Mixing of fast random walks on dynamic random permutations

Luca Avena [1]
Remco van der Hofstad [2]
Frank den Hollander [3]
Oliver Nagy [3]

March 25, 2024

Abstract

We analyse the mixing profile of a random walk on a dynamic random permutation, focusing on the regime where the walk evolves much faster than the permutation. Two types of dynamics generated by random transpositions are considered: one allows for coagulation of permutation cycles only, the other allows for both coagulation and fragmentation. We show that for both types, after scaling time by the length of the permutation and letting this length tend to infinity, the total variation distance between the current distribution and the uniform distribution converges to a limit process that drops down in a single jump. This jump is similar to a one-sided cut-off, occurs after a random time whose law we identify, and goes from the value 1 to a value that is a strictly decreasing and deterministic function of the time of the jump, related to the size of the largest component in Erdős-Rényi random graphs. After the jump, the total variation distance follows this function down to 0.

Keywords. Dynamic random permutation, random transposition, coagulation, fragmentation, Erdős-Rényi random graph, random walk, mixing time, cut-off, split-merge dynamics.
MSC2010. 05C81, 37A25, 60K37, 82C27.
Acknowledgment. The work in this paper was supported by the Netherlands Organisation for Scientific Research (NWO) through Gravitation-grant NETWORKS-024.002.003. The authors thank Nathanaël Berestycki and Sam Olesker-Taylor for useful discussions. The paper was finalized while the authors were participating in the ICTS-NETWORKS workshop "Challenges in Networks" (ICTS/NETWORKS2024/01) at the International Centre for Theoretical Sciences (ICTS) in Bengaluru, India. The authors are thankful to ICTS for hospitality.

[1]Dipartimento di Matematica e Informatica 'Ulisse Dini', Università degli Studi di Firenze, Viale Morgagni 67/a, 50134, Firenze, Italy
[2]Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
[3]Mathematical Institute, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

Contents

1 Introduction and main results **3**
- 1.1 Target . . . 3
- 1.2 Background and earlier work . . . 3
- 1.3 Setting, definitions and notation . . . 5
- 1.4 Preliminaries for Erdős-Rényi random graphs . . . 7
- 1.5 Associated graph process . . . 8
- 1.6 Main results . . . 9
- 1.7 Discussion . . . 10

2 Coagulative dynamic permutations **12**
- 2.1 Representation via associated graph process . . . 12
- 2.2 Connected components of the cycle-free Erdős-Rényi graph process . . . 12
- 2.3 Drop-down time and mixing profile . . . 16

3 Coagulative-fragmentative dynamic permutations **18**
- 3.1 Drop-down time . . . 18
- 3.2 Drop-down in a single permutation cycle . . . 18
- 3.3 Local mixing upon drop-down . . . 20
- 3.4 Mixing profile . . . 25

A Infinite-speed random walk as limit of finite-speed random walk **26**

B Normalisation of the jump-time distribution **27**

C Cycle structure of coagulative-fragmentative dynamic permutations and Schramm's coupling **28**
- C.1 Short summary of Schramm's coupling . . . 28
- C.2 Recurrence of large cycles . . . 31

D Mixing upon dropdown on the largest cycle **34**

E Mixing in dynamic degree-two graphs **35**
- E.1 Permutations and degree-two graphs . . . 35
- E.2 ISRW mixing on degree-two graphs . . . 36

1 Introduction and main results

1.1 Target

The goal of this paper is to identify the mixing profile of a fast random walk on a *dynamic random permutation*, where fast means that the random walk instantly achieves local equilibrium, i.e., fully mixes on the cycle of the permutation it sits on before the next change in the permutation occurs. The focus is on two types of dynamics for the permutation, both starting from the identity permutation and consisting of successive applications of *random transpositions*. The first type – called *coagulative dynamics* – imposes the constraint that transpositions leading to fragmentation of a permutation cycle are ignored. The second type – called *coagulative-fragmentative dynamics* – does not impose this constraint. A major feature of dynamic random permutations is that they represent a *disconnected* geometry, which marks a departure from the setting that was considered in earlier work.

We show that for both dynamics, after scaling time by the length of the permutation and letting this length tend to infinity, the total variation distance between the current distribution and the uniform distribution converges to a limit process that makes a *single jump* down from the value 1 to a value on a deterministic curve and subsequently follows this curve on its way down to 0. The aforementioned curve is strictly decreasing in time and is related to the size of the largest component in the Erdős-Rényi random graph. The jump down to this curve, which is similar to a *one-sided cut-off*, occurs after a *random time* whose law we identify. This type of mixing profile is *different* from that of previously studied models. The law of the drop-down time and the function describing the deterministic curve are different for the two types of dynamics. Visual representations of the mixing profiles are given in Figs. 1 and 3, while simulations are shown in Figs. 2 and 4.

The model analysed in this paper is a first step towards understanding the behaviour of *simple random walk on a dynamic permutation*. This process is, despite its apparent simplicity, hard to analyse in detail, especially when the stepping rate of the random walk is *commensurate* with the transposition rate of the dynamic permutation. (In Appendix A we show that our model indeed is the limit of the simple random walk model as the step rate tends to infinity.) A key tool in our analysis is Schramm's coupling [38]. While this coupling was used previously to study the cycle structure of a dynamic permutation at a *fixed* time, we adapt the arguments in a way that allows us to study the evolution of cycles over a time interval of *diverging length*.

The remainder of this section is organised as follows. Section 1.2 provides background and recalls earlier work. Section 1.3 fixes the setting and introduces relevant definitions and notations. Section 1.4 lists some preliminaries for Erdős-Rényi random graphs that are needed along the way. Section 1.5 introduces a graph process associated with the dynamics that serves as a tool for analysing the dynamics. Section 1.6 contains two main theorems, one for each type of dynamics, describing the evolution of the total variation distance between the current distribution of the random walk and its equilibrium distribution, which is the uniform distribution on $[n] = \{1, \dots, n\}$. Section 1.7 discusses the importance of the main theorems, places them in their proper context, and provides an outline of the remainder of the paper.

1.2 Background and earlier work

While over the past years random walks on *static* random graphs have received a lot of attention, and the scaling properties of quantities like mixing times, cover times and metastable crossover times have been identified, much less is known about random walks on *dynamic* random graphs. In the static setting, a two-sided cut-off on scale $\log n$ has been established for a general class of *undirected* sparse graphs with good expansion properties [9,11,28]. Similar results have been obtained for *directed* sparse graphs [14, 15] and for graphs with a *community structure* [8].

In the dynamic setting, predominantly the focus has been on dynamic percolation, Erdős-Rényi random graphs with edges switching on and off randomly, and configuration models with random rewiring of edges. Both directed and undirected graphs have been considered, as well as backtracking and non-backtracking random walks. In [22, 32, 33] random walks on *dynamic percolation* clusters on a d-dimensional *discrete torus* were considered. Mixing times were identified for several parameter regimes controlling the rate of the random walk and the rate of the random graph dynamics. Similar

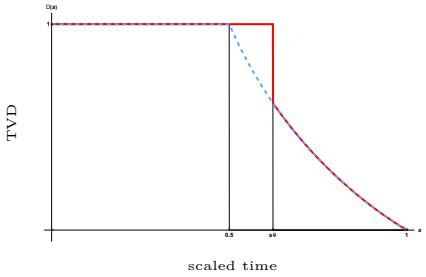

Figure 1: The red curve is a typical evolution of the total variation distance for an infinitely fast random walk on a *coagulative dynamic permutation*. The blue curve is a plot of the deterministic function of the scaled time to which the total variation distance drops at a random time and subsequently sticks to.

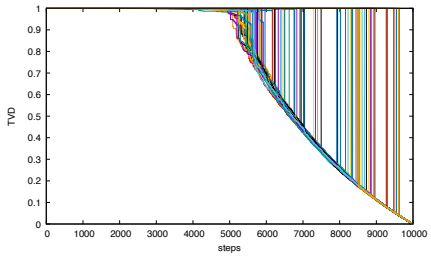

Figure 2: Simulations of the evolution of the total variation distance for 10^2 different realisations of a *coagulative dynamic permutation* of 10^4 elements and an infinitely fast random walk on top. Each simulation run corresponds to a single coloured curve.

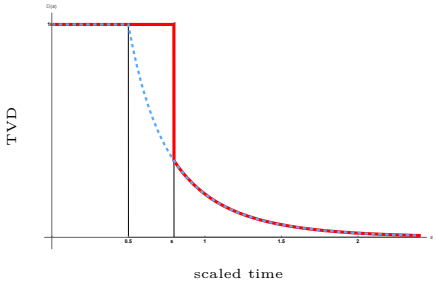

Figure 3: The same as Fig. 1 for a *coagulative-fragmentative dynamic permutation*.

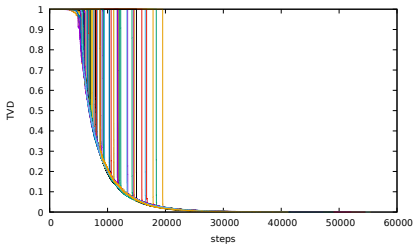

Figure 4: The same as Fig. 2 for a *coagulative-fragmentative permutation*.

results were obtained for dynamic percolation on the *complete graph* [37, 39]. Some further advances were achieved in [7], where general bounds on mixing times, hitting times, cover times and return times were derived for certain classes of dynamic random graphs under appropriate *expansion assumptions*. Non-backtracking random walks on configuration models that with high probability are connected were studied in a series of papers [4,5,6], which culminated in a general framework for studying mixing times of non-backtracking random walks on dynamic random graphs subject to mild regularity conditions. Mixing of random walks on *directed* configuration models was treated in [16].

Random permutations generated by random transpositions have attracted plenty of interest as well. An important starting point is [17], where a *cut-off* in the total variation distance was established after the application of $\frac{1}{2}n \log n + O(n)$ random transpositions. The *sharp* constant in front of the leading-order term was achieved with the help of *representation theory* for the symmetric group. This paper led to a flurry of follow-up work, of which we mention [38], where the structure of large cycles of an evolving random permutation was studied. Similar results were obtained in [10], including sharp control on the number of observed fragmentations. An important aspect of both [38] and [10] is the representation of an evolving random permutation, starting from the identity permutation, in terms of a random graph process that can be studied by using the theory of random graphs. For the *coagulative-fragmentative* dynamics considered in the present paper, also called *transposition dynamics*, this graph process representation yields a graph-growth model that at every step adds an edge drawn uniformly at random. This graph-growth model is closely related to the standard "combinatorial" Erdős-Rényi model, whose study is by now a classical topic in the theory of random graphs (see, for example, [19] or [23]). Yet another important feature of [38] is the introduction of *Schramm's coupling* as a tool to study the cycle structure of evolving random permutations. In a follow-up article [12], a modified version of this coupling is used to study the mixing of dynamic permutations endowed with a more general dynamics, of which the transposition dynamics is a special case. We also mention [13], which contains a detailed account of Schramm's coupling. The works cited above each highlight one particular facet of the random transposition model, but close relatives have been studied extensively under different names: *mean-field Tóth model* [41], the *interchange process* on the complete graph (see [3, 13, 21] and references therein), or *multi-urn Bernoulli–Laplace diffusion models* [36], where our setting corresponds to a particular choice of the model parameters.

The *coagulative* dynamics considered in the present paper can be recast, in the spirit of [38], as a graph-valued random process that starts with an empty graph on n vertices and describes a forest that progressively merges into a spanning tree on n vertices through the addition of edges that do not create a cycle. The study of this process and its close relatives has a somewhat twisted history. It is similar to the *standard additive coalescent* (see [2] for an overview), but it is also interesting in its own right (see [1, 29]). Finally, there is a wealth of results on *minimal spanning trees* and *Kruskal's algorithm*, which is another closely related process. In particular, we mention [27], since this work implies some facts that we list in Section 2. We derive these facts independently, using different techniques in a different setting.

1.3 Setting, definitions and notation

For $n \in \mathbb{N}$, let S_n denote the set of permutations of $[n]$, i.e., bijections from $[n]$ to itself. Recall that S_n endowed with the operation of permutation composition $\circ$ forms a group. Write $\gamma_v(\pi)$ to denote the cycle of the permutation π that contains the element v.

Definition 1.1 [Dynamic permutation] A sequence of permutations of $[n]$, denoted by $\Pi_n = (\Pi_n(t))_{t=0}^{t^{\max}}$ with $t^{\max} \in \mathbb{N}_0 \cup \{\infty\}$, is called a dynamic permutation. ♠

Example 1.2 [Transpositions may fragment cycles or coagulate cycles] Pick $n = 7$ and consider the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 5 & 6 & 7 & 1 \end{pmatrix} \quad \text{with cycle structure} \quad (1, 2, 3, 4, 5, 6, 7).$$

The transposition $(1,5)$ turns this into the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 3 & 4 & 5 & 2 & 7 & 1 \end{pmatrix} \quad \text{with cycle structure} \quad (1,6,7)(2,3,4,5).$$

Another application of the same transposition acts in reverse. Note that S_n is a *non-commutative group* for any $n \geq 3$. ■

We consider two types of dynamic permutations:

Definition 1.3 [Coagulative dynamic permutation] $\Pi_n = (\Pi_n(t))_{t=0}^{n-1}$ is called a *coagulative dynamic permutation* (CDP) when $\Pi_n(0) = \mathrm{Id}$ (i.e., the identity permutation) and

$$\Pi_n(t) = \Pi_n(t-1) \circ (a,b), \qquad t \in [n-1], \tag{1.1}$$

where, for each $t \in [n-1]$, (a,b) is a random transposition sampled uniformly at random from the set of all transpositions of $[n]$ that satisfy the *constraint*

$$\gamma_a(\Pi_n(t-1)) \neq \gamma_b(\Pi_n(t-1)). \tag{1.2}$$

The latter guarantees that no cycle of $\Pi_n(t-1)$ is fragmented by the transposition (a,b). ♠

Definition 1.4 [Coagulative-fragmentative dynamic permutation] $\Pi_n = (\Pi_n(t))_{t=0}^{\infty}$ is called a *coagulative-fragmentative dynamic permutation* (CFDP) when the same holds as in Definition 1.3, but without the constraint in (1.2). ♠

Remark 1.5 [Time horizon for dynamic permutations and cycle structure]. Since CDP starts from the identity permutation, it becomes a permutation with a single cycle after exactly $n-1$ steps. Once this happens, there is no permutation that satisfies (1.2) and the dynamics is trapped. CFDP has no traps and can evolve forever. The structure of cycles is random and the sizes of the large cycles, properly rescaled, converge in distribution to the Poisson-Dirichlet distribution with parameter 1 (see e.g. [38, Theorem 1.1] for a precise statement). ■

Our aim is to study mixing of fast random walks on both CDP and CFDP. To simplify our analysis, we work with *infinite-speed* random walks, as defined next:

Definition 1.6 [Infinite-speed random walk on Π_n] Fix Π_n and an element $v_0 \in [n]$. Recall that $\gamma_v(\Pi_n(t))$ is the cycle of $\Pi_n(t)$ that contains v. Formally, the infinite-speed random walk (ISRW) starting from v_0 is a sequence of probability distributions $(\mu^{n,v_0}(t))_{t\in\mathbb{N}_0}$ supported on $[n]$, with initial distribution at time $t=0$ given by

$$\mu^{n,v_0}(0) = \left(\mu_w^{n,v_0}(0)\right)_{w\in[n]}, \tag{1.3}$$

where $\mu_w^{n,v_0}(0)$, the mass at $w \in [n]$ at time $t=0$, is given by

$$\mu_w^{n,v_0}(0) = \begin{cases} \frac{1}{|\gamma_w(\Pi_n(0))|}, & w \in \gamma_{v_0}(\Pi_n(0)), \\ 0, & w \notin \gamma_{v_0}(\Pi_n(0)), \end{cases} \tag{1.4}$$

and with distribution at a later time $t \in \mathbb{N}$ given by

$$\mu^{n,v_0}(t) = \left(\mu_w^{n,v_0}(t)\right)_{w\in[n]}, \tag{1.5}$$

where

$$\mu_w^{n,v_0}(t) = \frac{1}{|\gamma_w(\Pi_n(t))|} \sum_{u\in\gamma_w(\Pi_n(t))} \mu_u^{n,v_0}(t-1). \tag{1.6}$$

Informally, ISRW *spreads infinitely fast over the cycle in the permutation it resides on.* ♠

In Appendix A we show that the infinite-speed random walk arises as the limit of a standard random walk whose stepping rate relative to the rate of the permutation dynamics tends to infinity. Note that the evolution of the ISRW distribution is fully determined by the initial position of the random walk and the realisation of the dynamic permutation. See Figure 5 for an illustration.

Remark 1.7 [ISRW as a mass-spreading process]. The reader may prefer to let go of the connection with the random walk and view the ISRW purely as a mass-spreading process. Such a change of perspective would change nothing in our arguments. ■

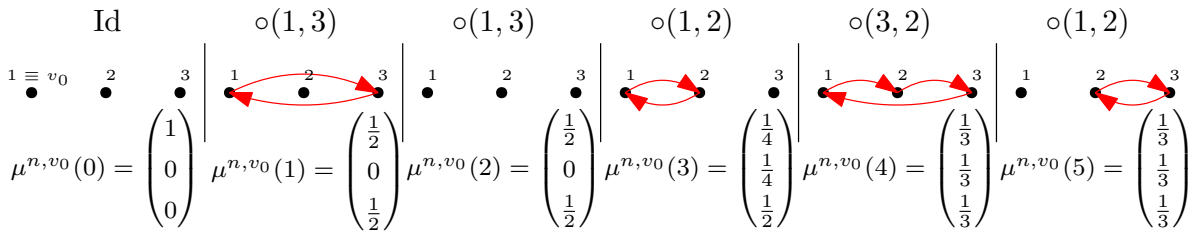

Figure 5: Example of an evolution of an ISRW on top of a CFDP with three elements starting from the identity permutation. The first row shows the transpositions that generate the next permutation. The second row is a graphical representation of the cycles of this permutation. The third row shows the evolution of the ISRW distribution, given that it started from the element 1.

1.4 Preliminaries for Erdős-Rényi random graphs

The arguments in this paper frequently make use of results on the structure of Erdős-Rényi random graphs. This section provides what is needed to state the main theorems in Section 1.6.

Definition 1.8 [Standard Erdős-Rényi multi-graph process] The *standard Erdős-Rényi multi-graph process* on n vertices is the discrete-time process $(G(n,t))_{t=0}^{t_{\max}}$ constructed as follows:

1. $G(n,0)$ is the graph with n vertices and no edges.
2. At each time $t \in \mathbb{N}_0$, pick an edge e_t uniformly at random from the $\binom{n}{2}$ possible edges, and let $G(n,t)$ be the graph obtained by adding e_t to $G(n,t-1)$.

Note that we do not allow for self-loops, but do allow for multiple edges. ♠

Remark 1.9 [Versions and asymptotic equivalence]. There are versions of the Erdős-Rényi multi-graph process that differ in how edges are deployed and whether or not multiple edges and self-loops are allowed. With respect to monotone properties, notably the expected size of connected components, the "random growth" $G(n,t)$ model described in Definition 1.8 is asymptotically equivalent to the "combinatorial" model $G(n,M)$ with $M = t$ edges at times $t = O(n)$, which in turn is asymptotically equivalent to the "bond percolation" model $G(n,p)$ with $p = M\binom{n}{2}^{-1}$. For details, see [19, Sections 1.1, 1.3]. Since we work on time scales of order n, we will use this asymptotic equivalence without further notice. ■

Definition 1.8 allows for some natural modifications, of which one is important for the study of CDP:

Definition 1.10 [Cycle-free Erdős-Rényi graph process] The *cycle-free Erdős-Rényi graph process* on n vertices is the graph process $(G(n,t))_{t=0}^{t_{\max}}$ starting from the empty graph with n vertices, such that at each time t an edge is added that is chosen uniformly at random from the set of edges that do not create a cycle, a multi-edge or a self-loop. Thus, $G(n,t)$ is a forest for all $0 \le t \le t_{\max}$. ♠

To understand the typical evolution of CDP, we make use of two *couplings*: one between CDP and cycle-free Erdős-Rényi graph processes, the other between cycle-free Erdős-Rényi graph processes and their standard counterparts. To explain how, we need to introduce three functions that describe key structural properties of these processes:

Definition 1.11 [Functions related to the structure of Erdős-Rényi random graphs]
(1) Define $\zeta\colon [0,\infty) \to [0,1)$ as $\zeta(u) = 0$ for $u \in [0, \frac{1}{2}]$ and as the unique positive solution of the equation $1 - \zeta(u) = \mathrm{e}^{-2u\zeta(u)}$ for $u \in (\frac{1}{2}, \infty)$. Note that ζ is non-decreasing and continuous on $[0,\infty)$, and analytic on $(\frac{1}{2}, \infty)$.
(2) Define $\phi\colon [0,\infty) \to [0,1)$ as

$$\phi(v) = \int_0^v \mathrm{d}u\,[1 - \zeta^2(u)], \qquad v \in [0,\infty). \tag{1.7}$$

Note that ϕ is strictly increasing and continuous on $[0,\infty)$, and hence has a well-defined inverse ϕ^{-1}. Furthermore, the function ϕ is properly normalised in the sense that $\phi(\infty)=1$ (see Appendix B).
(3) Define $\eta\colon\ [0,1)\to[0,1)$ as

$$\eta(w)=\zeta(\phi^{-1}(w)),\qquad w\in[0,1). \tag{1.8}$$

♠

The functions defined in Definition 1.11 are illustrated in Fig. 6 and have the following interpretation:

(1) $\zeta(u)$ describes the expected size of the largest component of the Erdős-Rényi random graph at time un. For $u\in[0,\infty)$, denote by $|\mathscr{C}^{\mathrm{ER}}_{\max}(n,un)|$ the size of the largest connected component in the Erdős-Rényi random graph with n vertices and un edges, and $|\mathscr{C}^{\mathrm{ER}}_{\mathrm{sec}}(n,un)|$ the size of the second-largest connected component. Then, as $n\to\infty$,

$$\frac{|\mathscr{C}^{\mathrm{ER}}_{\max}(n,un)|}{n}\xrightarrow{\mathbb{P}}\zeta(u),\qquad \frac{|\mathscr{C}^{\mathrm{ER}}_{\mathrm{sec}}(n,un)|}{n}\xrightarrow{\mathbb{P}}0. \tag{1.9}$$

(2) The function ϕ provides the link between the standard and the cycle-free Erdős-Rényi graph process (see Lemmas 2.4–2.5 below).

