Universiteit Leiden
The Netherlands

Opinion dynamics on random graphs

Capannoli, F.

CHAPTER 2

Meeting, coalescence and consensus time on random directed graphs

This chapter is based on the following paper:

Luca Avena, Federico Capannoli, Rajat S. Hazra, and Matteo Quattropani. Meeting, coalescence and consensus time on random directed graphs. *The Annals of Applied Probability*, 34(5):4940–4997, 2024

Abstract

We consider the so-called Directed Configuration Model (DCM), that is, a random directed graph with prescribed in- and out-degrees. In this random geometry, we study the meeting time of two random walks starting at stationarity, the coalescence time for a system of coalescent random walks, and the consensus time of the classical voter model. Indeed, it is known that the latter three quantities are related to each other under certain *mean field conditions* requiring fast enough mixing time and not too concentrated stationary distribution. Our main result shows that, for a typical large graph from the DCM ensemble, the meeting time is well-approximated by an exponential random variable for which we provide the first-order asymptotics of its expectation, showing that the latter is linear in the size of the graph, and its pre-constant depends on some explicit statistics of the degree sequence. As a byproduct, we explore the effect of the degree sequence in changing the meeting, coalescence and consensus time by discussing several classes of examples of interest also from an applied perspective. Our approach follows the classical idea of converting meeting into hitting times of a proper collapsed chain, which we control by the so-called First Visit Time Lemma. The main technical challenge is related to the fact that in such a directed setting the stationary distribution is random, and it depends on the whole realization of the graph. As a consequence, a good share of the proof focuses on showing that certain functions of the stationary distribution concentrate around their expectations, and on their characterization, via proper annealing arguments.

§2.1 Introduction

The voter model represents a classical interacting particle system on graphs, which has been used to mathematically model the formation of consensus across a given discrete geometry. In the classical model, each vertex is initially assigned either one of two opinions. Then, at exponential random times, a vertex randomly chooses one of its neighbors to adopt its opinion. Usually, one is interested in understanding the distribution of the so-called *consensus time*, i.e., the first time at which all the vertices share the same opinion. Voter models were introduced in the seminal works [CS73] and [HL75], and the analysis of the consensus times on finite graphs was first conducted in [DW83] and [Cox89]. In particular, in [Cox89] the author provides the exact first-order approximation of the consensus time on the torus of $\mathbb{Z}^d$ in the limit as the size of the graph tends to infinity. The list of examples in which such precise asymptotics can be provided is not very long. In [CFR10], the authors compute the exact constant in the case of a random regular graph. Clearly, in the latter case, the law of the voter dynamics depends on the specific realization of the graph. Nevertheless, it is possible to show that the expected consensus time (properly rescaled) converges in probability to a constant as the size of the graph goes to infinity. In the same spirit, in [HLYZ22] the authors show that for a configuration model with good expansion properties, the limiting constant can be interpreted in terms of an annealed observable of the random walk on the local weak limit of the random graph under consideration. However, such a constant remains implicit in their work. [FO23] studied the asymptotics of the consensus time on inhomogeneous random graph models, such as the Chung-Lu model and the Norros-Reitu model. In their work, the analysis provides the order of magnitude for the consensus time but not the precise preconstant. More generally, in [Dur07], the author presents several heuristic arguments showing that in many classical random graph models, the order of magnitude of the expected coalescence time is the same as that of the meeting time of two independent random walks (see also [Ald13]). The rationale underlying these ideas has been made rigorous by Oliveira in [Oli13]. Indeed, the author determines a set of *mean field conditions* on the underlying graph under which the first-order properties of the consensus time can be reduced to those of the meeting time of two independent stationary random walks. Under such conditions, the consensus time can be shown to converge to an explicit random variable in the Wasserstein-1 sense when the size of the graph grows to infinity. Similarly, in [CCC16], the authors show that a set of similar conditions is sufficient to show that the density of one of the two opinions converges to a Wright-Fisher diffusion in the Skorohod topology.

A good way to grasp the heuristic arguments behind the "mean field picture" is to think in terms of a system of *coalescent random walks*, which is the stochastic dual process of the voter model. Dual systems are commonly used to study various

interacting particle systems, and for more information, we refer the reader to the classical books [Lig85, Lig99]. As a consequence of such duality, it is possible to study the behavior of the consensus time by analyzing the coalescence time, which is the first time at which all walks coalesce into one. The precise connection between the voter model and coalescent random walks will be covered in detail in Section 2.2. The idea behind the study of the coalescence time on general geometries goes back to Kingman's coalescence [Kin82] on finite partitions of n elements and translates into the complete graph as a pure death process that jumps from k particles to $k-1$, with $k \in 2, \dots, n$, in an exponential time of rate $\binom{k}{2}$. This dynamics reflects the complete absence of geometry of the complete graph.

In [Oli13] Oliveira shows that there is a large class of geometries characterized by some minimal conditions that resemble their *mean field* nature. In this case, one can observe a very similar behavior to that of the complete graph. As a consequence, starting with $2 \leq k \leq n$ particles, if the random walk on the graph mixes fast and the stationary distribution is roughly uniform, then it is reasonable to expect that the first coalescence event is approximately exponentially distributed with mean $\mathbf{E}[\tau_{\text{meet}}]/\binom{k}{2}$, where τ_{meet} is the meeting time of two independent stationary random walks. By iterating this sort of argument, Oliveira is able to conclude that the coalescence time converges to an infinite sum of exponential random variables with the aforementioned expectations.

§2.1.1 Our contribution

The picture depicted in [Oli13] provides a clear recipe to compute the first-order asymptotic of the consensus time on large graphs: check that the *mean field conditions* are satisfied and then compute the first-order asymptotic of the meeting time of two independent walks. In this chapter, we follow such a path in the special case in which the underlying sequence of graphs are random and directed, sampled from the so-called Directed Configuration Model (DCM). The DCM is a natural generalization of the classical configuration model, in which the out- and in-degree of every vertex are prescribed as parameters of the model, and the graph is constructed through a uniform matching between the out- and in-stubs. Initially introduced in [CF04], in the last few years, a number of works have shed some light on the geometry of these random directed graphs and on the behavior of the random walk on them, see [BCS18, CP20, CP21, CQ20, CQ21a, CQ21b, CCPQ23]. It is crucial to realize that, unlike the classical *undirected* configuration model, the stationary distribution of the random walk on these graphs is a delicate random object, depending on the global realization of the graph, not only on its local features.

By invoking the above-mentioned results, we will show that the *mean field conditions* are satisfied by a typical realization of the directed graph. Doing so, we reduce

the analysis of the consensus time to that of the meeting time of two stationary random walks on the graph. A key technical tool for this purpose is the so-called First Visit Time Lemma (see [CFR10, CF08, MQS21]). The latter is a powerful instrument, particularly suited to show that, under certain assumptions very similar in spirit to those in [Oli13], the hitting time of a target vertex by a stationary random walk is (asymptotically) exponentially distributed. Moreover, the First Visit Time Lemma provides a computationally tractable expression for its expectation. In our setting, as we are interested in the meeting time, the underlying *graph* is the product of a directed graph from the DCM with itself, and the *target* is actually a set, i.e., the diagonal of the product graph. Following the approach of [CFR10, MQS21], in Section 2.5, we explain how to deal with the First Visit Time Lemma in such a setting.

On the technical side, the major effort of this work lies in computing the exact asymptotic preconstant of the meeting time. To carry out the approach mentioned in the paragraph above, beyond the classical tree approximation of sparse random graphs and the techniques introduced in the aforementioned papers, it is required the analysis of a new process, that we call *random walks with reset*. The process can be described as follows: two walks evolves independently up to their meeting, and when sitting on the same vertex they are "reset" somewhere else on the graph. However, such a reset distribution depends on the stationary distribution of the random walk, specifically on its square, π^2. As mentioned earlier, the stationary distribution is a complex random object that depends on the entire realization of the graph. In particular, resetting the walks according to π^2 imposes some *a priori* limitations on the use of *annealing* arguments. Roughly speaking, it is impossible to generate the *random walks with reset* together with the realization of the graph by letting the walks create the matching along their exploration. Indeed, when the two walks meet for the first time, in order to perform the reset, they need to know the *entire* realization of the graph. To overcome this difficulty, we study the random walks with reset in the case where the reset distribution, μ, is prescribed and does not depend on the graph. By doing so, we show that the quantities we are interested in only depend on a few features of the distribution μ. *A posteriori*, we demonstrate that such features concentrate for the distribution π^2 and, by means of some continuity argument, we can translate the case of a prescribed μ to the case $\mu = \pi^2$. A more detailed discussion on this and the other technical novelties is postponed to Sections 2.5, 2.6, and 2.7.

On the applied side, our study of the voter model on random graphs from the DCM ensemble is mostly motivated by the literature in network science and complex systems. In the last 20 years, physicists and computer scientists have produced an incredible amount of research about multi-agent systems on complex networks. A natural first question in such research programs is related to the effect of *first-order conditions* (i.e., the degrees of the network) on the speed at which information, such as news, infectious diseases, and opinions, travels along the network. These reasons

have made the study of the configuration models very popular in that community, and rigorous mathematical results have nowadays confirmed many of the physicists' predictions regarding these models. In this scenario, directed graph models are still less understood compared to their undirected counterparts. As mentioned above, the technical complications introduced as a byproduct of the edge orientations might be among the causes of this lack of results up until very recent times. On the other hand, it is clear that directed graphs constitute the natural model for those real-world networks in which directionality plays a prominent role, such as the World Wide Web or social networks like Twitter or Instagram.

Our main results show an explicit characterization of the expected consensus time as a function of a few simple statistics of the degree distributions. In other words, with our results at hand, it is possible to carry out a complete and rigorous analysis of the effect of the first-order properties on the speed of information diffusion in large directed networks, using the DCM as a natural benchmark. In Section 2.3.1, we invest some time in discussing the results of this analysis with several relevant examples. Finally, in Section 2.3.1, we provide the reader with an easily readable take-home message.

§2.1.2 Outline of the chapter

Before delving into the core of this chapter, we conclude this section with an outline of the chapter, presenting the entire structure and how the flow of arguments is articulated throughout the rest of the work.

(a) Section 2.2 is devoted to the presentation of the models of interest, the required notation, and the aforementioned *mean field conditions*. In particular, in Section 2.2.1, we introduce the random walk on a directed graph. Section 2.2.2 contains a formal description of the voter model, while the duality with coalescent random walks is introduced in Section 2.2.3. In Section 2.2.4, we formally introduce the *mean field conditions*, recalling the results in [Oli13]. Finally, in Section 2.2.5, we rigorously define the Directed Configuration Model and recall its main properties.

(b) In Section 2.3, we state our main technical contribution in Theorem 2.3.1 and its consequence on the consensus time in the subsequent Corollary 2.3.2. In Section 2.3.1, we utilize these results to analyze the consensus time as a function of the degree sequences.

(c) Section 2.4 describes the geometry of the DCM and the properties of the random walk on such random graphs, presenting the results from the literature that will be useful in the rest of the analysis.

(d) In Section 2.5, we provide a complete account of the proof strategy. In particular,

after recalling the First Visit Time Lemma, we argue that the main result in Theorem 2.3.1 boils down to three main claims: Propositions 2.5.5, 2.5.6, and 2.5.7. These propositions will be proved later in Sections 2.6, 2.7, and 2.8. A detailed discussion on the organization of the latter three sections is postponed to Section 2.5.2.

(e) Finally, Section 2.9 concludes the chapter with some open problems and possible future directions of research.

§2.2 Models and background

Before introducing formally the models of interest, we describe the general geometric setup that will considered in the whole chapter, introducing our notations for directed graphs and random walks on them.

We will be interested in finite directed multigraphs (from now on simply *graphs*) with $n \in \mathbb{N}$ labeled vertices, $G([n], E)$, where $[n] = \{1, \dots, n\}$ and E is a multiset with elements in $[n]^2$. We associate to G its adjacency matrix

$$A(x, y) \coloneqq |\{e \in E \mid e = (x, y)\}| = \#\text{ edges from } x \text{ to } y\,, \tag{2.1}$$

where, in this notation, we often refer to x as the *source* and to y as the *destination* of the edge $e = (x, y)$. For each $x \in [n]$, we let

$$d_x^+ = \sum_{y \in [n]} A(x, y)\,, \qquad d_x^- = \sum_{y \in [n]} A(y, x)\,,$$

denote the out- and the in-degree of x, respectively. Notice that we allow multiple edges with the same source and destination, as well as self-loops, i.e., edges in which the source and the destination coincide.

§2.2.1 Random walks on directed graphs

We let $(X_t)_{t \geq 0}$ denote the *continuous-time random walk* on G, which is the Markov process with state space $[n]$ and infinitesimal generator given by

$$L_{\mathrm{rw}} f(x) = \sum_{y \in [n]} \frac{A(x, y)}{d_x^+} \left[f(y) - f(x) \right], \qquad f : [n] \to \mathbb{R}\,. \tag{2.2}$$

Let $\mathcal{P}([n])$ denote the set of probability distributions on $[n]$. Fixed an initial distribution $\mu \in \mathcal{P}([n])$ for the random walk, we let $\mathbf{P}_\mu$ and $\mathbf{E}_\mu$ denote the law and the expectation on the space of trajectories of $(X_t)_{t \geq 0}$ with $X_0 \sim \mu$. Moreover, when μ is concentrated on a single vertex $x \in [n]$, we simply write $\mathbf{P}_x$ (respectively, $\mathbf{E}_x$).

In the following, we will assume G to be *ergodic*, that is, G admits a strongly connected component such that every vertex that is not in the component has at least a (directed) path leading to it. The latter requirement immediately implies that the random walk on G admits a unique stationary distribution, which we denote by π, such that

$$\lim_{t\to\infty} \mathbf{P}_\mu(X_t = x) = \pi(x)\,, \qquad \forall x \in [n]\,,\ \mu \in \mathcal{P}([n])\,. \tag{2.3}$$

Notice that, if G is ergodic but not strongly connected, then we have $\mathrm{supp}(\pi) \subsetneq [n]$.

With the aim of quantifying the speed of the convergence in (2.3), we define the worst-case total-variation distance at time t as

$$d_{\mathrm{TV}}(t) \coloneqq \max_{x\in[n]} \|\mathbf{P}_x(X_t = \cdot) - \pi\|_{\mathrm{TV}} = \frac{1}{2}\max_{x\in[n]} \sum_{y\in[n]} |\mathbf{P}_x(X_t = y) - \pi(y)|\,, \tag{2.4}$$

and we consider the *mixing time*

$$t_{\mathrm{mix}} \coloneqq \inf\left\{t \geq 0 \,:\, d_{\mathrm{TV}}(t) \leq \frac{1}{2e}\right\}. \tag{2.5}$$

In this work we will be particularly interested in considering a system of two independent random walks, that is, the continuous time Markov process $(X_t, Y_t)_{t\geq 0}$ on $[n]^2$ associated to the generator

$$L_{\mathrm{rw}}^{\otimes 2} = L_{\mathrm{rw}} \otimes \mathrm{Id} + \mathrm{Id} \otimes L_{\mathrm{rw}}\,.$$

We will consider the stopping time

$$\tau_{\mathrm{meet}} \coloneqq \inf\{t \geq 0 \mid X_t = Y_t\}\,, \tag{2.6}$$

called *meeting time*, representing the first time in which the two independent walks meet. Clearly, the law of τ_{meet} strongly depends on the initial distribution of the two walks.

Now that we set up the geometric framework and all the required preliminary notations, we are in shape to introduce the two models of interested for this work: the *voter model*, in Section 2.2.2, and the *coalescent random walks*, in Section 2.2.3.

§2.2.2 Voter model & consensus time

Let $G([n], E)$ be an ergodic graph and call the *voter model* on G the continuous-time Markov process $(\eta_t)_{t\geq 0}$ with state space $\{0,1\}^{[n]}$, and infinitesimal generator given by

$$L_{\mathrm{voter}} f(\eta) = \sum_{x\in[n]} \sum_{y\in[n]} \frac{A(x,y)}{d_x^+} \left[f(\eta^{x\to y}) - f(\eta)\right]\,, \qquad f : \{0,1\}^{[n]} \to \mathbb{R}\,, \tag{2.7}$$

where

$$\eta^{x\to y}(z) := \begin{cases} \eta(y), & \text{if } z = x\,, \\ \eta(z), & \text{otherwise}\,. \end{cases}$$

In words, the variable $\eta_t(x)$ represents the state of node x at time t, being either 0 or 1, to be interpreted as the binary *opinion* of the individual x at time t. The Markov evolution encoded in (2.7) can be phrased as follows. Each vertex $x \in [n]$ has an exponential clock of rate 1, when such a clock rings, vertex x chooses one of its out-edges at random and adopts the opinion of the vertex at the other extreme of the edge. See also Figure 2.1 to help visualization.

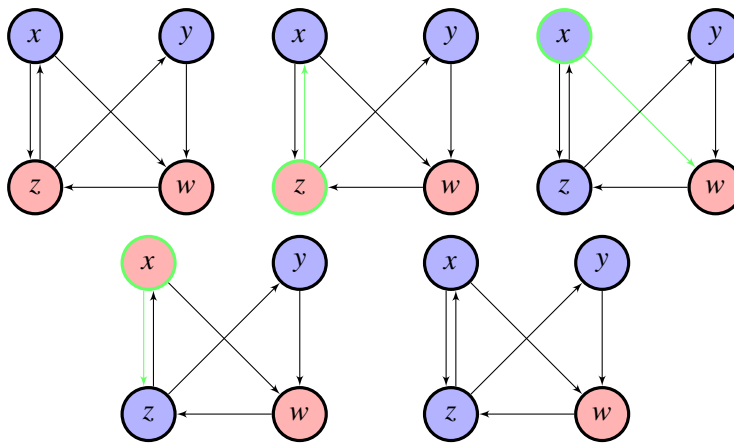

Figure 2.1: From left to right, up-to-down, the pictures describe a possible evolution of the voter model with generator as in (2.7) *on a directed graph with $n = 4$ vertices and initial opinions as in the first picture. In particular, at each step, the vertices with green boundary are the ones whose exponential clock rings first, while the corresponding green edges are the randomly selected ones among the out-neighbour on the green vertex.*

Under the ergodicity assumption, the voter model is a Markov chain with only two absorbing states, i.e., the monochromatic configurations $\bar{0}$ and $\bar{1}$ consisting of all 0's and 1's, respectively. As a consequence, regardless of the initial configuration of opinions, almost surely the system reaches in finite time one of these absorbing state. This naturally leads to the question of understanding the distribution of the *consensus time*, defined as

$$\tau_{\text{cons}} := \inf\left\{t \geq 0 :\ \eta_t \in \{\bar{1}, \bar{0}\}\right\}, \tag{2.8}$$

as a function of the underlying geometry.

§2.2.3 Coalescent random walks

As we will now argue, it is convenient to construct the trajectories of the process $(\eta_t)_{t\geq 0}$ by using a collection of independent Poisson processes indexed by the directed

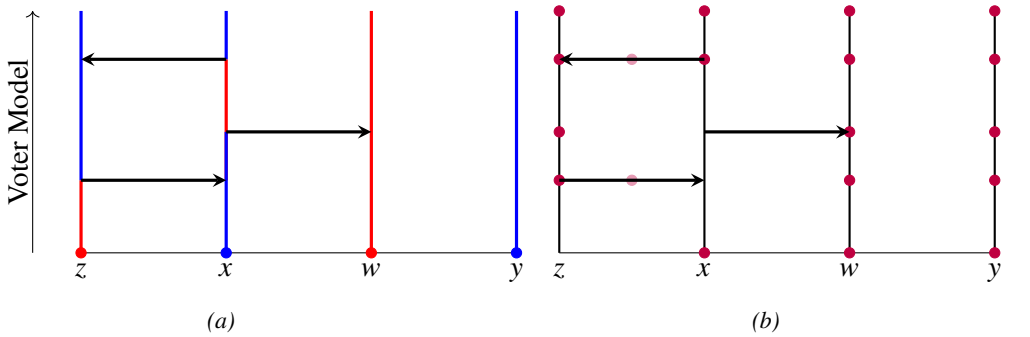

Figure 2.2: Picture (a) is the graphical representation of the voter model on the graph shown in Figure 2.1; while picture (b) represents its dual time reversal consisting of a system of coalescing random walks (CRWs). The two side arrows in the pictures represent the direction of time: in the voter model it runs upwards, while in the CRWs systems it runs downwards. In the voter model, picture (a), when the exponential clock associated to the edge $x \to y$ rings, we attach an arrow from x to y. We let opinions spread vertically until they reach the tail of a arrow. In that case the opinion changes according to the one sitting on the head of the arrow. In order to trace back in time the evolution of the opinions, at the initial time (corresponding to the final time for the voter model) we put independent random walks on each vertex and let them evolve back in time, following the same black arrows, from the tail to the head, with the added coalescing feature. As we can observe in (b), the two walks starting in x and z have coalesced into one particle that ended up in vertex x. This means that in the voter model the vertices x and z share the same opinion at the final time, as it can be checked in (a), and such opinion comes from the original opinion of x.

edges, $x \to y$, with corresponding intensities $1/d_x^+$. When the clock associated to an edge rings, say $x \to y$, the vertex at the origin of the edge, x, adopts the opinion of the vertex at the destination of the edge, y.

As we will now recall, this Poissonian construction allows to couple the dynamics of the voter model with that of a *dual process*, known as *coalescent random walks*. The latter is a continuous time stochastic process on $[n]^n$ which can be described as follows: the process starts with a random walk sitting on each vertex of the graph. Marginally, each random walk evolves according to the generator (2.2), with the only difference that, when two (or more) walks sit on the same vertex, then they stick together and continue their trajectory as a single walk. Clearly, under the ergodicity assumption, in finite time all the walks will coalesce on a single walk, which will then continue its trajectory according to the generator in (2.2). To simplify the reading, let $(X_t^x)_{t\geq 0}$ denote the trajectory of the walk starting at $x \in [n]$, and define the *coalescence time* as the first time at which all the walks are on the same site, i.e.,

$$\tau_{\text{coal}} := \inf\{t \geq 0 \,:\, X_t^x = X_t^y \,, \forall x, y \in [n]\} \,. \tag{2.9}$$

Now observe that the graphical representation used for the construction of the voter

model dynamics can be used also to sample the trajectory of the coalescent random walks: reversing the direction of time, when the clock associated to the edge $x \to y$ rings, then the walk(s) sitting at x (if any), move to y. In other words, the voter model, the random walk and the system of coalescing random walks can be sampled using the same graphical construction as a source of randomness. Moreover, the same conclusion still holds true for a system of two independent random walks, as soon as we restrict to $t \geq 0$ smaller than the stopping time τ_{meet} in (2.6). Therefore, there is no ambiguity in denoting all their laws and expectations with the same symbols, $\mathbf{P}$ and $\mathbf{E}$, respectively.

The beauty of the mentioned graphical construction can be realized by observing that it can be used to describe the ancestral history of the opinions on our graph. Indeed, for any collection $x_1, \dots, x_k \in [n]$ of distinct vertices and a prescribed time t, it is not hard to realize that

$$(\eta_t(x_1), \dots, \eta_t(x_k)) \stackrel{d}{=} (\eta_0\left(X_t^{x_1}\right), \dots, \eta_0(X_t^{x_k})\big) \,, \tag{2.10}$$

as exemplified in Figure 2.2. It is common to refer to the distributional relation in (2.10) by saying that the two models, voter and coalescent random walks, are *dual* of each other.

An immediate consequence of (2.10) is that the coalescence time in (2.9) is related to the consensus time in (2.8) by the stochastic domination

$$\tau_{\text{cons}} \preceq \tau_{\text{coal}} \,,$$

which holds regardless of the initial configuration η_0.

Having introduced the models, their properties, and all the required notations, we are now ready to discuss in some detail part of the literature about them. In particular, we will focus on presenting some condition on the underlying sequence of graphs which ensure that the asymptotic behaviour of the two processes is determined by some easier graph feature, i.e., the meeting time of two independent random walks started at stationarity.

