} \theta^{\mathcal{C}(X)} \mathcal{B}(dX)}.$$

By employing (B.4) we can bound the right hand side from above by

$$\frac{\int_{\mathfrak{X}} \mathbb{1}_{\{|X_{t,t+s}| \geq k\}} \theta^{\mathcal{C}(X \setminus X_{t,t+s})} \Theta^{|X_{t,t+s}|} \mathcal{B}(dX)}{\int_{\mathfrak{X}} \theta^{\mathcal{C}(X \setminus X_{t,t+s})} \Theta^{-|X_{t,t+s}|} \mathcal{B}(dX)} = \frac{\int_{\mathfrak{X}} \mathbb{1}_{\{|X_{t,t+s}| \geq k\}} \Theta^{|X_{t,t+s}|} \mathcal{B}(dX)}{\int_{\mathfrak{X}} \Theta^{-|X_{t,t+s}|} \mathcal{B}(dX)},$$

where in the equality we used the fact that the integrals factorize due to the independence property of the Poisson point process  $\mathcal{B}$  for disjoint time intervals.

To estimate the integrals, we observe that under  $\mathcal{B}$  the variable  $|X_{t,t+s}|$  has Poisson distribution with parameter  $\lambda = s\beta n^2$ . Thus we can write

$$\int_{\mathfrak{X}} \Theta^{-|X_{t,t+s}|} \mathcal{B}(dX) = \mathbb{E} \Theta^{-|X_{t,t+s}|} = e^{\lambda(\frac{1}{\Theta} - 1)},$$

and

$$\int_{\mathfrak{X}} \mathbb{1}_{\{|X_{t,t+s}| \geq k\}} \Theta^{|X_{t,t+s}|} \mathcal{B}(dX) = \sum_{i=k}^{\infty} e^{-\lambda} \frac{(\Theta\lambda)^i}{i!} = e^{\lambda(\Theta - 1)} \mathbb{P}(Y \geq k). \quad (\text{B.5})$$

where  $Y$  is a Poisson variable with parameter  $\Theta\lambda = s\Theta\beta n^2$ . Thus we obtain

$$\mathbb{P}(|X_{t,t+s}| \geq k) \leq e^{\lambda(\frac{1}{\Theta} - 1)} e^{\lambda(\Theta - 1)} \mathbb{P}(Y \geq k) = e^{\beta(\Theta - \frac{1}{\Theta})sn^2} \mathbb{P}(Y \geq k).$$

The proof of the second estimate is analogous.

For the case of  $k_n = o(n^2)$  we use Bennett's inequality – if  $Y$  is a Poisson variable with parameter  $\lambda$ , then for any  $0 \leq x \leq \lambda$  we have

$$\mathbb{P}(X \leq \lambda - x) \leq \exp \left\{ -\frac{x^2}{2\lambda} \psi \left( -\frac{x}{\lambda} \right) \right\},$$

where

$$\psi(t) = \frac{(1+t) \log(1+t) - t}{t^2/2}$$

for  $t \geq -1$ .

Writing  $k_n = \varepsilon_n \lambda$ , with  $\lambda = s\beta\Theta n^2$  and  $\varepsilon_n \rightarrow 0$  as  $n \rightarrow \infty$ , we have

$$\mathbb{P}(Y \leq \varepsilon_n \lambda) \leq e^{-\lambda(\varepsilon_n \log \varepsilon_n + 1 - \varepsilon_n)}$$

and thus

$$\mathbb{P}(|X \cap (E \times [t, t+s])| \leq k_n) \leq e^{\lambda - \frac{\lambda}{\Theta^2}} e^{-\lambda(\varepsilon_n \log \varepsilon_n + 1 - \varepsilon_n)} = e^{-\lambda(\frac{1}{\Theta^2} - \varepsilon_n + \varepsilon_n \log \varepsilon_n)}.$$

Since  $\varepsilon_n - \varepsilon_n \log \varepsilon_n \rightarrow 0$  as  $n \rightarrow \infty$ , the right hand side is at most  $C e^{-cn^2}$  for some  $C, c > 0$ .  $\square$

The following lemma will be useful in the proof of Lemma 3.6.

**Lemma B.3.** *As in the previous lemma, let  $X_{t,t+h} = X \cap (E \times [t, t+h])$ . For any  $k \in \mathbb{N}$  and  $t \in [0, 1)$  we have*

$$\int_{\mathfrak{X}} \mathbb{1}_{\{|X_{t,t+h}| \geq k\}} |X_{t,t+h}| \Theta^{|X_{t,t+h}|} \mathcal{B}(dX) = O(h^k),$$

where the implicit constant may depend on  $\Theta, \beta, n$  and  $k$ .