(3) $\eta(u)$ is the analogue of $\zeta(u)$ for the cycle-free Erdős-Rényi graph process at time un, $u\in[0,1]$ (see Lemma 2.6 below).

Note the change in behaviour of $\zeta,\phi,\phi^{-1},\eta$ at $\frac{1}{2}$. Note that ϕ^{-1} blows up at 1.

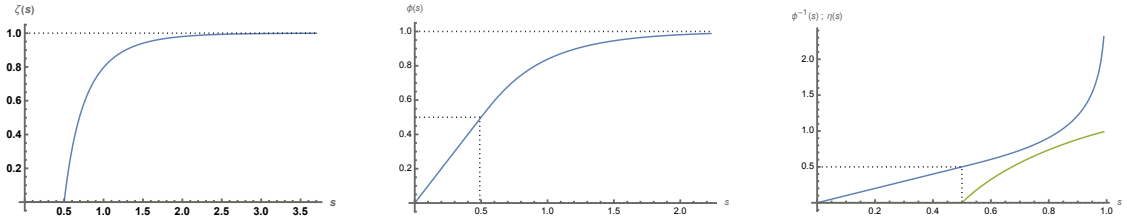

Figure 6: Graphs of the functions introduced in Definition 1.11: ζ, ϕ, respectively, ϕ^{-1} (upper curve), η (lower curve).

1.5 Associated graph process

For any dynamic permutation starting from the identity permutation, define the associated graph process as follows:

Definition 1.12 [Graph process associated with Π_n] Let $\Pi_n=(\Pi_n(t))_{t=0}^{t_{\max}}$ with $t_{\max}\in\mathbb{N}\cup\{\infty\}$ be a dynamic permutation starting from the identity permutation. Construct the *associated graph process*, denoted by A_{Π_n}, as follows:

1. At time $t=0$, start with the empty graph on the vertex set $\mathcal{V}=[n]$.
2. At times $t\in\mathbb{N}$, add the edge $\{a,b\}$, where a,b are such that $\Pi_n(t)=\Pi_n(t-1)\circ(a,b)$.

♠

Associated graph processes were used in [10, 38] and follow-up articles to represent the evolution of a general dynamic permutation in terms of a dynamics generated by applying a single transposition at every time step.

A crucial role will be played by the first time when the support of the random walk distribution *intersects the largest connected component of the associated graph process*:

Definition 1.13 [Largest component of the associated graph process] Denote by $\mathscr{C}_{\max}(A_{\Pi_n}(t))$ the set of vertices in the largest connected component in the associated graph process at time t. If such a connected component is not unique, then take all the vertices in all the largest connected components. ♠

Remark 1.14 [Possible non-uniqueness of the largest connected component]. In situations where we employ Definition 1.13, the largest connected component is unique with high probability. Situations where it is not unique will be of no importance. ■

Definition 1.15 [Drop-down time] Fix any $\varepsilon_n > 0$ such that $\varepsilon_n = \omega(n^{-1/3})$ and $\varepsilon_n = o(1)$ as $n \to \infty$. The *drop-down time* is defined as

$$T^{\Downarrow}_{n,v_0} = \inf\Big\{t > \tfrac{n}{2}[1+\varepsilon_n] \colon \operatorname{supp}(\mu^{n,v_0}(t)) \cap \mathscr{C}_{\max}(A_{\Pi_n}(t)) \neq \emptyset\Big\}. \tag{1.10}$$

♠

Remark 1.16 [Drop-down time and hitting time of the largest permutation cycle]. At first sight it might seem unintuitive that the time $T^{\Downarrow}_{n,v_0}$ from Definition 1.15 plays an important role. Given the diffusive nature of ISRW, an arguably more natural candidate would be the first time when the ISRW is supported on the largest *permutation cycle*. However, the above definition in terms of the associated graph process allows for a unified presentation of our results in different settings, even when the associated graph process at a single time *does not* provide all the information about the structure of permutation cycles. ■

For CDP, the drop-down time is the first time when the cycle that contains v_0 merges with the largest cycle. For CFDP, however, this is not necessarily true because cycles fragment. We therefore define the drop-down time to be the first time when the random walk 'sees' the maximal component, see (1.10). Later, we will see that, in fact, afterwards the mass spreads over $\mathscr{C}_{\max}(A_{\Pi_n}(t))$ quickly.

Remark 1.17 [Properties of drop-down time]. Clearly, $T^{\Downarrow}_{n,v_0}$ is random. However, if we condition on a particular realisation of Π_n, then $T^{\Downarrow}_{n,v_0}$ is a deterministic function of the starting point of the random walk. The role of ε_n is to ensure that $T^{\Downarrow}_{n,v_0}$ represents the first time in the *supercritical regime* when the *largest component* in the associated *Erdős-Rényi graph process* coincides with the support of the ISRW, see Section 2.1. The choice of ε_n ensures that the drop-down time *avoids the critical window*, which corresponds to $\frac{n}{2} + O(n^{2/3})$, yet covers the entire supercritical regime. ■

1.6 Main results

For convenience, we introduce the following shorthand notation:

Definition 1.18 [Total variation distance away from equilibrium] For $v \in [n]$, define

$$\mathcal{D}^{v_0}_n(t) = d_{\mathrm{TV}}\left(\mu^{n,v_0}(t), \mathsf{Unif}([n])\right), \qquad t \in \mathbb{N}_0. \tag{1.11}$$

♠

Our main results are the following two theorems ($\xrightarrow{d}$ denotes convergence in distribution):

Theorem 1.19 [Mixing profile for ISRW on CDP]

(1) *Uniformly in $v_0 \in [n]$,*

$$\frac{T^{\Downarrow}_{n,v_0}}{n} \xrightarrow{d} s^{\Downarrow}, \tag{1.12}$$

where $s^{\Downarrow}$ is the $[0,1]$-valued random variable with distribution (recall (1.8))

$$\mathbb{P}(s^{\Downarrow} \leq s) = \eta(s), \qquad s \in [0,1]. \tag{1.13}$$

(2) *Uniformly in* $v_0 \in [n]$,

$$(\mathcal{D}_n^{v_0}(sn))_{s\in[0,1]} \xrightarrow{d} \big(1-\eta(s)\mathbb{1}_{\{s>s^{\Downarrow}\}}\big)_{s\in[0,1]} \quad \textit{in the Skorokhod } M_1\textit{-topology.} \tag{1.14}$$

Theorem 1.20 [Mixing profile for ISRW on CFDP]

(1) *Uniformly in* $v_0 \in [n]$,

$$\frac{T^{\Downarrow}_{n,v_0}}{n} \xrightarrow{d} u^{\Downarrow}, \tag{1.15}$$

where $u^{\Downarrow}$ *is the non-negative random variable with distribution (recall Definition* 1.11*(1))*

$$\mathbb{P}(u^{\Downarrow} \le u) = \zeta(u), \qquad u \in [0,\infty). \tag{1.16}$$

(2) *Uniformly in* $v_0 \in [n]$,

$$(\mathcal{D}_n^{v_0}(un))_{u\in[0,\infty)} \xrightarrow{d} \big(1-\zeta(u)\mathbb{1}_{\{u>u^{\Downarrow}\}}\big)_{u\in[0,\infty)} \quad \textit{in the Skorokhod } M_1\textit{-topology.} \tag{1.17}$$

The proofs of these theorems are given in Sections 2 and 3, respectively.

1.7 Discussion

1. Despite the similarity of Theorems 1.19–1.20, the latter is *far more delicate*. For CDP, mixing is simply induced by the ISRW entering the ever-growing largest cycle. For CFDP, the presence of fragmentations breaks the direct link between the dynamic permutation and its associated graph process: *a single connected component may carry more than one permutation cycle*. Specifically, the largest component of the associated graph process carries a large number of permutation cycles and, at the drop-down time, the distribution of the ISRW is supported on only one of them. It is not a priori clear how many steps the dynamics needs to spread out the ISRW distribution over all the elements that lie on the largest component of the associated graph process. Therefore, a major hurdle in the proof of Theorem 1.20 is to show that such *local mixing* happens on time scale $o(n)$. We actually show a stronger statement, namely, that local mixing occurs on an arbitrarily small but diverging time scale (see Section 3.3 for details). The core of the proof is to show that on the largest component over time there is a diverging count of appearances of permutation cycles that span almost the entire largest component of the associated graph process.

2. Theorems 1.19–1.20 extend our earlier results for the total variation distance of a (non-backtracking) random walk on a configuration model subject to random rewirings [6]. There we assumed that *all the degrees are at least three*, which corresponds to a *supercritical* configuration model that with high probability is connected (see [24, Chapter 4]). Our model with evolving permutation cycles is closely related to the setting where *all the degrees are two*, which in turn corresponds to a special kind of configuration model that with high probability is disconnected (see Fig. 7). In this setting, even small perturbations of the degree sequence can lead to significantly different behaviour (see [18] for details). In Appendix E we comment further on the connection between permutations and degree-two graphs. More concretely, we show that in the setting of dynamic degree-two graphs with rewiring, we obtain an ISRW-mixing profile analogous to the one described in Theorem 1.20 (see Theorem E.3). We stress that in the present work the starting configuration is fixed to be the identity permutation, which would correspond to a graph with only self-loops, whereas in our previous work the starting configuration was sampled from the configuration model.

3. The mixing profile in Theorems 1.19–1.20 is unusual: the total variation distance makes a *single jump* down from the value 1 to a value on a deterministic curve and subsequently follows this curve on its way down to 0. This jump, which is similar to a *one-sided cut-off*, occurs after a *random time*. The

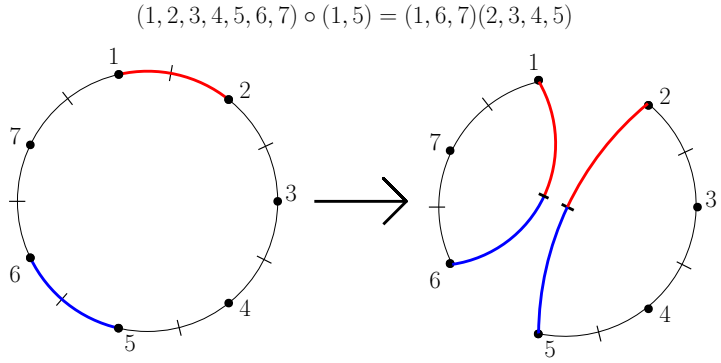

Figure 7: Dynamic permutations are similar to rewirings in the configuration model, where all degrees are two. Recall Example 1.2. Consider the permutation $\Pi(0) = (1, 2, 3, 4, 5, 6, 7)$, which consists of a single cycle and corresponds to a degree-two graph that has a single connected component. Apply the transposition $(1, 5)$ to get a new permutation $\Pi(1) = \Pi(0) \circ (1, 5) = (1, 6, 7)(2, 3, 4, 5)$, which consists of two cycles and corresponds to a degree-two graph that has two connected components, obtained by sampling the edges $(1, \Pi(0)(1)) = (1, 2)$ and $(5, \Pi(0)(5)) = (5, 6)$ and rewiring them.

law of the drop-down time and the function describing the deterministic curve depend on the choice of dynamics.

4. The pathwise statements in part (2) of Theorems 1.19–1.20 imply the following pointwise statements ($\sim$ denotes equality in distribution):

$$\begin{aligned} \mathcal{D}_n^{v_0}(sn) &\xrightarrow{d} 1 - \eta(s)Y(s), \quad s \in [0,1], \quad \text{with } Y(s) \sim \mathsf{Bernoulli}(\eta(s)), \\ \mathcal{D}_n^{v_0}(sn) &\xrightarrow{d} 1 - \zeta(s)\bar{Y}(s), \quad s \in [0,\infty), \quad \text{with } \bar{Y}(s) \sim \mathsf{Bernoulli}(\zeta(s)). \end{aligned} \tag{1.18}$$

Through the function ϕ plotted in Fig. 6, we can view the two mixing profiles as a continuous deformation of one another. Slower mixing for CFDP is intuitive: fragmentation slows down the mixing, while coagulation induces it.

5. Note the similarities between the mixing profiles described by Theorems 1.19–1.20. Both feature a single macroscopic jump at a random time to a deterministic curve that depends on the choice of the dynamics. We expect this type of behaviour to occur for any permutation dynamics whose associated graph process exhibits scaling behaviour similar to that of the Erdős-Rényi graph process. A class of graph processes that fits this criterion is the class of *Achlioptas processes with bounded-size rules* (see [35] or [40]).

6. We can formulate conjectures about finite-speed random walks as well. Settings where the random walk rate dominates are easy to handle. If the random walk is fast enough to ensure local mixing (e.g. $\gg n^2$ steps of the random walk occur for every step of the random permutation), then our theorems should remain the same with negligible error terms. In this regime, the mixing is fully driven by the underlying geometry. However, once these rates are commensurate, we would have to deal with random walk distributions that are *partially mixed over cycles*, meaning that the distribution of the random walk would not be uniform over its supporting cycle before this cycle is affected by the permutation dynamics.

7. Dynamic permutations are a natural model for discrete dynamic random environments, which typically are *disconnected* but nonetheless allow for interaction between their constitutive elements. We believe this setting to be interesting for other stochastic processes on random graphs as well, such as the *voter model* or *the contact process*.

Organisation of the paper. Section 2 starts by establishing a link between CDP and cycle-free Erdős-Rényi random graphs. A *coupling* construction is employed to describe the *cyclic structure* of a typical CDP. These results are used to prove Theorem 1.19. Section 3 deals with CFDP, where the main problem is that the associated graph process provides weaker control over permutation cycles

than for CDP. After this discrepancy is settled, we employ arguments analogous to those in Section 2 to prove Theorem 1.20.

Appendices A–B contain supplementary material that is not needed in Sections 2–3. Namely, Appendix A shows that the ISRW arises as a fast-speed limit of the standard random walk. Appendix B proves that the laws of the jump-down times in Theorems 1.19–1.20 are properly normalised. Appendix C contains the key coupling that is used to study the cycle structure of CFDP, which is technical and of interest in itself. This coupling is needed in Section 3. Appendix D contains a technical computation that is needed in Section 3 as well. Finally, Appendix E elucidates the connection between random permutations and graphs with all the degrees equal to two and extends Theorem 1.20 to the setting of dynamic degree-two graphs.

2 Coagulative dynamic permutations

In this section, we establish a link between dynamic permutations and evolving graphs. To do so, we couple a CDP with a cycle-free Erdős-Rényi graph process (Section 2.1), and couple the latter with the standard Erdős-Rényi graph process (Section 2.2) by making use of well-known results on the structure of connected components of Erdős-Rényi random graphs (recall Section 1.4). We use the couplings to prove Theorem 1.19 (Section 2.3).

2.1 Representation via associated graph process

Note that for the dynamics generated by transpositions sampled uniformly at random from the set of all transpositions of n elements, the associated graph process is equal in distribution to the Erdős-Rényi process defined in Definition 1.8. In the setting of *coagulative* dynamic transpositions, this leads us to the following observation:

Lemma 2.1 [Representation of CDP as cycle-free graph process] *If Π_n is a CDP, then its associated graph process A_{Π_n} is the cycle-free Erdős-Rényi graph process defined in Definition* 1.10*.*

Proof. Recall that the change between two successive permutations in a CDP is generated by applying a *single* transposition. Furthermore, note that the only transpositions causing a split of a permutation cycle are the ones that transpose two elements from the same cycle. Recall Definition 1.12, and note that if $A_{\Pi_n}(t)$ is a forest, then its connected components correspond to cycles of $\Pi_n(t)$. Furthermore, observe that cycle-splitting transpositions correspond to edges that join two vertices from the same connected component. Thus, if $A_{\Pi_n}(t)$ is a forest, then any transposition causing a fragmentation of a permutation cycle corresponds to an edge that creates a cycle in the associated graph process.

Observe that the associated graph process always starts as a forest. Since fragmentations of permutation cycles are not allowed, there can be no edges that lead to graph cycles in the associated graph process. Since the associated graph process for a dynamic permutation with no constraints is the Erdős-Rényi graph process, the associated graph process for a CDP is the Erdős-Rényi graph process constrained to be a forest (see Definition 1.10). □

2.2 Connected components of the cycle-free Erdős-Rényi graph process

• **Coupling of Erdős-Rényi graph processes.** We construct a coupling of the standard and the cycle-free Erdős-Rényi graph process that allows us to study the structure of the connected components of the cycle-free process.

Definition 2.2 [Coupling between cycle-free and standard Erdős-Rényi graph process] Let $G_n = (G_n(t))_{t\in\mathbb{N}_0}$ be the Erdős-Rényi graph process on $[n]$ defined in Definition 1.8, and denote the edge set of $G_n(t)$ by $\mathcal{E}_{G_n(t)}$. Based on G_n, construct a graph-valued process $F_n = (F_n(t))_{t\in\mathbb{N}_0}$ as follows:

1. $F_n(0)$ is the empty graph with vertex set $[n]$.
2. At times $t \in \mathbb{N}$, define $e^\star(t) = \mathcal{E}_{G_n(t)} \setminus \mathcal{E}_{G_n(t-1)}$, which is the edge added at time t to $G_n(t)$.

(a) Construct the candidate graph at time t, defined as $F_n^\star(t) = (\mathcal{V}, \mathcal{E}_{F_n(t-1)} \cup \{e^\star(t)\})$.

(b) If $F_n^\star(t)$ is a forest, then set $F_n(t) = F_n^\star(t)$.

(c) Otherwise, set $F_n(t) = F_n(t-1)$.

Define the *effective time* $\tau_n(t)$ of the coupled process $(F_n(t))_{t\in\mathbb{N}_0}$ by setting $\tau_n(0) = 0$ and, recursively for $t \in \mathbb{N}$,

$$\tau_n(t) = \begin{cases} \tau_n(t-1) + 1, & \text{if } F_n(t) \neq F_n(t-1), \text{ i.e., the proposed edge has been accepted,} \\ \tau_n(t-1), & \text{if } F_n(t) = F_n(t-1), \text{ i.e., the proposed edge has been rejected.} \end{cases} \tag{2.1}$$

Note that $\tau_n(t)$ is a random variable because it is a function of a random graph process. We suppress the dependence of $\tau_n(t)$ on F_n, since we will never work with more than one set of coupled processes at a time. ♠

Remark 2.3 [Relation between F_n and A_{Π_n}]. By the definition of the coupling, if there are edge-rejections at times $\{t, t+1, t+2, \ldots, t+k\}$, then a string of $k+1$ copies of the same graph is observed in F_n, i.e., $F_n(t-1) = F_n(t) = F_n(t+1) = \cdots = F_n(t+k)$. On the other hand, the associated graph process A_{Π_n} is by construction a sequence of graphs such that no two graphs are the same. To recover A_{Π_n} from F_n, from every string of copies of the same graph choose only one copy of that graph. ■

The reason why this construction is useful to control the connected components of the cycle-free Erdős-Rényi graph process is stated in the following lemma:

Lemma 2.4 [Connected components of F_n] *Let H be a graph with vertex set $\mathcal{V}$, and define $\boldsymbol{\mathcal{CC}}(H)$ to be the partition of $\mathcal{V}$ induced by the connected components of H. Let G_n, F_n be as in Definition 2.2. Then, at every time $t \in \mathbb{N}_0$, $\boldsymbol{\mathcal{CC}}(G_n(t)) = \boldsymbol{\mathcal{CC}}(F_n(t))$.*

Proof. Note that any edge creating a cycle does not influence the size of the connected components. □

• **Effective time.** To use the above observation, we need to control the effective time $\tau_n(t)$. The following lemma shows that with high probability and after scaling by $1/n$, there is a simple relation between the standard time t and the effective time $\tau_n(t)$:

Lemma 2.5 [Effective time of a cycle-free Erdős-Rényi graph process] *Let G_n, F_n and τ_n be as in Definition 2.2, and ϕ as in Definition 1.7. Then, for any for $u \in [0, \infty)$,*

$$\frac{\tau_n(un)}{n} \xrightarrow{\mathbb{P}} \phi(u). \tag{2.2}$$

Proof. The proof of Lemma 2.5 consists of two separate lines of argument. First, we show that the left-hand side in (2.2) concentrates around a deterministic quantity. Afterwards, the value of this quantity is computed.