§2.2.4 *Mean field conditions* for coalescence and consensus

As mentioned in the Introduction, the very first example in which a detailed analysis of the behaviour of the voter and the coalescence dynamics can be carried out is that of a complete graph, which is particularly simple due to the absence of geometry. In such a case, the distribution of τ_{coal} can be computed explicitly, see [AF02, Ch. 14] and [Oli13], obtaining

$$\frac{\tau_{\text{coal}}}{(n-1)/2} \stackrel{d}{=} \sum_{i=2}^{n} Z_i \,, \tag{2.11}$$

where the Z_i's are independent random variables with law

$$Z_i \stackrel{d}{=} \mathrm{Exp}\left(\binom{i}{2}\right), \qquad i \geq 2\,. \tag{2.12}$$

In particular, by taking the expectation one can conclude that $\mathbf{E}[\tau_{\mathrm{coal}}] = n-1$. To interpret the denominator in (2.11), notice that, recalling the definition of τ_{meet} in (2.6),

$$\mathbf{E}_{\pi\otimes\pi}[\tau_{\mathrm{meet}}] \sim \frac{1}{2}\,(n-1)\,,$$

where the notation $\pi \otimes \pi$ stands for the fact that we assume the two walks to start independently with law π, where in this case π is uniform over $[n]$. In other words, the distributional identity in (2.11) can be rephrased by saying that τ_{coal} converges in distribution to the sum of exponential random variables when scaled accordingly to $\mathbf{E}_{\pi\otimes\pi}[\tau_{\mathrm{meet}}]$.

A natural question is whether this picture is true in more general setups. In fact, one would expect that, if the random walk on the graph mixes quickly and the stationary distribution is not too concentrated, then the coalescence time behaves as in (2.11) after a proper rescaling. Indeed, the above scenario can be made rigorous in the case of d-dimensional tori (with $d \geq 3$), [Cox89], and for *most* regular graphs, [CFR10]. In this spirit, in [Oli13] Oliveira determines a set of *mean field conditions* on the underlying graph sequence ensuring the convergence in distribution of the rescaled coalescence time to the sum of random variables in (2.11). In such a general framework, the scaling factor will be given by $\mathbf{E}_{\pi\otimes\pi}[\tau_{\mathrm{meet}}]$.

Before stating Oliveira's results we point out that, to ease the reading, we specialize the statements to the setting of random walks on directed graphs introduced above, so that they look simpler (and less general) than in the original paper [Oli13]. We also remark that here and throughout the whole chapter, when considering distances between probability measures, we will adopt the standard abuse of notation of identifying random variables with their law.

Theorem 2.2.1 (Theorem 1.2 in [Oli13]). *Consider a sequence of ergodic graphs, $G = G_n([n], E)$, and let $\pi = \pi_n$ be the stationary distribution of the random walk on G. Let $\pi_{\max} := \max_{x\in[n]} \pi(x)$ and assume that*

$$\lim_{n\to\infty} t_{\mathrm{mix}}\, \pi_{\max}\, \log^5(n) = 0\,. \tag{2.13}$$

Then

$$\lim_{n\to\infty} d_W\left(\frac{\tau_{\mathrm{coal}}}{\mathbf{E}_{\pi\otimes\pi}[\tau_{\mathrm{meet}}]}\,,\, \sum_{k\geq 2} Z_k\right) = 0\,, \tag{2.14}$$

where d_W denotes the L^1 Wasserstein distance, and the Z_k's are independent random variables with law as in (2.12)*. In particular,*

$$\lim_{n\to\infty} \frac{\mathbf{E}[\tau_{\mathrm{coal}}]}{\mathbf{E}_{\pi\otimes\pi}[\tau_{\mathrm{meet}}]} = 2\,. \tag{2.15}$$

In words, the latter theorem tells us that as soon as we are able to verify (2.13), then the behavior of the coalescence time is the same as in the complete graph, and only the scaling factor, $\mathbf{E}_{\pi\otimes\pi}[\tau_{\text{meet}}]$, needs to be determined. A similar picture holds for the voter dynamics. Let $u \in (0,1)$ and consider again G to be the complete graph.

Indeed, it turns out that the asymptotic law of the consensus time can be expressed similarly to that of τ_{coal}, as pointed out by Oliveira in the following result.

Theorem 2.2.2 (Theorem 1.3 in [Oli13]). *Under the same assumption of Theorem 2.2.1, the consensus time of the voter model started with a product of i.i.d. Bernoulli opinions of parameter $u \in (0,1)$ satisfies*

$$\lim_{n\to\infty} d_W\left(\frac{\tau_{\text{cons}}}{\mathbf{E}_{\pi\otimes\pi}[\tau_{\text{meet}}]}, \sum_{k>K} Z_k\right) = 0\,, \tag{2.16}$$

where

$$K \overset{d}{=} UA + (1-U)B\,, \qquad U \overset{d}{=} \text{Bern}(u)\,, \qquad A \overset{d}{=} \text{Geom}(1-u)\,, \qquad B \overset{d}{=} \text{Geom}(u)\,, \tag{2.17}$$

and the Z_k's are independent random variables with law as in (2.12)*. In particular,*

$$\lim_{n\to\infty} \frac{\mathbf{E}_u[\tau_{\text{cons}}]}{\mathbf{E}_{\pi\otimes\pi}[\tau_{\text{meet}}]} = -2\left[(1-u)\log(1-u) + u\log(u)\right]\,. \tag{2.18}$$

Notice that the function of u on the right-hand side of (2.18) is symmetric around $\frac{1}{2}$ in the interval $[0,1]$ and it is maximal for $u = \frac{1}{2}$ where it attains the value $2\log(2) \approx 1.38$. Thanks to Theorems 2.2.1 and 2.2.2, one can conclude that an asymptotic analysis of the consensus time and of the coalescence time of can be derived, under the *mean field conditions*, by a precise asymptotic of the expected meeting time along the graph sequence.

§2.2.5 Directed configuration model (DCM)

In this section we formally introduce the random graph model that we will consider throughout all the chapter and state some of its typical properties. The Directed Configuration Model (DCM) is a natural generalization to the directed setting of the classical configuration model introduced by Bollobas [Bol80] in the early '80s . See also [Hof17] for a modern introduction to the topic. For each $n \in \mathbb{N}$ fix two finite sequences $\mathbf{d}^+ = \mathbf{d}_n^+ = (d_x^+)_{x\in[n]} \in \mathbb{N}^n$ and $\mathbf{d}^- = \mathbf{d}_n^- = (d_x^-)_{x\in[n]} \in \mathbb{N}_0^n$ such that

$$m = m_n := \sum_{x\in[n]} d_x^+ = \sum_{x\in[n]} d_x^-\,,$$

and let $d_{\min}^{\pm} = \min_{x\in[n]} d_x^{\pm}$ and $d_{\max}^{\pm} = \max_{x\in[n]} d_x^{\pm}$. We will work under the following assumptions on the degree sequences $(\mathbf{d}^+, \mathbf{d}^-)$:

Assumption 2.A (Degree assumptions). There exists some constant $C \geq 2$ such that for all $n \in \mathbb{N}$

$$\begin{aligned} &\text{(a):} \quad d^+_{\min} \geq 2\,, \\ &\text{(b):} \quad d^+_{\max} \leq C\,, \\ &\text{(c):} \quad d^-_{\max} \leq C\,. \end{aligned} \tag{2.19}$$

■

Remark 2.2.3. As we will discuss more extensively in Section 2.5.1, it is worth to notice that most of the proofs in this chapter work under weaker assumptions. We decided to state our results under the above degree assumptions so to rely on previous results, [CP20], on the minimum value of the stationary distribution π. However, we believe that the result in [CP20] could be extended to much weaker assumptions. Nevertheless, an extension of the result in [CP20] is out of the scope of this work.

Notice that assumptions (b) and (c) imply that our graphs are *sparse*, in the sense that $m \asymp n$, and that there are no restrictions on the minimal in-degree, therefore vertices with in-degree 0 are allowed. Assign to each vertex $x \in [n]$, d^-_x labeled *heads* and d^+_x labeled *tails*, denoting respectively the in- and out-stubs of x. Call E^-_x and E^+_x the sets of labeled heads and tails of x, respectively. Further, let $E^\pm = \cup_{x\in[n]} E^\pm_x$. Let $\omega = \omega_n$ be a uniformly random bijection $\omega : E^+ \to E^-$, viewed as a matching between tails and heads. The latter bijection can be projected to produce a directed graph $G = G_n([n], E)$, obtained by adding a directed edge $x \to y$ for every $f \in E^-_y$ and $e \in E^+_x$ such that $\omega(e) = f$. In what follows we let $\mathbb{P}$ (resp. $\mathbb{E}$) denote the probability law (resp. the expectation) of the sequence of random bijections $(\omega_n)_{n\in\mathbb{N}}$, and we will say that a sequence of graphs is sampled from $\mathrm{DCM}(\mathbf{d}^+, \mathbf{d}^-)$ to mean that for every n the graph $G = G_n$ is sampled according to the procedure above. We will be interested in studying the asymptotic regime in which $n \to \infty$, and we will say that G has a certain property *with high probability (w.h.p.)*, if the probability that G_n has such a property goes to 1 as n goes to infinity.

Being G random, so it is the law of the random walk on it. In particular, the stationary distribution π is a non-trivial random variable, and the same holds for $\mathbf{E}_{\pi\otimes\pi}[\tau_{\text{meet}}]$. Nevertheless, as we will formalize in Section 2.4.2, it is known that under Assumption 2.B (see [CP21, CCPQ23, BCS19]), w.h.p.,

- G is ergodic and $|\mathrm{supp}(\pi)| \asymp n$, hence there exists a unique stationary distribution, π, of random walk on G,
- the invariant distribution π is not too concentrated, i.e., $\pi_{\max} = \frac{\log^{O(1)}(n)}{n}$,
- the random walk on G mixes fast, i.e., $t_{\text{mix}} = \Theta(\log(n))$.

Hence, the mean field conditions in (2.13) are satisfied w.h.p.. To the aim of controlling the expected meeting time, it is worth to introduce the following quantities which will play a key role in our analysis. It is first convenient to define the probability distribution

$$\mu_{\mathrm{in}}(x) = \mu_{\mathrm{in},n}(x) \coloneqq \frac{d_x^-}{m}\,, \qquad x \in [n]\,, \tag{2.20}$$

that is, the law of a vertex sampled with probability proportional to its in-degree. We will further consider the following functions of the degree sequences

$$\begin{aligned} \delta = \delta_n &\coloneqq \frac{m}{n}\,, & \beta = \beta_n &\coloneqq \frac{1}{m}\sum_{x\in[n]} (d_x^-)^2\,, \\ \rho = \rho_n &\coloneqq \sum_{x\in[n]} \mu_{\mathrm{in}}(x)\frac{1}{d_x^+}\,, & \gamma = \gamma_n &\coloneqq \sum_{x\in[n]} \mu_{\mathrm{in}}(x)\,\frac{d_x^-}{d_x^+}\,, \end{aligned} \tag{2.21}$$

Notice that under Assumption 2.A we have that all the four quantities are of order $\Theta(1)$ and, moreover, satisfy the bounds

$$\rho \le \frac{1}{2}\,, \qquad \gamma \ge 1\,, \qquad \beta \ge 2\gamma\,, \qquad \delta \ge 2\,. \tag{2.22}$$

For the sake of intuition, let us provide an interpretation for the quantities in (2.21). In particular, it is immediate to interpret the first two parameters in a graph theoretical way: δ is simply the mean degree (equivalently, in- or out-) of the graph, while β is the ratio between the second and the first moment of the in-degree distribution. On the other hand, it will be convenient to interpret ρ and γ simply as expectations with respect to μ_{in}.

As we will show in the next section, the typical asymptotic behavior of the expected meeting time on a graph G sampled with the above procedure depends on the parameters of the model, $\mathbf{d} = (\mathbf{d}^+, \mathbf{d}^-)$, only through the quantities in (2.21).

§2.3 Main results

We are now ready to state our main results. Throughout this section $\mathbb{P}$ will represent the law of the sequence of graphs sampled, for all $n \in \mathbb{N}$, according to $\mathrm{DCM}(\mathbf{d}_n^+, \mathbf{d}_n^-)$, for some prescribed (sequence of) degree sequences $(\mathbf{d}_n^+, \mathbf{d}_n^-)_{n\ge1}$ satisfying Assumption 2.A. Provided that it exists and is unique, let π denote the unique stationary distribution of the random walk on G. To avoid degeneracy, in what follows we assume that, in case π is not well defined, then the same symbol π denotes uniform distribution on $[n]$.

In order to state the main result we need to introduce the following quantities which depend on the degree sequences through the functions defined in (2.21):

$$\mathfrak{p} = \mathfrak{p}_n(\mathbf{d}^+, \mathbf{d}^-) \coloneqq \frac{1}{\delta}\left(\frac{\gamma-\rho}{1-\rho} + \beta - 1\right) \ge 1\,, \tag{2.23}$$

$$\mathfrak{q} = \mathfrak{q}_n(\mathbf{d}^+, \mathbf{d}^-) := \frac{\gamma - \rho}{\gamma - \rho + (\beta - 1)(1 - \rho)} = \left(1 + (\beta - 1)\frac{1 - \rho}{\gamma - \rho}\right)^{-1} \leq 1\,, \quad (2.24)$$

and

$$\mathfrak{r} = \mathfrak{r}_n(\mathbf{d}^+, \mathbf{d}^-) := \frac{\rho}{\rho - \mathfrak{q}\left(1 - \sqrt{1 - \rho}\right)} \geq 1\,. \quad (2.25)$$

To enhance clarity, we will now offer a heuristic interpretation of the variables $\mathfrak{p}$, $\mathfrak{q}$, and $\mathfrak{r}$, deferring a more comprehensive explanation of their significance to Section 2.5.

As we will show in Section 2.7, the first two quantities can be interpreted as expectations in the probability space $\mathbb{P}$ of functions depending on the stationary distribution of the random walk on G. In particular, when $n \to \infty$,

$$\mathfrak{p} \approx \mathbb{E}\left[n \sum_{x \in [n]} \pi(x)^2\right], \qquad \mathfrak{q} \approx \mathbb{E}\left[\frac{n}{\mathfrak{p}} \sum_{x \in [n]} \pi^2(x) \frac{1}{d_x^+}\right]. \quad (2.26)$$

In words, $\mathfrak{p}$ is the (rescaled) expected sum of the entries of π^2, while $\mathfrak{q}$ is the expectation of the average inverse-out degree of a vertex sampled with probability proportional to π^2. As we will show below, the random variables within the parenthesis in (2.26) concentration around their expectation, that are asymptotically equal to $\mathfrak{p}$ and $\mathfrak{q}$, respectively. As for the quantity $\mathfrak{r}$, we will see in Section 2.7 that it can be seen as the inverse of the probability that two random walks on an certain infinite Galton-Watson tree never meet, assuming that one walk starts at the root, and the other one at one of its children.

The next theorem, which is the main technical contribution of the chapter, shows the weak convergence in probability of the meeting time to an exponential random variable. As a consequence, the random variable $\mathbf{E}_{\pi\otimes\pi}[\tau_{\text{meet}}]/n$ converges in probability to a constant, depending only on the degree sequences, which we explicitly characterize.

Theorem 2.3.1 (Meeting times on the DCM). *Let* $(\mathbf{d}^+, \mathbf{d}^-)$ *satisfy Assumption 2.A and* G *be sampled from DCM*$(\mathbf{d}^+, \mathbf{d}^-)$*. Then, letting* $\tau_{\text{meet}}^{\pi\otimes\pi}$ *denote the first meeting time of two independent stationary random walks, it holds*

$$d_W\left(\frac{\tau_{\text{meet}}^{\pi\otimes\pi}}{\frac{1}{2}\,\vartheta \times n}, \text{Exp}(1)\right) \xrightarrow{\mathbb{P}} 0\,, \quad (2.27)$$

with

$$\vartheta = \vartheta_n(\mathbf{d}^+, \mathbf{d}^-) := \frac{\mathfrak{r}}{\mathfrak{p}}\,, \quad (2.28)$$

where $\mathfrak{p}$ *and* $\mathfrak{r}$ *are defined as in* (2.23)*, and* (2.25)*, respectively.*

Before discussing in details the above result, we first state its immediate consequences for the voter model and coalescent random walks, based on the discussion of the *mean field conditions* provided in Subsection 2.2.4.

Corollary 2.3.2 (Coalescence and consensus time on the DCM). *Let* $(\mathbf{d}^+, \mathbf{d}^-)$ *satisfy Assumption 2.A and* G *be sampled from DCM*$(\mathbf{d}^+, \mathbf{d}^-)$. *Then,*

(i) Recalling the definition of τ_{coal} *in* (2.9)*,*

$$d_W\left(\frac{\tau_{\text{coal}}}{\frac{1}{2}\,\vartheta \times n}, \sum_{k\geq 2} Z_k\right) \xrightarrow{\mathbb{P}} 0\,. \tag{2.29}$$

where the Z_k*'s are independent random variables with law as in* (2.12)*.*

(ii) Consider the voter model on G *with* $\eta_0 = \bigotimes_{x\in[n]} \text{Bern}(u)$ *and* $u \in (0,1)$*. Using the definitions in Theorem 2.2.2,*

$$d_W\left(\frac{\tau_{\text{cons}}}{\frac{1}{2}\,\vartheta \times n}, \sum_{k\geq K} Z_k\right) \xrightarrow{\mathbb{P}} 0\,, \tag{2.30}$$

where K *is independent from the collection* $(Z_k)_{k\geq 2}$ *and is defined as in* (2.17)*.*

Remark 2.3.3. It is worth to remark that the result in Theorem 2.2.2 can be extend to an arbitrary number of opinions, as soon as the initial distribution can be expressed a product of i.i.d. random variables. As a consequence, the convergence in Corollary 2.3.2(ii) extend as well. Although, we prefer to present our results on the simpler case of the two-opinion model, in order to not deviate focus from the general structure and the novelties of our results.

Remark 2.3.4. It is known that, beyond the distribution of the coalescence and the consensus time, the precise knowledge of the expected meeting time can be used to determine the scaling limit of certain real valued processes. In particular, in [CCC16] the authors show that rescaling time by $\mathbf{E}_{\pi\otimes\pi}[\tau_{\text{meet}}]$, the weighted average of vertices having opinion 1 converges to the celebrated Wright-Fisher diffusion (see, e.g., [Lig85]) in the Skorohod topology. Moreover, in [Che13, Che18, Che23, CC18], similar results are obtained for some modification of the voter model related to the theory of evolutionary games.

§2.3.1 Discussion and examples

In Theorem 2.3.1 and Corollary 2.3.2 we provide a complete characterization of the distribution of meeting, coalescence and consensus time on a typical random graph as a function of a single quantity, ϑ.

Let us now discuss the property of the function ϑ in (2.28), and its dependence on the degree sequences. Being ϑ a function of only four parameters (those in (2.21)), it is possible to understand how changes in the degree sequences affect ϑ, and thus the distribution of our stopping times. For the sake of comparison, we can for instance fix δ, namely, the total number of edges, and see how do the other parameters in (2.21) influence the function ϑ. In particular, we are interested in answering the following kind of questions, which are of evident interests for applied network science:

- If the out-degrees (resp. in-degrees) are constant, increasing the variability of the in-degrees (resp. out-degrees) implies a speed-up or a slow-down in the consensus time? Which is of the range ϑ in these cases?
- What is the effect of positive/negative correlation between in- and out-degrees of the vertices on the consensus time? E.g., is consensus reached faster on an Eulerian digraph, or in one in which d_x^-/d_x^+ is typically far from 1?
- In the Eulerian setup, does the variability of the degrees speed-up the consensus time?
- If some features of the degree sequences are constrained and one is free to choose the exact degree sequence under such constraints, which are the guiding principles to minimize/maximize the consensus time?

In the rest of this section we will answer such questions and provide some simplified formulas for ϑ in some special cases.

The regular case.

The easiest model is the one in which every vertex as in- and out-degree equal to some constant $d \geq 2$. First notice that in the d-regular case the following simplifications of the parameters in (2.21) hold

$$\rho = \frac{1}{d}, \quad \delta = d, \quad \gamma = 1, \quad \beta = d\,,$$

so that

$$\mathfrak{p} = 1 \quad \text{and} \quad \mathfrak{q} = \frac{1}{d}\,. \tag{2.31}$$

Hence, it is immediate to deduce the following result.

Corollary 2.3.5. *Fix $d \geq 2$ and let $\vartheta(d)$ denote the quantity in* (2.28) *when $d_x^+ = d_x^- = d$ for all $x \in [n]$. Then*

$$\vartheta(d) = \sqrt{\frac{d}{d-1}} \in \left(1\,, \sqrt{2}\,\right]. \tag{2.32}$$

	ρ	γ	β	$\mathfrak{p}$	$\mathfrak{p}_{\text{emp}}$	$\mathfrak{q}$	$\mathfrak{q}_{\text{emp}}$	ϑ
Model 1	$\frac{1}{3}$	1	3	1	1	0.3333	0.3333	1.2247
Model 2	$\frac{1}{3}$	$\frac{10}{9}$	$\frac{10}{3}$	1.1667	1.1678±0.008	0.3333	0.3333	1.0500
Model 3	$\frac{3}{8}$	$\frac{9}{8}$	3	1.067	1.067±0.0050	0.3750	0.3753±0.0021	1.1859
Model 4	$\frac{1}{3}$	1	$\frac{10}{3}$	1.1111	1.1111	0.3000	0.3000	1.0780
Model 5	$\frac{5}{12}$	$\frac{3}{2}$	$\frac{10}{3}$	1.3968	1.4010±0.0331	0.4432	0.4436±0.0028	0.9562

Figure 2.3: We fix $n = 1000$ and consider a sample of G_n with degree distributions of the type $(\mathbf{d}^+, \mathbf{d}^-)$ where $d_x^- = a$ and $d_x^+ = b$ for all $x \leq 500$ and $d_x^- = c$ and $d_x^+ = d$ for all $x \geq 500$. To simplify the reading, we write [a,b]-[c,d] to refer to such models. For the sake of comparison, for all the models, we let $a + c = b + d = 6$, i.e., in all the cases there are $3n$ edges in total, hence $\delta = 3$. More precisely:
Model 1: [3,3]-[3,3], i.e., the regular case.
Model 2: [3,4]-[3,2], i.e., the out-regular case.
Model 3: [4,3]-[2,3], i.e., the in-regular case.
Model 4: [2,2]-[4,4], i.e., the Eulerian case.
Model 5: [2,4]-[4,2], i.e., the "alternate" model discussed in Subsection 2.3.1 .
The table reports the value of the quantities β, ρ and γ in (2.21)*, as well as the values of $\mathfrak{p}$, $\mathfrak{q}$ and ϑ (as defined in* (2.23), (2.24) *and* (2.28)*, respectively) associated to the 5 degree sequences. As mentioned in Section 2.3, the quantities $\mathfrak{q}$ and $\mathfrak{p}$ admit an interpretation in terms of expectation of the random variables (with respect to the generation of the graph) appearing in* (2.26)*. Therefore, we report here also the average of such random variables with respect to 100 random generations of the graph. In these cases, the ± error stands for the empirical standard deviation. To simplify the comparison, we express both the theoretical and the empirical value of $\mathfrak{p}$ and $\mathfrak{q}$ to the first 4 digits.*

Remark 2.3.6. Notice that, in the d-regular *undirected* configuration model, one has (see, e.g., [Che21, ABH+24])

$$\frac{\mathbf{E}_{\pi\otimes\pi}[\tau_{\text{meet}}]}{\frac{1}{2}\,\frac{d-1}{d-2}\times n} \xrightarrow{\mathbb{P}} 1\,, \qquad d\geq 3\,. \tag{2.33}$$

Hence, it is possible to compare the expected meeting time for the undirected d-regular case with that of the directed one. As an effect of the directionality, the random walks result to meet faster. Indeed,

$$\sqrt{\frac{d}{d-1}} < \frac{d-1}{d-2}\,, \qquad d\geq 3\,. \tag{2.34}$$

One may argue that the above comparison is improper, since the total number of neighbors of a vertex in a random d-regular directed graph is actually $2d$ (d in-neighbors and d-out-neighbors). Nevertheless, it is also true that

$$\sqrt{\frac{d}{d-1}} < \frac{2d-1}{2d-2}\,, \qquad d\geq 2\,. \tag{2.35}$$

Therefore, the speed-up experienced in the directed setting takes place whatever term of comparison we choose.