**Proof.** As under  $\mathcal{B}$  the variable  $|X_{t,t+h}|$  has Poisson distribution with parameter  $h\beta n^2$ , we have for small enough  $h > 0$

$$\int_{\mathfrak{X}} \mathbb{1}_{\{|X_{t,t+h}| \geq k\}} |X_{t,t+h}| \Theta^{|X_{t,t+h}|} \mathcal{B}(dX) = \sum_{i=k}^{\infty} e^{-h\beta n^2} i \frac{(\Theta h\beta n^2)^i}{i!} \leq h^k e^{-h\beta n^2} \sum_{i=k}^{\infty} \frac{(\Theta \beta n^2)^i}{(i-1)!} = O(h^k) \quad \square$$

### Appendix C: Proof of Lemma 3.6

**Proof of Lemma 3.6.** Recall from the introduction the definition of the canonical probability space  $(\mathfrak{X}, S, \mathcal{B})$ . Fix  $A \in \mathcal{F}_s$  and define for  $t > s$

$$f(t) = \mathbb{E} J_t \mathbb{1}_A.$$

Note that by Lebesgue's dominated convergence theorem and the fact that with probability one there are no jumps at a prescribed deterministic moment in time,  $f$  is a continuous function.

In what follows we will denote  $t' = t \bmod 1$ . Consider any  $t_1, t_2$  such that  $t_1 < t_2 < \lceil t_1 \rceil$ . Note that  $f(t_2) - f(t_1)$  is bounded from above by the mean number of bridges in  $E \times (t'_1, t'_2]$ . Using (2) and independence properties of the Poisson process we thus get

$$\begin{aligned} |f(t_2) - f(t_1)| &\leq \frac{1}{Z_{\beta, \theta, C}} \int_{\mathfrak{X}} |X \cap (E \times (t'_1, t'_2])| \theta^{C(X)} \mathcal{B}(dX) \\ &\leq \frac{1}{Z_{\beta, \theta, C}} \int_{\mathfrak{X}} |X \cap (E \times (t'_1, t'_2])| \Theta^{|X \cap (E \times (t'_1, t'_2])|} \Theta^{C(\emptyset) + |X \cap (E \times (t'_1, t'_2])^c|} \mathcal{B}(dX) \\ &= \frac{1}{Z_{\beta, \theta, C}} \int_{\mathfrak{X}} \Theta^{C(\emptyset) + |X \cap (E \times (t'_1, t'_2])^c|} \mathcal{B}(dX) \int_{\mathfrak{X}} |X \cap (E \times (t'_1, t'_2])| \Theta^{|X \cap (E \times (t'_1, t'_2])|} \mathcal{B}(dX) \\ &\leq K \int_{\mathfrak{X}} |X \cap (E \times (t'_1, t'_2])| \Theta^{|X \cap (E \times (t'_1, t'_2])|} \mathcal{B}(dX) \\ &= K \Theta \beta n^2 |t_2 - t_1| e^{(\Theta-1)\beta n^2 |t_2 - t_1|}, \end{aligned}$$



where  $K$  is some constant (depending on  $n$  and the parameters of the process but not on  $t_i$ ). Thus  $f$  is locally Lipschitz which implies that  $f'$  exists almost everywhere and  $f$  satisfies the fundamental theorem of calculus.

Consider any differentiability point  $t > s$  of  $f$  and small  $h > 0$ , in particular small enough so that  $t' + h < 1$ . As with probability one there is no jump at time  $t$ , by using Lemma B.3 and exploiting the properties of the process  $Q$  (specifically the fact that it may jump only when  $\mathcal{X}$  jumps and that it is càdlàg) we can write for  $h \searrow 0$

$$\begin{aligned} f(t+h) - f(t) &= \mathbb{E}(J_{t+h} - J_t) \mathbb{1}_A = \mathbb{E} \mathbb{1}_A \mathbb{1}_{\{J_{t+h} - J_t = 1\}} \mathbb{1}_{\{|X \cap (E \times (t', t'+h))| = 1\}} + o(h) \\ &= \sum_{v, w \in V} \mathbb{P}(A \cap \{\mathcal{X}_t = v\} \cap B_{v, w, h} \cap \{w \in Q_t\}) + o(h), \end{aligned}$$

where

$$B_{v, w, h} = \{\text{there is a unique bridge in } E \times (t', t' + h], \text{ it is unexplored at time } t \text{ and joins } v \text{ with } w\}.$$

Consider an additional event

$$C_{v, w, h} = \{v, w\} \times [t', t' + h] \text{ has not been visited by } \mathcal{X} \text{ before time } t\}.$$