Part 1: Concentration of the associated martingale. Observe that

$$\tau_n(t) = t - \sum_{s=0}^{t} \mathbb{1}_{R(s)}, \qquad t \in \mathbb{N}_0, \tag{2.3}$$

where $\mathbb{1}_{R(s)}$ is the edge-rejection indicator at time s. Let $\mathcal{F} = (\mathcal{F}_n(t))_{t\in\mathbb{N}_0}$ with $\mathcal{F}_n(t) = \sigma((G_n(q))_{q=0}^t)$ be the natural filtration with respect to the Erdős-Rényi graph process. By the construction of the coupling, a rejection occurs whenever there is an edge that creates a cycle within a connected component. Therefore

$$\mathbb{1}_{R(0)} = 0, \qquad \big[\mathbb{1}_{R(s)} \mid \mathcal{F}_n(s-1)\big] \sim \mathsf{Bernoulli}(p_s), \qquad s \in \mathbb{N}, \tag{2.4}$$

where the success probabilities are given by

$$p_s = \sum_{\mathscr{C} \in \boldsymbol{\mathcal{CC}}(G_n(s-1))} \frac{|\mathscr{C}|(|\mathscr{C}|-1)}{n(n-1)}, \qquad s \in \mathbb{N}, \tag{2.5}$$

i.e., the number of edges that can join two vertices from the same connected component at time $s-1$ divided by the total number of edges. Introduce the shorthand notation

$$\mathbb{E}_t[\cdot] = \mathbb{E}\left[\,\cdot \mid \mathcal{F}_n(t)\right], \tag{2.6}$$

and define two sequences of random variables $(D_t)_{t\in\mathbb{N}_0}$ and $(S_t)_{t\in\mathbb{N}_0}$ such that

$$\begin{aligned} S_0 &= D_0 = 0, \\ D_t &= \mathbb{1}_{R(t)} - \mathbb{E}_{t-1}\left[\mathbb{1}_{R(t)}\right], \qquad t\in\mathbb{N}, \\ S_t &= \sum_{s=0}^{t} D_s = \sum_{s=0}^{t} \mathbb{1}_{R(s)} - \sum_{s=0}^{t} \mathbb{E}_{s-1}\left[\mathbb{1}_{R(s)}\right], \qquad t\in\mathbb{N}. \end{aligned} \tag{2.7}$$

Note that, for any $t\in\mathbb{N}_0$,

$$\begin{aligned} \mathbb{E}[S_t] &\le t < \infty, \\ \mathbb{E}_t\left[S_{t+1}\right] &= \mathbb{E}_t\left[\mathbb{1}_{R(t+1)} - \mathbb{E}_t\left[\mathbb{1}_{R(t+1)}\right]\right] + \mathbb{E}_t\left[S_t\right] = S_t, \\ |S_t - S_{t-1}| &= |D_t| \le 1. \end{aligned} \tag{2.8}$$

Hence, $(S_t)_{t\in\mathbb{N}_0}$ is a martingale with bounded differences with respect to the natural filtration of the Erdős-Rényi graph process. Using the Azuma-Hoeffding inequality, we can estimate

$$\mathbb{P}\left(|S_t| \ge \varepsilon\right) \le 2\exp\left(-\frac{\varepsilon^2}{2t}\right). \tag{2.9}$$

Pick $t = un$, $u\in[0,\infty)$, and $\varepsilon_n = n^{\frac{1+\delta}{2}}$, $c>0$, $\delta\in(0,1)$. Introduce the event

$$\Xi(un) = \{|S_{un}| < n^{\frac{1+\delta}{2}}\}. \tag{2.10}$$

By (2.9),

$$\mathbb{P}\left(\Xi^{\mathrm{c}}(un)\right) \le 2\exp\left(-\frac{c^2 n^{\delta}}{2u}\right) = o(1) \tag{2.11}$$

and hence $\mathbb{P}\left(\Xi(un)\right) = 1 - o(1)$. By the definition of S_t in (2.7), we see that, on the event $\Xi(un)$,

$$\mathbb{1}_{\Xi(un)}\left[\frac{1}{n}\sum_{s=0}^{un}\mathbb{1}_{R(s)} - \frac{1}{n}\sum_{s=0}^{un}\mathbb{E}_{s-1}\left[\mathbb{1}_{R(s)}\right]\right] = \mathbb{1}_{\Xi(un)}\,o(1), \qquad \text{with } |o(1)| \le n^{\frac{1+\delta}{2}}, \tag{2.12}$$

which establishes the concentration of $\tau_n(un)/n$ as $n\to\infty$.

Part 2: Computation of limit. We compute

$$\frac{1}{n}\sum_{s=0}^{un}\mathbb{1}_{R(s)} = \frac{\left(\mathbb{1}_{\Xi(un)} + \mathbb{1}_{\Xi^{\mathrm{c}}(un)}\right)}{n}\sum_{s=0}^{un}\mathbb{1}_{R(s)} \tag{2.13}$$

$$= \frac{1}{n}\left[\mathbb{1}_{\Xi(un)}\left(\sum_{s=0}^{un}\mathbb{E}_{s-1}\left[\mathbb{1}_{R(i)}\right] + o(n)\right) + \mathbb{1}_{\Xi^{\mathrm{c}}(un)}\sum_{s=0}^{un}\mathbb{1}_{R(s)}\right]. \tag{2.14}$$

1. Observe that

$$\frac{1}{n}\mathbb{1}_{\Xi^{\mathrm{c}}(un)}\sum_{s=0}^{un}\mathbb{1}_{R(s)} \xrightarrow{\mathbb{P}} 0, \tag{2.15}$$

because $\sum_{s=0}^{un}\mathbb{1}_{R(s)} \le un$ and $\mathbb{1}_{\Xi^{\mathrm{c}}(un)} \xrightarrow{\mathbb{P}} 0$. To understand the first summand in (2.14), we need to introduce another event. Fix a sequence ε_n such that $\varepsilon_n = o(1)$ and $\varepsilon_n = \omega(n^{-1/3})$. For any $n\in\mathbb{N}$, $t\in\mathbb{N}_0$, define the Erdős-Rényi *typicality* event

$$\Omega_n(un) = \left\{|\mathscr{C}^{\mathrm{ER}}_{\max}(n,vn)| \in (n[\zeta(v)-\varepsilon_n], n[\zeta(v)+\varepsilon_n])\ \forall\, v\in[\tfrac{1}{2},u]\right\} \cap \left\{0 \le \max_{s\le un}|\mathscr{C}^{\mathrm{ER}}_{\mathrm{sec}}(n,s)| \le n\varepsilon_n\right\}. \tag{2.16}$$

Then we can write

$$\mathbb{1}_{\Xi(un)} \sum_{s=0}^{un} \mathbb{E}_{s-1}\left[\mathbb{1}_{R(s)}\right] = \mathbb{1}_{\Xi(un)} \left(\mathbb{1}_{\Omega_n(un)} + \mathbb{1}_{\Omega_n^c(un)}\right) \sum_{s=0}^{un} \mathbb{E}_{s-1}\left[\mathbb{1}_{R(s)}\right]. \tag{2.17}$$

Again, we see that

$$\frac{1}{n} \mathbb{1}_{\Omega_n^c(un)} \sum_{s=0}^{un} \mathbb{E}_{s-1}\left[\mathbb{1}_{R(s)}\right] \xrightarrow{\mathbb{P}} 0, \tag{2.18}$$

because $\sum_{s=0}^{un} \mathbb{E}_{s-1}[\mathbb{1}_{R(s)}] \le un$ and $\mathbb{1}_{\Omega_n^c(un)} \xrightarrow{\mathbb{P}} 0$.

2. It remains to compute $\frac{1}{n}\mathbb{1}_{\Xi(un)}\mathbb{1}_{\Omega_n(un)} \sum_{s=0}^{un} \mathbb{E}_{s-1}[\mathbb{1}_{R(s)}]$, which is the only term that will be non-zero after we take the limit $n \to \infty$ in (2.13). Recall that $\mathbb{E}_{s-1}[\mathbb{1}_{R(s)}] \sim \mathsf{Bernoulli}(p_s)$, where the success probabilities p_s were introduced in (2.5).

Since $|\mathcal{V}| = n$, we have the following bounds:

$$p_s = \sum_{\mathscr{C} \in \boldsymbol{\mathcal{CC}}(G_n(s-1))} \frac{|\mathscr{C}|(|\mathscr{C}|-1)}{n(n-1)} \le \begin{cases} \frac{|\mathscr{C}^{\mathrm{ER}}_{\max}(n,n/2)|}{n}, & 0 \le s \le \frac{1}{2}n, \\ \frac{|\mathscr{C}^{\mathrm{ER}}_{\max}(n,s)|^2}{n^2} + \max\limits_{0 \le s \le un} \frac{|\mathscr{C}^{\mathrm{ER}}_{\mathrm{sec}}(n,s)|}{n} & \frac{1}{2}n < s < un, \end{cases} \tag{2.19}$$

where the last bound is uniform in $s \le un$. The first line, for times below $n/2$, holds since

$$\sum_{\mathscr{C} \in \boldsymbol{\mathcal{CC}}(G_n(s-1))} \frac{|\mathscr{C}|(|\mathscr{C}|-1)}{n(n-1)} \le \frac{|\mathscr{C}^{\mathrm{ER}}_{\max}(n,s)|}{n} \le \frac{|\mathscr{C}^{\mathrm{ER}}_{\max}(n,n/2)|}{n}. \tag{2.20}$$

The second line separates the contribution of the maximal component and all the other components, and bound the non-maximal component similarly as in the first line.

In the *supercritical* regime, we separately describe the contribution of the unique largest component and give an upper bound only on the probability of rejection due to the other components. On the Erdős-Rényi typicality event $\Omega_n(un)$ (recall (2.16)), the size of all the components in the subcritical and critical regime and all the components but the largest one in the supercritical regime can be uniformly bounded by $|\mathscr{C}| \le Zn^{2/3}$, where Z is a positive random variable. From (2.4), (2.16) and (2.19), it follows that, on the event $\Omega_n(un)$,

$$\mathbb{1}_{\Omega_n(un)} \sum_{s=0}^{un} \mathbb{E}_{s-1}\left[\mathbb{1}_{R(s)}\right] = \mathbb{1}_{\Omega_n(un)} \sum_{s=0}^{un} p_s \begin{cases} \le u|\mathscr{C}^{\mathrm{ER}}_{\max}(n,n/2)| \le \varepsilon_n u, & 0 \le u \le \frac{1}{2}, \\ = \sum_{s=\frac{n}{2}}^{un} (\zeta(i/n) + \varepsilon_n)^2 + \mathcal{R}(un), & u > \frac{1}{2}, \end{cases} \tag{2.21}$$

where we use that $\zeta(\frac{1}{2}) = 0$. Since $|\mathscr{C}^{\mathrm{ER}}_{\max}(n,n/2)| \le \varepsilon_n n$ on $\Omega_n(un)$, the remainder term $\mathcal{R}(un)$ can be bounded as (recall (2.19))

$$\mathcal{R}(un) \le un \max_{0 \le s \le un} \frac{|\mathscr{C}^{\mathrm{ER}}_{\mathrm{sec}}(n,s)|}{n} \le \varepsilon_n un = o(n). \tag{2.22}$$

3. Before wrapping up, let us note that

$$\frac{o(n) + \sum_{s=0}^{un} \zeta(s/n)^2}{n} \to \int_0^u \mathrm{d}v\, \zeta^2(v), \tag{2.23}$$

because ζ^2 is bounded and hence Riemann integrable over compact intervals $[0,u]$, and $\frac{1}{n}\sum_{s=0}^{un} \zeta(s/n)^2$ is a Riemann sum of ζ^2 over a regular partition of $[0,u]$ into subintervals of length $1/n$. This allows us to finish our previous computation, namely,

$$\begin{aligned} &\frac{1}{n} \sum_{s=0}^{un} \mathbb{1}_{R(s)} \\ &\quad = \frac{1}{n} \left[\mathbb{1}_{\Xi(un)} \left(\mathbb{1}_{\Omega_n(un)} + \mathbb{1}_{\Omega_n^c(un)}\right) \left(\sum_{s=0}^{un} \mathbb{E}_{s-1}\left[\mathbb{1}_{R(s)}\right] + o(n) \right) + \mathbb{1}_{\Xi^c(un)} \sum_{s=0}^{un} \mathbb{1}_{R(s)} \right] \end{aligned} \tag{2.24}$$

$$= \frac{1}{n}\left(o(n) + \mathbb{1}_{\Xi(un)}\mathbb{1}_{\Omega_n(un)}\sum_{s=0}^{un}\mathbb{E}_{s-1}\left[\mathbb{1}_{R(s)}\right]\right) \xrightarrow{d} \int_0^u \mathrm{d}v\,\zeta^2(v) = u - \phi(u)$$

(recall (1.7)), from which the desired result follows. □

• **Mapping between times.** The main purpose of Lemma 2.5 is to show that on time scales of order n there is a function ϕ (recall Definition 1.11) capturing the correspondence, in the limit as $n \to \infty$, between the times at which the standard Erdős-Rényi graph process and its cycle-free counterpart have certain quantities distributed equally, notably, the sizes of their connected components. Since ϕ is strictly monotone, it admits a proper inverse, which allows us to relate the cycle-free graph process to the standard Erdős-Rényi graph growth process.

• **Largest component.** To conclude the analysis of the cycle-free graph process, we combine the above results to obtain a characterisation of the largest component of the cycle-free Erdős-Rényi graph process:

Lemma 2.6 [Size of the largest component] *For $s \in [0,1]$, let $|\mathscr{C}_{\max}^{\mathrm{cfER}}(n, sn)|$ be the size of the largest component of the cycle-free Erdős-Rényi graph process on n vertices at time sn. Then*

$$\frac{|\mathscr{C}_{\max}^{\mathrm{cfER}}(n, sn)|}{n} \xrightarrow{\mathbb{P}} \eta(s), \qquad \eta(s) = \zeta(\phi^{-1}(s)). \tag{2.25}$$

Proof. In Lemma 2.4 we have established that for every $n \in \mathbb{N}$ the connected components of the cycle-free graph $G_n^{\mathrm{cf}}(t)$ correspond exactly to the connected components of the standard Erdős-Rényi graph process at some random time, denoted by $\tau_n^{-1}(t)$. In Lemma 2.5 we have established that $\tau_n^{-1}(t) = n\phi^{-1}(t/n) + o_{\mathbb{P}}(n)$ and $|\mathscr{C}_{\max}^{\mathrm{ER}}(n, sn)| = n\zeta(s) + o_{\mathbb{P}}(n)$. Hence

$$|\mathscr{C}_{\max}^{\mathrm{cfER}}(n, sn)| \stackrel{d}{=} |\mathscr{C}_{\max}^{\mathrm{ER}}(n, n(\phi^{-1}(s) + o_{\mathbb{P}}(1)))| = n\zeta(\phi^{-1}(s) + o_{\mathbb{P}}(1)) + o_{\mathbb{P}}(n), \tag{2.26}$$

from which the claim follows. □

2.3 Drop-down time and mixing profile

As stated in Theorem 1.19, for CDP the mixing profile exhibits a cut-off at a *random* time. From that moment onwards, the total variation distance follows a *deterministic* curve that is related to the typical structure of CDP. The following lemma gives the distribution of the *drop-down* time and settles Theorem 1.19(1):

Lemma 2.7 [Limit distribution of drop-down time for ISRW on CDP] *Recall $T_{n,v_0}^{\Downarrow}$ from Definition* 1.15. *There exists a $[0,1]$-valued random variable $s^{\Downarrow}$ with a distribution function $\mathbb{P}(s^{\Downarrow} \leq s) = \eta(s)$, $s \in [0,1]$, such that*

$$\frac{T_{n,v_0}^{\Downarrow}}{n} \xrightarrow{d} s^{\Downarrow}. \tag{2.27}$$

Proof. By the arguments in the proof of Lemma 2.1, the sizes of the connected components of the cycle-free associated graph process exactly correspond to the sizes of the permutation cycles of the CDP at a given time t. Therefore we must study the probability that a uniform vertex lies on the largest component of a cycle-free Erdős-Rényi graph process.

Let $\mathbb{P}_n$ denote the law of CDP on $[n]$ and $\mathbb{P}_n^{\mathrm{cfER}}$ the law of the associated graph process, which is a cycle-free Erdős-Rényi graph process (recall Definition 1.12 and Lemma 2.1). Fix a sequence ε_n such that $\varepsilon_n = o(1)$ and $\varepsilon_n = \omega(n^{-1/3})$, and for $n \in \mathbb{N}$, $t \geq 0$ define the *typicality* event

$$\Omega_n^{\mathrm{cfER}}(t) = \left\{|\mathscr{C}_{\max}^{\mathrm{cfER}}(n,t)| \in n(\eta(t/n) - \varepsilon_n, \eta(t/n) + \varepsilon_n,)\right\} \cap \left\{|\mathscr{C}_{\mathrm{sec}}^{\mathrm{cfER}}(n,t)| \leq n\varepsilon_n\right\}. \tag{2.28}$$

For $s \in [0, \frac{1}{2})$, we have $\mathbb{P}_n(T_{n,v_0}^{\Downarrow} \leq sn) = 0$ by the definition of $T_{n,v_0}^{\Downarrow}$. For $s \in [\frac{1}{2}, 1]$, by Lemma 2.1,

$$\begin{aligned}\mathbb{P}_n(T_{n,v_0}^{\Downarrow} \leq sn) &= \mathbb{P}_n^{\mathrm{cfER}}\big(v_0 \in \mathscr{C}_{\max}^{\mathrm{cfER}}(sn)\big) \\ &= \mathbb{P}_n^{\mathrm{cfER}}\big(\{v_0 \in \mathscr{C}_{\max}^{\mathrm{cfER}}(sn)\} \cap \Omega_n^{\mathrm{cfER}}(sn)\big) + \mathbb{P}_n^{\mathrm{cfER}}\big(\{v_0 \in \mathscr{C}_{\max}^{\mathrm{cfER}}(sn)\} \cap [\Omega_n^{\mathrm{cfER}}(sn)]^{\mathrm{c}}\big).\end{aligned} \tag{2.29}$$

Since the event $\Omega_n^{\mathrm{cfER}}(sn)$ occurs with probability $1-o(1)$ and $\mathbb{P}_n^{\mathrm{cfER}}(\{v_0 \in \mathscr{C}_{\max}^{\mathrm{cfER}}(sn)\} \cap \Omega_n^{\mathrm{cfER}}(sn)) = \eta(s) + o(1)$, we see that, for any $s \in [0,1]$,

$$\mathbb{P}_n(T^{\Downarrow}_{n,v_0} \le sn) \overset{n\to\infty}{\to} \eta(s). \tag{2.30}$$

Since η is continuous, non-negative and non-decreasing on $[0,1]$ such that $\int_0^1 \mathrm{d}u\, \eta(u) = 1$, (2.30) defines a proper distribution function (recall Definition 1.11). See Appendix B for a detailed computation. □

With the above results in hand, we are ready to prove the pointwise version of Theorem 1.19(2), characterising the mixing profile of ISRW on CDP:

Lemma 2.8 [Pointwise limit of mixing profile for ISRWs on CDP] *Uniformly in $v_0 \in [n]$,*

$$\mathcal{D}_n^{v_0}(sn) \overset{d}{\to} 1 - \eta(s)Y(s), \qquad s \in [0,1), \tag{2.31}$$

where $Y(s) \sim \mathsf{Bernoulli}(\eta(s))$.

Proof. Given a permutation π, let

$$|\gamma_{\max}(\pi)| = \max\{|\gamma_v(\pi)|\colon\ v \in [n]\} \tag{2.32}$$

denote the size of the largest cycle of π. For every $n \in \mathbb{N}$,

$$\mathcal{D}_n^{v_0}(sn) = d_{\mathrm{TV}}\Big(\mathsf{Unif}([n]), \mathsf{Unif}([\gamma_{\max}(sn)])\Big) = 1 - \frac{|\gamma_{sn}(v_0)|}{n} \tag{2.33}$$

by Definition 1.6 and the definition of total variation distance. Using Lemma 2.1, we see that

$$\begin{aligned}\mathcal{D}_n^{v_0}(sn) &\overset{d}{=} 1 - \frac{1}{n}\left(|\mathscr{C}_{v_0}(sn)|\mathbb{1}_{\Omega_n^{\mathrm{cfER}}(sn)} + |\mathscr{C}_{v_0}(sn)|\mathbb{1}_{[\Omega_n^{\mathrm{cfER}}(sn)]^{\mathrm{c}}}\right)\\ &= 1 - \frac{1}{n}\left(|\mathscr{C}_{v_0}(sn)|\mathbb{1}_{\Omega_n^{\mathrm{cfER}}(sn)}\left(\mathbb{1}_{\{T^{\Downarrow}_{n,v_0}\le sn\}} + \mathbb{1}_{\{T^{\Downarrow}_{n,v_0}> sn\}}\right) + |\mathscr{C}_{v_0}(sn)|\mathbb{1}_{[\Omega_n^{\mathrm{cfER}}(sn)]^{\mathrm{c}}}\right).\end{aligned} \tag{2.34}$$

Standard results for the size of Erdős-Rényi connected components (recall (1.9)) imply that

$$\mathbb{1}_{[\Omega_n^{\mathrm{cfER}}(sn)]^{\mathrm{c}}} \overset{\mathbb{P}}{\longrightarrow} 0, \qquad \mathbb{1}_{\Omega_n^{\mathrm{cfER}}(sn)} \overset{\mathbb{P}}{\longrightarrow} 1, \qquad \frac{|\mathscr{C}_{v_0}(sn)|\mathbb{1}_{\Omega_n^{\mathrm{cfER}}(sn)}\mathbb{1}_{\{T^{\Downarrow}_{n,v_0}> sn\}}}{n} \overset{\mathbb{P}}{\longrightarrow} 0, \tag{2.35}$$

where the last limit follows from the fact that at times $0 \le sn < T^{\Downarrow}_{n,v_0}$ the initial vertex v_0 lies on a non-largest component, and hence the numerator scales as $o(n)$. Similarly, the size of the largest Erdős-Rényi component has a well-known limit (recall (1.9)), on the event $\{T^{\Downarrow}_{n,v_0} < sn\}$, namely,

$$\frac{|\mathscr{C}_{v_0}(sn)|\mathbb{1}_{\Omega_n^{\mathrm{cfER}}(sn)}}{n} \overset{\mathbb{P}}{\longrightarrow} \eta(s). \tag{2.36}$$

Finally, the only random variable that converges to a non-degenerate random variable is

$$\mathbb{1}_{\{T^{\Downarrow}_{n,v_0}\le sn\}} \overset{d}{\to} Y(s), \qquad Y(s) \sim \mathsf{Bernoulli}(\eta(s)), \tag{2.37}$$

which follows from Lemma 2.7. Note that $\mathcal{D}_n^{v_0}(sn)$ is a sum of several random variables, and we established the convergence of each in (2.34)–(2.37). Hence, the claim follows via Slutsky's theorem. □

To conclude this section, we use Lemma 2.8 to prove the pathwise convergence part of Theorem 1.19:

Proof of Theorem 1.19*(2).* Observe that, for every $n \in \mathbb{N}$, any realisation of $\mathcal{D}_n^{v_0}(\cdot)$ is a monotone càdlàg path on the compact set $[0,1]$. In this special situation, the pointwise convergence proven in Lemma 2.8 implies pathwise convergence in the Skorokhod M_1-topology. For details, see [43, Corollary 12.5.1]. □

3 Coagulative-fragmentative dynamic permutations

In Section 2, for CDP it took effort to control the structure of the associated graph process, while the mixing profile was obtained via an easy argument. For CFDP the opposite is true: the associated graph process, introduced in Section 2.1, is the Erdős-Rényi graph process defined in Definition 1.8 (which is one of the key facts used in [38]), while the link between the cycles of the underlying permutation and the connected components of the associated graph process is far less clear. Indeed, each non-tree connected component of the associated graph process may represent *multiple* permutation cycles, which brings a *substructure* into the problem that needs to be controlled. Moreover, it is not a priori clear whether or not this substructure influences the mixing profile, since *immediately* after the drop-down time the distribution of the ISRW is uniform over a component that spans only a random fraction of the largest component of the associated graph process.