The out-regular case

Another model that is natural to investigate is the one in which all the vertices share the same out-degree, d. In this case the following holds.

Corollary 2.3.7. *Let $d\geq 2$ and $(\mathbf{d}^+,\mathbf{d}^-)$ satisfying Assumption 2.A and such that, for all $n\in\mathbb{N}$, $d_x^+ = d$ for all x. Call $\vartheta(d,\mathbf{d}^-)$ the quantity in* (2.28) *in this case. Then*

$$\vartheta(d,\mathbf{d}^-) = \frac{\sqrt{d(d-1)}}{\beta-1} \in \left(0\,,\sqrt{\frac{d}{d-1}}\right]. \tag{2.36}$$

Proof. In the out-regular case we have

$$\delta = \frac{1}{\rho} = d\,, \qquad \beta = d\gamma\,. \tag{2.37}$$

After some algebraic manipulation we get

$$\mathfrak{p} = \frac{\beta-1}{d-1}\,, \qquad \mathfrak{q} = \frac{1}{d}\,, \qquad \mathfrak{r} = \sqrt{\frac{d}{d-1}}\,, \tag{2.38}$$

from which the result follows. □

It is worth noting that, as soon as $d\geq 3$ it is possible to choose $\mathbf{d}^-$ so to have β arbitrarily large. This implies that, for a typical random out-regular directed graph, the larger is the variance in the in-degree distribution, the smaller is the value of the expected meeting time. In other words, among the out-regular random digraphs, the regular is the one in which the meeting time is the largest.

The in-regular case

We now show that, for a typical in-regular random directed graph the situation looks much more similar to the regular case than to the out-regular one.

Corollary 2.3.8. *Let* $(\mathbf{d}^+, \mathbf{d}^-)$ *satisfying Assumption 2.A and such that, for all* $n \in \mathbb{N}$, $d_x^- = d$ *for all* x. *Call* $\vartheta(\mathbf{d}^+, d)$ *the quantity in* (2.28) *in this case. Then, fixed* $d \geq 3$, *it holds*

$$\vartheta(\mathbf{d}^+, d) = \frac{d\sqrt{1-\rho}}{d-1} \in \left(\frac{1}{\sqrt{2}}\frac{d}{d-1}, \sqrt{\frac{d}{d-1}}\right], \tag{2.39}$$

while

$$\vartheta(\mathbf{d}^+, 2) = \sqrt{2}\,. \tag{2.40}$$

Proof. Since $\rho \geq \frac{1}{d}$ by Jensen inequality, we have

$$\vartheta(\mathbf{d}^+, d) \leq \sqrt{\frac{d}{d-1}}.$$

On the other hand, for $d \geq 3$,

$$\vartheta(\mathbf{d}^+, d) > \frac{1}{\sqrt{2}}\frac{d}{d-1}$$

since $\rho \leq \frac{1}{2}$. Moreover, such a lower bound is sharp. To see this, fix $C = d^+_{\max}$ and consider the degree sequence in which the first $(1-\varepsilon)n$ vertices have $d^+ = 2$, and the last εn vertices have degree C. Then, ε is fixed by the equation

$$2(1-\varepsilon) + \varepsilon C = d$$

from which it follows that

$$\varepsilon = \frac{d-2}{C-2}. \tag{2.41}$$

Then

$$\rho = (1-\varepsilon)\frac{1}{2} + \frac{\varepsilon}{C} = \frac{C^2 - dC + 2(d-2)}{2C^2 - 4C}. \tag{2.42}$$

In conclusion, for any $d \geq 3$ and $\eta > 0$ one can get C large enough so to have $\rho > \frac{1}{2} - \eta$. □

Roughly speaking, the previous result shows that, for a fixed value of the average degree, the meeting time on a typical out-regular graphs (regardless of the choice of the in-degrees) is smaller than in a regular graph with the same number of edges, regardless of the out-degrees. Nevertheless, contrarily to the out-regular case, the enhancement obtained by taking the out-degrees non-regular is bounded and the value of the constant ϑ cannot go below the threshold $1/\sqrt{2}$.

The Eulerian case

Random Eulerian graphs are a particularly relevant subclass of the DCM, corresponding to the case in which the in-degree and the out-degree of each vertex coincide. Indeed, it is the only class of directed graphs with the following property: as soon as the graph is strongly connected, π is proportional to the degree of the vertices, exactly as in an undirected graph. For this reason, it is the model that it is easier to compare to the classical undirected Configuration Model. In this setting, it is worth to realize that the quantity β/δ represents the ratio between the second moment and the squared first moment of the degree sequence, and it is therefore a signature of variability of the degrees. As the next corollary shows, the value of ϑ in this case depends essentially just on such a ratio, and higher variability implies a faster meeting time.

Corollary 2.3.9. *Let* $(\mathbf{d}^+, \mathbf{d}^-)$ *be such that, for all* $n \in \mathbb{N}$, $d_x^+ = d_x^-$ *for all* x. *Call* $\vartheta(\mathbf{d})$ *the quantity in* (2.28) *in this case. Then*

$$\vartheta(\mathbf{d}) = \left(\frac{\beta}{\delta} - 1 + \sqrt{1 - \frac{1}{\delta}}\right)^{-1} \in \left(0, \sqrt{\frac{\delta}{\delta - 1}}\right]. \tag{2.43}$$

Proof. In this case we have

$$\rho = \frac{1}{\delta}, \qquad \gamma = 1, \tag{2.44}$$

from which we get

$$\mathfrak{p} = \frac{\beta}{\delta}, \qquad \mathfrak{q} = \frac{1}{\beta}, \qquad \mathfrak{r} = \frac{\beta/\delta}{\beta/\delta - 1 + \sqrt{1 - 1/\delta}}, \tag{2.45}$$

thus the validity of (2.43). □

It is worth to remark that, also in this case, the random regular graph is the one with the slowest meeting time among all the random Eulerian graphs with the same number of edges.

The "alternate" case

A first toy model to check for the effect of correlations between in- and out-degrees is what we call *alternate* model, that is, the vertices are divided in two even groups: in one group the degrees are $d^+ = a$ and $d^- = b$, and the vice versa for the other group. Indeed, in this model it is natural to see that the farther a/b is from 1, the more anti-correlated are the degrees.

Corollary 2.3.10. *Let* $b \geq a \geq 2$ *and* $(\mathbf{d}^+, \mathbf{d}^-)$ *be such that for all* $n \in 2\mathbb{N}$ $d_x^+ = a$ *and* $d_x^- = b$ *for all* $x \in [n/2]$, *while* $d_x^+ = b$ *and* $d_x^- = a$ *for* $x \in [n] \setminus [n/2]$. *Let* $d = \frac{a+b}{2}$

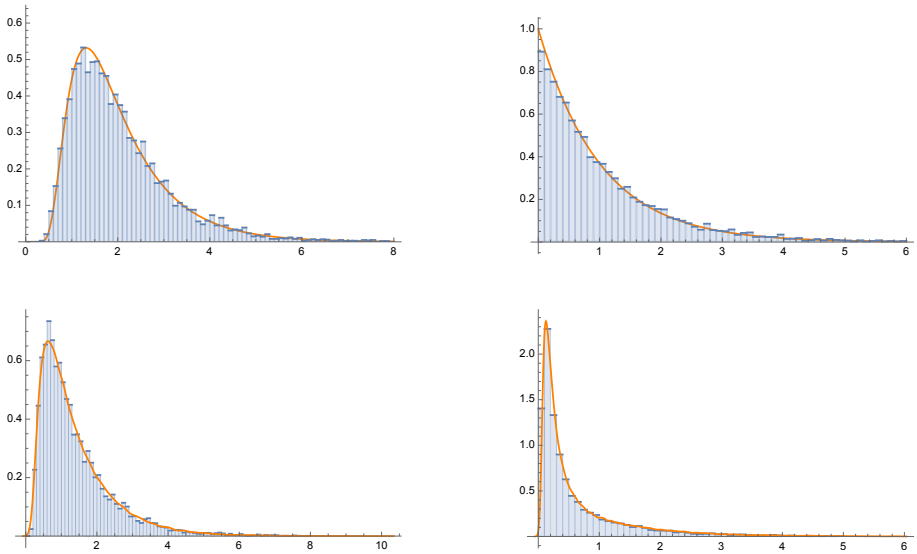

Figure 2.4: The pictures refer to the statistics of τ_{coal} and τ_{meet} for 10^5 runs of the dynamics on the same (quenched) realization of the graph, with $n = 1000$, for the "alternate" model described in Subsection 2.3.1, with $a = 2$ and $b = 4$.
Top left: in blue, the discretized empirical density of $\frac{2\tau_{\text{coal}}}{n\vartheta}$. In orange, the numerical approximation of the PDF of the infinite sum in (2.29).
Top right: in blue, the discretized empirical density of $\frac{2\tau_{\text{meet}}}{n\vartheta}$ where the initial distribution is $\pi \otimes \pi$. In orange, the PDF of an exponential distribution of mean 1.
Bottom left: in blue, the discretized empirical density of $\frac{2\tau_{\text{cons}}}{n\vartheta}$ with initial density $u = 1/2$. In orange, the numerical approximation of the PDF infinite sum in (2.30).
Bottom right: same as in the bottom left picture, but with $u = 1/10$.

and call $\vartheta(d, a)$ the quantity in (2.28) in this case. Then, for all $n \in 2\mathbb{N}$ and any fixed $d \geq 2$ the function

$$\vartheta(d, \cdot) : \{2, \dots, d\} \to \left(0, \sqrt{\frac{d}{d-1}}\right]$$

is monotone increasing.

Proof. Rewrite

$$\delta = d\,, \qquad \beta = \frac{1}{d}\frac{a^2+b^2}{2}\,, \qquad \rho = \frac{1}{d}\frac{a^2+b^2}{2ab}\,, \qquad \gamma = \frac{1}{d}\frac{a^3+b^3}{2ab}\,, \tag{2.46}$$

with $b = 2d - a$. It is immediate ti check that, fixed d, β, ρ and γ are decreasing in $\{2, \dots, d\}$. Similarly,

$$\mathfrak{p} = \frac{1}{d}\frac{\gamma-\rho}{1-\rho} + \frac{\beta-1}{d}\,, \tag{2.47}$$

is decreasing. On the other hand, it is possible to check that $\frac{(\beta-1)(1-\rho)}{\gamma-\rho}$ is increasing in a, writing

$$\mathfrak{q} = \left(1 + \frac{(\beta-1)(1-\rho)}{\gamma-\rho}\right)^{-1}, \tag{2.48}$$

we deduce that $\mathfrak{q}$ is monotone decreasing in a. Similarly, writing

$$\mathfrak{r} = \left(1 - \frac{\mathfrak{q}(1-\sqrt{1-\rho})}{\rho}\right)^{-1}, \tag{2.49}$$

by taking derivatives one can conclude that $\mathfrak{r}$ is monotone decreasing in a, hence the desired result follows from (2.47), (2.49) and the definition of ϑ in (2.28). □

Some further considerations about network design

Configuration models are particularly well studied in network science, due to the fact that they are *maximum likelihood* ensembles under the degree constraints. To see this in practice, suppose that a network designer is asked to design a large network in which the expected consensus time is in a certain target window. If the designer has control only on the degree sequence but no prior on the way in which the network will eventually be constructed, then the (Bayesian) designer will try to optimize the degrees so that a typical DCM with those degrees has an expected consensus time as close a possible to the given target. The reader should be convinced at this point that our results provide to the designer the opportunity to solve such a task. Going beyond the sub-models investigated above, we conclude this section with a take-home message concerning the effect of degree variability and degree correlations in affecting our stopping times.

On the one hand, it is clear from the special cases investigated above that the variability in the in-degree sequence plays a crucial role in determining the value of ϑ. On the other hand, with the general result of Theorem 2.3.1 at hand, we propose as a measure of correlation between in- and out-degrees the quantity

$$\alpha := \frac{\gamma-\rho}{1-\rho} \geq 1\,. \tag{2.50}$$

To see that the latter can be thought of as a measure of correlation it is helpful to note that it equals one in the Eulerian case and diverges when $\frac{1}{n}\sum_x \frac{d_x^-}{d_x^+}$ diverges. It is worth to introduce the increasing function

$$\epsilon : [0, 1/2] \to [0.5, 0.59)\,, \qquad x \mapsto \frac{1-\sqrt{1-x}}{x}\,, \tag{2.51}$$

Indeed, with this notation we get

$$\mathfrak{p} = \frac{\alpha+\beta-1}{\delta}\,, \qquad \mathfrak{q} = \left(1 + \frac{\beta-1}{\alpha}\right)^{-1}, \qquad \mathfrak{r} = \big(1 + \epsilon(\rho)\,\mathfrak{q}\big)^{-1}, \tag{2.52}$$

which result in

$$\vartheta = \frac{\delta}{\big(1-\epsilon(\rho)\big)\alpha+\beta-1}\,. \tag{2.53}$$

The advantage of (2.53) is that it immediately shows the effect of α and β on the function ϑ. Indeed, fixed the average degree δ, and realizing that the effect of the ϵ-term is somehow negligible due to its small range, we see that ϑ is essentially inverse proportional to α and β. In other words, the larger the variability of the in-degree distribution, and the more anti-correlated are the out- and in-degree distributions, the faster is the voter model in reaching consensus. It is worth to point out that in [SAR08] the authors predict, by means of some non-rigorous computation, that in the case of an *undirected* random graph from the configuration model one should have $\vartheta \approx \delta/\beta$. Notice that in the undirected setting (by Eulerianity) one has $\alpha = 1$ and therefore the formula in (2.53) coincides with the prediction in [SAR08] up to the quantity $\epsilon(\rho)$ at the denominator. Notice also that their prediction actually fails to capture the exact constant in the case of regular undirected graphs, in which ϑ is given by (2.33) while $\delta/\beta = 1$ in that case.

§2.4 Geometry of the DCM

In this section we provide the prerequisite knowledge on the typical feature of a graph sampled from the DCM and of the random walk on it. In particular in Section 2.4.1 we discuss the classical Breadth First (BF) construction of the graph and its coupling with the construction of a Galton-Watson tree. In Section 2.4.2 we introduce the discrete-time random walk on a typical realization of G, and recall some recent result about its mixing time and the shape of its stationary distribution.

§2.4.1 Locally tree like vertices

Recall from Section 2.2 the sequential construction used for generating the environment ω with n vertices and degree sequences $\mathbf{d} = \mathbf{d}_n = (\mathbf{d}^-, \mathbf{d}^+)$. In the following section we propose a well-known coupling describing the locally-tree-like structure of the (sparse) directed configuration model.

For any fixed $v \in [n]$ and any $h = h_n > 0$, define $\mathcal{B}_v^+(h)$, the h-out-neighborhood of vertex v, to be the set of paths starting from v of length at most h; where a path is a sequence of directed edges $(e_1 f_1, \dots, e_\ell f_\ell)$, $\ell \le h$, such that $v_{f_i} = v_{e_{i+1}}$ for all $i = 1, \dots, \ell - 1$, and v_f (resp. v_e) is the vertex incident to the head $f \in E^-$ (resp. tail $e \in E^+$). In order to generate $\mathcal{B}_x^+(h)$ we use the breadth-first procedure (BF) starting from v as priority rule, iterating the following steps:

(a) pick the first available unmatched tail $e \in E^+$ according to BF starting from v;

(b) pick uniformly at random an *unmatched* head $f \in E^-$;

(c) draw the resulting directed edge ef. Continue until the graph distance from an unmatched tail in Item (a) to v exceeds h.

We want to compare the exploration process of a neighborhood of G with an exploration process of a marked Galton-Watson tree. To this aim, for any fixed $v \in [n]$, let us define a marked (out-directed) random tree $\mathcal{T}_v^+$ rooted at v as follows: the root is assigned mark v, and all other vertices an independent mark $x \in [n]$ with probability $\frac{d_x^-}{m}$. Each vertex with mark $x \in [n]$ has d_x^+ children. Note that $\mathcal{T}_v^+$ is obtained by gluing together d_v^+ independent Galton-Watson trees with offspring distribution

$$\lambda_n^{\text{biased}}(k) = \lambda^{\text{biased}}(k) := \sum_{x\in[n]} \frac{d_x^-}{m} \mathbb{1}(d_x^+ = k)\,, \qquad k \in \mathbb{N}\,. \tag{2.54}$$

Let $\mathcal{T}_v^+(h)$ be a subtree of $\mathcal{T}_v^+$ given by its truncation up to generation h. We give a classical description of the coupling between $\mathcal{B}_v^+(h)$ and $\mathcal{T}_v^+(h)$ that can also be found e.g. in [CCPQ23, Sec. 2.2].
Consider the steps (a)-(c) for $\mathcal{B}_v^+(h)$, and change Item (b) into

(2') Pick uniformly at random a head $f \in E^-$ among *all possible* ones, rejecting the proposal if it was already matched, and resampling the head.

The subtree $\mathcal{T}_v^+(h)$ can be generated by iterating essentially the same steps with some minor differences. In item (2') we never reject the proposal. The head chosen in item (2') belongs to E_x^- for some $x \in [n]$. In item (3) we add a new leaf and provide the mark x to it. Note that this implies that the new leaf will have d_x^+ many children.

Lemma 2.4.1. *Assume the degree sequence satisfy Assumption 2.A. Let $v \in [n]$ and let $\hat{\mathbb{P}}$ be the law of the coupling between $\mathcal{B}_v^+(h)$ and $\mathcal{T}_v^+(h)$ for any $h > 0$. It holds that*

$$\hat{\mathbb{P}}\left(\mathcal{B}_v^+(\hbar) \neq \mathcal{T}_v^+(\hbar)\right) = o(1)\,,$$

where

$$\hbar = \hbar_n := \frac{\log(n)}{5\,\log(d_{\max}^+)}\,. \tag{2.55}$$

Proof. The coupling fails only in one of the following two cases: either the sampled head $f \in E^-$ coincides with one of the heads already used in a previous sample (the "rejection" condition of $\mathcal{B}_v^+(h)$), or f is not sampled yet, but it is incident to a vertex already present in the tree. Let τ be the first time such that an uniform random choice among all heads gives $f \in E_x^-$, for some mark x which is already present in the tree. By construction, the out-neighborhood and the tree coincide up to time τ, and it holds

$$\hat{\mathbb{P}}(\tau = t) \leq \frac{t\, d_{\max}^-}{m - t}, \quad t \geq 0\,.$$

Therefore

$$\hat{\mathbb{P}}(\tau \leq t) \leq \frac{t^2\, d_{\max}^-}{m - t}\,.$$

Note that, a.s., $t = (d^+_{\max})^{\hbar+1}$ steps are sufficient to explore the whole $\mathcal{B}^+_v(\hbar)$. It follows that

$$\hat{\mathbb{P}}\left(\mathcal{B}^+_v(\hbar) \neq \mathcal{T}^+_v(h)\right) \leq \hat{\mathbb{P}}(\tau \leq (d^+_{\max})^{h+1}) \leq \frac{(d^+_{\max})^{2\hbar+2}\, d^-_{\max}}{m - (d^+_{\max})^{\hbar+1}} = o\left(\frac{(d^+_{\max})^{2\hbar}}{\sqrt{n}}\right), \tag{2.56}$$

and the conclusion follows by the definition of $\hbar$. □

In what follows, we will to refer to the following corollary of Lemma 2.4.1.

Corollary 2.4.2 (LTL vertices). *Assume the degree sequence satisfies Assumption 2.A. Let*

$$V_\star = V_{\star,n} := \{v \in [n] \mid \mathcal{B}^+_v(\hbar) \text{ is a tree}\}, \tag{2.57}$$

where $\hbar$ is defined as in (2.55)*, then*

$$\frac{|V_\star|}{n} \xrightarrow{\mathbb{P}} 1\,. \tag{2.58}$$

§2.4.2 (Discrete-time) Random walk on the DCM

For a given realization of the *environment*, ω, we will consider the discrete time simple random walk $(X_t)_{t\geq 0}$ described by the transition matrix

$$P(x,y) = \frac{|\{e \in E^+_x \mid \omega(e) \in E^-_y\}|}{d^+_x}\,. \tag{2.59}$$

Recall here that given a tail $e \in E^+$, $\omega(e) \in E^-$ represents the head at which the tail e have been matched in the uniform matching generated by the environment ω. The latter describes a Markov chain on the vertex set, in which the walker chooses one of the out-going edges of the vertex it is currently visiting and moves to the vertex attached to the matched head.

Despite π being random, it is possible to show a very precise result for the mixing time of the random walk on the DCM under Assumption 2.A. In particular, in [BCS18] the authors show that the mixing time is logarithmic and the total-variation distance decays abruptly to zero at a precise spot on the time line, an instance of the so-called *cutoff at the entropic time* (see also [BCS19, CCPQ23]). More precisely, their result reads as follows.

Theorem 2.4.3 (Mixing time). *Assume the degree sequence satisfy Assumption 2.A, and let*

$$H = H_n := \sum_{x\in[n]} \frac{d^-_x}{m} \log(d^+_x) \quad \text{and} \quad t_{\text{ent}} = t_{\text{ent},n} := \frac{\log(n)}{H}\,.$$

For all $\varepsilon \neq 1$, it holds that

$$\max_{x\in[n]} |\|P^{\lfloor \varepsilon t_{\text{ent}} \rfloor}(x,\cdot) - \pi\|_{\text{TV}} - \mathbb{1}_{\varepsilon<1}| \xrightarrow{\mathbb{P}} 0\,. \tag{2.60}$$

Beyond the mixing time result, we will need to establish a control over the maximal and minimal values of π within its support as part of our proof. In [CQ20, CP20, CCPQ23] the authors deal exactly with this problem under different assumptions on the degree sequence. We collect their results in the following theorem.

Theorem 2.4.4 (Extremal values of π). *Assume the degree sequence satisfy Assumption 2.A. Then there exists some $\varepsilon > 0$ and $C > 1$ such that*

$$\frac{\max_{x\in[n]} \pi(x)}{n^{-1+\varepsilon}} \xrightarrow{\mathbb{P}} 0\,, \qquad \frac{n^{-C}}{\min_{x\in\mathrm{supp}(\pi)} \pi(x)} \xrightarrow{\mathbb{P}} 0\,. \tag{2.61}$$

§2.5 Strategy of proof

As mentioned in the Introduction, the core contribution of this work lies in showing that the expected meeting time of two independent random walks evolving on a typical random directed graph is well concentrated around a deterministic quantity (which we provide explicitly) depending only on the parameters of the model, that is, on the degree sequences. To this aim, we follow the strategy depicted by [CFR10], which consists in using the so-called *First Visit Time Lemma (FVTL)*. Essentially, the FVTL states that, given a mixing chain and a target state ∂, as soon as the chain mixes sufficiently fast compared to the stationary mass of ∂, $\pi(\partial)$, then the hitting time of the target is well approximated (uniformly in time) by a geometric random variable whose parameter depends only on $\pi(\partial)$ and on the "local geometry" around ∂. The FVTL has been introduced by Cooper and Frieze in a series of works (see, among others, [CF04, CF05, CF07, CF08]) in which the authors characterize the first order of the cover time of random walks in various random graphs. Recently, a simplified probabilistic proof has been provided by [MQS21]. For convenience, we refer here to the version in [QS23, Appendix A], in which the authors adopt the same notation. Nevertheless, the general idea that hitting times are essentially exponential under some *fast mixing* condition has a long history, tracing back to the work of Aldous and Brown, see, e.g., [AB92, Bro99, Ald89, FL14].