Note that  $B_{v, w, h} \cap \{\mathcal{X}_t = v\} \cap C_{v, w, h}^c = \emptyset$ . Moreover  $B_{v, w, h} \cap \{\mathcal{X}_t = v\} \subset \{w \in A_t\}$  (recall that  $A_t$  is the set of vertices which at time  $t$  are available to the CRW by a fresh jump). Thus, we have

$$\begin{aligned} \mathbb{E}(J_{t+h} - J_t) \mathbb{1}_A &= \sum_{v, w \in V} \mathbb{E} \mathbb{1}_{\{\mathcal{X}_t = v, w \in A_t \cap Q_t\}} \mathbb{1}_{A \cap B_{v, w, h} \cap C_{v, w, h}} + o(h) \\ &= Z_{\beta, \theta, C}^{-1} \sum_{v, w \in V} \int_{\mathfrak{X}} \mathbb{1}_{\{\mathcal{X}_t = v, w \in A_t \cap Q_t\}} \mathbb{1}_{A \cap B_{v, w, h} \cap C_{v, w, h}} \theta^{C(X)} \mathcal{B}(dX) + o(h). \end{aligned}$$

Denote a summand above by  $I_{v, w, h}$ . Since for a while we will be working with fixed  $v, w$  denote for simplicity  $e = \{v, w\}$ .

Note that for any  $U \in \mathcal{F}_t$ , the event  $C_{v, w, h} \cap U$  is measurable with respect to the restricted process  $X'_h := X \setminus (\{e\} \times (t', t' + h])$ .

Denote also

$$D_{v, w, h} = \{X \cap ((E \setminus \{e\}) \times (t', t' + h)) = \emptyset\}.$$

Recall that conditionally on having just one point of a Poisson process in an interval, its position is distributed uniformly. Combining this with the independence properties of Poisson processes we get

$$I_{v, w, h} = \int_{\mathfrak{X}} \mathbb{1}_{\{\mathcal{X}_t = v, w \in A_t \cap Q_t\}} \mathbb{1}_{A \cap D_{v, w, h} \cap C_{v, w, h}} \kappa e^{-\kappa h} \int_{t'}^{t'+h} \theta^{C(X'_h \cup \{(e, u)\})} du \mathcal{B}(dX),$$

where  $\kappa = \frac{\beta}{n-1}$ . Note that on  $D_{v, w, h} \cap \{X \cap (E \times \{t'\}) = \emptyset\}$ , the function  $[t', t' + h] \ni u \mapsto C((X'_h \cup \{(e, u)\}))$  is constant, so using the fact that almost surely there are no bridges at height  $t'$  we can further write

$$I_{v, w, h} = \int_{\mathfrak{X}} \mathbb{1}_{\{\mathcal{X}_t = v, w \in A_t \cap Q_t\}} \mathbb{1}_{A \cap D_{v, w, h} \cap C_{v, w, h}} \kappa e^{-\kappa h} h \theta^{C(X'_h \cup \{(e, t')\})} \mathcal{B}(dX).$$

By Lemma B.3 we also have

$$\int_{\mathfrak{X}} \mathbb{1}_{D_{v, w, h}^c} \theta^{C(X'_h \cup \{(e, t')\})} \mathcal{B}(dX) \leq \int_{\mathfrak{X}} \mathbb{1}_{D_{v, w, h}^c} \Theta^{C(\emptyset) + |X| + 1} \mathcal{B}(dX) = O(h),$$

so we get

$$I_{v, w, h} = \int_{\mathfrak{X}} \mathbb{1}_{\{\mathcal{X}_t = v, w \in A_t \cap Q_t\}} \mathbb{1}_{A \cap C_{v, w, h}} \kappa e^{-\kappa h} h \theta^{C(X'_h \cup \{(e, t')\})} \mathcal{B}(dX) + o(h).$$

Similarly, up to an error of order  $o(h)$  we can restrict the integration to the set  $\{X'_h = X\}$ , replace  $\theta^{C(X'_h \cup \{(e, t')\})}$  by  $\theta^{C(X \cup \{(e, t')\})}$  and then again return to integration over the whole space  $\mathfrak{X}$ , obtaining

$$I_{v, w, h} = \int_{\mathfrak{X}} \mathbb{1}_{\{\mathcal{X}_t = v, w \in A_t \cap Q_t\}} \mathbb{1}_{A \cap C_{v, w, h}} \kappa e^{-\kappa h} h \theta^{C(X \cup \{(e, t')\})} \mathcal{B}(dX) + o(h).$$