The key result in this section is that ISRW on CFDP exhibits *fast local mixing on the largest component of the associated graph process upon drop-down*. After scaling, this leads to results that are qualitatively similar to those obtained for CDP, namely, the occurrence of a *single* jump in the total variation distance, from 1 to a deterministic value on a curve related to the largest component of the associated graph process, at a *random time* whose distribution is again connected to the largest component of the associated graph process. The scaled time now takes values in $[0, \infty)$ instead of $[0, 1]$.

In Section 3.1 we identify the drop-down time and prove Theorem 1.20(1). In Section 3.2 we show that the support of ISRW lies on a single permutation cycle before the drop-down time. In Section 3.3 we prove fast local mixing after the drop-down time. In Section 3.4 we identify the mixing profile and prove Theorem 1.20(2).

Remark 3.1 [Permutation elements and graph vertices representing them]. Throughout this section we will (with a slight abuse of notation) identify the vertices in the associated graph process with the permutation elements they represent. ■

3.1 Drop-down time

Recall that the central object for the identification of the limit distribution of the drop-down time for CDP in Section 2.3 was the function η (recall Definition 1.11), which describes the size of the largest component in the cycle-free Erdős-Rényi graph process. In the setting of CFDP, we formulate a result for $T^{\Downarrow}_{n,v_0}$ analogous to Lemma 2.7, with the role of η taken over by ζ, which describes the size of the largest component in the standard Erdős-Rényi graph process:

Lemma 3.2 [Limiting distribution of drop-down time for ISRW on CFDP] *Recall $T^{\Downarrow}_{n,v_0}$ from Definition 1.15. There exists a $[0,\infty)$-valued random variable $u^{\Downarrow}$ with distribution function $\mathbb{P}(u^{\Downarrow} \leq u) = \zeta(u)$, $u \in [0,\infty)$, such that*

$$\frac{T^{\Downarrow}_{n,v_0}}{n} \xrightarrow{d} u^{\Downarrow}. \tag{3.1}$$

Proof. The proof is the same as that of Lemma 2.7, but uses the laws of CFDP and its associated graph processes, and uses ζ in place of η. □

3.2 Drop-down in a single permutation cycle

In principle, it could happen that the ISRW support has experienced fragmentation before the drop-down time, which would significantly complicate our analysis. The main point of this section is to show that, with high probability, this *does not* occur.

Lemma 3.3 [ISRW support lies on a single permutation cycle before $T^{\Downarrow}_{n,v_0}$] *Fix $\varepsilon_n > 0$ such that $\varepsilon_n = \omega(n^{-1/3})$ and $\varepsilon_n = o(1)$ as $n \to \infty$. Let $\Omega^{(\mathrm{SC})}(t)$ denote the event that the support of the ISRW at time t lies on a single permutation cycle. Then, uniformly in v_0 and $t = cn$ with $c \in (\frac{1}{2}, \infty)$,*

$$\mathbb{P}\left(\Omega^{(\mathrm{SC})}(t) \mid T^{\Downarrow}_{n,v_0} \geq t\right) = 1 - o(1). \tag{3.2}$$

Proof. Recall the associated graph process introduced in Definition 1.12, and the fact that the associated graph process of CFDP is equal in distribution to the standard Erdős-Rényi graph process. As explained in the proof of Lemma 2.1, tree components in the associated graph process correspond to permutation cycles that have never experienced fragmentation. The idea of the proof is to show that, conditionally on the event $\{T^{\Downarrow}_{n,v_0} \geq t\}$, the event $\Omega^{\text{tree}}(t)$ that ISRW at time t is supported on a single tree-component in the associated graph process occurs with high probability. Observe that $\Omega^{\text{tree}}(t) \subseteq \Omega^{(\text{SC})}(t)$. First we condition on the event $\{T^{\Downarrow}_{n,v_0} > t\}$. Afterwards we extend to the event $\{T^{\Downarrow}_{n,v_0} \geq t\}$.

Recall that in Lemma 3.2 we identified the limiting distribution of $T^{\Downarrow}_{n,v_0}/n$. Since

$$\mathbb{P}\left(\Omega^{\text{tree}}(t) \mid T^{\Downarrow}_{n,v_0} > t\right) = 1 - \mathbb{P}\left(\left[\Omega^{\text{tree}}(t)\right]^{\text{c}} \mid T^{\Downarrow}_{n,v_0} > t\right) = 1 - \frac{\mathbb{P}\left(\left[\Omega^{\text{tree}}(t)\right]^{\text{c}} \cap \{T^{\Downarrow}_{n,v_0} > t\}\right)}{\mathbb{P}(T^{\Downarrow}_{n,v_0} > t)}, \tag{3.3}$$

and the denominator is bounded away from 0 (recall Lemma 3.2), it suffices to show that $\mathbb{P}([\Omega^{\text{tree}}(t)]^{\text{c}} \cap \{T^{\Downarrow}_{n,v_0} > t\}) = o(1)$. By the law of total probability, we can take the sum over all possible realisations of the underlying dynamics to obtain

$$\mathbb{P}\left(\left[\Omega^{\text{tree}}(t)\right]^{\text{c}} \cap \{T^{\Downarrow}_{n,v_0} > t\}\right) = \mathbb{E}\left[\mathbb{P}\left(\left[\Omega^{\text{tree}}(t)\right]^{\text{c}} \cap \{T^{\Downarrow}_{n,v_0} > t\} \mid (\Pi_n(t))_{s=0}^{t}\right)\right]. \tag{3.4}$$

By [26, Theorem 5.10], with high probability the connected components of the associated graph process at time t consist of the *unique largest component, unicyclic connected components and trees.* By Definition 1.15, conditionally on $\{T^{\Downarrow}_{n,v_0} > t\}$, the support of the ISRW in the associated graph process *does not* lie on the largest component. It therefore lies, with high probability, on either a unicyclic component or a tree. Denote by $N^{\text{uc}}(t)$ the number of vertices in an Erdős-Rényi graph process that are in unicyclic connected components at time t, and recall from (1.9) that $\mathscr{C}^{\text{ER}}_{\max}(n,t)$ denotes the size of the largest component of an Erdős-Rényi graph on n vertices with t edges. It follows that

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{P}\left(\left[\Omega^{\text{tree}}(t)\right]^{\text{c}} \cap \{T^{\Downarrow}_{n,v_0} > t\} \;\middle|\; (\Pi_n(t))_{s=0}^{t}\right)\right] \\
&= \mathbb{E}\left[\mathbb{P}\left(\left[\Omega^{\text{tree}}(t)\right]^{\text{c}} \cap \{v_0 \notin \mathscr{C}_{\max}(A_{\Pi_n}(t))\} \;\middle|\; (\Pi_n(t))_{s=0}^{t}\right)\right] + o(1) \\
&\leq \mathbb{E}\left[\min\left(\frac{N^{\text{uc}}(t)}{n - |\mathscr{C}^{\text{ER}}_{\max}(n,t)|}, 1\right)\right] + o(1).
\end{aligned} \tag{3.5}$$

From [26, Theorem 5.11] it follows that $N^{\text{uc}}(t) = O_{\mathbb{P}}(n^{2/3})$ uniformly in $t = cn$ with $c \in (\frac{1}{2}, \infty)$. Since the size of the largest component is $\zeta(\frac{t}{n})n + o_{\mathbb{P}}(n)$ (recall Definition 1.11) and the number of vertices is n, it follows that the number of vertices outside the largest component at time t is $(1 - \zeta(\frac{t}{n}))n + o_{\mathbb{P}}(n) = \Theta_{\mathbb{P}}(n)$, again uniformly in $t = cn$ with $c \in (\frac{1}{2}, \infty)$. This gives

$$\mathbb{E}\left[\min\left(\frac{N^{\text{uc}}(t)}{n - |\mathscr{C}^{\text{ER}}_{\max}(n,t)|}, 1\right)\right] = o(1). \tag{3.6}$$

Putting the above estimates together, we get

$$\mathbb{P}\left(\Omega^{(\text{SC})}(t) \mid T^{\Downarrow}_{n,v_0} > t\right) = 1 - o(1). \tag{3.7}$$

It remains to show that this estimate holds not only conditionally on $\{T^{\Downarrow}_{n,v_0} > t\}$, but also conditionally on $\{T^{\Downarrow}_{n,v_0} \geq t\}$. Given that the ISRW support lies on a single cycle before time $T^{\Downarrow}_{n,v_0}$, at time $T^{\Downarrow}_{n,v_0}$ this cycle merges with exactly one other cycle whose elements are represented by the vertices in the largest component of the associated graph process, and so the ISRW remains supported on a single permutation cycle. It therefore follows that

$$\mathbb{P}\left(\Omega^{(\text{SC})}(t) \mid T^{\Downarrow}_{n,v_0} \geq t\right) = 1 - o(1). \tag{3.8}$$

□

3.3 Local mixing upon drop-down

The main difference with the setting in Section 2 is that each non-tree connected component of the associated graph process of CFDP may represent *multiple* permutation cycles. We show that, after scaling of time, this fine structure is not felt because the distribution of the ISRW rapidly becomes uniform over the elements of the permutation represented by the vertices of the relevant connected component of the associated graph process. A consequence of this *fast mixing* is the occurrence of the same phenomenon as observed for CDP, namely, at time $T^{\Downarrow}_{n,v_0}$ there is a single drop in the total variation distance.

• **Local mixing.** To formalise the arguments, we first introduce the notion of local mixing on the largest component of the associated graph process:

Definition 3.4 [Local mixing time] Consider an ISRW with distribution μ^{n,v_0} started from the element v_0 and running on top of CFDP Π_n (recall Definition 1.4), and let A_{Π_n} be the associated graph process. For $\varepsilon \in (0,1)$, define the stopping time

$$T^{(\mathrm{LM}),\varepsilon}_{n,v_0} = \min\left\{t > T^{\Downarrow}_{n,v_0}\colon\ d_{\mathrm{TV}}\left(\mu^{n,v_0}(t), \mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t)))\right) < \varepsilon\right\}. \tag{3.9}$$

♠

At time $T^{(\mathrm{LM}),\varepsilon}_{n,v_0}$, the ISRW is well mixed on the giant $\mathscr{C}_{\max}(A_{\Pi_n}(t))$. In the following statements we illustrate and quantify the influence of large-enough permutation cycles on ISRW-mixing. Below we play with three parameters n, ε, δ and take limits in the order $n \to \infty$, $\varepsilon \downarrow 0$ and $\delta \downarrow 0$. We also play with a time scale a_n satisfying $\lim_{n\to\infty} a_n = \infty$ and $a_n = o(n)$. Along the way we need some facts established in Appendices C and D that require more stringent conditions on a_n, namely, $a_n = o(n^{1/26})$, respectively, $a_n = o(n^{1/3})$. We summarise this by saying that a_n grows slowly enough.

We will often use the following Erdős-Rényi typicality event, which occurs with high probability:

Definition 3.5 [Erdős-Rényi typicality event] Take ε_n such that $\varepsilon_n = o(1)$ and $\varepsilon_n = \omega(n^{-1/3})$. Define the following event

$$\Omega^{(\mathrm{ER})}_n(t) = \left\{|\mathscr{C}^{\mathrm{ER}}_{\max}(n,t)| = n(\zeta(\tfrac{t}{n}) - \varepsilon_n, \zeta(\tfrac{t}{n}) + \varepsilon_n,)\right\} \cap \left\{|\mathscr{C}^{\mathrm{ER}}_{\mathrm{sec}}(n,t)| \le n\varepsilon_n\right\}. \tag{3.10}$$

Note that this event is different from the event $\Omega_n(un)$ defined in (2.16). ♠

Definition 3.6 [Events $\mathcal{M}_1(\varepsilon,\delta)$, $\mathcal{M}_2(\varepsilon)$] Denote by $\mathfrak{X}^{(n)}_1(t)$ the normalised size of the largest cycle at time t (see (C.1)). Recall the event $\Omega^{(\mathrm{SC})}(t)$ from Lemma 3.3, the Erdős-Rényi typicality event $\Omega^{(\mathrm{ER})}_n(t)$ from Definition 3.5 (which both occur with high probability for any $t = cn$ with $c \in (\frac{1}{2}, \infty)$), and introduce the abbreviation $M = |\mathscr{C}_{\max}(A_{\Pi_n}(T^{\Downarrow}_{n,v_0}))|$. Define the events

$$\begin{aligned}\mathcal{M}_1(\varepsilon,\delta) &= \left\{|\operatorname{supp}(\mu^{n,v_0}(T^{\Downarrow}_{n,v_0}))| > \varepsilon M\right\} \cap \Omega^{(\mathrm{SC})}(T^{\Downarrow}_{n,v_0}) \cap \Omega^{(\mathrm{ER})}_n(T^{\Downarrow}_{n,v_0}),\\ \mathcal{M}_2(\varepsilon) &= \left\{\exists\, t_L \in (T^{\Downarrow}_{n,v_0}, T^{\Downarrow}_{n,v_0} + a_n)\colon\ \mathfrak{X}^{(n)}_1(t_L) > 1 - \varepsilon^2\right\}.\end{aligned} \tag{3.11}$$

♠

Lemma 3.7 [Mixing induced by a single large cycle] *Recall Definition* 3.6. *Let* $(a_n)_{n\in\mathbb{N}}$ *be such that* $\lim_{n\to\infty} a_n = \infty$ *slowly enough. Then*

$$\left\{T^{(\mathrm{LM}),\varepsilon}_{n,v_0} \in (T^{\Downarrow}_{n,v_0}, T^{\Downarrow}_{n,v_0} + a_n)\right\} \supseteq \mathcal{M}_1(\varepsilon,\delta) \cap \mathcal{M}_2(\varepsilon). \tag{3.12}$$

Furthermore, on the event $\mathcal{M}_1(\varepsilon,\delta) \cap \mathcal{M}_2(\varepsilon) \cap \Omega^{(\mathrm{SC})}(T^{\Downarrow}_{n,v_0})$, *there exists a* $t_L \in (T^{\Downarrow}_{n,v_0}, T^{\Downarrow}_{n,v_0} + a_n)$ *such that*

$$1 - \frac{1}{n}|\mathscr{C}_{\max}(A_{\Pi_n}(t_L))| - \varepsilon \le \mathcal{D}^{v_0}_n(t_L) \le 1 - \frac{1}{n}|\mathscr{C}_{\max}(A_{\Pi_n}(t_L))| + \varepsilon. \tag{3.13}$$

Proof. Recall that the event $\Omega^{(\mathrm{SC})}(T^{\Downarrow}_{n,v_0}) \subset \mathcal{M}_1(\varepsilon,\delta)$ implies that all the mass of the ISRW-distribution enters the giant component on a single cycle. Therefore the event $\mathcal{M}_1(\varepsilon,\delta)$ implies that

$$\forall\, u \in \mathrm{supp}(\mu^{n,v_0}(T^{\Downarrow}_{n,v_0}))\colon \mu^{n,v_0}_u(T^{\Downarrow}_{n,v_0}) \le \frac{1}{\varepsilon M}, \tag{3.14}$$

where we recall that $M = |\mathscr{C}_{\max}(A_{\Pi_n}(T^{\Downarrow}_{n,v_0}))|$. The event $\mathcal{M}_1(\varepsilon,\delta) \cap \mathcal{M}_2(\varepsilon)$ indicates that a cycle of size at least $(1-\varepsilon^2)|\mathscr{C}_{\max}(A_{\Pi_n}(T^{\Downarrow}_{n,v_0}))|$ has appeared by time $T^{\Downarrow}_{n,v_0} + a_n$. We denote this large permutation cycle by $\mathfrak{X}^{(n)}_1(t_L)$. This cycle necessarily contains some mass of the ISRW-distribution because, due to the event $\mathcal{M}_1(\varepsilon,\delta)$, the mass was initially spread out over a cycle that is larger than the region not covered by $\mathfrak{X}^{(n)}_1(t_L)$. We compute the effect of the event $\mathcal{M}_1(\varepsilon,\delta)\cap\mathcal{M}_2(\varepsilon)$ on the decay of the total variation distance. The worst possible scenario is when the $\varepsilon^2 M$ elements not covered by $\mathfrak{X}^{(n)}_1(t_L)$ each carry mass $1/(\varepsilon M)$. Note that the definition of ISRW requires that the remaining mass is spread out uniformly over $\mathfrak{X}^{(n)}_1(t_L)$. A simple calculation (see Appendix D) shows that, at time t_L (introduced in the definition of the event $\mathcal{M}_2(\varepsilon)$) and for n large enough,

$$d_{\mathrm{TV}}\Big(\mu^{n,v_0}(t_L), \mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L)))\Big) < \varepsilon, \tag{3.15}$$

and (3.12) follows.

To prove (3.13) we use that, for probability mass functions $p = (p_x)_{x\in\mathcal{X}}$ and $q = (q_x)_{x\in\mathcal{X}}$,

$$d_{\mathrm{TV}}(p,q) = \sum_{x\in\mathcal{X}} [\, p_x - (p_x \wedge q_x)\,]. \tag{3.16}$$

• For the upper bound in (3.13), we use the triangle inequality to estimate

$$\begin{aligned} &d_{\mathrm{TV}}\left(\mu^{n,v_0}(t_L), \mathsf{Unif}([n])\right) \\ &\le d_{\mathrm{TV}}\left(\mu^{n,v_0}(t_L), \mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L)))\right) + d_{\mathrm{TV}}\big(\mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L))), \mathsf{Unif}([n])\big). \end{aligned} \tag{3.17}$$

Note that, by (3.16),

$$d_{\mathrm{TV}}\big(\mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L))), \mathsf{Unif}([n])\big) = 1 - \frac{1}{n}|\mathscr{C}_{\max}(A_{\Pi_n}(t_L))|, \tag{3.18}$$

while, by (3.15),

$$d_{\mathrm{TV}}\left(\mu^{n,v_0}(t_L), \mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L)))\right) \le \varepsilon. \tag{3.19}$$

Combing (3.18)–(3.19) with (3.17), we get the upper bound in (3.13).

• For the lower bound in (3.13), we note that

$$\begin{aligned} &d_{\mathrm{TV}}\left(\mu^{n,v_0}(t_L), \mathsf{Unif}([n])\right) \\ &\ge d_{\mathrm{TV}}\left(\mu^{n,v_0}(t_L), \mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L)))\right) - d_{\mathrm{TV}}\big(\mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L))), \mathsf{Unif}([n])\big). \end{aligned} \tag{3.20}$$

Combining (3.18)–(3.19) with (3.20), we get the lower bound in (3.13). □

Before proceeding we make the following observation. Since $\frac{1}{n}T^{\Downarrow}_{n,v_0} \xrightarrow{d} u^{\Downarrow}$ by Lemma 3.2, and $\mathbb{P}(u^{\Downarrow} \le \frac{1}{2}+\delta) = 2\delta(1+o(1))$ for $\delta > 0$ small enough, we note that $\mathbb{P}(\frac{1}{n}T^{\Downarrow}_{n,v_0} \le \frac{1}{2}+\delta)$ can be made arbitrarily small by picking $\delta > 0$ small. Furthermore, for $\delta > 0$ small enough, $\mathscr{C}_{\max}(A_{\Pi_n}((\frac{1}{2}+\delta)n)) \ge \delta n$ with high probability, which follows from the properties of the Erdős-Rényi giant, specifically from the fact that $\zeta'(\frac{1}{2}) = 2$. Thus, (3.14) implies that, on the event

$$\Big\{T^{\Downarrow}_{n,v_0} > (\tfrac{1}{2}+\delta)n,\ |\mathscr{C}_{\max}(A_{\Pi_n}((\tfrac{1}{2}+\delta)n))| \ge \delta n\Big\}, \tag{3.21}$$

we have

$$\begin{aligned} &\forall\, u \in \mathrm{supp}(\mu^{n,v_0}(T^{\Downarrow}_{n,v_0}))\colon \quad \mu^{n,v_0}_u(T^{\Downarrow}_{n,v_0}) \le \frac{1}{\varepsilon\delta n}, \\ &\forall\, w \notin \mathscr{C}_{\max}(A_{\Pi_n}(T^{\Downarrow}_{n,v_0}))\colon \quad \mu^{n,v_0}_w(T^{\Downarrow}_{n,v_0}) = 0, \end{aligned} \tag{3.22}$$

where the last equality holds on the event $\Omega^{\text{tree}}(T^{\Downarrow}_{n,v_0})$.