Theorem 2.5.1 (First Visit Time Lemma). *For every $N \in \mathbb{N}$ consider a discrete time ergodic Markov chain $(X_t)_{t\in\mathbb{N}}$ on $[N]$ with transition matrix $Q = Q_N$ and stationary distribution $\pi = \pi_N$. Consider further a target state $\partial \in [N]$. Call*

$$t_{\mathrm{mix}} = t_{\mathrm{mix},N} \coloneqq \inf\left\{t \geq 0 \mid \max_{x\in[N]} \|Q^t(x,\cdot) - \pi\|_{\mathrm{TV}} \leq \frac{1}{2e}\right\}, \tag{2.62}$$

and

$$\mathfrak{T} = \mathfrak{T}_N \coloneqq t_{\mathrm{mix}} \times \log\left(\left(\min_{x\in\mathrm{supp}(\pi)} \pi(x)\right)^{-1}\right), \tag{2.63}$$

and assume that

$$\lim_{N\to\infty} \pi(\partial)\, \mathfrak{T} = 0\,. \tag{2.64}$$

Fix any $T = T_N$ such that

$$T \geq 2\,\mathfrak{T}\,, \qquad \limsup_{N\to\infty} \pi(\partial)\, T = 0\,, \tag{2.65}$$

and call

$$R_T(\partial) = \sum_{t=0}^{T} Q^t(\partial,\partial)\,. \tag{2.66}$$

Then, there exists a sequence $\lambda = \lambda_N$ such that

$$\lim_{N\to\infty} \frac{\lambda}{\pi(\partial)/R_T(\partial)} = 1\,, \tag{2.67}$$

and

$$\lim_{N\to\infty} \sup_{t\geq 0} \left| \frac{\mathbf{P}_\pi(\tau_\partial > t)}{(1-\lambda)^t} - 1 \right| = 0\,. \tag{2.68}$$

In particular,

$$\lim_{N\to\infty} \frac{\mathbf{E}_\pi[\tau_\partial]}{R_T(\partial)/\pi(\partial)} = 1\,. \tag{2.69}$$

Remark 2.5.2 (Continuous vs discrete time). Despite the fact that our main results in Section 2.3 are written for continuous-time processes, the First Visit Time Lemma above is written for a general discrete time chain. Let us clarify now that there is no difference between the two, due to the fact that the exit rates are normalized to 1, that is

$$L_{\mathrm{rw}} = P - \mathrm{Id}\,. \tag{2.70}$$

Therefore, by a Poissonization argument, it follows immediately that neither π nor the order of t_{mix} depend on the fact that time is discrete or continuous. Similarly, the convergence in (2.68) can be rephrased in continuous time as

$$\lim_{N\to\infty} \sup_{t\geq 0} \left| \frac{\mathbf{P}_\pi(\tau_\partial > t)}{e^{-\lambda t}} - 1 \right| = 0\,. \tag{2.71}$$

We aim to compute the first order asymptotic of the meeting time of independent random walks on a graph G. To do so, the idea is to apply the FVTL to the product chain on $[n]\times[n]$ associated to the transition matrix

$$P^{\otimes 2} = \frac{1}{2}\,(P\otimes I + I\otimes P)\,. \tag{2.72}$$

Notice that the latter corresponds to the evolution of two independent asynchronous random walks on G where at each step a walk is selected u.a.r. to perform a single

random walk step. Clearly, the stationary distribution of $P^{\otimes 2}$ is given by $\pi \otimes \pi$. It follows that, in our setting, the role of the *target* in the FVTL is played by the diagonal set

$$\Delta := \{(x, y) \in [n] \times [n] \mid x = y\}. \tag{2.73}$$

Nonetheless, being the FVTL suited for a single target, an extra step is needed when the target is a non-trivial set. As one can imagine, this extra step lies in considering a chain in which the target set is merged into a single state. In other words, we consider the state space

$$\tilde{V} := ([n]^2 \setminus \Delta) \cup \{\partial\}, \tag{2.74}$$

where the state ∂ represents the merged set Δ. Then, we are interested in constructing a process $\tilde{P}$ on $\tilde{V}$ having the following properties:

(a) The transitions from $\tilde{V} \setminus \{\partial\}$ to $\tilde{V} \setminus \{\partial\}$ should coincide with those from $[n]^2 \setminus \Delta$ to $[n]^2 \setminus \Delta$ for the product chain $P^{\otimes 2}$.

(b) The transitions from any $\boldsymbol{x} \in \tilde{V} \setminus \{\partial\}$ to ∂ should coincide with the cumulative transitions from $\boldsymbol{x} \in [n]^2 \setminus \Delta$ to Δ for the product chain $P^{\otimes 2}$.

(c) The stationary distribution of the new chain, say $\tilde{\pi}$, satisfies

$$\left\langle \pi^2 \right\rangle := \sum_{v \in [n]} \pi(v)^2 = \tilde{\pi}(\partial), \qquad \pi(x)\pi(y) = \tilde{\pi}((x, y)), \quad x, y \in [n] \text{ s.t. } x \neq y. \tag{2.75}$$

Indeed, if the conditions (1), (2) and (3) are satisfied, then we have the identity

$$\tilde{\mathbf{E}}_{\tilde{\pi}}[\tau_{\partial}] = \mathbf{E}^{\otimes 2}_{\pi \otimes \pi}[\tau_{\text{meet}}], \tag{2.76}$$

thus it is enough to check that the assumption of the FVTL apply to the chain $\tilde{P}$ to conclude that the meeting time occurs at a geometric time of rate given by (2.67). As pointed out in [MQS21], it is immediate to check that all the three conditions can be satisfied by defining $\tilde{P}$ as follows: call

$$\tilde{\mu}(v) := \frac{\pi(v)^2}{\left\langle \pi^2 \right\rangle}, \qquad v \in [n] \tag{2.77}$$

and for every $\boldsymbol{x}, \boldsymbol{y} \in \tilde{V}$ define

$$\tilde{P}(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} \frac{1}{2} P(x_1, y_1) & \boldsymbol{x} = (x_1, z),\ \boldsymbol{y} = (y_1, z),\ z \neq x_1, y_1, \\ \frac{1}{2} P(x_2, y_2) & \boldsymbol{x} = (z, x_2),\ \boldsymbol{y} = (z, y_2),\ z \neq x_2, y_2, \\ \frac{1}{2} \left(P(x_1, x_2) + P(x_2, x_1)\right) & \boldsymbol{x} = (x_1, x_2),\ \boldsymbol{y} = \partial\,, x_1 \neq x_2, \\ \frac{1}{2} \left(\tilde{\mu}(y_1) P(y_1, y_2) + \tilde{\mu}(y_2) P(y_2, y_1)\right) & \boldsymbol{x} = \partial,\ \boldsymbol{y} = (y_1, y_2),\ y_1 \neq y_2, \\ \sum_{z \in [n]} \tilde{\mu}(z) P(z, z) & \boldsymbol{x} = \partial,\ \boldsymbol{y} = \partial. \end{cases} \tag{2.78}$$

In other words, $\tilde{P}$ can be thought as the product chain $P^{\otimes 2}$ with the addition of a *reset step* at Δ: when the random walks on G meet, they are instantaneously "reset" to another vertex (possibly the same) sampled according to the probability distribution $\tilde{\mu}$.

As it will become clear later, one of the challenges in the application of the First Visit Time Lemma lies in the explicit computation of the truncated Green function $R_T(\partial)$. The main idea, that will be fully developed in Section 2.6, is based on the intuition that the graph of the Markov chain $\tilde{Q}$, *seen from* ∂, looks like a infinite transient (and directed) graph. Based on this intuition, we will couple the quantity $R_T(\partial)$ with the Green function of the random walk on this infinite graph rooted at ∂. By transience, the latter can be computed as the escape probability of the random walk, i.e., the probability that the walk never returns to the root. More precisely, let us call $\mathcal{T}_{\tilde{\mu}}$ an out-directed Galton-Watson tree in which the root has k children with probability $\sum_{x\in V} \tilde{\mu}(x)\mathbb{1}_{d_x^+=k}$, while every other vertex has offspring distribution λ^{biased}. In view of the construction of the process $\tilde{P}$, the escape probability described above will be shown to coincide with the (annealed) probability that two walks starting at the root of $\mathcal{T}_{\tilde{\mu}}$ never meet anymore after time $t = 0$.

As the reader can imagine, such an approach can be worked out similarly also in the undirected setup, as already heuristically explained, e.g., in [AF02, Dur07]. Nevertheless, the explicit computation of the escape probability requires a detailed combinatorial analysis of the possible trajectories on the random limiting graph, which is much easier in our directed setting where backtracking is not allowed.

Remark 2.5.3. If we desire that only conditions (1) and (2) are satisfied by $\tilde{P}$, then the reset distribution $\tilde{\mu}$ can be replaced by any probability distribution μ on $[n]$. In what follows we will need to adopt such a general perspective. In fact, one should keep in mind that we are interested in *random* directed graphs. Therefore, the distribution $\tilde{\mu}$ in (2.77), which depends on π, is a *random* probability distribution depending on the realization of the graph G. This adds a further level of complication, which distinguishes the directed setting from the classical *undirected* configuration model. For this reason, as will be explained in Section 2.5.2, we will start by analyzing chains with a fixed distribution μ (which does not depend on the graph), and subsequently prove the result for $\tilde{\mu}$ as in (2.77) by means of some concentration arguments.

Remark 2.5.4 (Continuous vs discrete time – continued). It is worth to stress that the discrete time processes in (2.72) and (2.78) are, in expectation, twice as slow as their continuous time counterparts. That is to say, the expected meeting time of two independent discrete-time random walks evolving accordingly to $P^{\otimes 2}$ is twice the expected meeting time of the same two particles evolving independently according to L_{rw}. It will be technically convenient in what follows to study the discrete-time processes introduced in this section, and then recall that the value of the constant λ

in Theorem 2.5.1 has to be doubled to recover the same result in the continuous time setting. Notice indeed that, a posteriori, $\lambda \sim (\vartheta n)^{-1}$, from which, in the continuous time setting of Theorem 2.3.1, we get the factor $1/2$ in the denominator of (2.27).

In light of Theorem 2.5.1 and (2.76), the proof of Theorem 2.3.1 relies on three main propositions, ensuring that conditions of Theorem 2.5.1 are satisfied w.h.p. by the (random) chain $\tilde{P}$ in (2.78):

Proposition 2.5.5 (Mass at the diagonal). *Assume the degree sequences satisfy Assumption 2.A. Let* $\mathfrak{p}$ *as in* (2.23)*. Then it holds that*

$$\frac{n\left\langle \pi^2 \right\rangle}{\mathfrak{p}} \xrightarrow{\mathbb{P}} 1. \tag{2.79}$$

Proposition 2.5.6 (Returns to the diagonal). *Assume the degree sequences satisfy Assumption 2.A. Let* $\mathfrak{q}$ *and* $\mathfrak{r}$ *be as in* (2.24) *and* (2.25) *respectively. Then, with* $T := \lfloor \log^5(n) \rfloor$*, it holds*

$$\frac{R_T(\partial)}{\mathfrak{r}} \xrightarrow{\mathbb{P}} 1\,. \tag{2.80}$$

Proposition 2.5.7 (Mixing time for the collapsed chain). *Assume the degree sequences satisfy Assumption 2.A. Then, the mixing time of the chain* $\tilde{P}$ *defined in* (2.78)*, i.e.,*

$$\tilde{t}_{\mathrm{mix}} := \inf\left\{ t \geq 0 \mid \max_{\boldsymbol{x} \in \tilde{V}} \|\tilde{P}^t(\boldsymbol{x}, \cdot) - \tilde{\pi}\|_{\mathrm{TV}} \leq \frac{1}{2e} \right\}, \tag{2.81}$$

satisfies

$$\frac{\tilde{t}_{\mathrm{mix}}}{\log(n)^3} \xrightarrow{\mathbb{P}} 0. \tag{2.82}$$

At this point, our main results follow at once.

Proof of Theorem 2.3.1 and Corollary 2.3.2. To see the validity of Theorem 2.3.1 it is enough to realize that, by Propositions 2.5.5, 2.5.6, 2.5.7 and Theorem 2.4.4, the convergences in (2.67), (2.68) and (2.69) hold in probability for the process $\tilde{\mathbf{P}}$. Hence, by the identity in (2.76), they hold in probability for the product chain $\mathbf{P}^{\otimes 2}$. The passage to continuous time follows immediately by Remarks 2.5.2 and 2.5.4. Then, the convergence in probability of the Wasserstein-1 distance in (2.27) follows by the continuous time version of (2.68) by working under the appropriate high-probability events.

As for Corollary 2.3.2, notice that the condition in (2.13) is satisfied w.h.p. thanks to Theorems 2.4.3 and 2.4.4. At this point, the convergences in probability in (2.29) and (2.30) follow by Theorems 2.2.1 and 2.2.2. □

§2.5.1 Heavy-tailed in-degrees

We presented our main the result under the strong constraint of uniform boundedness of out- and in-degrees. Nevertheless, it is natural to conjecture that the result still holds true in a more general setup. In particular, in [CCPQ23] the authors show that Theorem 2.4.3 holds under the following weaker assumption

Assumption 2.B (Relaxed degree assumptions). Assume the same constraints (a) and (b) as in Assumption 2.A, with (c) replaced by: there exists some $\epsilon > 0$ such that for all $n \in \mathbb{N}$

$$\text{(c')}: \qquad \sum_{x\in[n]} (d_x^-)^{2+\epsilon} \le C\, n\,. \tag{2.83}$$

■

Notice that under Assumption 2.B the quantities in (2.21) are still $\Theta(1)$, even though one can have any $d_{\max}^- \asymp n^{1/2-o(1)}$. Notice indeed that the latter condition still ensures that the quantities in (2.21) are all of order 1. In fact, our technical results are proved under such weaker assumptions. The only technical lemma in which we need to impose a stronger assumption is Lemma 2.7.5, where we essentially show the concentration of the quantities in (2.26). Therein, our techniques require the following bound

$$\text{(d)}: \qquad d_{\max}^- \le C\, n^{\frac{1}{3}-\epsilon}\,. \tag{2.84}$$

In particular, our proof of Proposition 2.5.7, works in the more general framework of Assumption 2.B, whereas Proposition 2.5.5 and 2.5.6 require in addition (2.84).

Nevertheless, to formally extend the result in Theorem 2.3.1 to any degree sequences satisfying Assumption 2.B and (2.84), it is further required a control on the minimum of π. In fact, the second result in Theorem 2.4.4 has been proved in [CP20] only in the bounded degree setting. Nevertheless, it is natural to expect that the effect of large in-degrees cannot impact the fact that $\pi_{\min} = n^{-O(1)}$. Investigating the extent to which the heuristic argument above can be made precise is not the focus of this work.

§2.5.2 Organization of the proof

Due to the point raised in Remark 2.5.3, we start by analyzing independent random walks undergoing the reset according to any fixed distribution μ. In particular, in Section 2.6 we show that the *annealed version* of such random walks can be successfully coupled with certain random rooted forests. In Section 2.6.2 we exploit such a computational tool to control the quantity $R_T(\partial)$ in Theorem 2.5.1, again in the case

in which the graph G is random but the auxiliary chain $\tilde{P}$ has a deterministic reset distribution μ in place of $\tilde{\mu}$. In doing so, we show that the dependence of $R_T(\partial)$ on μ is weak, meaning that it depends only on the expectation with respect to μ of certain bounded functions of the out-degrees. Subsequently, in Section 2.7, we show that the expectation of such bounded functions with respect to the random distribution $\tilde{\mu}$ concentrates around a function of the degree sequence, see Proposition 2.7.3. In particular, as a consequence of such a general result, we extend the result obtained for deterministic reset distributions in Section 2.6.2 to the case of $\tilde{\mu}$-resets, thus deducing the validity of Propositions 2.5.5 and 2.5.6. Finally, Section 2.8 is devoted to the proof of Proposition 2.5.7.

§2.6 Random walks with μ-reset

Throughout this section, we fix a probability distribution $\mu = \mu_n$ on $[n]$ and consider, for a fixed realization of the environment ω, the discrete time Markov chain $\tilde{P} = \tilde{P}^{\omega}_{\mu}$ on $\tilde{V}$, defined as in (2.78) with μ in place of $\tilde{\mu}$. It is not hard to see that for each choice of μ the chain $\tilde{P}_{\mu}$ will have a unique stationary distribution $\tilde{\pi}_{\mu}$. For a prescribed distribution μ, recall that the transition probabilities from ∂ are as follows:

$$\tilde{P}_{\mu}(\partial, \boldsymbol{x}) = \begin{cases} \frac{1}{2}\left(\mu(x)P(x,y) + \mu(y)P(y,x)\right) & \boldsymbol{x} = (x,y)\,, \\ \sum_{z\in[n]} \mu(z)P(z,z) & \boldsymbol{x} = \partial\,, \end{cases} \qquad \boldsymbol{x} \in \tilde{V}\,.$$

Moreover, we will use the symbols $\tilde{\mathbf{P}}^{\mu}_{\partial} = \tilde{\mathbf{P}}^{\omega,\mu}_{\partial}$ and $\tilde{\mathbf{E}}^{\mu}_{\partial} = \tilde{\mathbf{E}}^{\omega,\mu}_{\partial}$ to denote the law and the expectation of the trajectories of the Markov chain $(\tilde{X}_t)_{t\geq 0}$ on $\tilde{V}$, with $\tilde{X}_0 = \partial$ and transition matrix $\tilde{P} = \tilde{P}^{\omega}_{\mu}$, on the quenched realization $G = G^{\omega}$ of the graph. Finally, we remind the reader that the symbols $\mathbb{P}$ and $\mathbb{E}$ are concerned uniquely with the probability space of the generation of the environment ω.

§2.6.1 Annealed random walks with μ-reset

In what follows we will be interested in computing expectations, with respect to $\mathbb{P}$, to the transition probabilities of the chain $\tilde{P}$. A very useful way to compute this quantities (see Remark 2.6.3 below for references) is to rewrite such expectations as probabilities of a non-Markovian process that we will call *annealed walks*. In other words, the annealed walks process is the one in which the environment ω is realized together with the random walks evolving on it. More precisely, for every $A \subseteq \tilde{V}$ and $t \geq 0$ one can write

$$\mathbb{E}\left[\tilde{\mathbf{P}}^{\mu}_{\partial}\left(\tilde{X}_t \in A\right)\right] = \tilde{\mathbb{P}}^{\mathrm{an},\mu}_{\partial}\left(\tilde{W}_t \in A\right), \tag{2.85}$$

where the law $\tilde{\mathbb{P}}^{\mathrm{an},\mu}_{\partial}$ is the law of the random variable $(\omega_t, \tilde{W}_t)_{t\geq 0}$, where ω_t is a partial matching of E^+ and E^- with at most t matchings, and $\tilde{W}_t \in \tilde{V}$ is the location of the annealed walk at time t.

The equality in (2.85) is a consequence of the exchangeability property of the DCM that translates into a sort of *spatial Markov property*. It resembles the fact that we can compute the left-hand side of (2.85) using a local exploration process. More explicitly, samples according to $\tilde{\mathbb{P}}_{\partial}^{\mathrm{an},\mu}$ can be obtained as result of the following randomized algorithm:

(a) set $\tilde{W}_0 = \partial$;

(b) for all $s \in \{0, \dots, t\}$, if $\tilde{W}_s = \partial$, sample a vertex x according to μ. Select one of the tails of x, $e \in E_x^+$ uniformly at random:

  (2a) if e is already matched to some $f \in E^-$, call v_f the vertex incident to the head f. If $v_f \neq x$, set either $\tilde{W}_{s+1} = (x, v_f)$ or $\tilde{W}_{s+1} = (v_f, x)$ w.p. $1/2$. Whereas, if $v_f = x$, set $\tilde{W}_{s+1} = \partial$.

  (2b) if e is unmatched, choose u.a.r. some $f \in E^-$ which is still unmatched, set $\omega(e) = f$. Call v_f the vertex incident to the head f. If $v_f \neq x$ set either $\tilde{W}_{s+1} = (x, v_f)$ or $\tilde{W}_{s+1} = (v_f, x)$ w.p. $1/2$. Whereas, if $v_f = x$, set $\tilde{W}_{s+1} = \partial$.

(c) if instead $\tilde{W}_s = (x, y)$ with $x \neq y$, select the coordinate to move with uniform probability. For the sake of illustration, let us as assume that this the first coordinate. Select u.a.r. one of the tails of vertex x (which is associated with the selected coordinate) and call it $e \in E_x^+$:

  (3a) if e was already matched at a previous step to some head $f \in E^-$, let $v_f \in [n]$ denote the vertex incident to the head f. Then, if $v_f \neq y$ let $\tilde{W}_{s+1} = (v_f, y)$, otherwise let $\tilde{W}_{s+1} = \partial$;

  (3b) if e is still unmatched, select uniformly at random a head, $f \in E^-$, among the unmatched ones, match it to e, and let $v_f \in [n]$ the associated vertex. As in the previous case, if $v_f \neq y$ let $\tilde{W}_{s+1} = (v_f, y)$, otherwise let $\tilde{W}_{s+1} = \partial$.

Remark 2.6.1 (Multiple annealed random walks). In what follows, it will be convenient to use the *annealed philosophy* to compute higher moments of the transition probabilities associated to the law $\tilde{\mathbf{P}}_{\partial}^{\mu}$. Indeed, for all $\kappa \in \mathbb{N}$, $A_1, \dots, A_\kappa \subset \tilde{V}$ and $t_1, \dots, t_\kappa \geq 0$ one can write

$$\mathbb{E}\left[\prod_{i=1}^{\kappa} \tilde{\mathbf{P}}_{\partial}^{\mu}\left(\tilde{X}_{t_i} \in A_i\right)\right] = \tilde{\mathbb{P}}_{\partial}^{\kappa\text{-an},\mu}\left(\tilde{W}_{t_i}^{(i)} \in A_i\,,\ \forall i \leq \kappa\right), \tag{2.86}$$

where the random variables $(\tilde{W}_s^{(i)})_{i \leq \kappa\,,\, s \leq t_i}$ can be sampled by means of the same construction described above, constructing the κ walks sequentially, with the i-th one evolving in the partial environment constructed by the first $i-1$.

Remark 2.6.2 (Annealed random walks without reset). As should be clear to the reader, an annealing formula is available also for the simple random walk, with law $\mathbf{P}$, and for the independent random walks with law $\mathbf{P}^{\otimes 2}$. Moreover, an annealing formula can be written also when conditioning on a partial realization of the underlying graph.

In general, called σ a partial matching of E^- and E^+ and fixed an initial distribution $\nu \in \mathcal{P}([n]^2)$ depending only on σ, any $\kappa \in \mathbb{N}$, a collection of times $t_1, \dots, t_\kappa \geq 0$ and some events $\mathcal{E}_1(X^{(1)}, X^{(2)}, t_1), \dots, \mathcal{E}_\kappa(X^{(1)}, X^{(2)}, t_\kappa)$ depending only on the trajectory of $(X^{(1)}, X^{(2)})$ up to time t_i, respectively, one can write

$$\mathbb{E}\left[\prod_{i=1}^{\kappa} \mathbf{P}_\nu^{\otimes 2}\left(\mathcal{E}_i(X^{(1)}, X^{(2)}, t_i)\right) \mid \sigma\right] = \mathbb{P}_\nu^{\otimes 2,\, \kappa\text{-an}|\sigma}\big(\cap_{i\leq\kappa}\, \mathcal{E}_i(W^{(i,1)}, W^{(i,2)}, t_i)\big), \tag{2.87}$$

where $\mathbb{P}_\nu^{\otimes 2,\kappa\text{-an}|\sigma}$ is the joint law of $(\omega_s)_{s\leq\kappa T}$ and $(W_s^{(i,1)}, W_s^{(i,2)})_{i\leq\kappa,\, s\leq t_i}$ can be sampled by constructing the κ couples of walks sequentially, with the first couple starting at ν with $\omega_0 = \sigma$ and, in general, the i-th couple starting independently at ν and evolving in the partial environment given by σ and the additional matchings constructed by the first $i-1$ couples. Differently from the construction of annealed walks with μ-reset, when a couple of annealed walks meet, they do not experience any reset but rather the construction continues as for any other point in $[n]^2$.

Remark 2.6.3 (Annealed simple random walks). The construction of the annealed random walk with μ-reset presented in this section—as well as its variant without reset introduced in Remark 2.6.2—is a natural variation of the original *annealed simple random walk*, which is a classical tool for the analysis of random walks on sparse random graphs. In particular, in the context of random directed graphs, it has been used extensively in [BCS18, BCS19, CQ20, CQ21a, CQ21b, CCPQ23].