Note also that on the event  $\{w \in A_t\} \cap \{\mathcal{X}_t = v\}$  neither  $(v, t')$  nor  $(w, t')$  could have been visited by  $\mathcal{X}$  before time  $t$ . Since with probability one there are only finitely many bridges, this implies that up to a set of probability zero  $\{w \in A_t\} \cap \{\mathcal{X}_t = v\} \cap C_{v,w,h} \nearrow \{w \in A_t\} \cap \{\mathcal{X}_t = v\}$  as  $h \searrow 0$ . Thus we get

$$\lim_{h \rightarrow 0^+} \frac{I_{v,w,h}}{h} = \kappa \int_{\mathfrak{X}} \mathbb{1}_{\{\mathcal{X}_t = v, w \in A_t \cap Q_t\}} \mathbb{1}_A \theta^{C(X \cup \{(e, t')\})} \mathcal{B}(dX),$$

which implies that

$$f'(t) = \mathbb{E} \mathbb{1}_A \kappa \sum_{w \in A_t \cap Q_t} Y_t^w = \mathbb{E} \mathbb{1}_A \kappa \sum_{w \in A_t \cap Q_t} \mathbb{E}(Y_t^w | \mathcal{F}_t),$$

where  $Y_t^w = \mathbb{1}_{\{\{\mathcal{X}_t, w\} \in E\}} \theta^{C(X \cup \{(\mathcal{X}_t, w), t')\}) - C(X)}$ .

Now for fixed  $w$  we have that  $Y_t^w : [0, \infty) \times \mathfrak{X} \rightarrow \mathbb{R}$  is measurable with respect to  $\text{Bor}([0, \infty)) \otimes \mathcal{S}$ , so by Corollary 2 in [26] we obtain that there is a choice of  $\mathbb{E}(Y_t^w | \mathcal{F}_t)$  which as a stochastic process is  $\mathcal{F}_t$ -progressively measurable. Set  $S_t = \kappa \sum_{w \in A_t \cap Q_t} \mathbb{E}(Y_t^w | \mathcal{F}_t)$  and define the progressively measurable process  $\lambda_t = (\kappa |A_t \cap Q_t| \Theta^{-1}) \vee (S_t \wedge (\kappa |A_t \cap Q_t| \Theta))$ . Note that by the Lipschitz condition (2) on  $\mathcal{C}$ , for every  $t$  we have  $S_t = \lambda_t$  almost surely.

Thus, by Fubini's theorem, we have for  $t > s$ ,

$$\mathbb{E}(J_t - J_s) \mathbb{1}_A = \int_s^t f'(u) du = \int_s^t \mathbb{E} \mathbb{1}_A S_u du = \mathbb{E} \mathbb{1}_A \int_0^t \lambda_u du - \mathbb{E} \mathbb{1}_A \int_0^s \lambda_u du,$$

which proves that  $J_t - \int_0^t \lambda_u du$  is indeed a martingale with respect to  $(\mathcal{F}_t)_{t \geq 0}$ .  $\square$

## Appendix D: Proof of Lemma 5.6

**Proof of Lemma 5.6.** By Lemma 4.1  $p_{u,e}$ , the conditional probability of an edge  $e$  being added to the graph process in the transition from  $u$  to  $u + 1$  belongs to  $[\Theta^{-2}/|E|, \Theta^2/|E|]$ . Let  $\{U_{u,e}\}_{u \in \mathbb{N}, e \in E}$  be i.i.d. random variables uniformly distributed on  $[0, 1]$  which are also independent of  $G^s$ . We define a coupled random graph process  $\tilde{G}^s$ . First, we set  $\tilde{G}_0^s := G_0^s$ , then in the transition from  $u$  to  $u + 1$  an edge  $e$  is added to the edge set of  $\tilde{G}_u^s$  if and only if it is added to  $G_u^s$  and  $U_{u,e} \leq (\Theta^2 |E| p_{u,e})^{-1}$ . Note that in the new process at each step there is probability  $1 - \Theta^{-2}$  of no new edge being added, and if a new edge is added, each one is chosen with probability  $1/|E|$ , independently of the previous steps.

Clearly, for any  $s, u, \ell$  we have  $\tilde{\mathcal{G}}_{s,u}(\ell) \subset \mathcal{G}_{s,u}(\ell)$ . As the process  $\tilde{G}^s$  is monotonic it is enough to prove the statement for the process  $\tilde{G}^s$  and  $u = (n^2/\sqrt{\ell}) \log n$ .

The proof is an implementation of the classical sprinkling argument, introduced in [1]. We will work conditionally on  $\tilde{G}_0^s$ . To shorten the notation we denote  $\mathbb{Q}(\cdot) = \mathbb{P}(\cdot | \tilde{G}_0^s, |\mathcal{G}_{s,0}(\ell)| \geq \delta n^2, |X| = h)$ . Also, for any two sets  $A, B \subset V$  by  $E(A, B)$  we will denote the number of edges  $\{v, w\} \in E$  such that  $v \in A, w \in B$ .