Note that for times $t \geq T^{\Downarrow}_{n,v_0}$, just as for CDP,

$$\mu_w^{n,v_0}(t) = 0 \quad \forall\, w \notin \mathscr{C}_{\max}(A_{\Pi_n}(t)) \qquad \forall\, t \geq T^{\Downarrow}_{n,v_0}, \tag{3.23}$$

since, by the construction of the associated graph process, the support of $\mu_u^{n,v_0}(T^{\Downarrow}_{n,v_0})$ always lies on a single connected component of the associated graph process. The uniform bounds above will prove to be essential below.

Lemma 3.7 allows us to quantify the probability of ε-mixing after a single appearance of a cycle of size $(1-\varepsilon^2)M$:

Proposition 3.8 *Fix $\delta > 0$. Let $(a_n)_{n\in\mathbb{N}}$ be such that $\lim_{n\to\infty} a_n = \infty$ slowly enough. Then there exists a function $\varepsilon \mapsto f(\varepsilon)$ satisfying $\lim_{\varepsilon\downarrow 0} f(\varepsilon) = 0$ such that*

$$\mathbb{P}\Big(T^{(\mathrm{LM}),\varepsilon}_{n,v_0} \notin (T^{\Downarrow}_{n,v_0}, T^{\Downarrow}_{n,v_0} + a_n), T^{\Downarrow}_{n,v_0} \geq (\tfrac{1}{2}+\delta)n\Big) \leq f(\varepsilon) + o(1), \qquad n \to \infty. \tag{3.24}$$

Consequently, on the event that $T^{\Downarrow}_{n,v_0} \geq (\frac{1}{2}+\delta)n$, the conclusion of (3.13) *fails with probability at most $f(\varepsilon)$.*

Proof. We will derive an upper bound for the probability of the event $(\mathcal{M}_1(\varepsilon,\delta)^{\mathrm{c}} \cup \mathcal{M}_2(\varepsilon)^{\mathrm{c}}) \cap \{T^{\Downarrow}_{n,v_0} \geq (\frac{1}{2}+\delta)n\}$, which by (3.12) includes the event in the left-hand side of (3.24). To do so, we will work with a further sub-event.

Denote the number of vertices in $\mathscr{C}_{\max}(A_{\Pi_n}(t))$ that are in cycles of size smaller than εn by $S(\varepsilon n, t)$. We use [13, Lemma 2.4], which states that for any $t > cn$ with $c > \frac{1}{2}$ there exists a $C > 0$ such that, for any $\varepsilon \in (0,1)$ and n large enough,

$$\mathbb{E}\left[S(\varepsilon n, t)\right] < C\varepsilon \log(\tfrac{1}{\varepsilon})\, n. \tag{3.25}$$

Define the event

$$\mathcal{M}_3(\varepsilon, t) = \{S(\varepsilon n, t) < \sqrt{\varepsilon}\, n\}. \tag{3.26}$$

Recall that the mass at $T^{\Downarrow}_{n,v_0}$ enters the largest component of the associated graph process on a cycle that belongs to a uniform element. Trivially,

$$\begin{aligned}(\mathcal{M}_1(\varepsilon,\delta)^{\mathrm{c}} \cup \mathcal{M}_2(\varepsilon)^{\mathrm{c}}) \cap \{T^{\Downarrow}_{n,v_0} \geq (\tfrac{1}{2}+\delta)n\} \\ \subseteq \Big((\mathcal{M}_1(\varepsilon,\delta)^{\mathrm{c}} \cup \mathcal{M}_2(\varepsilon)^{\mathrm{c}}) \cap \mathcal{M}_3(\varepsilon, T^{\Downarrow}_{n,v_0}) \cap \{T^{\Downarrow}_{n,v_0} \geq (\tfrac{1}{2}+\delta)n\}\Big) \cup \mathcal{M}_3(\varepsilon, T^{\Downarrow}_{n,v_0})^{\mathrm{c}}.\end{aligned} \tag{3.27}$$

We estimate the probability of these events one by one. First, use the Markov inequality and (3.25) to estimate, for n large enough,

$$\mathbb{P}\left(\mathcal{M}_3(\varepsilon, T^{\Downarrow}_{n,v_0})^{\mathrm{c}}\right) \leq C\sqrt{\varepsilon}\log(\tfrac{1}{\varepsilon}). \tag{3.28}$$

Second, estimate

$$\mathbb{P}\left(\mathcal{M}_1(\varepsilon,\delta)^{\mathrm{c}} \cap \mathcal{M}_3(\varepsilon, T^{\Downarrow}_{n,v_0})\right) \leq \mathbb{P}\left(\mathcal{M}_1(\varepsilon,\delta)^{\mathrm{c}} \mid \mathcal{M}_3(\varepsilon, T^{\Downarrow}_{n,v_0})\right) \leq \sqrt{\varepsilon}, \tag{3.29}$$

with the last inequality following from the definition of $\mathcal{M}_1(\varepsilon,\delta)$. Third, the key estimate stated in Proposition C.10, whose proof turns out to be rather delicate, yields that, for $\delta > 0$ fixed,

$$\mathbb{P}\left(\mathcal{M}_2(\varepsilon)^{\mathrm{c}}, T^{\Downarrow}_{n,v_0} \geq (\tfrac{1}{2}+\delta)n\right) = o(1). \tag{3.30}$$

Indeed, the key event that is estimated in Proposition C.10 is

$$\mathcal{E}_n(c,\varepsilon,\kappa) = \left\{\exists\,(t_k)_{k=1}^{\kappa} \in (cn, cn+a_n)\colon\ \mathfrak{X}_1^{(n)}(t_k - 1) < 1-\varepsilon,\ \mathfrak{X}_1^{(n)}(t_k) \geq 1-\varepsilon\right\}, \tag{3.31}$$

which states that there are at least $\kappa \in \mathbb{N}$ times in the interval $(cn, cn+a_n)$ such that the size of the maximal cycle crosses $(1-\varepsilon)$ upwards, i.e., $\mathfrak{X}_1^{(n)}(t_k) \geq 1-\varepsilon$. Proposition C.10 states that $\mathcal{E}_n(c,\varepsilon,\kappa)$

occurs with high probability for all $c \in (1/2, \infty)$, $\kappa \in \mathbb{N}$ and $\varepsilon > 0$. We apply Corollary C.12, which is a consequence of Proposition C.10, to obtain (3.30).

Combining (3.27)–(3.30), we find that there exists a $C > 0$ such that

$$\begin{aligned}
&\mathbb{P}\big((\mathcal{M}_1(\varepsilon,\delta)^{\mathrm{c}} \cup \mathcal{M}_2(\varepsilon)^{\mathrm{c}}) \cap \{T^{\Downarrow}_{n,v_0} \geq (\tfrac{1}{2}+\delta)n\}\big) \\
&\quad \leq \mathbb{P}\left(\mathcal{M}_3(\varepsilon, T^{\Downarrow}_{n,v_0})^{\mathrm{c}}\right) + \mathbb{P}\left(\mathcal{M}_1(\varepsilon,\delta)^{\mathrm{c}} \cap \mathcal{M}_3(\varepsilon, T^{\Downarrow}_{n,v_0})\right) + \mathbb{P}\left(\mathcal{M}_2(\varepsilon)^{\mathrm{c}}, T^{\Downarrow}_{n,v_0} \geq (\tfrac{1}{2}+\delta)n\right) \qquad (3.32) \\
&\quad \leq C\sqrt{\varepsilon}\log(\tfrac{1}{\varepsilon}) + \sqrt{\varepsilon} + o(1) < \varepsilon^{1/3}
\end{aligned}$$

for ε small enough, which in turn decays to 0 as $\varepsilon \to 0$. □

• **Adaptation of Lemma** 3.7 **and Proposition** 3.8**.** Finally, we adapt Lemma 3.7 and Proposition 3.8. Note that Lemma 3.7 is true at time $t = cn$ when we replace the events $\mathcal{M}_1(\varepsilon,\delta), \mathcal{M}_2(\varepsilon)$ by (compare with (3.11))

$$\begin{aligned}
\mathcal{M}'_1(cn,\varepsilon,\delta) &= \big\{|\operatorname{supp}(\mu^{n,v_0}(cn))| > \varepsilon n\big\} \cap \Omega^{(\mathrm{SC})}(T^{\Downarrow}_{n,v_0}), \\
\mathcal{M}'_2(cn,\varepsilon,\delta) &= \big\{\exists\, t_L \in (cn, cn+a_n)\colon\ \mathfrak{X}^{(n)}_1(t_L) > 1 - \tfrac{\varepsilon^2}{\delta}\big\}.
\end{aligned} \qquad (3.33)$$

Here, we recall the event $\Omega^{(\mathrm{SC})}(t)$ from Lemma 3.3 (which occurs with high probability for any $t = cn$ with $c \in (\frac{1}{2}, \infty)$), and the extra factor $1/\delta$ is added to accommodate the extra factor $1/\delta$ in the first line of (3.22). It remains to redo the calculations in the proofs of Lemma 3.7 and Proposition 3.8 with these modified events. Take $t = cn$ with $c \in (\frac{1}{2}, \infty)$, and define

$$T^{(\mathrm{LM}),\varepsilon}_{n,v_0}(t) = \min\left\{s > t\colon\ d_{\mathrm{TV}}\big(\mu^{n,v_0}(s), \mathsf{Unif}(\mathscr{C}_{\max}(t))\big) < \varepsilon\right\}. \qquad (3.34)$$

We start by adapting Lemma 3.7:

Lemma 3.9 [Mixing induced by a single large cycle] *Let $(a_n)_{n\in\mathbb{N}}$ be such that $\lim_{n\to\infty} a_n = \infty$ slowly enough, and let $c \in (1/2, \infty)$. Then, for any $\delta \in (0, c - \frac{1}{2})$,*

$$\begin{aligned}
&\big\{T^{(\mathrm{LM}),\varepsilon}_{n,v_0}(cn) \in (cn, cn+a_n)\big\} \cap \{(\tfrac{1}{2}+\delta)n \leq T^{\Downarrow}_{n,v_0} \leq cn\} \qquad (3.35) \\
&\qquad \supseteq \mathcal{M}'_1(cn,\varepsilon,\delta) \cap \mathcal{M}'_2(cn,\varepsilon,\delta) \cap \{(\tfrac{1}{2}+\delta)n \leq T^{\Downarrow}_{n,v_0} \leq cn - a_n\}.
\end{aligned}$$

Furthermore, on the event $\mathcal{M}_1(\varepsilon,\delta) \cap \mathcal{M}_2(\varepsilon) \cap \{(\frac{1}{2}+\delta)n \leq T^{\Downarrow}_{n,v_0} \leq cn - a_n\} \cap \Omega^{(\mathrm{SC})}(T^{\Downarrow}_{n,v_0})$ there exists a $t_L \in (cn, cn+a_n)$ such that

$$1 - \frac{1}{n}|\mathscr{C}_{\max}(A_{\Pi_n}(t_L))| - \varepsilon \leq \mathcal{D}^{v_0}_n(t_L) \leq 1 - \frac{1}{n}|\mathscr{C}_{\max}(A_{\Pi_n}(t_L))| + \varepsilon. \qquad (3.36)$$

Proof. The main ingredient in the proof of Lemma 3.7 was (3.14). Recall the extension of (3.14) in (3.22). With (3.22) in hand, we can simply follow the proof of Lemma 3.7. □

We continue by adapting Proposition 3.8:

Proposition 3.10 *Let $(a_n)_{n\in\mathbb{N}}$ be such that $\lim_{n\to\infty} a_n = \infty$ slowly enough, and let $c > \frac{1}{2}$. Then, for any $\delta \in (0, c - \frac{1}{2})$, with $\varepsilon \mapsto f(\varepsilon)$ as in Proposition* 3.8*,*

$$\mathbb{P}\Big(T^{(\mathrm{LM}),\varepsilon}_{n,v_0}(cn) \notin (cn, cn+a_n), (\tfrac{1}{2}+\delta)n \leq T^{\Downarrow}_{n,v_0} \leq cn - a_n\Big) \leq f(\varepsilon) + o(1), \qquad n \to \infty. \qquad (3.37)$$

Consequently, on the event that $(\frac{1}{2}+\delta)n \leq T^{\Downarrow}_{n,v_0} \leq cn - a_n$, the conclusion of (3.36) *fails with probability at most $f(\varepsilon)$.*

Proof. We follow the proof of Proposition 3.8, which relies on the inclusion in Lemma 3.7. Instead, we now rely on the inclusion in Lemma 3.9. Recall from the proof of Lemma 3.7 that $S(\varepsilon n, t)$ denotes the number of vertices in $\mathscr{C}_{\max}(A_{\Pi_n}(t))$ that are in cycles of size smaller than εn, and that, by (3.25), $\mathbb{E}\left[S(\varepsilon n, t)\right] < C\varepsilon\log(\frac{1}{\varepsilon})\, n$.

Recall $T^{(\mathrm{LM}),\varepsilon}_{n,v_0}$ from (3.9). Define the event

$$\mathcal{M}'_3(\varepsilon) = \{T^{(\mathrm{LM}),\varepsilon}_{n,v_0} \in (T^{\Downarrow}_{n,v_0}, T^{\Downarrow}_{n,v_0} + a_n)\}. \tag{3.38}$$

Trivially,

$$\begin{aligned}
&\Big(\mathcal{M}'_1(cn,\varepsilon,\delta)^{\mathrm{c}} \cup \mathcal{M}'_2(cn,\varepsilon,\delta)^{\mathrm{c}}\Big) \cap \Big\{(\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\Big\} \\
&\quad \subseteq \Big(\mathcal{M}'_3(\varepsilon) \cap \Big(\mathcal{M}'_1(cn,\varepsilon,\delta)^{\mathrm{c}} \cup \mathcal{M}'_2(cn,\varepsilon,\delta)^{\mathrm{c}}\Big) \cap \Big\{(\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\Big\}\Big) \\
&\qquad \cup \Big(\mathcal{M}'_3(\varepsilon)^{\mathrm{c}} \cap \{(\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\}\Big).
\end{aligned} \tag{3.39}$$

We estimate the probability of these events one by one. First, for n large enough,

$$\mathbb{P}\Big(\mathcal{M}'_3(\varepsilon)^{\mathrm{c}}, (\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\Big) \le \mathbb{P}\Big(\mathcal{M}'_3(\varepsilon)^{\mathrm{c}}, (\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0}\Big) \le f(\varepsilon) + o(1), \tag{3.40}$$

where the last inequality follows from Proposition 3.8. Second, if $T^{(\mathrm{LM}),\varepsilon}_{n,v_0} \in (T^{\Downarrow}_{n,v_0}, T^{\Downarrow}_{n,v_0} + a_n)$ and $T^{\Downarrow}_{n,v_0} \ge (\frac{1}{2}+\delta)n$, then

$$\mathbb{P}\Big(\mathcal{M}'_1(cn,\varepsilon,\delta)^{\mathrm{c}} \;\Big|\; \mathcal{M}'_3(\varepsilon), (\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\Big) = 0. \tag{3.41}$$

Indeed, $\mathcal{M}'_1(cn,\varepsilon,\delta)^{\mathrm{c}}$ and $T^{\Downarrow}_{n,v_0} \le cn - a_n$ imply that $|\operatorname{supp}(\mu^{n,v_0}(T^{\Downarrow}_{n,v_0} + a_n))| \le \varepsilon n$. By an application of (3.16) with $\mathcal{X} = [n]$, $p = \mathsf{Unif}([n])$ (for which $p_v = \frac{1}{n}$ for all $v \in [n]$) and $q_v = \mu^{n,v_0}(T^{\Downarrow}_{n,v_0} + a_n)$, this implies that

$$\mathcal{D}^{v_0}_n(T^{\Downarrow}_{n,v_0} + a_n) \ge 1 - \varepsilon. \tag{3.42}$$

However, the latter is incompatible with (3.13) when $T^{\Downarrow}_{n,v_0} \ge (\frac{1}{2}+\delta)n$, since

$$\begin{aligned}
\mathcal{D}^{v_0}_n(T^{\Downarrow}_{n,v_0} + a_n) \le \mathcal{D}^{v_0}_n(t_L) &\le 1 - \frac{1}{n}|\mathscr{C}_{\max}(A_{\Pi_n}(t_L))| + \varepsilon \\
&\le 1 - \frac{1}{n}|\mathscr{C}_{\max}(A_{\Pi_n}((\tfrac{1}{2}+\delta)n))| + \varepsilon \\
&\le 1 - 2\delta + o(1) + \varepsilon < 1 - \varepsilon,
\end{aligned} \tag{3.43}$$

where the second inequality uses the definition of the event $\mathcal{M}'_3(\varepsilon)$, and the last inequality is valid for ε small enough depending on δ. Third, apply the key estimate stated in Proposition C.10 (see the explanation below (3.31)), to get

$$\mathbb{P}\Big(\mathcal{M}'_2(cn,\varepsilon,\delta), (\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\Big) = o(1). \tag{3.44}$$

Combining (3.35), (3.39)–(3.41) and (3.44), we obtain

$$\begin{aligned}
&\mathbb{P}\Big(T^{(\mathrm{LM}),\varepsilon}_{n,v_0}(cn) \notin (cn, cn+a_n), (\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\Big) \\
&\quad \le \mathbb{P}\Big((\mathcal{M}'_1(cn,\varepsilon,\delta)^{\mathrm{c}} \cup \mathcal{M}'_2(cn,\varepsilon,\delta)^{\mathrm{c}}) \cap \{(\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\}\Big) \\
&\quad \le \mathbb{P}\Big(\mathcal{M}'_3(\varepsilon)^{\mathrm{c}}, (\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\Big) \\
&\qquad + \mathbb{P}\Big(\mathcal{M}'_1(cn,\varepsilon,\delta)^{\mathrm{c}} \cap \mathcal{M}'_3(\varepsilon), (\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\Big) \\
&\qquad + \mathbb{P}\Big(\mathcal{M}'_2(cn,\varepsilon,\delta)^{\mathrm{c}}, (\tfrac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le cn - a_n\Big) \\
&\quad \le f(\varepsilon) + o(1) + 0 + o(1) = f(\varepsilon) + o(1),
\end{aligned} \tag{3.45}$$

where the first inequality uses the inclusion in Lemma 3.9. □

3.4 Mixing profile

Like in the case of CDP, the results on the mixing profile are established in two steps. First we establish pointwise convergence, afterwards we extend to process convergence. The following lemma settles Theorem 1.20(2):

Lemma 3.11 [Pointwise convergence of the mixing profile for ISRW on CFDP] *Uniformly in $v_0 \in [n]$,*

$$\mathcal{D}_n^{v_0}(sn) \xrightarrow{d} 1 - \zeta(s)W(s), \qquad s \in [0,\infty), \tag{3.46}$$

where $W(s) \sim \mathsf{Bernoulli}(\zeta(s))$.

Proof. Fix $s \in [0,\infty)$ and split the random variable $\mathcal{D}_n^{v_0}(sn) - 1$ as

$$\begin{aligned}\mathcal{D}_n^{v_0}(sn) - 1 = [\mathcal{D}_n^{v_0}(sn) - 1] \left(\mathbb{1}_{\{T^{\Downarrow}_{n,v_0} > sn\}} + \mathbb{1}_{\{T^{\Downarrow}_{n,v_0} \le sn\}} \right) \\ \times \left(\mathbb{1}_{\Omega_n^{(\mathrm{ER})}(sn)} + \mathbb{1}_{[\Omega_n^{(\mathrm{ER})}(sn)]^{\mathrm{c}}} \right) \left(\mathbb{1}_{\Omega^{\mathrm{tree}}(sn)} + \mathbb{1}_{[\Omega^{\mathrm{tree}}(sn)]^{\mathrm{c}}} \right),\end{aligned} \tag{3.47}$$

where the event $\Omega^{\mathrm{tree}}(sn)$ is defined in the proof of Lemma 3.3, and we recall the Erdős-Rényi typicality event (see (3.10))

$$\Omega_n^{(\mathrm{ER})}(t) = \left\{ |\mathscr{C}^{\mathrm{ER}}_{\max}(n,t)| = n(\zeta(\tfrac{t}{n}) - \varepsilon_n, \zeta(\tfrac{t}{n}) + \varepsilon_n,) \right\} \cap \left\{ |\mathscr{C}^{\mathrm{ER}}_{\mathrm{sec}}(n,t)| \le n\varepsilon_n \right\}. \tag{3.48}$$

Because $\Omega_n^{(\mathrm{ER})}(sn)$ and $\Omega^{\mathrm{tree}}(sn)$ both occur with high probability, the terms containing the indicators $\mathbb{1}_{[\Omega_n^{\mathrm{ER}}(sn)]^{\mathrm{c}}}$ and $\mathbb{1}_{[\Omega^{\mathrm{tree}}]^{\mathrm{c}}}$ converge to 0 in probability, and hence

$$\mathcal{D}_n^{v_0}(sn) - 1 = [\mathcal{D}_n^{v_0}(sn) - 1] \left(\mathbb{1}_{\{T^{\Downarrow}_{n,v_0} > sn\}} + \mathbb{1}_{\{T^{\Downarrow}_{n,v_0} \le sn\}} \right) \mathbb{1}_{\Omega_n^{(\mathrm{ER})}(sn)} \mathbb{1}_{\Omega^{\mathrm{tree}}(sn)} + o_{\mathbb{P}}(1). \tag{3.49}$$

To deal with the first term in (3.49), we note that

$$\begin{aligned}&[\mathcal{D}_n^{v_0}(sn) - 1] \mathbb{1}_{\{T^{\Downarrow}_{n,v_0} > sn\}} \mathbb{1}_{\Omega_n^{(\mathrm{ER})}(sn)} \mathbb{1}_{\Omega^{\mathrm{tree}}(sn)} \\ &\quad \stackrel{d}{=} \left[\left(1 - \frac{O_{\mathbb{P}}(n^{2/3})}{n} \right) - 1 \right] \mathbb{1}_{\{T^{\Downarrow}_{n,v_0} > sn\}} \mathbb{1}_{\Omega_n^{(\mathrm{ER})}(sn)} \mathbb{1}_{\Omega^{\mathrm{tree}}(sn)} \xrightarrow{\mathbb{P}} 0,\end{aligned} \tag{3.50}$$

since, on the above events, the distribution of ISRW is uniform over a single permutation cycle *outside* of the largest component of the associated graph process, whose size is $O_{\mathbb{P}}(n^{2/3})$.