In the *annealed simple random walk* construction, there is a single annealed walk W that replaces either $\tilde{W}$ or $(W^{(1)}, W^{(2)})$. Nevertheless, we postpone a detailed definition of the algorithm to the proof of Lemma 2.7.4. In order to introduce the notation that will be used later, call σ a partial matching of E^- and E^+ and fix an initial distribution $\mu \in \mathcal{P}([n])$ depending only on σ, any $\kappa \in \mathbb{N}$, a collection of times $t_1, \dots, t_\kappa \geq 0$ and some events $\mathcal{E}_1(X, t_1), \dots, \mathcal{E}_\kappa(X, t_k)$ depending only on the trajectory of X up to time t_i, respectively. We write

$$\mathbb{E}\left[\prod_{i=1}^{\kappa} \mathbf{P}_\mu(\mathcal{E}_i(X, t_i)) \mid \sigma\right] = \mathbb{P}_\nu^{\kappa\text{-an}|\sigma}\big(\cap_{i\leq\kappa}\, \mathcal{E}_i(W^{(i)}, t_i)\big), \tag{2.88}$$

where the $\mathbb{P}_\nu^{\kappa\text{-an}|\sigma}$ is the joint law of $(\omega_s)_{s\leq\kappa T}$ and $(W_s^{(i)})_{i\leq\kappa,\, s\leq t_i}$ can be sampled by constructing the κ walks sequentially, with the first one starting at μ with $\omega_0 = \sigma$ and, in general, the i-th walk independently starting at ν and evolving in the partial environment given by σ and the additional matchings constructed by the first $i-1$ walks.

In the same spirit of Section 2.4.1 we now define the generation process of a random forest which will be coupled with the construction of the annealed walk just described. More precisely, our aim is to show that, under some weak assumption on the distribution μ, the part of graph explored by $\tilde{W}$ within time $t = t_n$, for t not too large, can be coupled at a polynomially small TV-cost to a random rooted forest $\mathcal{F}$, i.e. a collection of independent, rooted, multi-type Galton-Watson trees, the construction of which we now describe.

The forest we construct is made by a random number of (out)-trees. Every vertex in each tree has a mark $v \in [n]$. Moreover, each vertex in each tree has at most two children and in particular at most one vertex in the whole forest has exactly two children. The construction of the forest goes as follows:

(a) At time $s = 0$:

- select the mark of the root according to μ, say $r \in [n]$, put a red flag on the root, set $\tilde{R}_0 = r$ and sample $\mathtt{red}_0 \in \{1, 2\}$ u.a.r.;
- sample $e \in E_r^+$ and $f \in E^-$ u.a.r.; create a new edge of the forest from the root to a new vertex with mark v_f, and give a label (e, f) to such a new edge; put a blue flag on the new vertex of the forest and set $\tilde{B}_0 = v_f$.

(b) Given the construction up to time $s \geq 0$, construct $(\mathcal{F}_{s+1}, \tilde{R}_{s+1}, \tilde{B}_{s+1}, \mathtt{red}_{s+1})$ as follows:

(a) If the red flag and the blue flag are on the same vertex, end the construction of the current tree and construct another tree of the forest starting as in Item (a).

(b) If the red flag and the blue flag are in different spots in the tree, select one of the two colors with uniform probability:

- If blue is selected, and v is the type of the vertex with the blue flag attached, sample $e \in E_v^+$ and $f \in E^-$ u.a.r., create a new edge of the forest from the vertex with the blue flag to a new vertex with mark v_f, and give label (e, f) to such a new edge. Move the blue flag to the new vertex of the forest, and set $\tilde{R}_{s+1} = \tilde{R}_s$, $\tilde{B}_{s+1} = v_f$ and $\mathtt{red}_{s+1} = \mathtt{red}_s$.
- If red is selected, and v is the mark of vertex with the red flag on top proceed as follows:
  - If v has a (unique) child: sample $e \in E_v^+$ u.a.r.
    (I) if the edge connecting v to its child has label (e, f) for some $f \in E^-$, move the red flag to the v_f and set $\tilde{R}_{s+1} = v_f$, $\tilde{B}_{s+1} = \tilde{B}_s$ and $\mathtt{red}_{s+1} = \mathtt{red}_s$.

(II) otherwise, sample $f \in E^-$ u.a.r., create a new edge of the forest from the vertex with the red flag to a new vertex with mark v_f. Label the new edges as (e, f), and move the red flag to such a new vertex. Set $\tilde{R}_{s+1} = v_f$, $\tilde{B}_{s+1} = \tilde{B}_s$ and $\mathtt{red}_{s+1} = \mathtt{red}_s$.

– If v has no children, sample $e \in E_v^+$ and $f \in E^-$ u.a.r., create a new edge of the forest from the vertex with the red flag to a new vertex with mark v_f, and give a label (e, f) to such a new edge. Move the red flag to the newly created vertex. Set $\tilde{R}_{s+1} = v_f$, $\tilde{B}_{s+1} = \tilde{B}_s$ and $\mathtt{red}_{s+1} = \mathtt{red}_s$.

Let us try to explain in plain English how the construction of the forest is performed: there are two particles sitting on the same spot r with mark distributed according to μ. As a result of a fair coin toss, one of the particles is given color red and the other color blue. The blue particle is the first to move, and creates a directed edge (of the forest) having the root as origin and as destination some vertex with mark v, sampled according to μ_{in}. At each subsequent step, a fair coin is tossed to decide which of the two colors has to move. The blue particle always creates a new directed edge as in the first step. On the other hand, the red particle, which sits on some vertex with mark v_r, can either follow the unique out-going edge already present at the vertex it is currently visiting (if any) w.p. $1/d^+_{v_r}$; or it can create a new edge from the vertex it is currently visiting to some new vertex in the forest which will have a mark sampled accordingly to μ^{in}. Notice that, in the latter scenario, the red and blue particles will never meet again. If the two particles meet (that is, the red particles reaches the blue one), then the process restarts as follows: the two particles will be placed on a new vertex that will be the root of a new tree in the forest; the mark of such vertex will be chosen accordingly to μ and the colors of the two particles are reinitialized. Notice also that the auxiliary alea needed to define $(\mathtt{red}_t)_{t\geq 0}$, plays essentially no role in the construction, but it will be useful in a while to define the coupling. Indeed, as one can imagine at this point, the quantity $\mathtt{red}_t$ will determine which of the two coordinates of $\tilde{W}$ is "escaping" and which is "chasing" at time t. Clearly such a construction leads to a collection $(\mathcal{F}_t, \tilde{R}_t, \tilde{B}_t, \mathtt{red}_t)_{t\geq 0}$, where $\tilde{R}_t$ (resp. $\tilde{B}_t$) is the mark at the vertex in which lies the red (resp. blue) particle at time t, and $\mathcal{F}_t$ is the rooted forest realized up to step t, which we identify with the collection of its labeled edges (with multiplicities). The next result shows that it is possible to couple the two constructions at a small TV-cost, as soon as t is sufficiently small and μ is not too concentrated.

Proposition 2.6.4. *Assume that, for some $\varepsilon > 0$*

$$\max_{x\in[n]} \mu(x) \leq n^{-\varepsilon} \,. \tag{2.89}$$

Call $(\tilde{\mathcal{L}}_s)_{s\leq t}$ the law of the annealed walks construction of $(\omega_s, \tilde{W}_s)_{s\leq t}$ and $(\tilde{\mathcal{L}}^{\mathcal{F}}_s)_{s\leq t}$ the law of the process $(\mathcal{F}_s, W(\tilde{R}_s, \tilde{B}_s, \mathtt{red}_s), \mathtt{red}_s)_{s\leq t}$ where $(\mathcal{F}_s, \tilde{R}_s, \tilde{B}_s, \mathtt{red}_s)_{s\leq t}$ are

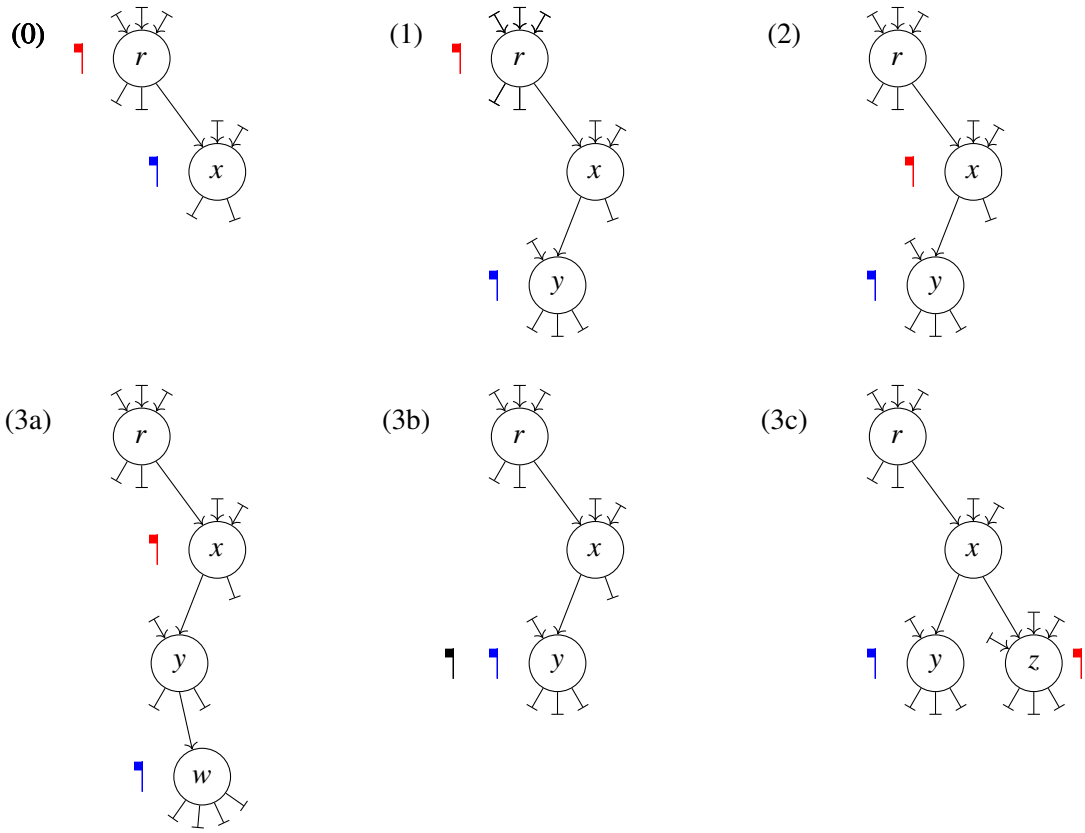

Figure 2.5: An example of a realization of the forest. As shown in picture (0), $\tilde{R}_0 = r$ is the mark, chosen according to μ, of the first root and $\tilde{B}_0 = x$. The forest at time $s = 1$ is presented in picture (1), while $\tilde{R}_1 = \tilde{R}_0 = r$, $\tilde{B}_1 = y$. Similarly, the forest at time 2 is presented in picture (2), where $\tilde{R}_2 = \tilde{B}_0 = x$ and $\tilde{B}_2 = \tilde{B}_1 = y$. Given $(\mathcal{F}_2, \tilde{R}_2, \tilde{B}_2)$, there are three possible scenarios that can happen: if the blue flag is selected, then it will be moved to a vertex having a random mark, say w, so that $\tilde{B}_3 = w$ and $\tilde{R}_3 = \tilde{R}_2 = x$, as in picture (3a); whereas, if the red flag is selected there are two possible scenarios: in the first one, described in picture (3b), the random selected tail from E_x^+ coincides with the one connecting the vertex with mark x to its unique child, thus the two flags will be in the same vertex and a new root needs to be sampled at the forthcoming step. It follows that $\tilde{R}_3 = \tilde{B}_3 = \tilde{B}_2 = y$. Instead, picture (3c) describes the situation in which the chosen tail does not coincide with the unique edge leaving x, therefore we need to select another random mark, say z, create a new labeled edge, and move the red flag to z. Hence the vertex with mark x will have two children and the two flags will never meet again. It follows that $\tilde{R}_3 = z$, and $\tilde{B}_3 = \tilde{B}_2 = y$.

sampled via the forest construction and $W : [n]^2 \times \{1,2\} \to \tilde{V}$ *is defined as*

$$W(R,B,\texttt{red}) := \begin{cases} (R,B) & \text{if } \texttt{red} = 1 \text{ and } R \neq B\,, \\ (B,R) & \text{if } \texttt{red} = 2 \text{ and } R \neq B\,, \\ \partial & \text{if } R = B\,. \end{cases} \tag{2.90}$$

Then

$$\max_{t \le n^{\varepsilon/3}} \|\tilde{\mathcal{L}}_t - \mathcal{L}_t^{\mathcal{F}}\|_{\mathrm{TV}} = o(n^{-\varepsilon/3})\,. \tag{2.91}$$

Proof. In the same vein of Section 2.4.1, consider the coupling between the two constructions in which in Step (3b) the sample of an head f is made *u.a.r. among all heads*, and rejecting the sample if f is already matched. In case of a rejection, we say that the coupling has *failed*, and the two constructions will then continue independently. Similarly, if the vertex x sampled in Step (2) of the construction has already a matched head or a matched tail, declare the coupling as failed and continue the two constructions independently. Let $\hat{\mathbb{P}}$ denote the law of such coupling. We now show that, if t is not too large, the probability that the coupling fails goes to zero as $n \to \infty$ at the speed in (2.91), which immediately implies the desired result. For a fixed $s \le t$, consider the following events:

$$A_s = \left\{ \begin{array}{c} s \text{ is the first time in which at step (2b) or (3b) is sampled some } f \in E^- \\ \text{that was already sampled in step (2b) or (3b) at some time} < s \end{array} \right\},$$
$$B_s = \left\{ \begin{array}{c} s \text{ is the first time in which a vertex } x \in [n] \text{ which has} \\ \text{already a matched tail or head is sampled in step (2)} \end{array} \right\},$$

and define, similarly to the proof of Lemma 2.4.1, the hitting time τ to be the the first time such that the coupling fails. Note that the only ways in which the coupling can fail at a fixed time $s \le t$ are the ones stated in the events A_s and B_s. Therefore, by the union bound we get

$$\begin{aligned} \hat{\mathbb{P}}(\tau \le t) &\le \hat{\mathbb{P}}\left(\bigcup_{s \le t}(A_s \cup B_s)\right) \le \sum_{s \le t} \hat{\mathbb{P}}\big(A_s \cup B_s\big) \\ &\le t \max_{s \le t} \hat{\mathbb{P}}(A_s) + t \max_{s \le t} \hat{\mathbb{P}}(B_s), \end{aligned} \tag{2.92}$$

The first term on the right-hand side of the latter inequality can be bounded as in (2.56), so that

$$t \max_{s \le t} \hat{\mathbb{P}}(A_s) \le t^2 \, \frac{d^-_{\max}}{m - t}\,. \tag{2.93}$$

As for the second term in (2.92), we have that

$$t \max_{s \le t} \hat{\mathbb{P}}(B_s) \le t \max_{s \le t} \sum_{x \in [n]} \mu(x) \mathbb{1}_{x \in \Gamma_s} \le t^2 \max_{x \in [n]} \mu(x)\,, \tag{2.94}$$

where Γ_s denotes the number of vertices explored by time s, thus $|\Gamma_s| \leq s$. Plugging the bounds in (2.93) and (2.94) into (2.92) we get

$$\hat{\mathbb{P}}(\tau \leq t) \leq t^2 \, \frac{d^-_{\max}}{m-t} + \; t^2 \max_{x \in [n]} \mu(x) \; . \tag{2.95}$$

Under our assumptions on t and μ the right-hand side of (2.95) vanishes as $n \to \infty$, and the desired result follows. □

Remark 2.6.5 (Coupling of multiple annealed random walks). Recall the multiple annealed walk $(\tilde{W}_t^{(1)}, \ldots, \tilde{W}_t^{(\kappa)})_{t \leq T}$ introduced in Remark 2.6.1 and assume that $\kappa \geq 2$ is bounded. The reader should be convinced at this point that the law of such iterated construction can be coupled with the iterated construction of κ i.i.d. forests (of polylogarithmic size) at a polynomially small TV-cost.

§2.6.2 Return to the diagonal for random walks with μ-reset

Throughout this section, we set

$$T = T_n \coloneqq \lfloor \log(n)^6 \rfloor \; . \tag{2.96}$$

For a fixed probability distribution $\mu = \mu_n$ on $[n]$ we aim at understanding the asymptotic behavior of the random variable with respect to the generation of ω given by

$$R^\mu = R_T^{\omega,\mu}(\partial) \coloneqq \sum_{t=0}^{T} \tilde{P}_\mu(\partial, \partial) \; . \tag{2.97}$$

Start by noting that its expectation can be expressed in terms of the annealed walks defined as in (2.85)

$$\mathbb{E}[R^\mu] = \sum_{t=0}^{T} \mathbb{E}[\tilde{P}_\mu(\partial, \partial)] = \sum_{t=0}^{T} \mathbb{P}^{\mathrm{an},\mu}(\tilde{W}_t = \partial) \; . \tag{2.98}$$

In what follows, it will turn out to be useful to introduce the following stopping times with respect to the annealed process

$$\begin{aligned} \tau_\partial^{(1)} &\coloneqq \inf\{t > 0 : \tilde{W}_t = \partial\} \; , \\ \tau_\partial^{(j)} &\coloneqq \inf\{t > \tau_\partial^{(j-1)} : \tilde{W}_t = \partial\} \; , \qquad j \geq 1 \; . \end{aligned} \tag{2.99}$$

Notice that, under the success of the coupling in Proposition 2.6.4 the stopping times $\tau_\partial^{(j)}$ can be interpreted at the successive times at which the red and the blue particles are on the same vertex. Exploiting the coupling with the random forest introduced in Proposition 2.6.4, we will start by showing the following quantitative estimate.

Lemma 2.6.6. *For every n fix a degree sequence satisfying Assumption 2.A. Fix $\varepsilon > 0$ and for each $q \in (0, 1/2]$ consider the compact set*

$$\mathcal{M}(q) = \mathcal{M}_n(q_n, \varepsilon) \coloneqq \left\{ \mu \in \mathcal{P}([n]) \mid \sum_{x\in[n]} \mu(x) \frac{1}{d_x^+} = q \,,\ \max_{x\in[n]} \mu(x) \le n^{-\varepsilon} \right\} . \tag{2.100}$$

Then,

$$\sup_{\mu\in\mathcal{M}(q)} \sup_{t\le \frac{T}{2}} \tilde{\mathbb{P}}_{\partial}^{\mathrm{an},\mu}(\tau_{\partial}^{(1)} = 2t+1) = o(n^{-\varepsilon/3}) \,, \tag{2.101}$$

and

$$\sup_{\mu\in\mathcal{M}(q)} \sup_{t\le \frac{T}{2}} \left| \tilde{\mathbb{P}}_{\partial}^{\mathrm{an},\mu}(\tau_{\partial}^{(1)} = 2t) - 2^{-2t+1} C_{t-1} \rho^{t-1} q \right| = o(n^{-\varepsilon/3}) \,, \tag{2.102}$$

where

$$C_s \coloneqq \frac{1}{s+1}\binom{2s}{s} , \qquad s \in \mathbb{N} \,, \tag{2.103}$$

is the s-th Catalan number and ρ and T are defined as in (2.21) *and* (2.96)*, respectively.*

Proof. Thanks to Proposition 2.6.4 it is enough to consider the same probabilities under the event that the coupling succeeds. For this reason we restrict the focus on the construction of the random forest, and let $\hat{\mathbb{P}}$ denote its law. Accordingly, the stopping time $\tau_{\partial}^{(1)}$ will be interpreted as explained below (2.99). Fix $t \le T/2$. Recall that on the forest process, the two particles are always on the same tree and their distance at time t, which we will refer to as $d(t)$, starts at $d(0) = 0$ and can only reduce by one unity (if the red particle is in the same ray as the blue particle and does a step in its direction), increase by a unity (if the red particle is in the same ray as the blue particle and the latter does a step forward), or become infinity and keep this value for the rest of the time (if the red particle is in the same ray as the blue particle and creates a new branch). Therefore, the two particles can meet only at even times, so that (2.101) follows immediately.

Recalling that $d(0) = 0$, we now need to compute the probability of the event

$$\mathcal{E}_t = \{d(s) > 0 \,, \forall s \in \{1, \dots, 2t-1\}\} \cap \{d(2t) = 0\} \,, \qquad t \le \frac{T}{2} \,. \tag{2.104}$$

Notice that, each simple event in $\mathcal{E}_t$ corresponds to an evolution in which the red particle follows the blue one up to reaching it at distance t from the root. Therefore, each simple events in $\mathcal{E}_t$ can be associated uniquely to a couple $(\mathfrak{D}_t, E_t)$ where

- $\mathfrak{D}_t$ is a Dyck path with $t-1$ upstrokes and $t-1$ downstrokes having ± 1 increments, representing the evolution of $d(s)$ for $s \le t$;

- $E_t = (x_0, \dots, x_{t-1}) \in [n]^t$ is the sequence of marks of all the vertices along the path of length t from the root to the vertex at which the particles meet (not included).

Notice that two simple events in $\mathcal{E}_t$ associated to the same sequence of marks but to different Dyck paths are equiprobable. Therefore, it is enough to take a representative Dyck path $\mathfrak{D}_t$ and write

$$\hat{\mathbb{P}}(\tau_\partial^{(1)} = 2t) = C_{t-1} \sum_{E_t \in [n]^t} \hat{\mathbb{P}}(\mathfrak{D}_t,\ E_t) = 2^{-2t+1} C_{t-1} \sum_{E_t \in [n]^t} \hat{\mathbb{P}}(E_t \mid \mathfrak{D}_t)\ , \quad (2.105)$$

where we exploited the fact that all Dyck paths are equiprobable, since they only depend on the color of the moving particle at each step, and the fact that the number of Dyck paths of length $2s$ is C_s, as in (2.103). A natural choice for the representative Dyck path consists in choosing $\mathcal{D}_t$ to by the Dyck path in which the walks move alternately, with the exception of the first two steps, in which the blue particle moves, and the last two step, in which the red particle moves twice. In such a way the distance between them oscillates between 1 and 2 until $\tau_\partial^{(1)}$. Recall the definition of λ^{biased} in (2.54) and consider the collection of independent random variables $D_0, D_1, \dots, D_{t-1}$ where $D_i \sim \lambda^{\text{biased}}$ for $i \geq 1$ while D_0 is distributed as

$$\lambda^\mu(d) \coloneqq \sum_{x \in [n]} \mu(x) \mathbb{1}(d_x^+ = d)\ , \qquad d \in \mathbb{N}\ . \quad (2.106)$$

the out degree of a vertex sample according to μ. Then

$$\begin{aligned} \sum_{E_t \in [n]^t} \hat{\mathbb{P}}\left(E_t \mid \mathfrak{D}_t\right) &= \sum_{d_0 \geq 2} \cdots \sum_{d_{t-1} \geq 2} \prod_{j=0}^{t-1} \frac{1}{d_j}\, \hat{\mathbb{P}}(D_0 = d_0, D_1 = d_1, \dots, D_t = d_{t-1}) \\ &= \sum_{d_0, d_1, \dots, d_{t-1} \geq 2} \prod_{j=0}^{t-1} \frac{1}{d_j}\, \hat{\mathbb{P}}(D_j = d_j) \\ &= \left(\sum_{d \geq 2} \frac{1}{d}\, \lambda^\mu(d) \right) \left(\sum_{d \geq 2} \frac{1}{d}\, \lambda^{\text{biased}}(d) \right)^{t-1} . \end{aligned} \quad (2.107)$$

More explicitly,

$$\sum_{d \geq 2} \frac{1}{d}\, \lambda^{\text{biased}}(d) = \sum_{d \geq 2} \frac{1}{d} \sum_{x \in [n]} \frac{d_x^-}{m} \mathbb{1}_{d_x^+ = d} = \frac{1}{m} \sum_x \frac{d_x^-}{d_x^+} = \rho\ , \quad (2.108)$$

and for all $\mu \in \mathcal{M}(q)$

$$\sum_{d \geq 2} \frac{1}{d}\, \lambda^\mu(d) = \sum_{d \geq 2} \frac{1}{d} \sum_{x \in [n]} \mu(x) \mathbb{1}_{d_x^+ = d} = \sum_{x \in [n]} \mu(x) \frac{1}{d_x^+} = q\ . \quad (2.109)$$

Thus the desired result follows by plugging (2.107), (2.109) and (2.108) into (2.105). □

Lemma 2.6.7. *In the same setting of Lemma 2.6.6,*

$$\sup_{\mu\in\mathcal{M}(q)}\sup_{t\le T/2}\tilde{\mathbb{P}}_{\partial}^{\mathrm{an},\mu}(\tilde{W}_{2t+1}=\partial)=o(n^{-\varepsilon/2})\,, \tag{2.110}$$

and

$$\sup_{\mu\in\mathcal{M}(q)}\sup_{t\le T/2}\left|\tilde{\mathbb{P}}_{\partial}^{\mathrm{an},\mu}(\tilde{W}_{2t}=\partial)-\sum_{k=1}^{t}2^{-2t+k}q^{k}\rho^{t-k}B(t,k)\right|=o(n^{-\varepsilon/2})\,, \tag{2.111}$$

where

$$B(t,k)\coloneqq\frac{k}{2t-k}\binom{2t-k}{t}\,,\quad 1\le k\le t\,. \tag{2.112}$$

Proof. As for Lemma 2.6.6, we notice that thanks to Proposition 2.6.4 it is enough to work under the event that the coupling succeeds. Hence, we let here $\hat{\mathbb{P}}$ denote the law of the forest construction and interpret the stopping time $\tau_{\partial}^{(j)}$ as explained below (2.99). Moreover, we will write for simplicity $\{\tilde{W}_{2s}=\partial\}$ to mean that at time s the blue and the red particles in the construction of the forest sit on the same vertex.