The event  $\{|\tilde{\mathcal{G}}_{s,u}(\delta n^2/8)| < \delta n^2/8\}$  (i.e., there is no component of size at least  $\delta n^2/8$  in  $\tilde{G}_u^s$ ) implies that  $\tilde{\mathcal{G}}_{s,0}(\ell)$  can be partitioned into two sets  $A$  and  $B$  such that each of them has size at least  $\delta n^2/4$ , each of them is a union of some connected components of  $\tilde{G}_0^s$ , and there are no paths joining  $A$  and  $B$  in  $\tilde{G}_u^s$ . We will show that such a partition is unlikely to exist in  $\tilde{\mathcal{G}}_{s,0}(\ell)$ .

Fix two sets  $A$  and  $B$  which partition  $\tilde{\mathcal{G}}_{s,0}(\ell)$  as above, each of size at least  $\delta n^2/4$ , and let  $\mathcal{C}_{A,B}(G)$  be the event that no path in a graph  $G \subset H_n$  has one endpoint in  $A$  and the other endpoint in  $B$ . We write simply  $\mathcal{C}_{A,B}$  for  $\mathcal{C}_{A,B}(\tilde{G}_u^s)$ . Let

$$D_{A,B} := \{v \in V : E(\{v\}, A) \geq \delta^2 n/64 \text{ and } E(\{v\}, B) \geq \delta^2 n/64\}$$

be the set of vertices that in the Hamming graph  $H_n$  have at least  $\delta^2 n/64$  neighbors both in  $A$  and in  $B$ . Let  $D_1 := \{v \in D_{A,B} : v \in A \cup B\}$  and  $D_2 := D_{A,B} \setminus D_1$ .

First we bound the size of  $D_{A,B}$ . Note that there are at least  $\delta^2 n^4/16$  paths of length 2 in  $H_n$  between  $A$  and  $B$ , since in  $H_n$  all vertices are connected by a path of length at most 2 and we assumed  $|A|, |B| \geq \delta n^2/4$ . On the other hand, for every  $v \notin D_{A,B}$ , there are at most  $\delta^2 n/64 \cdot 2(n-1) \leq \delta^2 n^2/32$  paths of length 2 between  $A$  and  $B$  with  $v$  as the midpoint. Every  $v \in D_{A,B}$  can be a midpoint in at most  $4n^2$  paths of length 2 between  $A$  and  $B$ .

Hence the total number of paths of length 2 between  $A$  and  $B$  is bounded from above by  $(\delta^2 n^2 / 32) |D_{A,B}^c| + 4n^2 |D_{A,B}|$ . Combining this with the lower bound we get that

$$\frac{\delta^2 n^2}{32} (n^2 - |D_{A,B}|) + 4n^2 |D_{A,B}| \geq \frac{\delta^2 n^4}{16}$$

and thus there exists a constant  $\rho > 0$  (depending only on  $\delta$ ) such that  $|D_{A,B}| \geq \rho n^2$ .

Throughout the rest of the proof it will be convenient to work with a random graph which has edges chosen independently. Since in the graph process  $\tilde{G}^s$  at each step with probability  $\Theta^{-2}$  a uniformly random edge is added, the graph  $\tilde{G}_u^s$  is obtained by adding  $k$  uniformly random edges (with multiple edges allowed) to  $\tilde{G}_0^s$ , where  $k$  has binomial distribution corresponding to  $u$  trials with success probability  $\Theta^{-2}$ . Now let  $\tilde{G}_{s,u}$  be a random graph obtained by adding each edge  $e \in E$  to  $\tilde{G}_0^s$  independently with probability  $p = \frac{u\Theta^{-2}}{2|E|}$ .

Let  $\bar{\mathbb{Q}}(P)$  denote the probability that the graph  $\tilde{G}_{s,u}$  satisfies property  $P$ . By using the second moment method one can see that with high probability after removing multiple edges from  $\tilde{G}_u^s$  we will still have (for  $n$  large enough) at least  $\frac{u\Theta^{-2}}{2}$  edges in the graph, distributed uniformly. Indeed, as  $u$  is small compared to  $|E|$ , the expected number of edges chosen at least once in  $\tilde{G}_u^s$  is at least, say,  $\frac{3}{4}u\Theta^{-2}$ . As the edges present in  $\tilde{G}_u^s$  are negatively correlated, the variance can be bounded from above by  $u\Theta^{-2}$ , which implies that with high probability we have at least  $\frac{u\Theta^{-2}}{2}$  distinct edges. Therefore by the equivalence of  $G(n, p)$  and  $G(n, M)$  random graph models with respect to monotone properties (see e.g., [21, Section 1.4])  $\bar{\mathbb{Q}}(P) \rightarrow 0$  as  $n \rightarrow \infty$  will imply  $\mathbb{Q}(P) \rightarrow 0$  for any decreasing graph property  $P$ .