To deal with the second term in (3.49), which only contributes when $s > \frac{1}{2}$, we use Lemma 3.2. For $\delta > 0$ sufficiently small and a_n as in Proposition 3.10, we split

$$\mathbb{1}_{\{T^{\Downarrow}_{n,v_0} \le sn\}} = \mathbb{1}_{\left\{ (\frac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le sn - a_n \right\}} + \mathbb{1}_{\left\{ sn - a_n < T^{\Downarrow}_{n,v_0} \le sn \right\}} + \mathbb{1}_{\left\{ T^{\Downarrow}_{n,v_0} < (\frac{1}{2}+\delta)n \right\}}. \tag{3.51}$$

We rely on (3.36) in Lemma 3.9, which holds with high probability due to Proposition 3.10. (It is here that we need $T^{\Downarrow}_{n,v_0} \ge (\frac{1}{2}+\delta)n$, since this appears as an assumption in Proposition 3.10.) We claim that

$$\sup_{t \in \mathbb{N}} \left| \frac{1}{n} |\mathscr{C}_{\max}(A_{\Pi_n}(t))| - \zeta(\tfrac{t}{n}) \right| = o_{\mathbb{P}}(1). \tag{3.52}$$

Indeed, (3.52) holds because $\frac{1}{n}|\mathscr{C}_{\max}(A_{\Pi_n}(sn))| \xrightarrow{\mathbb{P}} \zeta(s)$ for all $s > 0$ fixed, $s \mapsto \zeta(s)$ is non-decreasing and continuous, and $s \mapsto \frac{1}{n}|\mathscr{C}_{\max}(A_{\Pi_n}(sn))|$ is non-decreasing. By (3.36) and (3.52), we obtain, for all $s > \frac{1}{2} + \delta$ and on the event $\{T^{(\mathrm{LM}),\varepsilon}_{n,v_0}(cn) \in (sn, sn + a_n), (\frac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le sn - a_n\}$, that there exists a $t_L \in (sn, sn + a_n)$ such that

$$1 - \zeta(\tfrac{t_L}{n}) - \varepsilon - o_{\mathbb{P}}(1) \le \mathcal{D}_n^{v_0}(t_L) \le 1 - \zeta(\tfrac{t_L}{n}) + \varepsilon + o_{\mathbb{P}}(1). \tag{3.53}$$

Since $\varepsilon > 0$ is arbitrary, we conclude that, on the event $\{T^{(\mathrm{LM}),\varepsilon}_{n,v_0}(sn) \in (sn, sn + a_n), (\frac{1}{2}+\delta)n \le T^{\Downarrow}_{n,v_0} \le sn - a_n\}$, there exists a $t_L \in (sn, sn + a_n)$ such that

$$\mathcal{D}_n^{v_0}(t_L) = 1 - \zeta(\tfrac{t_L}{n}) + o_{\mathbb{P}}(1). \tag{3.54}$$

Since the above is true for all $s > \frac{1}{2} + \delta$, and $t \mapsto \mathcal{D}_n^{v_0}(t)$ is non-increasing, while $s \mapsto 1 - \zeta(s)$ is non-increasing and continuous, (3.54) implies that, for all $c > \frac{1}{2} + \delta$ and on the event $\{(\frac{1}{2} + \delta)n \leq T^{\Downarrow}_{n,v_0} \leq sn - a_n\}$,

$$\mathcal{D}_n^{v_0}(sn) = 1 - \zeta(s) + o_{\mathbb{P}}(1). \tag{3.55}$$

Since $\mathbb{1}_{\Omega_n^{(\mathrm{ER})}(sn)} \mathbb{1}_{\Omega^{\mathrm{tree}}(sn)} \xrightarrow{\mathbb{P}} 1$, it follows that

$$[1 - \mathcal{D}_n^{v_0}(sn)] \mathbb{1}_{\{(\frac{1}{2}+\delta)n \leq T^{\Downarrow}_{n,v_0} \leq sn - a_n\}} \mathbb{1}_{\Omega_n^{(\mathrm{ER})}(sn)} \mathbb{1}_{\Omega^{\mathrm{tree}}(sn)} - \zeta(s) \mathbb{1}_{\{\frac{1}{2}+\delta \leq \frac{1}{n} T^{\Downarrow}_{n,v_0} \leq s - \frac{a_n}{n}\}} = o_{\mathbb{P}}(1). \tag{3.56}$$

By Lemma 3.2 and Slutsky's theorem, we thus conclude that

$$[1 - \mathcal{D}_n^{v_0}(sn)] \mathbb{1}_{\{T^{\Downarrow}_{n,v_0} \leq sn - a_n\}} \mathbb{1}_{\Omega_n^{(\mathrm{ER})}(sn)} \mathbb{1}_{\Omega^{\mathrm{tree}}(sn)} \xrightarrow{d} \zeta(s) \mathbb{1}_{\{\frac{1}{2}+\delta \leq u^{\Downarrow} \leq s\}}. \tag{3.57}$$

Finally, by Lemma 3.2,

$$\mathbb{P}(sn - a_n < T^{\Downarrow}_{n,v_0} \leq sn) + \mathbb{P}(\tfrac{1}{n} T^{\Downarrow}_{n,v_0} < \tfrac{1}{2} + \delta) \to \zeta(\tfrac{1}{2} + \delta), \tag{3.58}$$

which tends to 0 as $\delta \downarrow 0$. The claim in (3.46) follows by combining (3.49)–(3.51) and (3.57)–(3.58). □

Finally, an argument based on monotonicity and a growing sequence of compact intervals settles Theorem 1.20 and concludes this section.

Proof of Theorem 1.20*(2).* Observe that for every $n \in \mathbb{N}$, any realisation of $\mathcal{D}_n^{v_0}(\cdot)$ is a monotone càdlàg path on the set $[0, \infty)$. The pointwise convergence proven in Lemma 3.11 implies, by [43, Corollary 12.5.1.], pathwise convergence in the Skorokhod M_1-topology on any compact set $[0, t]$ such that $t > 0$ is with probability 1 a continuity point of the limiting process. But the latter is true for any $t > 0$ because the limiting process has almost surely *one* point of discontinuity, whose position is distributed randomly according to the non-atomic distribution identified in Lemma 3.2. Taking a sequence $(t_k)_{k \in \mathbb{N}}$ of such continuity points with $t_k \to \infty$ as $k \to \infty$, we also obtain pathwise convergence in the Skorokhod M_1-topology on the non-compact set $[0, \infty)$. For details, see [43, p. 414]. □

A Infinite-speed random walk as limit of finite-speed random walk

It is natural to ask what the relation is between an infinite-speed random walk (ISRW) and a finite-speed standard random walk.

Definition A.1 [Finite-speed random walk on Π_n] Let Π_n be a dynamic permutation. Fix an element v_0. Recall that $\gamma_{v_0}(0)$ is the cycle of $\Pi_n(0)$ that contains v_0. Pick $\rho \in \mathbb{N}$ as the speed ratio between the evolution of the random walk and the random graph. More formally, let Π_n satisfy the condition

$$\forall\, n \in \mathbb{N}_0\ \forall\, i\ \in \{0, \ldots, \rho - 1\} \colon\ \Pi_n(\rho n) \equiv \Pi_n(\rho n + i). \tag{A.1}$$

Denote by $\Pi_n(t)[i]$ the image of the element i under the permutation $\Pi_n(t)$. The finite-speed random walk is the random process $(Y_n^{v_0}(t))_{t \in \mathbb{N}_0}$ on $[n]$ that has $Y_n^{v_0}(0) = v_0$ and, for any time $t > 0$,

$$\begin{aligned} &\mathbb{P}(Y_n^{v_0}(t) = i \,|\, Y_n^{v_0}(t-1) = j) \\ &= \begin{cases} 1, & \text{if } \Pi_n(t)[i] = j \text{ and } \Pi_n(t)[j] = i \text{ (which is the case for 1- and 2-cycles),} \\ \frac{1}{2}, & \text{if either } \Pi_n(t)[i] \neq j \text{ and } \Pi_n(t)[j] = i \text{ or } \Pi_n(t)[i] = j \text{ and } \Pi_n(t)[j] \neq i, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{A.2}$$

i.e., a simple symmetric random walk on the elements of any given cycle, where the underlying permutation changes once every ρ steps. ♠

With this definition in hand, we can explain what we mean by saying that ISRW has infinite speed:

Proposition A.2 [ISRW arises as limit] *Let Π_n be a dynamic permutation. Recall $(\mu^{n,v_0}(t))_{t\in\mathbb{N}_0}$ defined in Definition* 1.6*, and let $Y_n^{v_0} = \{Y_n^{v_0}(t)\}_{t\in\mathbb{N}_0}$ be a finite-speed random walk with speed ratio ρ on the dynamic permutation Π_n^Y, with distribution $\mu^{Y_n^{v_0}}(t)$ at time t. Note that $v_0 \in [n]$ is the starting element for both the ISRW and the finite-speed random walk. Finally, suppose that the following relation holds between Π_n and Π_n^Y:*

$$\forall\, n \in \mathbb{N}_0\ \forall\, i\ \in \{0,\dots,\rho-1\}\colon\ \Pi_n(n+i) \equiv \Pi_n^Y(n). \tag{A.3}$$

Then, with high probability, for any $i \in \mathbb{N}$,

$$d_{\mathrm{TV}}\left(\mu^{n,v_0}(i), \mu^{Y_n^{v_0}}(\rho i)\right) \overset{\rho\to\infty}{\to} 0. \tag{A.4}$$

Proof. It is well-known that a simple random walk on a circle of size m mixes in $O(m^2)$ steps. Pick $\varepsilon_{RW} > 0$, and pick $\rho(\varepsilon_{RW})$ larger than the ε_{RW}-mixing time of a simple random walk on a circle of size n. Denote by $Y_n^{v_0}(\rho(\varepsilon_{RW})i)|_{\mathcal{C}_{\rho i}^Y(v)}$ the restriction of $Y_n^{v_0}(\rho(\varepsilon_{RW})i)$ to the cycle $\mathcal{C}_{\rho i}^Y(v) \in \Pi_n^Y(\rho i)$, and let $c\,\mathsf{Unif}(S)$ be the uniform distribution multiplied component-wise by the constant c. Then

$$d_{\mathrm{TV}}\left(Y_n^{v_0}(\rho(\varepsilon_{RW})i)|_{\mathcal{C}_{\rho i}^Y(v)}, \frac{1}{\sum_{v\in\mathcal{C}_i(v)}\mu_v^{n,v_0}(i)}\mathsf{Unif}([|\mathcal{C}_i(v)|])\right) < \varepsilon_{RW}. \tag{A.5}$$

Since, by definition,

$$\mu^{n,v_0}(i)|_{\mathcal{C}_i(v)} = \frac{1}{\sum_{v\in\mathcal{C}_i(v)}\mu_v^{n,v_0}(i)}\,\mathsf{Unif}([|\mathcal{C}_i(v)|]), \tag{A.6}$$

the claim follows. □

B Normalisation of the jump-time distribution

In this appendix we return to Definition 1.11 and show that $\lim_{s\to\infty}\phi(s) = 1$, i.e., the laws of the jump-down times in Theorems 1.19–1.20 are normalised. The proof amounts to showing that

$$\int_{\frac12}^{\infty} \mathrm{d}s\,[1-\zeta^2(s)] = \tfrac12, \tag{B.1}$$

where $\zeta(u)$ is the unique solution of the equation

$$\mathrm{e}^{-2s\zeta(s)} = 1-\zeta(s), \qquad s \in \left[\tfrac12,\infty\right). \tag{B.2}$$

To evaluate (B.1), introduce a new variable $u = \zeta(s)$ and write

$$\int_{\frac12}^{\infty} \mathrm{d}s\,[1-\zeta^2(s)] = \int_0^1 \mathrm{d}u\,\frac{\mathrm{d}s}{\mathrm{d}u}\,(1-u^2). \tag{B.3}$$

To compute the Jacobian $\frac{\mathrm{d}s}{\mathrm{d}u}$, we take the logarithm of (B.2) and differentiate:

$$s = -\frac{1}{2u}\log(1-u), \qquad \frac{\mathrm{d}s}{\mathrm{d}u} = \frac{1}{2u(1-u)} + \frac{\log(1-u)}{2u^2}. \tag{B.4}$$

This gives

$$\begin{aligned}\int_0^1 \mathrm{d}u\,\frac{\mathrm{d}s}{\mathrm{d}s}\,(1-u^2) &= \int_0^1 \mathrm{d}u\,(1+u)\frac{1}{2u}\left[1-\frac{1-u}{u}\sum_{k\in\mathbb{N}}\frac{u^k}{k}\right]\\ &= \frac12\int_0^1 \mathrm{d}u\,\left[(1+u)-(1-u^2)\sum_{l\in\mathbb{N}_0}\frac{u^l}{l+2}\right] = \frac12\int_0^1 \mathrm{d}u\sum_{l\in\mathbb{N}_0} c_l u^l\end{aligned} \tag{B.5}$$

with the coefficients c_l given by

$$c_l = \delta_{0l} + \delta_{1l} - \frac{1}{l+2}\mathbb{1}_{\{l\geq 0\}} + \frac{1}{l}\mathbb{1}_{\{l\geq 2\}}, \qquad l \in \mathbb{N}_0. \tag{B.6}$$

The sum has radius of convergence 1, and so term by term integration gives

$$\begin{aligned}\frac{1}{2}\int_0^1 \mathrm{d}u \sum_{l\in\mathbb{N}_0} c_l u^l &= \frac{1}{2}\left[\frac{1}{2}\times 1 + \frac{2}{3}\times\frac{1}{2} + \sum_{l\geq 2}\frac{1}{l+1}\left(\frac{1}{l}-\frac{1}{l+2}\right)\right] \\ &= \frac{1}{4}+\frac{1}{6}+\frac{1}{2}\sum_{l\geq 2}\left[\frac{1}{l(l+1)}-\frac{1}{(l+1)(l+2)}\right] = \frac{1}{4}+\frac{1}{6}+\frac{1}{12} = \frac{1}{2},\end{aligned} \tag{B.7}$$

as required.

C Cycle structure of coagulative-fragmentative dynamic permutations and Schramm's coupling

In [38], Schramm introduced a remarkable coupling to study the cycle structure of CFDP. The coupling is realised between the part of the dynamic permutation that is supported on elements that lie in the largest component of the associated graph process and an independent PoiDir(1)-sample, both seen as partitions of the unit interval $[0,1]$ evolving via coagulative-fragmentative dynamics. The main idea of the coupling is to evolve both partitions according to the same dynamics, but to rearrange and *match* components of the partition that become close in size. A detailed account of Schramm's coupling is given in [13, Section 5], to which we refer the reader. Below we provide a short overview of the main features.

Appendix C.1 provides a short summary of Schramm's coupling and collects a number of estimates that are needed later on. Appendix C.2 proves a key proposition (Proposition C.10 below) stating that large cycles recur arbitrarily often in any time interval of diverging length.

C.1 Short summary of Schramm's coupling

Schramm's coupling is defined for a pair $\vec{Y}(t), \vec{Z}(t)$ of evolving partitions of the unit interval $[0,1]$ into countable many subintervals. The dynamics of the evolution is chosen such that the Poisson-Dirichlet distribution PoiDir(1) is invariant under the dynamics (see e.g. [42]). As we will see below, such a dynamics corresponds to the coagulative-fragmentative dynamics considered in this paper.

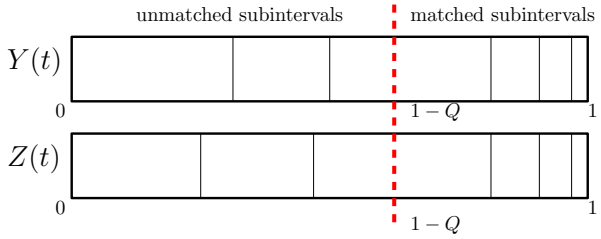

Figure 8: Illustration of the setup for Schramm's coupling defined in Definition C.1.

Definition C.1 [Abstract version of Schramm's coupling] Take two partitions $\vec{Y}(0), \vec{Z}(0)$ of the unit interval $[0,1]$ into countably many subintervals. A single step of the coupling proceeds as follows:

1. If there is a subinterval $a \in \vec{Y}(0)$ that has the same length as some other subinterval $b \in \vec{Z}(0)$, then declare a and b to be *matched*. (Note that the relation of being matched is symmetric.)

2. Reorder the subintervals within $\vec{Y}(0)$ and $\vec{Z}(0)$ as follows:

(a) Let Q be the total length of all the matched intervals. In both $\vec{Y}(0)$ and $\vec{Z}(0)$, place within $(1-Q,1]$ all the *matched* subintervals, ordered by their size such that the longest matched subinterval is on the left.

(b) In both $\vec{Y}(0)$ and $\vec{Z}(0)$, place within $[0,1-Q)$ the respective *unmatched* subintervals from the respective partitions, once again ordered by their size such that the longest matched subinterval is on the left.

3. Sample $U,U' \sim \mathsf{Unif}([0,1))$ and use these random variables to evolve the partitions $\vec{Y}(0),\vec{Z}(0)$ as follows:

(a) Call the subintervals $h_1 \in \vec{Y}(0)$ or $h_2 \in \vec{Z}(0)$ *highlighted* if U falls into these subintervals after the reordering described above.

(b) If U' falls into a subinterval $g_1 \in \vec{Y}(0)$ such that $g_1 \neq h_1$, then merge h_1 and g_1. Do the same for h_2 and $\vec{Z}(0)$.

(c) If U' falls into a highlighted subinterval $h_1 \in \vec{Y}(0)$, then split h_1 at U'. Do the same for h_2 and $\vec{Z}(0)$.

4. Call $\vec{Y}(1)$ and $\vec{Z}(1)$ the new partitions that are obtained by the reordering and the application of the dynamics, and repeat.

See Fig. 8 for an illustration. ♠

The construction specified in Definition C.1 has to be modified slightly to fit the setting of CFDP. First, we need to convert a permutation into a partition of the unit interval.

Definition C.2 [Cycle structure on a dynamic permutation] Denote by $|\gamma^{(i)}(t)|$ the size of the i^{th}-largest permutation cycle at time $t \geq 0$, with the following conventions: if there is only a finite number $k \in \mathbb{N}$ of permutation cycles, then $|\gamma^{(j)}(t)| = 0$ for all $j > k$, while if two or more elements have the same size, then their ordering (among each other) is arbitrary. Define

$$\vec{\mathfrak{X}}^{(n)}(t) = \left(\mathfrak{X}_i^{(n)}(t)\right)_{i\in\mathbb{N}} = \left(\frac{|\gamma^{(i)}(t)|}{|\mathscr{C}_{\max}(A_{\Pi_n}(t))|}\right)_{i\in\mathbb{N}}, \tag{C.1}$$

i.e., the ordered sequence of sizes of permutation cycles normalised by the size of $\mathscr{C}_{\max}(A_{\Pi_n}(t))$. Abusing notation, we denote by $\vec{\mathfrak{X}}^{(n)}(t)$ also the partition of the interval $[0, n/|\mathscr{C}_{\max}(A_{\Pi_n}(t)|)]$ into subintervals induced by $\vec{\mathfrak{X}}^{(n)}(t)$. ♠

Remark C.3 [Normalisation of $\vec{\mathfrak{X}}^{(n)}(t)$]. Note that, with this normalisation, the sum of all the elements in $\vec{\mathfrak{X}}^{(n)}(t)$ is $n/|\mathscr{C}_{\max}(A_{\Pi_n}(t))|$, which is roughly $\frac{1}{\zeta(t/n)}$, and the permutation cycles with elements on the largest component of the associated graph process (which is unique with high probability) correspond approximately to the subinterval $[0,1]$ of the partition induced by $\vec{\mathfrak{X}}^{(n)}(t)$. ■

All the details required to modify Schramm's coupling to the setting of CFDP are explained in [13, Sections 5.2 and 5.3], and we will not reproduce them here in full. The main modifications of the coupling, this time between $\vec{\mathfrak{X}}^{(n)}(t)$ seen as a partition of $[0, n/|\mathscr{C}_{\max}(A_{\Pi_n}(t))|]$ and a random partition of the interval $[0,1]$ distributed according to $\mathsf{PoiDir}(1)$, are as follows:

1. Allow for *approximate* matching of components, i.e., allows for a margin of order $O(n^{-1/2})$.

2. Markers U,U' that generate the dynamics are sampled uniformly from $[0, n/|\mathscr{C}_{\max}(A_{\Pi_n}(t))|]$, but if one of them falls outside the interval $[0,1]$, then the move is *not carried out* in the initially $\mathsf{PoiDir}(1)$-distributed partition.

3. A *forbidden* set $F(t)$ is the set of points that U,U' must avoid for the coupling to be successful. This set takes care of the possible errors that may arise due to, for example, the discrete nature of the permutations or the growing size of the largest component in the associated graph process.