Again, we can immediately rule out the odd time cases, meaning that (2.110) follows immediately. Fix $t\le T/2$ and define α_t as the number of times in which the red and the blue particle meet within time $2t$. More precisely

$$\alpha_t=\max\{k\ge 0\mid\tau_{\partial}^{(k)}\le 2t\}=\sum_{s=1}^{t}\mathbb{1}(\tilde{W}_{2s}=\partial)\,. \tag{2.113}$$

Since $1\le\alpha_t\le t$, we have that

$$\hat{\mathbb{P}}(\tilde{W}_{2t}=\partial)=\sum_{k=1}^{t}\hat{\mathbb{P}}(\tilde{W}_{2t}=\partial,\alpha_t=k)=\sum_{k=1}^{t}\hat{\mathbb{P}}(\tau_{\partial}^{(k)}=2t,\alpha_t=k)\,. \tag{2.114}$$

If $k=1$ the latter expression boils down to (2.102) in Lemma 2.6.6. Consider $k\ge 2$. The first hitting time of ∂ has support on the interval $\{2,4,\dots,2(t-k+1)\}$. Similarly, the j-th one, $\tau_{\partial}^{(j)}$, belongs to $\{\tau_{\partial}^{(j-1)}+2,\dots,2(t-k+j+1)\}$. Therefore, defining

$s_0 = 0$, $s_k = t$, and $\tau_\partial^{(0)} = 0$, it holds

$$\begin{aligned}
&\hat{\mathbb{P}}(\tau_\partial^{(k)} = 2t, \alpha_t = k) \\
&= \sum_{s_1=1}^{t-k+1} \sum_{s_2=s_1+1}^{t-k+2} \cdots \sum_{s_{k-1}=s_{k-2}+1}^{t-1} \prod_{j=1}^{k} \hat{\mathbb{P}}(\tau_\partial^{(j)} = 2s_j, \alpha_t = k \mid \tau_\partial^{(j-1)} = 2s_{j-1}) \\
&= \sum_{s_1=1}^{t-k+1} \sum_{s_2=s_1+1}^{t-k+2} \cdots \sum_{s_{k-1}=s_{k-2}+1}^{t-1} \prod_{j=1}^{k} \hat{\mathbb{P}}(\tau_\partial^{(j)} - \tau_\partial^{(j-1)} = 2(s_j - s_{j-1}), \alpha_t = k) \\
&= \sum_{s_1=1}^{t-k+1} \sum_{s_2=s_1+1}^{t-k+2} \cdots \sum_{s_{k-1}=s_{k-2}+1}^{t-1} \prod_{j=1}^{k} \hat{\mathbb{P}}(\tau_\partial^{(1)} = 2(s_j - s_{j-1})) \\
&= \sum_{s_1=1}^{t-k+1} \sum_{s_2=s_1+1}^{t-k+2} \cdots \sum_{s_{k-1}=s_{k-2}+1}^{t-1} \prod_{j=1}^{k} 2^{-2(s_j - s_{j-1})+1} C_{s_j - s_{j-1}-1}\, q\, \rho^{s_j - s_{j-1}-1} \\
&= 2^{-2t+k}\, q^k \rho^{t-k} B(t,k)\,,
\end{aligned} \tag{2.115}$$

where the third equality comes from the renewal structure of the annealed process at each visit to ∂ under the success of the coupling; the fourth equality holds thanks to Lemma 2.6.6; and the last equality follows from

$$\begin{aligned}
&\sum_{s_1=1}^{t-k+1} \sum_{s_2=s_1+1}^{t-k+2} \cdots \sum_{s_{k-1}=s_{k-2}+1}^{t-1} \prod_{j=1}^{k} C_{s_j - s_{j-1}-1} \\
&\qquad = \sum_{l_1=0}^{t-k} \sum_{l_2=0}^{t-k-l_1} \cdots \sum_{l_{k-1}=0}^{t-k-\sum_{j=1}^{k-2} l_j} \sum_{l_k=t-k-\sum_{j=1}^{k-1} l_j}^{t-k-\sum_{j=1}^{k-1} l_j} \prod_{j=1}^{k} C_{l_j} \\
&\qquad = \sum_{\substack{l_1+\cdots+l_k=t-k \\ l_1,\ldots,l_k \geq 0}} \prod_{j=1}^{k} C_{l_j} \;=\; \frac{k}{2t-k} \binom{2t-k}{t}\,,
\end{aligned}$$

where the last equality comes from the Catalan k-fold self-convolution identity, see e.g. [FL03, Eq. (19)], and equals the so-called "ballot numbers", $B(t,k)$, defined in (2.112). The desired result now follows by plugging (2.115) into (2.114). □

Thanks to (2.97), as a corollary of Lemma 2.6.7, we deduce the following result.

Corollary 2.6.8. *In the same setting of Lemma 2.6.6, defined* $\Phi : [0, 1/2]^2 \to \mathbb{R}$ *as*

$$\Phi(\rho, q) \coloneqq \sum_{t=1}^{\infty} \sum_{k=1}^{t} 2^{-2t+k}\, q^k\, \rho^{t-k}\, B(t,k)\,, \tag{2.116}$$

it holds

$$\lim_{n\to\infty} \sup_{\mu \in \mathcal{M}(q)} |\mathbb{E}[R^\mu] - (1 + \Phi(\rho, q))| = 0\,, \tag{2.117}$$

where $\rho = \rho_n$ and Φ are defined as in (2.112)*,* (2.21) *and* (2.116)*, respectively.*

Proof. The only thing that is worth to notice is that Φ is bounded since $q, \rho \leq 1/2$. □

At this point, it is enough to use Corollary 2.6.8 and some concentration result for the random variable R^μ to conclude the validity of the following proposition, which is the main result of this section.

Proposition 2.6.9. *In the same setting of Lemma 2.6.6,*

$$\sup_{\mu \in \mathcal{M}(q)} \left| R^\mu - (1 + \Phi(\rho, q)) \right| \xrightarrow{\mathbb{P}} 0\,, \tag{2.118}$$

where $\rho = \rho_n$ and Φ are defined as in (2.21) *and* (2.116)*, respectively.*

Proof. We only need to study the second moment of R^μ. We can rewrite

$$\mathbb{E}\left[(R^\mu)^2\right] = \sum_{s,t \leq T} \mathbb{E}\left[\tilde{\mathbf{P}}_\partial(\tilde{X}_{2t} = \partial)\tilde{\mathbf{P}}_\partial(\tilde{X}_{2s} = \partial)\right] = \sum_{s,t \leq T} \tilde{\mathbb{P}}^{2\text{-an},\mu}\left(\tilde{W}^{(1)}_{2t} = \partial,\ \tilde{W}^{(2)}_{2s} = \partial\right)\,,$$

where the law $\tilde{\mathbb{P}}^{2\text{-an},\mu}$ is defined as in Remark 2.6.1. As pointed out in Remark 2.6.5, we can couple at a polynomially small TV-cost the construction of the two annealed walks with that of two i.i.d. forests. Therefore,

$$\begin{aligned}
&\sum_{s,t \leq T} \tilde{\mathbb{P}}^{2\text{-an},\mu}\left(\tilde{W}^{(1)}_{2t} = \partial,\ \tilde{W}^{(2)}_{2s} = \partial\right) \\
&\qquad \leq o(T^2 n^{-\frac{\varepsilon}{3}}) + \sum_{s,t \leq T} 2^{-(2t+2s)} \sum_{k_1=1}^{t} \sum_{k_2=1}^{s} q^{k_1+k_2}\, \rho^{t+s-k_1-k_2}\, B(t,k_1)\, B(s,k_2) \\
&\qquad = (1 + o(1))\, \mathbb{E}[R^\mu]^2\,.
\end{aligned}$$

It follows that

$$\mathbb{E}\left[(R^\mu)^2\right] \sim \mathbb{E}\left[R^\mu\right]^2\,,$$

thus $\mathrm{Var}(\mathrm{R}^\mu) = \mathrm{o}(1)$. Therefore, by Chebyshev inequality, we can conclude that for any $\delta > 0$

$$\mathbb{P}\left(|R^\mu - \mathbb{E}[R^\mu]| > \delta\right) \leq \frac{\mathrm{Var}(\mathrm{R}^\mu)}{\delta^2} = o(1)\,,$$

Hence, $|R^\mu - \mathbb{E}[R^\mu]| \xrightarrow{\mathbb{P}} 0$ and the desired conclusion follows by Corollary 2.6.8. □

To conclude the section, we show that the function Φ admits a closed form expression.

Lemma 2.6.10. *For all $\rho, q \in (0, 1/2]$, recalling the definition of Φ in (2.116), it holds*

$$1 + \Phi(\rho, q) = \frac{\rho}{\rho - q\left(1 - \sqrt{1-\rho}\right)}. \tag{2.119}$$

Proof. Start by rewriting

$$\begin{aligned}\sum_{t=1}^{\infty}\sum_{k=1}^{t} 2^{-2t+k}\ \frac{k}{2t-k}\binom{2t-k}{t} q^k\ \rho^{t-k} &= \sum_{k=1}^{\infty} 2^k \left(\frac{q}{\rho}\right)^k \sum_{t=k}^{t} 2^{-2t}\ \frac{k}{2t-k}\binom{2t-k}{t}\rho^t \\ &= \sum_{k=1}^{\infty} 2^k \left(\frac{q}{\rho}\right)^k \sum_{t=k}^{t} \left(\frac{\rho}{4}\right)^t\ \frac{k}{2t-k}\binom{2t-k}{t-k} \\ &= \sum_{k=1}^{\infty} 2^{-k} q^k \sum_{\ell=0}^{\infty} \left(\frac{\rho}{4}\right)^{\ell}\ \frac{k}{2\ell+k}\binom{2\ell+k}{\ell}.\end{aligned} \tag{2.120}$$

For any $x \in (0, 1)$, call

$$G(x) := \frac{1 - \sqrt{1-4x}}{2x} \tag{2.121}$$

the generating function of the Catalan numbers. Notice that the inner series in the last row of (2.120) corresponds to the generating function of the generalized binomial, evaluated at $\rho/4$. Indeed, for any fixed $k \in \mathbb{N}$, it holds (see, e.g., [GKP94, Ch. 5.4] or [FL03, Eq. 24])

$$G(x)^k = \sum_{\ell=0}^{\infty} \frac{k}{2\ell+k}\binom{2\ell+k}{\ell} x^{\ell}. \tag{2.122}$$

In our setting this fact reads

$$\sum_{\ell=0}^{\infty} \left(\frac{\rho}{4}\right)^{\ell}\ \frac{k}{2\ell+k}\binom{2\ell+k}{\ell} = G^k(\rho/4) = \frac{\left(1-\sqrt{1-\rho}\right)^k}{2^{-k}\rho^k}.$$

Therefore we can conclude plugging the latter equality in (2.120), as

$$\begin{aligned}1 + \sum_{k=1}^{\infty} 2^{-k} q^k \sum_{\ell=0}^{\infty} \left(\frac{\rho}{4}\right)^{\ell}\ \frac{k}{2\ell+k}\binom{2\ell+k}{\ell} &= \sum_{k=0}^{\infty} \left(\frac{q}{\rho}\left(1-\sqrt{1-\rho}\right)\right)^k \\ &= \frac{1}{1 - \frac{q}{\rho}\left(1-\sqrt{1-\rho}\right)} = \frac{\rho}{\rho - q\left(1-\sqrt{1-\rho}\right)},\end{aligned}$$

where in the second equality we used that, since $q < \frac{1}{2}$,

$$\frac{q}{\rho}\left(1-\sqrt{1-\rho}\right) \le \frac{1-\sqrt{1-\rho}}{2\rho} = \frac{1}{2(1+\sqrt{1-\rho})} < \frac{1}{2}. \tag{2.123}$$

□

§2.7 Return to the diagonal for random walks with $\tilde{\mu}$-reset

From Proposition 2.6.9, we see that when considering the random walk with μ-reset the only feature of the law μ that determines the value of R^{μ} is the expectation of the inverse out-degree of a vertex sampled according to μ. In this section, we show how to adapt this result to the case in which the distribution μ is replaced by the random distribution $\tilde{\mu}$ defined in (2.77). The argument can be ideally divided into two parts: first, we show that the expectation of the inverse out-degree of a vertex sampled according to $\tilde{\mu}$ converges in probability to the value $\mathfrak{q}$ in (2.24); second, we use a continuity argument to conclude that $R^{\tilde{\mu}}$ converges in probability to $\mathfrak{r} = 1 + \Phi(\rho, \mathfrak{q})$, where Φ is defined as in (2.116). This concludes the proof of Proposition 2.5.6.

Concerning the first part of the above mentioned argument, we prove a more general result, showing that actually the expectation of any bounded function of the degrees of a vertex sampled according to $\tilde{\mu}$ converges in probability to an explicit function of the degree sequences. This fact is formalized in Proposition 2.7.3. Before proceeding with the statement, it is convenient to provide the following definition.

Definition 2.7.1. For every $n \in \mathbb{N}$ let $(\mathbf{d}^+, \mathbf{d}^-)$ be a degree sequence satisfying Assumption 2.A. Consider the function $g = g_n : \mathbb{N}_0 \times \mathbb{N}_{\geq 2} \to \mathbb{R}$ defined as follows: for all $d^- \geq 0$ and $d^+ \geq 2$

$$g(d^-, d^+) = \frac{d^- |V_{d^-,d^+}|}{m^2} \left(d^- + \frac{\gamma - \rho}{1 - \rho} - 1 \right) , \tag{2.124}$$

where $\rho = \rho_n$ and $\gamma = \gamma_n$ are defined in (2.21), and V_{d^-,d^+} denotes the set of vertices having in- and out-degree equal to d^- and d^+, respectively. ■

Remark 2.7.2. Notice that, thanks to Assumption 2.A

$$n \sum_{d^- \geq 0} \sum_{d^+ \geq 2} g(d^-, d^+) = \mathfrak{p} = \Theta(1) , \tag{2.125}$$

where $\mathfrak{p}$ has been defined in (2.23). Moreover, if the graph is Eulerian, then $\gamma = 1$, and the function is non zero only on the diagonal, where

$$g_n(d, d) = \frac{d^2 |V_d|}{m^2} . \tag{2.126}$$

If in particular the graph is regular of degree d, the latter is the indicator function at d, divided by n.

Proposition 2.7.3. *For every $n \in \mathbb{N}$ let $(\mathbf{d}^+, \mathbf{d}^-)$ be a degree sequence satisfying Assumption 2.A. Then, for every bounded function $f : \mathbb{N}_0 \times \mathbb{N}_{\geq 2} \to \mathbb{R}$ it holds that*

$$\left| n \sum_{x \in [n]} \pi^2(x) f(d_x^-, d_x^+) - n \sum_{d^- \geq 0} \sum_{d^+ \geq 2} g(d^-, d^+) f(d^-, d^+) \right| \xrightarrow{\mathbb{P}} 0 , \tag{2.127}$$

where $g = g_n$ is as in Definition 2.7.1.

The proof of Proposition 2.7.3 is postponed to Section 2.7. We now show how to derive from it the proof of Proposition 2.5.6.

Proof of Propositions 2.5.5 and 2.5.6. By Proposition 2.7.3, choosing $f(d^-, d^+) \equiv 1$ (cf. (2.125)) we have

$$\left| n \sum_{x \in [n]} \pi^2(x) - \mathfrak{p} \right| \xrightarrow{\mathbb{P}} 0 \,. \tag{2.128}$$

Similarly, choosing $f(d^-, d^+) = \mathbf{1}_{d^+ \geq 2}/d^+$,

$$\left| n \sum_{x \in [n]} \pi^2(x) \frac{1}{d_x^+} - n \sum_{d^- \geq 1} \sum_{d^+ \geq 2} g(d^-, d^+) \frac{1}{d^+} \right| \xrightarrow{\mathbb{P}} 0 \,. \tag{2.129}$$

In conclusion, by Definition 2.7.1 and (2.24) we have

$$\left| \frac{1}{\langle \pi^2 \rangle} \sum_{x \in [n]} \pi^2(x) \frac{1}{d_x^+} - \mathfrak{q} \right| \xrightarrow{\mathbb{P}} 0 \,. \tag{2.130}$$

Recall the definition of the sets $(\mathcal{M}(q))_{q \geq 0}$ in (2.100). Similarly, for any $\varepsilon > 0$, define the set

$$\mathcal{M}_\varepsilon(q) := \left\{ \mu \in \mathcal{P}([n]) \mid \sum_{x \in [n]} \mu(x) \frac{1}{d_x^+} \in [q - \varepsilon, q + \varepsilon] \right\} \,. \tag{2.131}$$

By (2.130), for any $\varepsilon > 0$, if

$$\mathcal{G}_\varepsilon = \{ \tilde{\mu} \in \mathcal{M}_\varepsilon(\mathfrak{q}) \} \,,$$

then we have $\mathbb{P}(\mathcal{G}_\varepsilon) = 1 - o(1)$. As a consequence,

$$\begin{aligned} \mathbb{P}\left(\left| R^{\tilde{\mu}} - (1 + \Phi(\rho, \mathfrak{q})) \right| > \varepsilon \right) &= \mathbb{P}\left(\left| R^{\tilde{\mu}} - (1 + \Phi(\rho, \mathfrak{q})) \right| > \varepsilon \,,\ \mathcal{G}_\varepsilon \right) + o(1) \\ &\leq \sup_{\mu \in \mathcal{M}_\varepsilon(\mathfrak{q})} \mathbb{P}\left(\left| R^{\mu} - (1 + \Phi(\rho, \mathfrak{q})) \right| > \varepsilon \right) + o(1) \,, \end{aligned}$$

hence, the desired result follows by the continuity of the function Φ in its second variable, which can be checked immediately thanks to representation in Lemma 2.10.1.

□

Proof of Proposition 2.7.3

We begin this section by explaining the strategy of proof and all the required technical ingredients. We take inspiration from [CQ21a, Lemma 4.1] and [QS23, Lemma 4], where the authors where interested only in determining a first order bound on the quantity $\langle \pi^2 \rangle$ for the random walk on the DCM and on the related random Deterministic Finite Automata (DFA) model, respectively.

The first ingredient in the proof is an approximation of π with the T (defined as in (2.96)) step evolution[1] of the simple random walk started at the uniform distribution,

$$\mu_T(x) = \mu_{n,T}(x) \coloneqq \frac{1}{n} \sum_{y \in [n]} P^T(y,x) \qquad x \in [n] \,. \tag{2.132}$$

Indeed, as a corollary of Theorem 2.4.3, we have (see also [CCPQ23, Corollary 3.7]) that the event

$$\mathcal{H} = \mathcal{H}_n \coloneqq \left\{ \max_{x \in [n]} \left| \pi(x) - \frac{1}{n} \sum_{y \in [n]} P^T(y,x) \right| \leq e^{-\log(n)^{3/2}} \right\} , \tag{2.133}$$

satisfies

$$\mathbb{P}(\mathcal{H}) = 1 - o(1) \,. \tag{2.134}$$

For each $d^- \geq 0$ and $d^+ \geq 2$ and any bounded function $f : \mathbb{N}_0 \times \mathbb{N}_{\geq 2} \to \mathbb{R}$ we consider the random variables

$$Z_{d^-,d^+} = \sum_{x \in V_{d^-,d^+}} \pi(x)^2 f(d^-,d^+) \,. \tag{2.135}$$

Notice that the quantity of interest in (2.127) can be rewritten as

$$\sum_{x \in [n]} \pi(x)^2 f(d_x^-, d_x^+) = \sum_{d^- \geq 0} \sum_{d^+ \geq 2} Z_{d^-,d^+} \,. \tag{2.136}$$

We also consider the collection of random variables

$$Y_x \coloneqq \frac{1}{n^2} \sum_{y,z \in [n]} P^T(y,x) P^T(z,x) f(d_x^-, d_x^+) \,, \qquad x \in [n] \,. \tag{2.137}$$

By definition of the event $\mathcal{H}$ in (2.133), for all n sufficiently large,

$$\mathbb{1}(\mathcal{H}) \left| \sum_{d^- \geq 0} \sum_{d^+ \geq 2} Z_{d^-,d^+} - \sum_{x \in [n]} Y_x \right| \leq e^{-\log(n)^{4/3}} \,, \qquad \mathbb{P}-\text{a.s.} \,. \tag{2.138}$$

[1] In [CQ21a, QS23], T is chosen $\log^3(n)$, which is enough to guarantee that the ℓ^∞ distance between π and λ is $o(n^{-c})$ for every $c > 0$. Since the bounds we are going to show are not affected by polylogarithmic factors, we prefer to use the same T as in (2.96), in order to avoid the introduction of further notation.