From now on we will work with the graph  $\tilde{G}_{s,u}$ . Let  $\tilde{G}_{s,u}(k)$  denote the set of vertices of  $\tilde{G}_{s,u}$  contained in connected components of size at least  $k$ . Let

$$E_1 = \{\{v, w\} \in E : v \in D_1, \{v, w\} \in E(A, B)\}$$

and let  $\mathcal{C}(E_1)$  denote the event that none of the edges from  $E_1$  are in  $\tilde{G}_{s,u}$ . Let  $\mathcal{C}(D_2)$  denote the event that none of the vertices in  $D_2$  have neighbors both in  $A$  and in  $B$  in  $\tilde{G}_{s,u}$ . Clearly we have

$$\bar{\mathbb{Q}}(\mathcal{C}_{A,B}(\tilde{G}_{s,u})) \leq \bar{\mathbb{Q}}(\mathcal{C}(E_1) \cap \mathcal{C}(D_2)) = \bar{\mathbb{Q}}(\mathcal{C}(E_1)) \bar{\mathbb{Q}}(\mathcal{C}(D_2)). \quad (\text{D.1})$$

We first estimate  $\bar{\mathbb{Q}}(\mathcal{C}(E_1))$ . Because of independence of the edges in  $\tilde{G}_{s,u}$  we have

$$\bar{\mathbb{Q}}(\mathcal{C}(E_1)) \leq (1-p)^{|E_1|}.$$

Since  $|E(\{v\}, B)| \geq \delta^2 n / 64$  for each  $v \in D_1 \cap A$  and likewise  $|E(\{v\}, A)| \geq \delta^2 n / 64$  for each  $v \in D_1 \cap B$ , we easily get  $|E_1| \geq \frac{1}{2} |D_1| \delta^2 n / 64$ , so

$$\bar{\mathbb{Q}}(\mathcal{C}(E_1)) \leq (1-p)^{|D_1| \delta^2 n / 128} \leq e^{-\frac{p|D_1| \delta^2 n}{128}}.$$

For the upper bound on  $\bar{\mathbb{Q}}(\mathcal{C}(D_2))$ , we note that by independence of the edges in  $\tilde{G}_{s,u}$

$$\begin{aligned} \bar{\mathbb{Q}}(\mathcal{C}(D_2)) &= \prod_{v \in D_2} (1 - \bar{\mathbb{Q}}(v \text{ has neighbors both in } A \text{ and } B)) \\ &= (1 - \bar{\mathbb{Q}}(v \text{ has a neighbor in } A) \bar{\mathbb{Q}}(v \text{ has a neighbor in } B))^{|D_2|}. \end{aligned}$$

Since  $|E(\{v\}, A)| \geq \delta^2 n / 64$  for  $v \in D_2$ , we have

$$\bar{\mathbb{Q}}(v \text{ has a neighbor in } A) \geq 1 - (1-p)^{\delta^2 n / 64} \geq 1 - e^{-\frac{p\delta^2 n}{64}}.$$

An analogous estimate holds for  $B$ , which gives

$$\bar{\mathbb{Q}}(\mathcal{C}(D_2)) \leq (1 - (1 - e^{-\frac{p\delta^2 n}{64}})^2)^{|D_2|}.$$

Recall that  $p = \frac{u}{2\Theta^2|E|}$  and  $|E| = n^2(n-1)$ . Since  $|D_{A,B}| \geq \rho n^2$ , we have  $|D_1| \geq \frac{\rho}{2} n^2$  or  $|D_2| \geq \frac{\rho}{2} n^2$ . In the first case we get

$$\bar{\mathbb{Q}}(\mathcal{C}(E_1)) \leq e^{-\frac{\delta^2}{128} \frac{u}{2\Theta^2|E|} \frac{\rho n^2}{2} n} \leq e^{-cu}$$

for some  $c > 0$  depending on  $\delta$  and  $\Theta$ . In the second case we have (exploiting  $u \leq n^2$ )

$$\bar{\mathbb{Q}}(\mathcal{C}(D_2)) \leq \left(1 - \left(1 - e^{-\frac{\delta^2}{64} n \frac{u}{2\Theta^2 |E|}}\right)^2\right)^{\frac{\rho}{2} n^2} \leq C e^{-c' \frac{u^2}{n^2}}$$

for some  $C, c' > 0$  depending on  $\delta$  and  $\Theta$ .