Remark C.4 [Limiting distribution of the cycle structure $\vec{\mathfrak{X}}^{(n)}$].
In [38, Theorem 1.1], Schramm's coupling in the form adapted to CFDP was used to show that, at any time $t \geq cn$ with $c > \frac{1}{2}$, the restriction of $\vec{\mathfrak{X}}^{(n)}(t)$ to the interval $[0,1]$ has the distributional limit

$$\vec{\mathfrak{X}}^{(n)}_{|[0,1]}(t) \xrightarrow{d} \mathsf{PoiDir}(1), \tag{C.2}$$

where $\mathsf{PoiDir}(1)$ is the Poisson-Dirichlet distribution with parameter 1, which is the *unique* invariant distribution of $\vec{\mathfrak{X}}^{(n)}$ w.r.t. the permutation dynamics [38]. ■

Note that the largest cycle $\mathfrak{X}_1^{(n)}(cn)$, for $c > 1/2$, is typically large (see e.g. [25]). In the lemma stated below, which is adapted from [13, Section 5.3], we collect results that give us control over $\vec{\mathfrak{X}}^{(n)}(t)$. To understand the limitations in the statement of this lemma, we need to introduce some extra notation:

Definition C.5 [Events relating to the success of Schramm's coupling] Let $c > \frac{1}{2}$, fix $\beta > 0$ and set $T_\beta = \lceil \beta^{-1/2} \rceil - 1$ and $\hat{I}_{c,\beta} = [cn, cn + T_\beta]$. Consider Schramm's coupling

$$\big(\vec{\mathfrak{X}}^{(n)}(t), \vec{Z}(t)\big)_{t \in \hat{I}_{c,\beta}}$$

with $\vec{Z}(0) \sim \mathsf{PoiDir}(1)$. For $t \in \hat{I}_{c,\beta}$, define

- $N_\beta(\vec{\mathfrak{X}}^{(n)}(t))$ to be the number of unmatched blocks in $\vec{\mathfrak{X}}^{(n)}(t)$ whose sizes are larger than β. Analogously for $N_\beta(\vec{Z}(t))$.
- $\overline{N}_\beta(t) = N_\beta(\vec{\mathfrak{X}}^{(n)}(t)) + N_\beta(\vec{Z}(t))$.
- $\sigma(\beta, \vec{\mathfrak{X}}^{(n)}(t))$ to be the total length of the blocks of size smaller than β in the partition $\vec{\mathfrak{X}}^{(n)}(t)$. Analogously for $\vec{Z}(t)$.
- $\overline{\beta}(t) = \beta + \sigma(\beta, \vec{\mathfrak{X}}^{(n)}(t)) + \sigma(\beta, \vec{Z}(t))$.
- $\mathfrak{X}_1^{(n,\mathrm{UM})}(t)$ to be the largest unmatched segment in $\vec{\mathfrak{X}}^{(n)}(t)$.

Using the above notation, define the following events:

$$\mathcal{A}_1(t) = \{\overline{\beta}(t) \leq \beta^{3/4}\}, \qquad \mathcal{A}_2(t) = \{\overline{N}_\beta(t) \leq \beta^{-1/4}\}. \tag{C.3}$$

♠

The next definition captures the key regularity event used in the upcoming arguments.

Definition C.6 [Dynamics regularity event for Schramm's coupling] Define the event[1] $\mathcal{A}_3(\hat{I}_{c,\beta})$ by requiring that for any time $t \in \hat{I}_{c,\beta}$ the following three facts hold:

1. U, U' sampled at time t do not fall in the forbidden set $F(t)$.
2. The dynamics does not split a component of $\vec{\mathfrak{X}}^{(n)}(t)$ of size $\leq \sqrt{n}/|\mathscr{C}_{\max}(A_{\Pi_n}(t))|$.
3. If the dynamics induces a fragmentation in one of the partitions, then it does so also in the other partition.

♠

Since the event $\mathcal{A}_3(\hat{I}_{c,\beta})$ is crucial for the success of Schramm's coupling, we will need the following quantitative estimate (see also [13, eq. (5.7)]) to get uniform control over the time scale on which Schramm's coupling remains successful with high probability.

[1] See [13, p. 44], where this event is denoted by $\mathcal{G}$ and the three defining properties correspond to the three items before Eq. (5.7) therein.

Lemma C.7 [Probability bound for $\mathcal{A}_3(\hat{I}_{c,\beta})$] *Let $c > \frac{1}{2}$ and $\beta > 0$. Then*

$$\mathbb{P}\left(\mathcal{A}_3^{\mathrm{c}}(\hat{I}_{c,\beta})\right) \leq 16\beta^{-1}n^{-1/13}. \tag{C.4}$$

With this notation and information in hand, we can now state the following key lemma, adapted from [13, Section 5.3], which gives us control over the cycle structure $\vec{\mathfrak{X}}^{(n)}(t)$:

Lemma C.8 [$\mathsf{PoiDir}(1)$ approximation of $\vec{\mathfrak{X}}^{(n)}(t)$] *Fix $c > \frac{1}{2}$ and $\beta > 0$. Let $\vec{\mathfrak{X}}^{(n)}(t)$ be as in Definition* C.2. *Consider $(\vec{Z}(t))_{t\in\hat{I}_{c,\beta}}$ such that $\vec{Z}(cn) \sim \mathsf{PoiDir}(1)$ is sampled independently of anything else, and at later times the evolution of $\vec{Z}(t)$ is governed by the dynamics of the underlying permutation (see [13, Section 5]). Consider Schramm's coupling $(\vec{\mathfrak{X}}^{(n)}(t), \vec{Z}(t))_{t\in\hat{I}_{c,\beta}}$, and let $q \sim \mathsf{Unif}([cn, \dots, cn + T_\beta])$ be a uniform random variable in $\hat{I}_{c,\beta}$, independent of anything else. For $\delta \in (0,1)$, define the event*

$$\mathscr{D}_{\delta,\beta} = \left\{\left\|\vec{\mathfrak{X}}^{(n)}(q) - \vec{Z}(q)\right\|_\infty < \delta\right\}. \tag{C.5}$$

Recall the events in (2.16) *and in Definition* C.5. *There exist constants $C, C' > 0$ such that, for n sufficiently large, the following estimate is valid uniformly in β and δ:*

$$\begin{aligned}\mathbb{P}(\mathscr{D}_{\delta,\beta}^{\mathrm{c}}) &\leq \mathbb{P}\left(\Omega_n^{\mathcal{C}}(cn)\right) + \mathbb{P}\left(\Omega_n(cn)\cap\mathcal{A}_1^{\mathrm{c}}(cn)\right) + \mathbb{P}\left(\Omega_n(cn)\cap\mathcal{A}_2^{\mathrm{c}}(cn)\right) + \mathbb{P}\left(\Omega_n(cn)\cap\mathcal{A}_3^{\mathrm{c}}(\hat{I}_{c,\beta})\right) \qquad \text{(C.6)}\\ &\quad + \mathbb{P}\left(\left\{\mathfrak{X}_1^{(n,\mathrm{UM})}(q) \geq \delta/2\right\}\cap\Omega_n(cn)\cap\mathcal{A}_1(cn)\cap\mathcal{A}_2(cn)\cap\mathcal{A}_3(\hat{I}_{c,\beta})\right)\\ &\leq o(1) + 3C\beta^{1/4}\log\left(\beta^{-1}\right) + 2C'\beta^{1/4}\log^2\left(\beta^{-1}\right) + 16\beta^{-1}n^{-1/13} + \frac{2C_1}{\delta\log(\beta^{-3/4})}. \qquad \text{(C.7)}\end{aligned}$$

The proof of this lemma can be extracted from the various arguments and statements presented in [38, Section 3]. Alternatively, the precise bounds in (C.6) and (C.7) are explicitly derived in [13, Section 5]. In particular, the first inequality in (C.6) is obtained by applying a union bound in the second term of the first bound in the last display of the proof of [13, Theorem 1.1, p. 45]. The quantitative bounds on the non-trivial terms in (C.7) can be found in, respectively, the second to last display in the proof of [13, Theorem 1.1, p. 45] for the terms involving $\mathcal{A}_1^{\mathrm{c}}$ and $\mathcal{A}_2^{\mathrm{c}}$, Lemma C.7 for the term involving $\mathcal{A}_3^{\mathrm{c}}$, and [13, Corollary 5.7] for the last term related to the maximal unmatched block $\mathfrak{X}_1^{(n,\mathrm{UM})}$.

C.2 Recurrence of large cycles

We will make use of Schramm's coupling in the proof of Proposition C.10 below, which is at the very core of our argument in the proofs of Propositions 3.8 and 3.10. To this aim, we first use the quantitative estimate in Lemma C.8 to derive the following technical lemma, stating that the Poisson-Dirichlet approximation of the underlying permutation dynamics is good with high probability over a properly chosen time interval of diverging length.

Lemma C.9 [Pathwise approximation property of Schramm's coupling] *Fix $\varepsilon \in (0, \frac{1}{2})$ and $c > \frac{1}{2}$. Take any time interval of the form $I_{c,n} = [cn, cn + a_n]$ with $\lim_{n\to\infty} a_n = \infty$ and $a_n = o(n^{1/26})$, and denote by $\mathcal{F}^{\mathrm{SC}} = (\mathcal{F}_t^{\mathrm{SC}})_{t\in I_c}$ the natural filtration of Schramm's coupling. For $a, b \in I_{c,n}$, define the event that for all $t \in [a,b]$ the cycle structure of the underlying permutation $\vec{\mathfrak{X}}^{(n)}(t)$ is well-approximated by the process $\vec{Z}(t)$ introduced above. More precisely, define*

$$\Omega^{\mathrm{WA}}(a,b,\varepsilon) = \left\{\forall\, t\in[a,b]\colon\ \left\|\vec{\mathfrak{X}}^{(n)}(t) - \vec{Z}(t)\right\|_\infty < \varepsilon\right\} \in \mathcal{F}_b^{\mathrm{SC}}, \tag{C.8}$$

pick a sequence $(d_n)_{n\in\mathbb{N}}$ such that $d_n = o(\log a_n)$ and $d_n \to \infty$ as $n \to \infty$. Then

$$\mathbb{P}\left(\Omega^{\mathrm{WA}}(q_n, q_n + \sqrt{d_n}, \varepsilon)\right) = 1 - o(1) \tag{C.9}$$

for $q_n \sim \mathsf{Unif}(I_{c,n})$.

Proof. We will use the estimates in Appendix C.1. In particular, with the choice of a_n and d_n as in the statement of the lemma, we set $\delta = \delta_n = \varepsilon/d_n$, and $\beta = \beta_n$ such that $a_n = \lceil \beta_n^{-1/2} \rceil - 1$, and note that Lemma C.7 with $\hat{I}_{c,\beta} = I_{c,n}$ guarantees that

$$\mathbb{P}\left(\mathcal{A}_3(I_{c,n})\right) \geq 1 - o(1) \tag{C.10}$$

as soon as $a_n = o(n^{1/26})$.

The key event $\mathcal{A}_3(I_{c,n})$, occurring with high probability, guarantees regularity of the dynamics uniformly over the time interval $I_{c,n}$. On this event, we can find a (much smaller) random subinterval $[q_n, q_n + \sqrt{d_n}\,]$, still of diverging length, for which the Poisson-Dirichlet approximation up to the threshold ε is valid with high probability. Indeed, with the notation as in (C.5), we can bound

$$\begin{aligned}
\mathbb{P}\big(\Omega^{\mathrm{WA}}(q_n, q_n + \sqrt{d_n}, \varepsilon)\big) &= \mathbb{P}\Big(\forall\, t \in [q_n, q_n + \sqrt{d_n}]\colon\ \Big\|\vec{\mathfrak{X}}^{(n)}(t) - \vec{Z}(t)\Big\|_\infty < \varepsilon\Big) \\
&\geq \mathbb{P}\big(\mathscr{D}_{\varepsilon/d_n,\beta_n} \cap \mathcal{A}_3(I_{c,n})\big) = \mathbb{P}\big(\mathscr{D}_{\varepsilon/d_n,\beta_n}\big) - \mathbb{P}\big(\mathscr{D}_{\varepsilon/d_n,\beta_n} \cap \mathcal{A}_3^{\mathrm{c}}(I_{c,n})\big) \\
&\geq 1 - \mathbb{P}\big(\mathscr{D}^{\mathrm{c}}_{\varepsilon/d_n,\beta_n}\big) - o(1) \geq 1 - o(1),
\end{aligned}$$

where the first inequality is true because, on the event $\mathcal{A}_3(I_{c,n})$, if at time $q_n \in I_{c,n}$ we have $\|\vec{\mathfrak{X}}^{(n)}(q_n) - \vec{Z}(q_n)\|_\infty < \varepsilon/d_n$, then the next $\sqrt{d_n}$ (coagulation or fragmentation) moves of the dynamics can only increase the sup-norm in a bounded way. To see why, observe that on the event $\mathcal{A}_3(I_{c,n})$ only the following can occur:

- Due to Definition C.6[1.], we can leave out the effect of fragmentations of tiny components. Furthermore, Definition C.6[3.] guarantees consistency of moves between the two coupled cycle structures.
- Coagulation or fragmentation within the *matched* components can increase the sup-norm by at most $\frac{C}{\sqrt{n}}$, which is the margin allowed by the approximate matching rule. This discrepancy can be made arbitrarily small by taking n large enough. This follows from Definition C.6[1.] and the definition of a forbidden set $F(t)$.
- On the event $\mathscr{D}_{\varepsilon/d_n,\beta_n}$, the largest unmatched component satisfies $\mathfrak{X}_1^{(n,\mathrm{UM})}(q_n) \leq \frac{\varepsilon}{2d_n}$. This guarantees that, during the interval $[q_n, q_n + \sqrt{d_n}]$, the coagulation or fragmentation moves involving at least one *unmatched* component can, in the worst case, increase the sup-norm by at most $\sum_{k=1}^{\sqrt{d_n}} \frac{k\varepsilon}{2d_n} \leq \varepsilon$.

The other inequalities follow, respectively, from (C.10) and (C.7), with the choice of parameters as here. □

Using the pathwise approximation result in Lemma C.9, we can finally state the core proposition of this appendix (which was used to obtain (3.41)):

Proposition C.10 [Infinite crossings of level $1-\varepsilon$ for $\mathfrak{X}_1^{(n)}$] *Fix $\varepsilon \in (0, \frac{1}{2})$ and $c > \frac{1}{2}$. Take any time interval of the form $I_{c,n} = [cn, cn + a_n]$ with $\lim_{n\to\infty} a_n = \infty$ and $a_n = o(n^{1/26})$. Consider the event $\mathcal{E}_n(c, \varepsilon, \kappa)$ that the process $(\mathfrak{X}_1^{(n)}(t) - (1-\varepsilon))_{t=cn}^{cn+a_n}$ changes sign $\kappa \in \mathbb{N}$ times along an increasing sequence of times $(t_k)_{k=1}^{\kappa}$, i.e.,*

$$\mathcal{E}_n(c, \varepsilon, \kappa) = \big\{\exists\,(t_k)_{k=1}^{\kappa}\colon\ \mathfrak{X}_1^{(n)}(t_k - 1) < 1 - \varepsilon,\ \mathfrak{X}_1^{(n)}(t_k) \geq 1 - \varepsilon\,\big\}. \tag{C.11}$$

For any $\kappa \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n(c, \varepsilon, \kappa)) = 1. \tag{C.12}$$

Proof. The idea behind the proof is that if the coupling described above holds, then the behaviour of the PoiDir(1)-sample will induce the desired large cycles in the dynamic permutation.

Let $I_{c,n} = [cn, cn+a_n]$, and denote by $\mathcal{F}^{\mathrm{SC}} = (\mathcal{F}^{\mathrm{SC}}_t)_{t\in I_c}$ the natural filtration of Schramm's coupling. For any $a, b \in I_{c,n}$, $\varepsilon \in (0, \frac{1}{2})$ and $\kappa \in \mathbb{N}$, define the event that large cycles in the Z process *recur* more that κ times:

$$\Omega^{\mathrm{ZR}}(a, b, \varepsilon, \kappa) = \left\{\#\{t \in [a, b] \colon\ Z_1(t) > 1 - \varepsilon\} > \kappa\right\} \in \mathcal{F}^{\mathrm{SC}}_b. \tag{C.13}$$

The key observation is that, for any $0 < \lambda < \varepsilon$,

$$\left\{\#\{t \in [a, b] \colon\ \mathfrak{X}^{(n)}_1(t) > 1 - \varepsilon\} > \kappa\right\} \supseteq \Omega^{\mathrm{WA}}(a, b, \lambda) \cap \Omega^{\mathrm{ZR}}(a, b, \varepsilon + \lambda, \kappa). \tag{C.14}$$

This is true because if the supremum norm between $\mathfrak{X}^{(n)}(t)$ and $Z(t)$ is at all times bounded by λ, then every occurrence of a cycle larger than $1 - \varepsilon + \lambda$ in $Z(t)$ induces a cycle of size at least $1 - \varepsilon$ in $\mathfrak{X}^{(n)}(t)$.

From Lemma C.9, it follows that for $q_n \sim \mathsf{Unif}(I_{c,n})$ with $d_n \to \infty$ and $d_n = o(\log a_n)$, the well-approximation event $\Omega^{\mathrm{WA}}(q_n, q_n + \sqrt{d_n}, \varepsilon/2)$ occurs with high probability. Recall that, at the beginning of the coupling, $Z(cn) \sim \mathsf{PoiDir}(1)$ independently of everything else, and the dynamics of the coupling is such that $\mathsf{PoiDir}(1)$ is invariant. Denote by $\mathcal{P}_1$ the space of size-ordered countable partitions of the unit interval $[0, 1]$ into subintervals, which is the space over which the measure $\nu = \mathsf{PoiDir}(1)$ is defined. Introduce the notation

$$L_\varepsilon = \{P \in \mathcal{P}_1 \colon\ P_1 > 1 - \varepsilon\}, \tag{C.15}$$

where P_1 denotes the first, and therefore the largest, element of P. By [30, Theorem 1.2], for any $\varepsilon \in (0, \frac{1}{2})$,

$$\nu(L_\varepsilon) = -\log(1 - \varepsilon) > 0. \tag{C.16}$$

Recall from [38] that the evolution of Z is a time-homogeneous Markov process and that $\nu = \mathsf{PoiDir}(1)$ is the unique invariant measure of this process. Since, for any $\varepsilon \in (0, \frac{1}{2})$, the set L_ε has a strictly positive ν-measure and the starting point of the process is sampled according to ν, it follows that the *hitting* time of L_ε is finite a.s. and has finite expectation. Furthermore, the *return* times to L_ε are all finite a.s. and, by the *Kac recurrence time lemma* [34, Theorem 4.6], have finite expectation, equal to $1/\nu(L_\varepsilon)$. Consequently, for any fixed number of occurrences $\kappa \in \mathbb{N}$, any fixed threshold for cycle sizes $\varepsilon \in (0, \frac{1}{2})$, random times $q_n \sim \mathsf{Unif}(I_{c,n})$ independent of anything else, and any $d_n \to \infty$,

$$\lim_{n\to\infty} \mathbb{P}\big(\Omega^{\mathrm{ZR}}(q_n, q_n + d_n, \varepsilon, \kappa)\big) = 1. \tag{C.17}$$

Therefore, for $q_n \sim \mathsf{Unif}(I_{c,n})$, any $\kappa \in \mathbb{N}$ and any sequence $(d_n)_{n\in\mathbb{N}}$ such that $d_n = o(\log a_n)$ and $d_n \to \infty$ as $n \to \infty$,

$$\mathbb{P}\Big(\Omega^{\mathrm{WA}}(q_n, q_n + \sqrt{d_n}, \varepsilon) \cap \Omega^{\mathrm{ZR}}(q_n, q_n + \sqrt{d_n}, \tfrac{\varepsilon}{2}, \kappa)\Big) = 1 - o(1), \tag{C.18}$$

which by (C.14) yields the desired result. □

Remark C.11 [Extension of Proposition C.10 to the degree-two graphs with rewiring]. In [20] and [13, Section 5.2] it is noted that Schramm's coupling can be adapted to the setting of coagulation-fragmentation dynamics that keep the measure $\mathsf{PoiDir}(\theta)$, $\theta \in (0, 1]$, invariant. An example is a dynamic graph model with all degrees equal to two and endowed with a rewiring dynamics, which corresponds to a coagulation-fragmentation dynamics with invariant measure $\mathsf{PoiDir}(1/2)$ (see [31]). Since $\nu_\theta(L_\varepsilon) > 0$ for any $\theta, \varepsilon \in (0, 1)$ and $\nu_\theta = \mathsf{PoiDir}(\theta)$, the proof of Proposition C.10 can be adapted to the aforementioned situation. ■

In conclusion of this section, we state a corollary of the previous proposition, which is useful in Section 3:

Corollary C.12 *In the setting of Proposition* C.10*,*

$$\mathbb{P}\left(\mathcal{E}^{\mathrm{c}}_n(\tfrac{T^{\Downarrow}_{n,v_0}}{n}, \varepsilon, \kappa) \cap \{T^{\Downarrow}_{n,v_0} \geq (\tfrac{1}{2} + \delta)n\}\right) = o(1). \tag{C.19}$$

Proof. First, let us rewrite the desired expression as

$$\mathbb{P}\left(\mathcal{E}_n^{\mathrm{c}}(\tfrac{T^{\Downarrow}_{n,v_0}}{n},\varepsilon,\kappa)\cap\{T^{\Downarrow}_{n,v_0}\geq(\tfrac{1}{2}+\delta)n\}\right) \tag{C.20}$$
$$=\mathbb{E}\left[\mathbb{P}\left(\mathcal{E}_n^{\mathrm{c}}(\tfrac{T^{\Downarrow}_{n,v_0}}{n},\varepsilon,\kappa)\mid T^{\Downarrow}_{n,v_0}\right)\mathbb{1}_{\{T^{\Downarrow}_{n,v_0}\geq(\frac{1}{2}+\delta)n\}}\right].$$

Since the inner part of the right-hand side is bounded between 0 and 1, by dominated convergence theorem it is enough to show that

$$\mathbb{P}\left(\mathcal{E}_n^{\mathrm{c}}(\tfrac{T^{\Downarrow}_{n,v_0}}{n},\varepsilon,\kappa)\mid T^{\Downarrow}_{n,v_0}\right)\mathbb{1}_{\{T^{\Downarrow}_{n,v_0}\geq(\frac{1}{2}+\delta)n\}}=o_{\mathbb{P}}(1). \tag{C.21}$$

The core observation is that arguments based on Schramm coupling from Proposition C.10 work even under the conditioning in (C.21). From the perspective of the cycle structure, on the event $\{T^{\Downarrow}_{n,v_0}\geq(\frac{1}{2}+\delta)n\}$, at time $T^{\Downarrow}_{n,v_0}$ the giant component of the associated graph process is enlarged by a connected component whose size can be, with high probability, uniformly bounded by $C\log^2 n$, for some $C>0$. But the effect of these moves is already captured by the forbidden set construction (recall Definition C.6 and the discussion below Remark C.3), and a quantitative bound on the effect of these moves is expressed in Lemma C.7. Therefore, by repeating the arguments from Lemma C.9 and Proposition C.10, we obtain (C.21), which yields the desired expression.