Hence, the proof of (2.127) comes as a consequence of the following convergence

$$\widehat{F} = \widehat{F}_n := \left| n \sum_{x\in[n]} Y_x - n \sum_{d^-\geq 0} \sum_{d^+\geq 2} g(d^-,d^+) f(d^-,d^+) \right| \xrightarrow{\mathbb{P}} 0\,. \tag{2.139}$$

Indeed, calling

$$F = F_n := \left| n \sum_{x\in[n]} \pi^2(x) f(d_x^-,d_x^+) - n \sum_{d^-\geq 0} \sum_{d^+\geq 2} g(d^-,d^+) f(d^-,d^+) \right|, \tag{2.140}$$

thanks to (2.134) and (2.138), we have for all $\varepsilon > 0$

$$\mathbb{P}(F>\varepsilon) \leq \mathbb{P}(F>\varepsilon\,,\mathcal{H}) + \mathbb{P}(\mathcal{H}^c) \leq \mathbb{P}\left(\widehat{F}>\varepsilon/2\,,\mathcal{H}\right) + o(1) \leq \mathbb{P}\left(\widehat{F}>\varepsilon/2\right) + o(1)\,. \tag{2.141}$$

Therefore, the rest of the proof is devoted to establishing the convergence in (2.139). To this aim, we will use the second moment method for the random variable $n\sum_{x\in[n]} Y_x$, showing that

$$\mathbb{E}\left[n \sum_{x\in[n]} Y_x \right] = (1+o(1))\, n \sum_{d^-\geq 1} \sum_{d^+\geq 2} g_n(d^-,d^+) f(d^-,d^+)\,, \tag{2.142}$$

and

$$\mathbb{E}\left[\left(n \sum_{x\in[n]} Y_x \right)^2 \right] = (1+o(1))\, \mathbb{E}\left[n \sum_{x\in[n]} Y_x \right]^2, \tag{2.143}$$

from which (2.139) follows immediately by Chebyshev inequality. Let us start by computing the expectation

$$\mathbb{E}\left[\sum_{x\in[n]} Y_x \right] = \frac{1}{n^2} \sum_{y,z\in[n]} \mathbb{E}\left[P^T(y,x)P^T(z,x)\right] f(d_x^-,d_x^+) = \sum_{x\in[n]} \mathbb{E}\left[\mu_T(x)^2\right] f(d_x^-,d_x^+)\,, \tag{2.144}$$

where we recall to the reader the definition of μ_T is in (2.132). Note that, due to the symmetry of the model, if $x, x' \in V_{d^-,d^+}$ for some $d^-\geq 0$ and $d^+\geq 2$, then

$$\mathbb{E}\left[\mu_T(x)\right] = \mathbb{E}\left[\mu_T(x')\right]. \tag{2.145}$$

Therefore, (2.144) can be rewritten as

$$\mathbb{E}\left[\sum_{x\in[n]} Y_x \right] = \sum_{d^-\geq 1} \sum_{d^+\geq 2} |V_{d^-,d^+}|\, f(d^-,d^+)\, \mathbb{E}\left[\mu_T\big(x(d^-,d^+)\big)^2 \right], \tag{2.146}$$

where $x(d^-, d^+)$ denotes an arbitrary representative of the vertices with degrees (d^-, d^+). Notice that we can neglect the case $d^- = 0$, since for every x such that $d_x^- = 0$ it holds $\mu_T(x) = 0$, $\mathbb{P}$ – a.s.

We will show the validity of the following lemma, from which (2.142) follows immediately due to (2.146) and the definition of g in (2.124).

Lemma 2.7.4. *Fixed $d^- \geq 1$, $d^+ \geq 2$ and $x \in V_{d^-,d^+}$ it holds*

$$\mathbb{E}\big[\mu_T(x)^2\big] = (1 + o(1))\,\frac{d^-}{m^2}\left(d^- + \frac{\gamma - \rho}{1 - \rho} - 1\right). \tag{2.147}$$

Proof. We use again an *annealing argument*, this time for the simple random walk. We start by defining the annealed walks. Notice that

$$\mathbb{E}\big[\mu_T(x)^2\big] = \mathbb{P}^{2\text{-an}}_{\text{unif}}(W_t^{(1)} = W_t^{(2)} = x)\,, \qquad x \in [n]\,, \tag{2.148}$$

where $\mathbb{P}^{2\text{-an}}_{\text{unif}}$ is the law of a two random walk, $(W_t^{(1)}, W_t^{(2)})_{t \leq T}$ running one after the other, both starting (independently) at a uniform vertex, and creating the environment along with their trajectory, as explained in Remark 2.6.3. More precisely, samples according to $\mathbb{P}^{2\text{-an}}_{\text{unif}}$ can be obtained as follows:

(1) sample $W_0^{(1)}$ u.a.r. in $[n]$;

(2) for all $s \in \{1, \dots, T\}$, choose $e \in E^+_{W_{s-1}}$ u.a.r. :

  (2a) if e is already matched to some $f \in E^-$, call v_f the vertex incident to the head f, and set $W_s^{(1)} = v_f$;

  (2b) if e is unmatched, choose u.a.r. some $f \in E^-$ which is still unmatched, set $\omega(e) = f$, call v_f the vertex incident to f and set $W_s^{(1)} = v_f$;

(3) once $(W_t^{(1)})_{t \leq T}$ has been sampled, sample $(W^{(2)})_{t \leq T}$ in the same way, starting from the partial environment constructed by trajectory of $W^{(1)}$.

Similar to what was done in Section 2.6.1, we now aim at coupling steps (1) and (2) of the previous randomized algorithm with an i.i.d. process.

Let $(U_t)_{t \leq T}$ be an independent sequence of random variables in $[n]$ such that U_0 is uniformly distributed on $[n]$, and for all $t \geq 1$, $U_t \overset{d}{=} \mu_{\text{in}}$, with μ_{in} as in (2.20). Such an i.i.d. process can be coupled with the annealed walk $(W_t^{(1)})_{t \leq T}$ in a joint probability space $\mathbb{Q}$ that can be described as follows:

(a) let $W_0^{(1)} = U_0$;

(b) for $t \geq 1$, sample $U_t \sim \mu_{\text{in}}$:

(ii a) if $U_t \neq U_j$ for all $j < t$, sample $e \in E^+_{W^{(1)}_{t-1}}$ u.a.r. and match it with a uniform random head $f \in E^-_{U_t}$ and declare $W^{(1)}_t = U_t$;

(ii b) if $U_t = U_s$ for some $s < t$, select a uniform random head $f \in E^-_{U_t}$:

- if f is already matched, declare the coupling as *failed*, and continue the two constructions independently;
- if f is still unmatched, sample a uniformly random tail $e \in E^+_{W^{(1)}_{t-1}}$, match it with f and set $W^{(1)}_t = W^{(1)}_s = U_t = U_s$, declare the coupling as *failed* and continue the two constructions independently.

After the realization of the coupled process just described, the trajectory of the second annealed walk, $(W^{(2)}_t)_{t\le T}$ will be sampled according to the usual procedure, conditioned on $(W^{(1)}_t)_{t\le T}$. We will use the same symbol $\mathbb{Q}$ the refer to such an enlarged probability space, sufficiently rich to include the coupled generation of $(U_t)_{t\le T}$ and $(W^{(1)}_t)_{t\le T}$ and that of $(W^{(2)}_t)_{t\le T}$. Call

$$\mathcal{D} := \{W^{(1)}_T = x\}\,, \tag{2.149}$$

and $\mathcal{F}$ the event in which the coupling fails. We start by showing that

$$\mathbb{Q}(\mathcal{D}) = (1 + o(1))\frac{d^-}{m}\,. \tag{2.150}$$

and

$$\mathbb{Q}(\mathcal{F} \mid \mathcal{D}) = O\left(\frac{T^3\, d^-_{\max}}{n}\right)\,. \tag{2.151}$$

Call $\mathcal{E}_x$ the event in which the vertex x is visited by the first annealed trajectory more than once, i.e.,

$$\mathcal{E}_x := \{\exists s, t \le T \text{ s.t. } W^{(1)}_s = W^{(1)}_t = x\,,\ s \neq t\}\,. \tag{2.152}$$

There are two possible ways in which the latter event can happen: either $W^{(1)}_0 = x$, and at time $t < T$ vertex x is visited again; this happens with probability at most

$$\frac{1}{n}\,\frac{d^-}{m - T}\,T\,.$$

On the other had, if the initial position of $W^{(1)}$ is different from x, in order to realize $\mathcal{E}_x$, x must be visited once by matching one of its heads and, subsequently, the walk has to visit again one of the vertices already visited. The probability of this event can be bounded by

$$\frac{d^-}{m - T}\,\frac{d^-_{\max} - 1}{m - T}\,T^2\,.$$

Combining the two bounds we get, for all n large enough,

$$\mathbb{P}^{\text{an}}_{\text{unif}}(\mathcal{E}_x) = O(p_1)\,, \qquad \text{where} \qquad p_1 \coloneqq \frac{T^2\, d^-\, d^-_{\max}}{n^2}\,. \tag{2.153}$$

Therefore

$$\mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D}) = \mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D} \cap \mathcal{E}_x^c) + O(p_1)\,. \tag{2.154}$$

Consider also the event $\bar{\mathcal{E}}_x$ in which there exists some other vertex that is visited more than once, i.e.,

$$\bar{\mathcal{E}}_x \coloneqq \cup_{y\in[n]\setminus x}\, \mathcal{E}_y\,.$$

Hence,

$$\mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D}) = \mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D} \cap \mathcal{E}_x^c \cap \bar{\mathcal{E}}_x^c) + \mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D} \cap \mathcal{E}_x^c \cap \bar{\mathcal{E}}_x) + O(p_1)\,. \tag{2.155}$$

We can further bound

$$\mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D} \cap \mathcal{E}_x^c \cap \bar{\mathcal{E}}_x) = O(p_2)\,, \qquad \text{where} \qquad p_2 \coloneqq \frac{T^3\, d^-\, d^-_{\max}}{n^2}\,. \tag{2.156}$$

Indeed, the probability that $y \neq x$ is visited at least twice by the annealed walk can be bounded analogously to (2.153). Moreover, under this event it has to visit x at time T for the first time. Hence by a union bound

$$\mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D} \cap \mathcal{E}_x^c \cap \bar{\mathcal{E}}_x) \le \sum_{y\in[n]} T\, \frac{d^-_y}{m-T} \times T^2\, \frac{d^-_{\max}}{m-T} \times \frac{d^-}{m-T}\,, \tag{2.157}$$

from which (2.156) follows.

It follows that

$$\mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D}) = \mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D}\,,\ \mathcal{E}_x^c \cap \bar{\mathcal{E}}_x^c) + O(p_1 + p_2)\,. \tag{2.158}$$

Finally, notice that

$$\mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D} \cap \mathcal{E}_x^c \cap \bar{\mathcal{E}}_x^c) \le \frac{d^-}{m-T} = (1+o(1))\frac{d^-}{m}\,, \tag{2.159}$$

and

$$\mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D} \cap \mathcal{E}_x^c \cap \bar{\mathcal{E}}_x^c) \ge \left(1 - \frac{T d^-_{\max}}{m-T}\right)^T \frac{d^-}{m-T} = (1+o(1))\frac{d^-}{m}\,. \tag{2.160}$$

From (2.159) and (2.160) it follows that

$$\mathbb{P}^{2\text{-an}}_{\text{unif}}(\mathcal{D} \cap \mathcal{E}_x^c \cap \mathcal{E}^c) = (1+o(1))\frac{d^-}{m}\,. \tag{2.161}$$

Then (2.150) follows by plugging (2.161) into (2.158) and noting that $p_1,\ p_2 = o(n^{-1})$.

At this point, to prove (2.151) it suffices to note that

$$\mathbb{Q}(\mathcal{F} \mid \mathcal{D}) = \frac{\mathbb{Q}(\mathcal{D} \cap (\mathcal{E}_x \cup \bar{\mathcal{E}}_x))}{\mathbb{Q}(\mathcal{D})} = O\left(\frac{n}{d^-}\ (p_1 + p_2)\right) = O\left(\frac{T^3\ d^-_{\max}}{n}\right), \quad (2.162)$$

where we used (2.150) to estimate the denominator and for the numerator we bounded

$$\mathbb{Q}(\mathcal{D} \cap (\mathcal{E}_x \cup \bar{\mathcal{E}}_x)) \le \mathbb{Q}(\mathcal{E}_x) + \mathbb{Q}(\mathcal{D} \cap \bar{\mathcal{E}}_x \cap \mathcal{E}_x^c), \quad (2.163)$$

and then used (2.153) and (2.156).

We go back to estimating the right hand side of (2.148) using the coupling. Call $\mathcal{G}$ the event in which the second walk enters the trajectory of the first one, follows it, and at a certain time exits it, i.e.

$$\mathcal{G} := \bigcup_{0 \le s < s' < s'' \le T} \left(\mathcal{A}_s \cap \mathcal{B}_{s'} \cap C_{s''}\right),$$

where

$$\mathcal{A}_s = \{W_s^{(2)} \in \{W_t^{(1)}\}_{t \le T}\},\ \mathcal{B}_{s'} = \{W_{s'}^{(2)} \notin \{W_t^{(1)}\}_{t \le T}\},\ C_{s''} = \{W_{s''}^{(2)} \in \{W_t^{(1)}\}_{t \le T}\}.$$

For any $\sigma \in \mathcal{D}$ we have that

$$\mathbb{Q}(\mathcal{G} \mid \sigma) \le \sum_{0 \le s < s' < s'' \le T} \mathbb{Q}(\mathcal{A}_s \mid \sigma)\ \mathbb{Q}(\mathcal{B}_{s'} \mid \mathcal{A}_s, \sigma)\ \mathbb{Q}(C_{s''} \mid \mathcal{B}_{s'}, \mathcal{A}_s, \sigma),$$

where the single probabilities can be bounded as follows

$$\begin{aligned} \mathbb{Q}(\mathcal{A}_s \mid \sigma) &\le \frac{T}{n} + T^2\ \frac{d^-_{\max}}{m - 2T}, \\ \mathbb{Q}(\mathcal{B}_{s'} \mid \mathcal{A}_s,\ \sigma) &\le 1, \\ \mathbb{Q}(C_{s''} \mid \mathcal{B}_{s'} \cap \mathcal{A}_s,\ \sigma) &\le T^2\ \frac{d^-_{\max}}{m - 2T}. \end{aligned}$$

Hence, taking the supremum over $\sigma \in \mathcal{D}$,

$$\sup_{\sigma \in \mathcal{D}} \mathbb{Q}(\mathcal{G} \mid \sigma) = O\left(\frac{T^5\ (d^-_{\max})^2}{n^2}\right). \quad (2.164)$$

To ease the reading, for $t \in [0, T]$ and $\sigma \in \mathcal{D}$ let us consider the function $\sigma \mapsto q_t(\sigma) \in [0, 1]$ such that

$$\begin{aligned} q_t(\sigma) &:= \mathbb{Q}(\mathcal{G}^c,\ W_{t-1}^{(2)} \ne W_{t-1}^{(1)},\ W_s^{(2)} = W_s^{(1)},\ s \in [t, T] \mid \sigma) \quad (2.165) \\ &= \mathbb{Q}(W_{s'}^{(2)} \notin \{W_t^{(1)}\}_{t \le T},\ s' \in [0, t) \text{ and } W_s^{(2)} = W_s^{(1)},\ s \in [t, T] \mid \sigma), \quad (2.166) \end{aligned}$$

where, by convention, we consider $\{W_{-1}^{(1)} \neq W_{-1}^{(2)}\}$ as the sure event. Further, define the same probability conditioned to the event $\mathcal{D}$ as

$$q_t := \mathbb{Q}(W_{s'}^{(2)} \notin \{W_t^{(1)}\}_{t\le T},\ s' \in [0,t) \text{ and } W_s^{(2)} = X_s^{(1)},\ s \in [t,T] \mid \mathcal{D})\,, \quad (2.167)$$

so that

$$q_t = \sum_{\sigma\in\mathcal{D}} \mathbb{Q}(\sigma \mid \mathcal{D})\, q_t(\sigma) = \hat{\mathbb{E}}[q_t(\sigma) \mid \mathcal{D}]\,, \quad (2.168)$$

where $\hat{\mathbb{E}}$ is the expectation in the conditional probability space $\mathbb{Q}(\,\cdot \mid \mathcal{D})$. As a consequence of (2.164), for every $\sigma \in \mathcal{D}$,

$$\mathbb{Q}(W_T^{(2)} = x \mid \sigma) = \sum_{t=0}^{T} q_t(\sigma) + O\left(\frac{T^5\,(d^-_{\max})^2}{n^2}\right). \quad (2.169)$$

Therefore, using Assumption 2.B we get

$$\mathbb{Q}(W_T^{(1)} = W_T^{(2)} = x) = \mathbb{Q}(\mathcal{D})\left(\hat{\mathbb{E}}\Big[\sum_{t=0}^{T} q_t(\sigma) \mid \mathcal{D}\Big] + o(n^{-1})\right). \quad (2.170)$$

Hence Lemma 2.7.4 follows at once by (2.150) and (2.170) as soon as we prove that

$$\sum_{t=0}^{T} q_t = (1+o(1))\frac{1}{m}\left(d^- - 1 + \frac{\gamma-\rho}{1-\rho}\right). \quad (2.171)$$

For every $t \le T$ and any $\sigma \in \mathcal{D}$, define

$$\delta_t(\sigma) := \mathbb{Q}\left(\bigcap_{s=0}^{t-1} W_{s'}^{(2)} \notin \{W_t^{(1)}\}_{t\le T} \mid \sigma\right) \ge \left(1-\frac{T}{n}\right)\left(1-\frac{3\,T\,d^-_{\max}}{m}\right)^T = 1-o(1)\,,$$

where the inequality holds for all n large enough. Notice also that, calling $d_s^{\pm} = d^{\pm}_{X_s^{(1)}}$ for all $s \in \{0,\dots,T\}$, one has

$$q_t(\sigma) = \begin{cases} \delta_T(\sigma)\frac{d^- -1}{m-2T} = (1+o(1))\,\frac{d^- -1}{m} & \text{if } t = T\,,\\ \delta_t(\sigma)\frac{d_t^- -1}{m-T-t}\prod_{s=t}^{T-1}\frac{1}{d_s^+} = (1+o(1))\,\frac{d_t^- -1}{m}\prod_{s=t}^{T-1}\frac{1}{d_s^+} & \text{if } t \in \{0,\dots,T-1\}\,,\end{cases}$$

in particular

$$q_T = \hat{\mathbb{E}}[q_T(\sigma) \mid \mathcal{D}] = (1+o(1))\,\frac{d^- -1}{m}\,. \quad (2.172)$$

For $t < T$, we split

$$\hat{\mathbb{E}}[q_t(\sigma) \mid \mathcal{D}] = \hat{\mathbb{E}}[q_t(\sigma)\mathbb{1}(\mathcal{F}^c) \mid \mathcal{D}] + \hat{\mathbb{E}}[q_t(\sigma)\mathbb{1}(\mathcal{F}) \mid \mathcal{D}]\,. \quad (2.173)$$

To bound the above two terms on the rhs we use (2.151) and Assumption 2.A. The first term is bounded as follows

$$\begin{aligned}\hat{\mathbb{E}}[q_t(\sigma)\mathbb{1}(\mathcal{F}) \mid \mathcal{D}] &\le \frac{d^-_{\max}}{m}\, 2^{-T+t}\, \mathbb{Q}(\mathcal{F} \mid \mathcal{D}) \\ &= O\left(T^3\, \frac{(d^-_{\max})^2}{n^2} 2^{-T+t}\right) = o\big(n^{-1} \times 2^{-T+t}\big)\,.\end{aligned} \tag{2.174}$$

The second term can be written as follows

$$\begin{aligned}\hat{\mathbb{E}}[q_t(\sigma)\mathbb{1}(\mathcal{F}^c) \mid \mathcal{D}] &= (1+o(1))\, \hat{\mathbb{E}}\left[\frac{d^-_{U_t}-1}{m} \prod_{s=t}^{T-1} \frac{1}{d^+_{U_s}} \mathbb{1}(\mathcal{F}^c) \mid \mathcal{D}\right] \\ &= (1+o(1))\left\{\hat{\mathbb{E}}\left[\frac{d^-_{U_t}-1}{m} \prod_{s=t}^{T-1} \frac{1}{d^+_{U_s}} \mid \mathcal{D}\right] - \hat{\mathbb{E}}\left[\frac{d^-_{U_t}-1}{m} \prod_{s=t}^{T-1} \frac{1}{d^+_{U_s}} \mathbb{1}(\mathcal{F}) \mid \mathcal{D}\right]\right\} \\ &= (1+o(1))\, \hat{\mathbb{E}}\left[\frac{d^-_{U_t}-1}{m} \prod_{s=t}^{T-1} \frac{1}{d^+_{U_s}}\right] - O\left(\frac{d^-_{\max}}{m}\, 2^{-T+t}\, \mathbb{Q}(\mathcal{F} \mid \mathcal{D})\right) \\ &= (1+o(1))\, \hat{\mathbb{E}}\left[\frac{d^-_{U_t}-1}{m} \prod_{s=t}^{T-1} \frac{1}{d^+_{U_s}}\right] - o\left(n^{-1} \times 2^{-T+t}\right)\,.\end{aligned} \tag{2.175}$$

It is easy to check that, if $U \sim \mu^{\mathrm{in}}$,

$$\hat{\mathbb{E}}\left[\frac{1}{d^+_U}\right] = \rho\,, \qquad \hat{\mathbb{E}}\left[\frac{d^-_U - 1}{d^+_U}\right] = \frac{1}{m} \sum_{x \in [n]} \frac{d^-_x (d^-_x - 1)}{d^+_x} = \gamma - \rho\,. \tag{2.176}$$

Therefore, putting together (2.172), (2.173), (2.174), (2.175) and (2.176), we deduce

$$\begin{aligned}\hat{\mathbb{E}}\left[\sum_{t=0}^{T} q_t(\sigma) \mid \mathcal{D}\right] &= (1+o(1))\left(\frac{d^- - 1}{m} + \hat{\mathbb{E}}\left[\sum_{t=0}^{T-1} \frac{d^-_{U_t}-1}{m} \prod_{s=t}^{T-1} \frac{1}{d^+_{U_s}}\right]\right) + o(n^{-1}) \\ &= (1+o(1))\left(\frac{d^- - 1}{m} + \frac{\gamma-\rho}{m} \sum_{t=0}^{T-1} \rho^t\right) \sim \frac{1}{m}\left(d^- - 1 + \frac{\gamma-\rho}{1-\rho}\right)\,,\end{aligned} \tag{2.177}$$

hence proving the validity of (2.171). □

To conclude the section, we are now left to show (2.143), which follows immediately by the following lemma.

Lemma 2.7.5. *Recall the definition of* $(Y_x)_{x\in[n]}$ *in* (2.137)*. It holds*

$$\lim_{n\to\infty} \frac{\mathbb{E}\left[\left(\sum_{x\in[n]} Y_x\right)^2\right]}{\mathbb{E}\left[\sum_{x\in[n]} Y_x\right]^2} = 1\,. \tag{2.178}$$

Proof. Start by noting that it suffices to show that

$$\mathbb{E}\Big[\Big(\sum_{x\in[n]} Y_x\Big)^2\Big] \le (1+o(1))\,\mathbb{E}\Big[\sum_{x\in[n]} Y_x\Big]^2, \tag{2.179}$$

since the other inequality is trivially true. Rewrite

$$\mathbb{E}\Big[\Big(\sum_{x\in[n]} Y_x\Big)^2\Big] = \sum_{x\in[n]}\sum_{y\in[n]} \mathbb{E}[Y_x Y_y] = \sum_{x\in[n]}\sum_{y\in[n]} \mathbb{E}[\mu_T(x)^2\mu_T(y)^2]\, f(d_x^-, d_x^+)\, f(d_y^-, d_y^+), \tag{2.180}$$

and consider the annealed process described in the proof of Lemma 2.7.4, but this time, consider also the two additional annealed trajectories $\{W_t^{(3)}\}_{t\le T}$ and $\{W_t^{(4)}\}_{t\le T}$, to be generated, sequentially, after the generation of $\{W_t^{(1)}\}_{t\le T}$ and $\{W_t^{(2)}\}_{t\le T}$. We have

$$\mathbb{E}[Y_x Y_y] = \mathbb{P}^{4\text{-an}}_{\text{unif}}\Big(W_T^{(1)} = W_T^{(2)} = x\,,\, W_T^{(3)} = W_T^{(4)} = y\Big)\; f(d_x^-, d_x^+)\; f(d_y^-, d_y^+)\,. \tag{2.181}$$

Let

$$\mathcal{N}_x := \Big\{W_T^{(1)} = W_T^{(2)} = x\Big\}\,, \quad \text{and} \quad \mathcal{I}_y := \Big\{W_T^{(3)} = W_T^{(4)} = y\Big\}\,. \tag{2.182}$$

Start by considering the case $x \neq y$, and let $\mathcal{J}$ be the event where the trajectories of $(W_t^{(1)}, W_t^{(2)})_{t\le T}$ hit the set of trajectories of $(W_t^{(3)}, W_t^{(4)})_{t\le T}$, which we formally write as follows:

$$\mathcal{J} := \Big\{\exists s,t \le T \text{ and } \exists i\in\{1,2\},\; j\in\{3,4\} \text{ s.t. } W_s^{(i)} = W_t^{(j)}\Big\}\,. \tag{2.183}$$

We claim that, uniformly in $x \neq y$

$$\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{I}_y \cap \mathcal{J}) = o\big(\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x)\,\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{I}_y)\big)\,, \tag{2.184}$$

and

$$\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{I}_y \cap \mathcal{J}^c) \le \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x)\,\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{I}_y)\,. \tag{2.185}$$

Assuming the validity of (2.184) and (2.185), from (2.180) and (2.181) we deduce

$$\begin{aligned}\mathbb{E}\Big[\Big(\sum_{x\in[n]} Y_x\Big)^2\Big] &\le (1+o(1)) \sum_{x\in[n]}\sum_{\substack{y\in[n]\\ y\neq x}} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x)\,\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{I}_y) f(d_x^-, d_x^+) f(d_y^-, d_y^+) \\ &\qquad + \sum_x \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{I}_x) f(d_x^-, d_x^+)^2 \\ &\le (1+o(1))\mathbb{E}\Big[\sum_{x\in[n]} Y_x\Big]^2 + O\left(\sum_x \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{I}_x)\right).\end{aligned} \tag{2.186}$$

To bound the error term in (2.186), notice that for all $x \in [n]$,

$$\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{I}_x) \leq \frac{T\, d_x^-}{m-T} \times \frac{T^2\, d_{\max}^-}{m-2T} \times \frac{2\,T^2\, d_{\max}^-}{m-3T} \times \frac{3\,T^2\, d_{\max}^-}{m-4T} = O\left(\frac{T^7\, d_x^-\, (d_{\max}^-)^3}{n^4}\right). \tag{2.187}$$

Thus the validity of (2.179) follows, since, thanks to the fact that f is bounded above,

$$\mathbb{E}\Big[\Big(\sum_{x\in[n]} Y_x\Big)^2\Big] \leq (1+o(1))\mathbb{E}\Big[\sum_{x\in[n]} Y_x\Big]^2 + O\left(\frac{T^7\,(d_{\max}^-)^3}{n^3}\right) \leq (1+o(1))\mathbb{E}\Big[\sum_{x\in[n]} Y_x\Big]^2, \tag{2.188}$$

where in the last step we used Assumption 2.A and $\mathbb{E}\big[\sum_{x\in[n]} Y_x\big] = O(n^{-1})$.