As  $u = (n^2/\sqrt{\ell}) \log n$  and  $\log n \leq \sqrt{\ell}$ , we have  $u^2/n^2 = (n^2/\ell) \log^2 n \leq u$ . Coming back to (D.1), we obtain for some  $C, c > 0$

$$\bar{\mathbb{Q}}(\mathcal{C}_{A,B}(\bar{G}_{s,u})) \leq C e^{-c \frac{n^2}{\ell} \log^2 n}.$$

Let  $\mathcal{C}(G)$  denote the event that  $\mathcal{C}_{A,B}(G)$  holds for some partition  $A, B$  of the set  $\mathcal{G}_{s,0}(\ell)$ . Notice that there are at most  $2^{n^2/\ell}$  such partitions, so by performing a union bound we obtain

$$\bar{\mathbb{Q}}(\mathcal{C}(\bar{G}_{s,u})) \leq 2^{\frac{n^2}{\ell}} \cdot C e^{-c \frac{n^2}{\ell} \log^2 n}.$$

Recalling that  $\{|\bar{\mathcal{G}}_{s,u}(\delta n^2/8)| < \delta n^2/8\} \subset \mathcal{C}(\bar{G}_{s,u})$ , we have that there exist constants  $C_1, c_1 > 0$  (depending only on  $\delta, \Theta$  and  $u$ ) such that

$$\bar{\mathbb{Q}}(|\bar{\mathcal{G}}_{s,u}(\delta n^2/8)| < \delta n^2/8) \leq C_1 \exp\left\{\frac{n^2}{\ell} \log 2 - c_1 \frac{n^2 \log^2 n}{\ell}\right\}.$$

As  $\ell \leq n^2$  and  $n \rightarrow \infty$ , we have that the probability above converges to zero. Since the property of having a component of size at least  $\delta n^2/8$  is increasing, the same holds with  $\bar{\mathbb{Q}}(\cdot)$  and  $\bar{\mathcal{G}}_{s,u}$  replaced by  $\mathbb{Q}(\cdot)$  and  $\mathcal{G}_{s,u}$ . Thus we have

$$\mathbb{Q}(|\mathcal{G}_{s,u}(\delta n^2/8)| < \delta n^2/8) = \mathbb{P}(|\mathcal{G}_{s,u}(\delta n^2/8)| < \delta n^2/8 | G_0^s, |\mathcal{G}_{s,0}(\ell)| \geq \delta n^2, |X| = h) \leq 1 - a_n$$

for some  $a_n \nearrow 1$ . By integrating this bound over all  $G_0^s$  satisfying  $|\mathcal{G}_{s,0}(\ell)| \geq \delta n^2$  we obtain the statement of the lemma.  $\square$

## Acknowledgements

We would like to thank Roman Kotecký and Daniel Ueltschi for useful discussions. We thank Wojtek Samotij for providing us with the idea of the proof of Lemma 3.19. Research partially supported by the National Science Centre, Poland, grants no. 2015/18/E/ST1/00214 (RA), no. 2014/15/B/ST1/02165 (PM) and no. 2019/32/C/ST1/00525 (MK).

## References

- [1] M. Ajtai, J. Komlós and E. Szemerédi. Largest random component of a  $k$ -cube. *Combinatorica* **2** (1982) 1–7. MR0671140 <https://doi.org/10.1007/BF02579276>
- [2] G. Alon and G. Kozma. The probability of long cycles in interchange processes. *Duke Math. J.* **162** (2013) 1567–1585. MR3079255 <https://doi.org/10.1215/00127094-2266018>
- [3] G. Alon and G. Kozma. The mean-field quantum Heisenberg ferromagnet via representation theory, 2018. Available at [arXiv:1811.10530](https://arxiv.org/abs/1811.10530).
- [4] G. Alon and G. Kozma. Comparing with octopi, 2018. Available at [arXiv:1811.10537](https://arxiv.org/abs/1811.10537).
- [5] O. Angel. Random infinite permutations and the cyclic time random walk. In *Discrete Random Walks 9–16. Paris, 2003. Discrete Math. Theor. Comput. Sci. Proc., AC. Assoc. Discrete Math. Theor. Comput. Sci. Nancy, 2003.* MR2042369
- [6] N. Berestycki. Emergence of giant cycles and slowdown transition in random transpositions and  $k$ -cycles. *Electron. J. Probab.* **16** (5) (2011) 152–173. MR2754801 <https://doi.org/10.1214/EJP.v16-850>
- [7] N. Berestycki and G. Kozma. Cycle structure of the interchange process and representation theory. *Bull. Soc. Math. France* **143** (2015) 265–280. MR3351179 <https://doi.org/10.24033/bsmf.2686>
- [8] J. E. Björnberg. Large cycles in random permutation related to the Heisenberg model. *Electron. Commun. Probab.* **20** (55) (2015) 11. MR3384113 <https://doi.org/10.1214/ECP.v20-4328>
- [9] J. E. Björnberg. The free energy in a class of quantum spin systems and interchange processes. *J. Math. Phys.* **57** (2016), 073303, 17. MR3529570 <https://doi.org/10.1063/1.4959238>
- [10] J. E. Björnberg, M. Kotowski, B. Lees and P. Miłoś. The interchange process with reversals on the complete graph. *Electron. J. Probab.* **24** (2019), 43 pp. MR4017126 <https://doi.org/10.1214/19-ejp366>
- [11] J. E. Björnberg and D. Ueltschi. Critical parameter of random loop model on trees. *Ann. Appl. Probab.* **28** (2018) 2063–2082. MR3843823 <https://doi.org/10.1214/17-AAP1315>
- [12] B. Bollobás, G. Grimmett and S. Janson. The random-cluster model on the complete graph. *Probab. Theory Related Fields* **104** (1996) 283–317. MR1376340 <https://doi.org/10.1007/BF01213683>