□

D Mixing upon dropdown on the largest cycle

Recall that for two laws μ,ν defined on the same countable probability space

$$d_{\mathrm{TV}}(\mu,\nu)=\frac{1}{2}\sum_v|\mu_v-\nu_v|. \tag{D.1}$$

Fix $n\in\mathbb{N}$. Recall the short-hand notation $M=|\mathscr{C}_{\max}(A_{\Pi_n}(T^{\Downarrow}_{n,v_0}))|$. We wish to show that, on the events (recall (3.11))

$$\begin{aligned}\mathcal{M}_1(\varepsilon,\delta)&=\left\{|\operatorname{supp}(\mu^{n,v_0}(T^{\Downarrow}_{n,v_0}))|>\varepsilon M\right\}\cap\Omega^{(\mathrm{SC})}(T^{\Downarrow}_{n,v_0})\cap\Omega^{(\mathrm{ER})}_n(T^{\Downarrow}_{n,v_0}),\\ \mathcal{M}_2(\varepsilon)&=\left\{\exists\, t_L\in(T^{\Downarrow}_{n,v_0},T^{\Downarrow}_{n,v_0}+a_n)\colon\ \mathfrak{X}_1^{(n)}(t_L)>1-\varepsilon^2\right\},\end{aligned} \tag{D.2}$$

the following estimate (recall (3.15)) is valid:

$$d_{\mathrm{TV}}\Big(\mu^{n,v_0}(t_L),\mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L)))\Big)<\varepsilon. \tag{D.3}$$

On the event $\mathcal{M}_1(\varepsilon,\delta)$, the ISRW-distributiuon at time $T^{\Downarrow}_{n,v_0}$ is uniform over a single cycle supported on the largest component of the associated graph process (due to $\Omega^{(\mathrm{SC})}(T^{\Downarrow}_{n,v_0})$), which by assumption is larger than εM (due to $\{|\operatorname{supp}(\mu^{n,v_0}(T^{\Downarrow}_{n,v_0}))|>\varepsilon M\}$). Therefore,

$$\mu_v^{n,v_0}(T^{\Downarrow}_{n,v_0})\begin{cases}\leq\frac{1}{\varepsilon M}, & v\in\operatorname{supp}(\mu^{n,v_0}(T^{\Downarrow}_{n,v_0})),\\ =0, & \text{otherwise.}\end{cases} \tag{D.4}$$

We will use (D.4) to provide us with an upper bound on the mass carried by the elements on the giant component of the associated graph process.

On the event $\mathcal{M}_1(\varepsilon,\delta)\cap\mathcal{M}_2(\varepsilon)$, the largest cycle at time t_L necessarily carries some of the mass of the ISRW-distribution, since the initial size of the support is larger than the size of the subset not covered by the largest cycle (see Fig. 9 for a visual explanation).

We proceed by computing the worst-case l^1-distance between the two distributions at time t_L, which will yield the desired inequality in (D.3). Put $\Delta=|\mathscr{C}_{\max}(A_{\Pi_n}(t_L))|-M$, which represents the

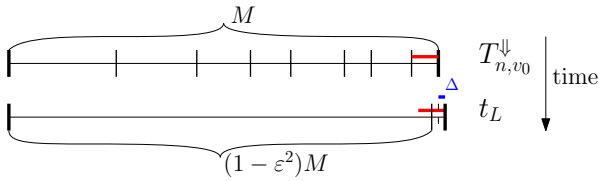

Figure 9: The first line segment represents the cycle structure of the permutation restricted to the giant component of the associated graph process at time $T^{\Downarrow}_{n,v_0}$. The red line represents the size of the single cycle that contains the full mass of the ISRW-distribution. The second line segment represents the same at time t_L: a cycle of size $(1-\varepsilon^2)M$ has appeared, which necessarily carries some of the mass of the ISRW-distribution.

growth of the largest component of the associated graph process compared to its size at time $T^{\Downarrow}_{n,v_0}$. Note that the uniform distribution in (D.3) gives the individual elements the following mass:

$$\mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L)))(v) = \begin{cases} \frac{1}{M+\Delta}, & v \in \mathscr{C}_{\max}(A_{\Pi_n}(t_L)), \\ 0, & \text{otherwise.} \end{cases} \tag{D.5}$$

To compute the worst-case l^1-distance at time t_L, we maximise the individual summands in (D.1). To do so, we work with a distribution $\mu^\dagger$ supported on the giant component of the associated graph process that puts mass $1/(\varepsilon M)$ (see (D.2)) on the elements outside the largest cycle of size $(1-\varepsilon^2)M$, and spreads the remaining mass $\Sigma = 1 - \frac{\varepsilon^2 M+\Delta}{\varepsilon M}$ over the largest cycle. In symbols,

$$\mu^\dagger_v = \begin{cases} \frac{\Sigma}{(1-\varepsilon^2)M}, & v \in \mathfrak{X}^{(n)}_1(t_L), \\ 0, & v \notin \mathscr{C}_{\max}(A_{\Pi_n}(t_L)), \\ \frac{1}{\varepsilon M}, & \text{otherwise.} \end{cases} \tag{D.6}$$

The measure $\mu^\dagger$ constructed in this way gives the bound

$$d_{\mathrm{TV}}\Big(\mu^{n,v_0}(t_L), \mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L)))\Big) \le d_{\mathrm{TV}}\Big(\mu^\dagger, \mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L)))\Big). \tag{D.7}$$

Finally, we compute, using (D.1) and (D.5)–(D.6),

$$\begin{aligned} d_{\mathrm{TV}}\Big(\mu^\dagger, \mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L)))\Big) &= \frac{1}{2}\left\|\mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L))) - \mu^\dagger\right\|_{l^1} \\ &= \frac{1}{2}(\varepsilon^2 M + \Delta)\left[\frac{1}{\varepsilon M} - \frac{1}{M+\Delta}\right] + \frac{1}{2}(1-\varepsilon^2)M\left[\frac{1}{M+\Delta} - \frac{\Sigma}{(1-\varepsilon^2)M}\right] \\ &= \varepsilon + \frac{\Delta}{\varepsilon M} - \frac{\varepsilon^2 M}{M+\Delta} - \frac{\Delta}{M+\Delta}. \end{aligned} \tag{D.8}$$

On the event $\Omega^{(\mathrm{ER})}_n(T^{\Downarrow}_{n,v_0})$, the sizes of all but the largest component are uniformly bounded by $Cn^{2/3}$ for some $C > 0$. Therefore, for a_n growing slowly enough, more concretely $a_n = o(n^{1/3})$, it follows that $\Delta = o(M) = o(n)$. Therefore

$$d_{\mathrm{TV}}\Big(\mu^{n,v_0}(t_L), \mathsf{Unif}(\mathscr{C}_{\max}(A_{\Pi_n}(t_L)))\Big) \le \varepsilon - \varepsilon^2 + o(1) < \varepsilon, \tag{D.9}$$

where the last inequality is true for n large enough.

E Mixing in dynamic degree-two graphs

E.1 Permutations and degree-two graphs

Let $\pi \in S_n$ be a permutation of $[n] = \{1, \dots, n\}$. Such a permutation admits a decomposition into distinct permutation *cycles*, and this decomposition can be used to create a mapping between

permutations and graphs whose vertex degrees are all equal to two. Note that this mapping is *not* a bijection between the set of all degree-two graphs on n vertices $\mathfrak{G}_n^{\deg 2}$ and the set of all permutations of n elements S_n. Such a bijection is not possible in general: while the connected components of degree-two graphs indeed are cycles, permutation cycles also carry information about the direction of traversal within them.

Let us nonetheless construct two mappings between these sets. There is a natural bijection between $[n]$ seen as the set of permutation *elements* and the same set seen as the set of graph *vertices*. Thanks to this bijection, we will (abusing notation) make no distinction between the two sets. Moreover, given a permutation $\pi \in S_n$, we can construct a degree-two graph with n vertices as follows:

1. Set e to be the least element of π, i.e., the element "1".
2. Add an edge between the vertices e and $\pi(e)$.
3. Set e to be $\pi(e)$, and repeat step 2 until the entire permutation cycle containing e is traversed.
4. Once the entire permutation cycle is traversed, set e to be the least non-traversed element of π and repeat from step 2 onwards. If there is no such element, then the entire permutation is traversed and the algorithm terminates.

Note that the mapping induced by this algorithm is a surjection, and hence the relation "being represented by the same degree-two graph" is an equivalence relation on S_n. Equivalence classes are formed by permutations whose cycle decomposition differs only in the reversal of the cyclic order between some of the cycles. For example, take $\pi_1, \pi_2 \in S_3$ such that

$$\pi_1 = (1, 2, 3), \qquad \pi_2 = (3, 2, 1). \tag{E.1}$$

By following the algorithm described above, we see that these two permutations are represented by the same degree-two graph.

When constructing the mapping from $\mathfrak{G}_n^{\deg 2}$ to S_n, the problem of cycle orientation manifests itself again. Since all the connected components of the degree-two graph are cycles, it is trivial to represent them as permutation cycles, but we are free to choose the direction of traversal of these connected components. These discrepancies are not relevant, since none of the mathematical objects used in the present paper depends on the direction of traversal of the permutation cycles. The most important property for our results is the size of the permutation cycles, or equivalently the size of the connected components.

E.2 ISRW mixing on degree-two graphs

To connect the result of the present paper to our previous work describing mixing of random walks on top of configuration models endowed with *rewiring* dynamics (see [4, 5, 6]), we note that rewiring of degree-two graphs and CFDP can easily be related to each other (recall Fig. 7). Our techniques and results can be adapted to the setting where the underlying geometry is modelled by a graph process starting from the configuration where *all the vertices have a self-loop*, equipped with rewiring dynamics. This process has been previously studied in [31].

To state these results, let us first define the process and the underlying geometry. The following definition are an adaptation of Definitions 1.4 and 1.6:

Definition E.1 [ISRW on a graph] Take a sequence of graphs $(G_n(t))_{t\in\mathbb{N}_0}$ and a vertex $v_0 \in [n]$. Denote by $\boldsymbol{\mathcal{CC}}(G_n(t), v)$ the connected component of the graph $G_n(t)$ that contains the vertex v. Formally, the infinite-speed random walk (ISRW) starting from v_0 is a sequence of probability distributions $(\mu^{n,v_0}(t))_{t\in\mathbb{N}_0}$ supported on $[n]$, with initial distribution at time $t = 0$ given by

$$\mu^{n,v_0}(0) = \left(\mu_w^{n,v_0}(0)\right)_{w\in[n]}, \tag{E.2}$$

where $\mu_w^{n,v_0}(0)$, the mass at $w \in [n]$ at time $t = 0$, is given by

$$\mu_w^{n,v_0}(0) = \begin{cases} \frac{1}{|\boldsymbol{\mathcal{CC}}(G_n(0),v)|}, & w \in \boldsymbol{\mathcal{CC}}(G_n(0), v), \\ 0, & w \notin \boldsymbol{\mathcal{CC}}(G_n(0), v), \end{cases} \tag{E.3}$$

and with distribution at later time $t \in \mathbb{N}$ given by

$$\mu^{n,v_0}(t) = (\mu_w^{n,v_0}(t))_{w\in[n]}\,, \tag{E.4}$$

where

$$\mu_w^{n,v_0}(t) = \frac{1}{|\mathcal{CC}(G_n(t),w)|}\sum_{u\in\mathcal{CC}(G_n(t),w)} \mu_u^{n,v_0}(t-1). \tag{E.5}$$

Informally, ISRW *spreads infinitely fast over the connected component it resides on.* ♠

Definition E.2 [Dynamic degree-two graph with rewiring dynamics] Fix the vertex set $\mathcal{V} = [n]$ and let $G_n(0)$ be the graph with all vertices having a single self-loop. At any later time $t \in \mathbb{N}$, $G_n(t)$ is obtained from $G_n(t-1)$ as follows:

1. Pick two edges e_1, e_2 uniformly at random without replacement from the set of edges within $G_n(t-1)$.
2. Break edge e_1 into half-edges h_1^A, h_1^B and edge e_2 into half-edges h_2^A, h_2^B (see [23, Section 7.2] for a definition of a half-edge).
3. Graph $G_n(t)$ has all the unbroken edges of $G_n(t-1)$, while the broken edges e_1, e_2 are replaced by a uniform choice from 2 possible sets of new edges: $\{(h_1^A, h_2^A), (h_1^B, h_2^B)\}, \{(h_1^A, h_2^B), (h_1^B, h_2^A)\}$.

We call the sequence $(G_n(t))_{t\in\mathbb{N}_0}$ the *dynamic degree-two graph with rewiring dynamics.*

In this setting, we state the following theorem, analogous to Theorem 1.20:

Theorem E.3 [Mixing profile for ISRW on dynamic degree-two graphs with rewiring]

(1) *Uniformly in* $v_0 \in [n]$,

$$\frac{T_{n,v_0}^{\Downarrow}}{n} \xrightarrow{d} u^{\Downarrow}, \tag{E.6}$$

where $u^{\Downarrow}$ *is the non-negative random variable with distribution (recall Definition* 1.11*(1))*

$$\mathbb{P}(u^{\Downarrow} \le u) = \zeta(u), \qquad u \in [0,\infty). \tag{E.7}$$

(2) *Uniformly in* $v_0 \in [n]$,

$$(\mathcal{D}_n^{v_0}(un))_{u\in[0,\infty)} \xrightarrow{d} \big(1-\zeta(u)\mathbb{1}_{\{u>u^{\Downarrow}\}}\big)_{u\in[0,\infty)} \quad \textit{in the Skorokhod } M_1\textit{-topology.} \tag{E.8}$$

Proof. Since Schramm's coupling and related arguments are fully applicable in this setting (see Remark C.11), the arguments from Section 3 can be used again with only slight modifications, as outlined below:

1. **Associated graph process and drop-down time:** The two constructions carry over without modifications. Even if a cycle-creating edge in the associated graph process produces a fragmentation only with probability 1/2 (since such an edge can with equal probability tear a segment of a cycle apart and create a new cycle or put the segment back into its original cycle in reversed order), this does not influence the arguments that underlie Lemma 3.2.
2. **Fast mixing upon drop-down:** Since the proof of Proposition 3.10 is based on Proposition C.10, which can be adapted to the alternative setting (see Remark C.11), we conclude that Proposition 3.10 also carries over.
3. **Drop-down in a single cycle:** Lemma 3.3 requires no adaptations.
4. **Mixing profile:** Since all the ingredients used in the proof of Lemma 3.11 and Theorem 1.20 carry over, the proofs themselves do likewise.

□

Remark E.4 [ISRW on a dynamic degree-two configuration model]. The setting of Theorem E.3 is *different* from a degree-two configuration model with rewiring, where the edges in the initial graph would be created by a random matching of half-edges. We conjecture that if the underlying geometry were modelled by a degree-two configuration model with rewiring, then ISRW would mix in $o_{\mathbb{P}}(n)$ steps and the mixing profile at scale n would be trivial. ■

References

[1] D. Aldous. A random tree model associated with random graphs. *Random Struct. Algorithms*, 1:383–402, 1990.

[2] D. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5:3–48, 1999.

[3] G. Alon and G. Kozma. The probability of long cycles in interchange processes. *Duke Math. J.*, 162(9):1567–1585, 2013.

[4] L. Avena, H. Güldaş, R. van der Hofstad, and F. den Hollander. Mixing times of random walks on dynamic configuration models. *Ann. Appl. Probab.*, 28:1977–2002, 2018.

[5] L. Avena, H. Güldaş, R. van der Hofstad, and F. den Hollander. Random walks on dynamic configuration models: A trichotomy. *Stoch. Proc. Appl.*, 129:3360–3375, 2019.

[6] L. Avena, H. Güldaş, R. van der Hofstad, F. den Hollander, and O. Nagy. Linking the mixing times of random walks on static and dynamic random graphs. *Stoch. Proc. Appl.*, 153:145–182, 2022.

[7] C. Avin, M. Koucký, and Z. Lotker. Cover time and mixing time of random walks on dynamic graphs. *Random Struct. Algorithms*, 52:576–596, 2018.

[8] A. Ben-Hamou. A threshold for cutoff in two-community random graphs. *Ann. Appl. Probab.*, 30:1824–1846, 2020.

[9] A. Ben-Hamou and J. Salez. Cutoff for nonbacktracking random walks on sparse random graphs. *Ann. Probab.*, 45:1752–1770, 2017.

[10] N. Berestycki and R. Durrett. A phase transition in the random transposition random walk. *Probab. Theory Relat. fields*, 136:203–233, 2006.

[11] N. Berestycki, E. Lubetzky, Y. Peres, and A. Sly. Random walks on the random graph. *Ann. Probab.*, 46:456–490, 2018.

[12] N. Berestycki, O. Schramm, and O. Zeitouni. Mixing times for random k-cycles and coalescence-fragmentation chains. *Ann. Probab.*, 39(5):1815–1843, 2011.

[13] J. E. Björnberg, M. Kotowski, B. Lees, and P. Miłoś. The interchange process with reversals on the complete graph. *Electron. J. Prob.*, 24:1–43, 2019.

[14] C. Bordenave, P. Caputo, and J. Salez. Random walk on sparse random digraphs. *Probab. Theory Relat. Fields*, 170:933–960, 2018.

[15] C. Bordenave, P. Caputo, and J. Salez. Cutoff at the "entropic time" for sparse Markov chains. *Probab. Theory Relat. Fields*, 173:261–292, 2019.

[16] P. Caputo and M. Quattropani. Mixing time trichotomy in regenerating dynamic digraphs. *Stoch. Proc. Appl.*, 137:222–251, 2019.

[17] P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Probab. Theory Relat. Fields*, 57:159–179, 1981.

[18] L. Federico. Almost-2-regular random graphs. *Australas. J. Comb.*, 86(1):76–96, 2023.

[19] A. Frieze and N. Karoński. *Introduction to Random Graphs*. Cambridge University Press, 2016.

[20] C. Goldschmidt, D. Ueltschi, and P. Windridge. Quantum Heisenberg models and their probabilistic representations. *Entropy and the quantum II, Contemp. Math*, 552:177–224, 2011.

[21] J. Hermon and J. Salez. The interchange process on high-dimensional products. *Ann. Appl. Probab.*, 31(1):84–98, 2021.

[22] J. Hermon and P. Sousi. A comparison principle for random walk on dynamical percolation. *Ann. Probab.*, 48:2952–2987, 2020.

[23] R. van der Hofstad. *Random Graphs and Complex Networks*, volume 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2016.

[24] R. van der Hofstad. *Random Graphs and Complex Networks*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2024.

[25] L. Holst. The Poisson-Dirichlet distribution and its relatives revisited. *Preprint of the Royal Institute of Technology, Stockholm, Sweden*, 2001.

[26] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 2011.

[27] S. Janson and G. Sorkin. Successive minimum spanning trees. *Random Struct. Algorithms*, 61:126–172, 2019.

[28] E. Lubetzky and A. Sly. Cutoff phenomena for random walks on random regular graphs. *Duke Math. J.*, 153:475–510, 2010.

[29] T. Łuczak and B. Pittel. Components of random forests. *Comb. Probab. Comput.*, 1:35–52, 1992.

[30] S. A. Molchanov and V. A. Panov. The Dickman–Goncharov distribution. *Russ. Math. Surv.*, 75(6):1089, 2020.

[31] S. Olesker-Taylor. Cutoff for rewiring dynamics on perfect matchings. *Ann. Appl. Probab.*, 33(1):641–676, 2023.

[32] Y. Peres, P. Sousi, and J. Steif. Mixing time for random walk on supercritical dynamical percolation. *Probab. Theory Relat. Fields*, 176:809–849, 2020.

[33] Y. Peres, A. Stauffer, and J. Steif. Random walks on dynamical percolation: mixing times, mean squared displacement and hitting times. *Probab. Theory Relat. Fields*, 162:487–530, 2015.

[34] K. E. Petersen. *Ergodic Theory*, volume 2. Cambridge University Press, 1989.

[35] O. Riordan and L. Warnke. The phase transition in bounded-size Achlioptas processes, 2017. arXiv:1704.08714.

[36] J. Salez. A sharp log-Sobolev inequality for the multislice. *Ann. H. Lebesgue*, 4:1143–1161, 2021.

[37] T. Sauerwald and L. Zanetti. Random walks on dynamic graphs: mixing times, hitting times, and return probabilities. In *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, volume 132, pages 93:1–93:15, 2019.

[38] O. Schramm. Compositions of random transpositions. *Israel J. Math.*, 147:221–243, 2005.

[39] P. Sousi and S. Olesker-Taylor. Cutoff for random walk on dynamical Erdős–Rényi graph. *Ann. Inst. H. Poincaré Probab. Statist.*, 56:2745–2773, 2020.

[40] J. Spencer and N. Wormald. Birth control for giants. *Combinatorica*, 27:587–628, 2007.

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

[42] N. V. Tsilevich. Stationary random partitions of positive integers. *Theory Probab.*, 44(1):60–74, 2000.

[43] W. Whitt. *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer Series in Operations Research and Financial Engineering. Springer New York, 2002.