- [13] P. Brémaud. *Point Processes and Queues. Springer Series in Statistics*. Springer, New York, 1981. [MR0636252](#)
- [14] P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** (1981) 159–179. [MR0626813](#) <https://doi.org/10.1007/BF00535487>
- [15] R. Durrett. *Random Graph Dynamics. Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2010. [MR2656427](#)
- [16] H.-O. Georgii, O. Häggström and C. Maes. The random geometry of equilibrium phases. *Phase Transit. Crit. Phenom.* **18** (2001) 1–142. [MR2014387](#) [https://doi.org/10.1016/S1062-7901\(01\)80008-2](https://doi.org/10.1016/S1062-7901(01)80008-2)
- [17] C. Goldschmidt, D. Ueltschi and P. Windridge. Quantum Heisenberg models and their probabilistic representations. In *Entropy and the Quantum II* 177–224. *Contemp. Math.* **552**. Amer. Math. Soc., Providence, RI, 2011. [MR2868048](#) <https://doi.org/10.1090/conm/552/10917>
- [18] A. Hammond. Infinite cycles in the random stirring model on trees. *Bull. Inst. Math. Acad. Sin. (N.S.)* **8** (2013) 85–104. [MR3097418](#)
- [19] A. Hammond. Sharp phase transition in the random stirring model on trees. *Probab. Theory Related Fields* **161** (2015) 429–448. [MR3334273](#) <https://doi.org/10.1007/s00440-013-0543-7>
- [20] A. Hammond and M. Hegde. Critical point for infinite cycles in a random loop model on trees. *Ann. Appl. Probab.* **29** (2019) 2067–2088. [MR3983335](#) <https://doi.org/10.1214/18-AAP1442>
- [21] S. Janson, T. Łuczak and A. Ruciński. *Random Graphs. Wiley Series in Discrete Mathematics and Optimization*. Wiley, New York, 2011. [MR1782847](#) <https://doi.org/10.1002/9781118032718>
- [22] R. Kotecký, P. Miłoś and D. Ueltschi. The random interchange process on the hypercube. *Electron. Commun. Probab.* **21** (2016). Paper No. 4, 9. [MR3485373](#) <https://doi.org/10.1214/16-ECP4540>
- [23] G. Lowther. *Compensators of Counting Processes*, 2011. Available at <https://almostsure.wordpress.com/2011/12/27/compensators-of-counting-processes/>.
- [24] C. McDiarmid. On the method of bounded differences. In *Surveys in Combinatorics, 1989: Invited Papers at the Twelfth British Combinatorial Conference* 148–188. *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1989. [MR1036755](#) <https://doi.org/10.1017/CBO9781107359949.008>
- [25] P. Miłoś and B. Şengül. Existence of a phase transition of the interchange process on the Hamming graph. *Electron. J. Probab.* **24** (2019), 21 pp. [MR3978214](#) <https://doi.org/10.1214/18-EJP171>
- [26] M. Ondreját and J. Seidler. On existence of progressively measurable modifications. *Electron. Commun. Probab.* **18** (20) (2013) 6. [MR3037218](#) <https://doi.org/10.1214/ECP.v18-2548>
- [27] O. Penrose. Bose-Einstein condensation in an exactly soluble system of interacting particles. *J. Stat. Phys.* **63** (1991) 761–781. [MR1115812](#) <https://doi.org/10.1007/BF01029210>
- [28] O. Schramm. Compositions of random transpositions. *Israel J. Math.* **147** (2005) 221–243. [MR2166362](#) <https://doi.org/10.1007/BF02785366>
- [29] B. Tóth. Phase transition in an interacting Bose system. An application of the theory of Ventsel’ and Freidlin. *J. Stat. Phys.* **61** (1990) 749–764. [MR1086297](#) <https://doi.org/10.1007/BF01027300>
- [30] B. Tóth. Improved lower bound on the thermodynamic pressure of the spin 1/2 Heisenberg ferromagnet. *Lett. Math. Phys.* **28** (1993) 75–84. [MR1224836](#) <https://doi.org/10.1007/BF00739568>
- [31] R. van der Hofstad and M. J. Łuczak. Random subgraphs of the 2D Hamming graph: The supercritical phase. *Probab. Theory Related Fields* **147** (2010) 1–41. [MR2594346](#) <https://doi.org/10.1007/s00440-009-0200-3>