We are left to show (2.184) and (2.185). To prove (2.184), call

$$\mathcal{Q}_y = \Big\{\exists s \leq T \text{ and } i \in \{1,2\} \text{ s.t. } X_s^{(i)} = y\Big\}, \tag{2.189}$$

let σ be a realization of $\{W_t^{(1)}\}_{t\leq T}$ and $\{W_t^{(2)}\}_{t\leq T}$ satisfying $\mathcal{N}_x$, and rewrite

$$\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{I}_y \cap \mathcal{J}) = \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{Q}_y \cap \mathcal{I}_y \cap \mathcal{J}) + \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{Q}_y^c \cap \mathcal{I}_y \cap \mathcal{J}). \tag{2.190}$$

The first term on the right-hand side of (2.190) can be bounded as follows

$$\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{Q}_y \cap \mathcal{I}_y \cap \mathcal{J}) \leq \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{Q}_y) \sup_{\sigma\in\mathcal{N}_x\cap\mathcal{Q}_y} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{I}_y \mid \sigma). \tag{2.191}$$

Notice that the event on the first probability on the right-hand side of (2.191) can be bounded by

$$\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{Q}_y) = O\left(\frac{T^4\, d_x^-\, d_y^-\, d_{\max}^-}{n^3}\right). \tag{2.192}$$

Similarly, the conditional probability on the right-hand side of (2.191) can be bounded uniformly in $\sigma \in \mathcal{N}_x \cap \mathcal{Q}_y$ by

$$\sup_{\sigma\in\mathcal{N}_x\cap\mathcal{Q}_y} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{I}_y \mid \sigma) = O\left(\frac{T^4\,(d_{\max}^-)^2}{m^2}\right). \tag{2.193}$$

Plugging (2.192) and (2.193) into (2.191) we deduce

$$\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{Q}_y \cap \mathcal{I}_y \cap \mathcal{J}) = O\left(\frac{T^8\, d_x^-\, d_y^-\, (d_{\max}^-)^3}{n^5}\right) = o\big(\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x)\, \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{I}_y)\big), \tag{2.194}$$

where in the last estimate we used Assumption 2.A and Lemma 2.7.4.

As for the second term on the right-hand side of (2.190), we can start by splitting

$$\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{Q}_y^c \cap \mathcal{I}_y \cap \mathcal{J}) \leq \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x) \sup_{\sigma\in\mathcal{N}_x\cap\mathcal{Q}_y^c} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{I}_y \cap \mathcal{J} \mid \sigma). \tag{2.195}$$

Then, to bound the latter conditional probability we argue as follow: either $W^{(3)}$ hits the trajectory of $W^{(1)} \cup W^{(2)}$ and y, and $W^{(4)}$ hits $W^{(1)} \cup W^{(2)} \cup W^{(3)}$; or $W^{(3)}$ hits y but not $W^{(1)} \cup W^{(2)}$ and y, and then $W^{(4)}$ hits $W^{(1)} \cup W^{(2)}$ and $W^{(3)}$. In conclusion,

$$\sup_{\sigma \in \mathcal{N}_x \cap Q_y^c} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{I}_y \cap \mathcal{J} \mid \sigma) = O\left(\frac{T^4\, d_y^-\, (d_{\max}^-)^2}{n^3}\right). \tag{2.196}$$

Plugging (2.196) into (2.195) and using Lemma 2.7.4 we conclude

$$\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap Q_y^c \cap \mathcal{I}_y \cap \mathcal{J}) = O\left(\frac{T^4\, d_x^-\, d_y^-\, (d_{\max}^-)^2}{n^5}\right) = o\big(\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x)\, \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{I}_y)\big). \tag{2.197}$$

Thus, (2.184) follows from (2.190), (2.194) and (2.197).

To show (2.185), it is enough to realize that, for each $\sigma \in \mathcal{N}_x$, called $(\mathcal{I}_y \cap \mathcal{J}^c)(\sigma)$ the set of trajectories for $W^{(3)}$ and $W^{(4)}$ that do not intersect σ and satisfy $W_T^{(3)} = W_T^{(4)} = y$, one has

$$\begin{aligned}
\mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x \cap \mathcal{I}_x \cap \mathcal{J}^c) &= \sum_{\sigma \in \mathcal{N}_x} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\sigma) \sum_{\eta \in (\mathcal{I}_y \cap \mathcal{J}^c)(\sigma)} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\eta \mid \sigma) \\
&= \sum_{\sigma \in \mathcal{N}_x} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\sigma) \sum_{\eta \in (\mathcal{I}_y \cap \mathcal{J}^c)(\sigma)} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\eta) \\
&\le \sum_{\sigma \in \mathcal{N}_x} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\sigma) \sum_{\eta \in \mathcal{I}_y} \mathbb{P}^{4\text{-an}}_{\text{unif}}(\eta) = \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{N}_x)\, \mathbb{P}^{4\text{-an}}_{\text{unif}}(\mathcal{I}_y),
\end{aligned} \tag{2.198}$$

where in the second equality in (2.198) we removed the conditioning on σ since, under $\mathcal{J}^c$, the evolution of $W^{(3)}$ and $W^{(4)}$ is not affected by the exact set of vertices that outline the trajectories of $W^{(1)}$ and $W^{(2)}$. This concludes the proof. □

§2.8 Mixing time of random walks with $\tilde{\mu}$-reset

In this section we prove that the mixing time of the auxiliary chain $\tilde{\mathbf{P}}^{\tilde{\mu}}$ on $\tilde{V}$ has polylogarithmic order with respect to the size n of the original network G. The proof technique adapts the one used in [QS23, Section 4.2], where the authors investigate the mixing time of two independent walks moving synchronously on DFA. Recall that $\tilde{P} = \tilde{P}^{\tilde{\mu}}$ refers to the transition matrix of $(\tilde{X})_{t \ge 0}$, the two walks process on G with $\tilde{\mu}$-reset, as defined in (2.77), and $\tilde{\pi}$ to its unique stationary distribution on $\tilde{V}$. Throughout the whole section we set

$$S = S_n := \lceil \log^3(n) \rceil. \tag{2.199}$$

Proposition 2.8.1. *Assume that the degree sequence satisfies Assumption 2.B. Then,*

$$\max_{\boldsymbol{x}\in\tilde{V}} \|\tilde{P}^S(\boldsymbol{x},\cdot) - \tilde{\pi}(\cdot)\|_{\mathrm{TV}} \xrightarrow{\mathbb{P}} 0\,, \tag{2.200}$$

where S is defined as in (2.199).

Recall the definition of the random rooted tree $\mathcal{T}_x^+$, $x \in [n]$, defined in Section 2.4.1, of $\hbar$ in (2.55) and of the set of locally-tree-like vertices $V_\star$ in (2.57). By using the same approach as in Proposition 2.7.3, we start by showing the following property of the distribution $\tilde{\mu}$.

Proposition 2.8.2. *Assume that the degree sequence satisfies Assumption 2.A, and let $V_\star$ as in* (2.57)*. It holds*

$$\tilde{\mu}(V_\star) = \frac{\sum_{x\in V_\star}\pi^2(x)}{\langle\pi^2\rangle} \xrightarrow{\mathbb{P}} 1\,.$$

Proof. It is enough to show the following

$$n\left|\sum_{x\in[n]}\pi^2(x) - \sum_{x\in[n]}\mathbb{1}_{\{x\in V_\star\}}\pi^2(x)\right| \xrightarrow{\mathbb{P}} 0, \tag{2.201}$$

as

$$\tilde{\mu}(V_\star) = 1 - \frac{n\sum_{x\in[n]}\left[\pi^2(x) - \mathbb{1}_{\{x\in V^\star\}}\pi^2(x)\right]}{n\sum_{x\in[n]}\pi^2(x)},$$

and $n\sum_{x\in[n]}\pi^2(x) = \Theta_{\mathbb{P}}(1)$, as proved in Proposition 2.5.5. Fix $x \in [n]$, and consider the approximation μ_T of $\pi(x)$ introduced in (2.132). Recall the definition of the event $\mathcal{H}$ in (2.133) and that, by (2.134), $\mathcal{H}$ holds with high probability. Fixed $\delta > 0$, we get

$$\mathbb{P}\left(n\sum_{x\notin V_\star}\pi^2(x) > \delta\right) = \mathbb{P}\left(n\left|\sum_{x\in[n]}\pi^2(x) - \sum_{x\in[n]}\mathbb{1}_{\{x\in V_\star\}}\pi^2(x)\right| > \delta\,,\,\mathcal{H}\right) + o(1)\,.$$

Consider the events

$$\mathcal{A} = \left\{n\sum_{x\in[n]}|\pi^2(x) - \mu_S^2(x)| > \frac{\delta}{3}\right\}, \quad \mathcal{B} = \left\{n\sum_{x\in[n]}\mathbb{1}_{x\in V_\star}|\pi^2(x) - \mu_S^2(x)| > \frac{\delta}{3}\right\},$$

$$\mathcal{C} = \left\{n\sum_{x\in[n]}\left[\mu_S^2(x) - \mathbb{1}_{x\in V_\star}\mu_S^2(x)\right] > \frac{\delta}{3}\right\}, \quad \mathcal{K} = \left\{n\sum_{x\in[n]}\left[\pi^2(x) - \mathbb{1}_{\{x\in V_\star\}}\pi^2(x)\right] > \delta\right\}.$$

By the triangular inequality we have that

$$\mathbb{P}(\mathcal{K}\,,\,\mathcal{H}) \le \mathbb{P}(\mathcal{A}\,,\,\mathcal{H}) + \mathbb{P}(\mathcal{B}\,,\,\mathcal{H}) + \mathbb{P}(\mathcal{C}\,,\,\mathcal{H}) \le 2\,\mathbb{P}(\mathcal{A}\,,\,\mathcal{H}) + \mathbb{P}(\mathcal{C}\,,\,\mathcal{H})\,. \tag{2.202}$$

The first term on the right hand side of (2.202) can be easily bounded by Markov inequality, indeed

$$\mathbb{P}(\mathcal{A}, \mathcal{H}) \leq \frac{3\,n}{\delta} \sum_{x\in[n]} \mathbb{E}\left[\mathbb{1}_{\mathcal{H}} |\pi(x) - \mu_S(x)||\pi(x) + \mu_S(x)|\right] \leq \frac{6}{\delta}\, n^2\, e^{-\log(n)^{3/2}} = o(1)\,.$$

Therefore we are left to prove that, for any $\delta > 0$,

$$\mathbb{P}(C) = o(1)\,. \tag{2.203}$$

In order to show (2.203), suppose that

$$\mathbb{E}[\mathbb{1}_{\{x\in V_\star\}}\mu_S^2(x)] = (1 - o(1))\mathbb{E}[\mu_S^2(x)] \tag{2.204}$$

for any $x \in [n]$. Then, summing over x and using Lemma 2.7.4, we deduce that $\mathbb{E}[\mu_S^2([n] \setminus V_\star)] = o(n^{-1})$, so that (2.203) follows by Markov inequality.

Therefore, we are left to prove (2.204). To this aim, note that it is enough to show the lower bound, as the upper bound is trivially satisfied.
Let $\gamma_x := \mathcal{B}_x^+(\hbar)$ denote any realization of the complete out-neighbourhood of x up to distance $\hbar$, defined in (2.55), and notice that, under $\mathbb{P}$, the event $\{x \in V_\star\}$ is measurable with respect to γ_x. Thus

$$\mathbb{E}\big[\mathbb{1}_{\{x\in V_\star\}}\mu_S^2(x)\big] = \mathbb{E}\left[\mathbb{1}_{\{x\in V_\star\}}\mathbb{E}[\mu_S^2(x) \mid \gamma_x]\right]. \tag{2.205}$$

In order to study $\mathbb{E}[\mu_S^2(x) \mid \gamma_x]$, we need to employ another annealing argument as in proof of Lemma 2.7.4, but this time the two annealed walks evolve in an environment that starts as γ_x.

Let $\mathbb{P}_{\text{unif}}^{2\text{-an}|\gamma_x}$ denote the law of such a process, as introduced in Remark 2.6.3. In particular,

$$\mathbb{E}[\mu_S^2(x) \mid \gamma_x] = \mathbb{P}_{\text{unif}}^{2\text{-an}|\gamma_x}(W_S^{(1)} = W_S^{(2)} = x)\,.$$

We aim to show that

$$\mathbb{P}_{\text{unif}}^{2\text{-an}|\gamma_x}(W_S^{(1)} = W_S^{(2)} = x) \geq \mathbb{P}_{\text{unif}}^{2\text{-an}}(W_S^{(1)} = W_S^{(2)} = x)(1 - o(1))\,, \tag{2.206}$$

where the annealing law on the right-hand side corresponds to the one in which the initial environment is given by the empty matching of the edges, i.e., that of the process described in the proof of Lemma 2.7.4. If (2.206) holds, then (2.205) reads

$$\mathbb{E}\left[\mathbb{1}_{\{x\in V_\star\}}\mathbb{E}[\mu_S^2(x) \mid \gamma_x]\right] \geq \mathbb{P}(x \in V_\star)\mathbb{E}[\mu_S^2(x)](1 - o(1)) = (1 - o(1))\mathbb{E}[\mu_S^2(x)]\,,$$

since $\mathbb{P}(x \in V_\star) = 1 - o(1)$, as shown in Lemma 2.4.1. To prove (2.206) we construct a coupling of the two annealed processes as follows:

(a) Let $(\bar{W}_t^{(1)}, \bar{W}_t^{(2)})_{t\le S}$ denote the annealed walks having as initial environment γ_x. On the other hand, let $(W_t^{(1)}, W_t^{(2)})_{t\le S}$ denote the annealed random walks having as initial environment the empty matching of the edges. Henceforth, set $\bar{\omega}_0 = \gamma_x$ and $\omega_0 = \emptyset$.

(b) For the first walk, at time $t = 0$, sample $W_0^{(1)} = \bar{W}_0^{(1)}$ according to the uniform distribution on $[n]$. If $W_0^{(1)} \in \gamma_x$ declare the coupling as *failed* and let the two processes evolve independently.

(c) For the first walk and $t \in (0, S)$, construct $W_t^{(1)}$ as for the annealed walk in the proof of Lemma 2.7.4, i.e., sample a tail $e \in E^+_{W_{t-1}^{(1)}}$ u.a.r.:

(c-1) If $\omega_{t-1}(e) = f \in E^-$ set $W_t^{(1)} = \bar{W}_t^{(1)} = v_f$, where v_f is the vertex incident to f;

(c-2) otherwise, select a head f u.a.r. among those that are not matched in ω_{t-1}:

- set $\omega_t(e) = f$ and $W_t^{(1)} = v_f$;
- if $\omega_{t-1}^{-1}(f) = \emptyset$ but $\bar{\omega}_{t-1}^{-1}(f) = e' \in E^+$ or if $v_f \in \gamma_x$, declare the coupling as *failed* and let the processes evolve independently;
- otherwise let $\bar{W}_t = W_t$ and $\bar{\omega}_t(e) = f$.

(d) Follow the same procedure for the walks $W^{(2)}$ and $\bar{W}^{(2)}$.

Call $\hat{\mathbb{P}}$ the joint probability space just described, and note that the marginals of the coupled process corresponds to the laws $\mathbb{P}^{2\text{-an}|\gamma_x}_{\text{unif}}$ and $\mathbb{P}^{2\text{-an}}_{\text{unif}}$, respectively. Let $\mathcal{F}$ be the event that the coupling fails. It holds that

$$\hat{\mathbb{P}}\left(\bigcap_{t=0}^{S}\bigcap_{i\in\{1,2\}}\{W_t^{(i)} = \bar{W}_t^{(i)}\} \mid \mathcal{F}^c\right) = 1.$$

Notice also that, at each step, in order for the coupling to fail, it must be the case that the sample at step (c-2) is a head of a vertex in γ_x. Hence, for each step of the construction, the failure probability of the coupling can be bounded uniformly by

$$\begin{aligned}&\max_{x\in[n]}\max_{\gamma_x}\max_{t\in[0,S]}\max_{i=1,2}\hat{\mathbb{P}}(\text{Coupling fails at step } t \text{ of the } i\text{-th walk}) \le \\ &\le \frac{(d^+_{\max})^\hbar}{n} + \frac{d^-_{\max}\,(d^+_{\max})^\hbar}{m - 2S - (d^+_{\max})^\hbar}\,,\end{aligned} \tag{2.207}$$

hence, by a union bound,

$$\hat{\mathbb{P}}(\mathcal{F}) \le 2S\left(\frac{(d^+_{\max})^\hbar}{n} + \frac{d^-_{\max}\,(d^+_{\max})^\hbar}{m - 2S - (d^+_{\max})^\hbar}\right) = o(1)\,,$$

as $S = \log^3(n)$, $d^-_{\max} = o(n^{1/2-\varepsilon})$, for some $\varepsilon > 0$ and $(d^+_{\max})^{\hbar} = O(n^{1/5})$. As a consequence

$$\mathbb{P}^{2\text{-an}|\gamma_x}_{\text{unif}}(X_S = Y_S = x) \geq \hat{\mathbb{P}}(W^{(1)}_S = W^{(2)}_S = x\,,\ \mathcal{F}^c) = (1 - o(1))\mathbb{P}^{2\text{-an}}_{\text{unif}}(X_S = Y_S = x)\,,$$

from which (2.204), and in turn (2.203), follow. □

Proof of Proposition 2.8.1

The general idea behind the proof is based on a comparison between the chain with $\tilde{\mu}$-reset, $\tilde{\mathbf{P}}^{\tilde{\mu}}$, and the product chain, $\mathbf{P}^{\otimes 2}$. Assuming that the two processes start at some (x, y) with $y \neq x$, they can be perfectly coupled up to the first meeting of the two walks, corresponding to the first hitting time of ∂ for the process with reset. Recall that the two stationary distributions, $\tilde{\pi}$ and $\pi^{\otimes 2}$, coincide for every couple $(x, y) \notin \Delta$ and $\pi^{\otimes 2}(\Delta) = \tilde{\pi}(\partial) = o_{\mathbb{P}}(1)$. Therefore, if the meeting happens after the mixing horizon T, we can rely on Theorem 2.4.3 and the fact that the mixing time of the product chain is at most a constant multiple of the one of a single walk. At that point, we are left to prove that the collapsed chain $\tilde{P}$, started at ∂ makes logarithmic number of steps, within time T, without visiting ∂ has probability $1 - o_{\mathbb{P}}(1)$.

We now start to make the above mentioned heuristic rigorous. Let us start by defining the surjective map $\varphi : V^2 \to \tilde{V}$ by

$$\varphi(\boldsymbol{x}) = \begin{cases} (x_1, x_2) & \boldsymbol{x} = (x_1, x_2) \text{ with } x_1, x_2 \in [n]\,,\ x_1 \neq x_2\,, \\ \partial & \boldsymbol{x} = (x, x) \text{ with } x \in [n]\,. \end{cases} \tag{2.208}$$

Moreover, recall that in our notation $\tilde{\mathbf{P}}^{\tilde{\mu}}_{\boldsymbol{x}}$ describes the quenched law of $\tilde{X}$, with $\tilde{X}_0 = \boldsymbol{x} \in \tilde{V}$ and $\tilde{\mu}$-reset, while $\mathbf{P}^{\otimes 2}_{(x,y)}$ the quenched law of two independent walks with $(x, y) \in [n]^2$ as initial position. By the triangular inequality and observing that $\|\tilde{\pi}(\cdot) - \pi^{\otimes 2}(\varphi^{-1}(\cdot))\|_{\text{TV}} = 0$ almost surely, it holds that

$$\begin{aligned} \max_{\boldsymbol{x}\in\tilde{V}} \|\tilde{P}^S(\boldsymbol{x}, \cdot) - \tilde{\pi}(\cdot)\|_{\text{TV}} &\leq \max_{(x,y)\in[n]^2} \| \left(P^{\otimes 2}\right)^S ((x, y), \cdot) - \pi^{\otimes 2}(\cdot)\|_{\text{TV}} \\ &+ \max_{(x,y)\in[n]^2} \sup_{A\subset\tilde{V}} \left| \tilde{\mathbf{P}}^{\tilde{\mu}}_{\varphi((x,y))} \left(\tilde{X}_S \in A\right) - \mathbf{P}^{\otimes 2}_{(x,y)} \left(X^{\otimes 2}_S \in \varphi^{-1}(A)\right) \right|. \end{aligned} \tag{2.209}$$

The first term on the right-hand side vanishes in probability thanks to Theorem 2.4.3 and the fact that $S \gg \log(n)$, while the second term can be rewritten as

$$\max_{(x,y)\in[n]^2} \sup_{A\subset\tilde{V}} \left| \tilde{\mathbf{P}}^{\tilde{\mu}}_{\varphi((x,y))} \left(\tilde{X}_S \in A, \tau_\partial < S\right) - \mathbf{P}^{\otimes 2}_{(x,y)} \left(X^{\otimes 2}_S \in \varphi^{-1}(A), \tau_{\text{meet}} < S\right) \right|, \tag{2.210}$$

since the product chain and the collapsed chain can be perfectly coupled until the first hitting time of the diagonal. In particular

$$\tilde{\mathbf{P}}^{\tilde{\mu}}_{\varphi((x,y))}\left(\tau_\partial = t\right) = \mathbf{P}^{\otimes 2}_{(x,y)}\left(\tau_{\text{meet}} = t\right), \quad \forall (x,y) \in [n]^2, x \neq y, t \in \mathbb{N}. \tag{2.211}$$

By the strong Markov property and (2.211) we get, uniformly in $(x,y) \in V^2$ and $A \subset \tilde{V}$,

$$\begin{aligned}
&\left|\tilde{\mathbf{P}}^{\tilde{\mu}}_{\varphi((x,y))}\left(\tilde{X}_S \in A, \tau_\partial < S\right) - \mathbf{P}^{\otimes 2}_{(x,y)}\left(X^{\otimes 2}_S \in \varphi^{-1}(A), \tau_{\text{meet}} < S\right)\right| \\
&= \Bigg| \sum_{t=0}^{S} \tilde{\mathbf{P}}^{\tilde{\mu}}_{\varphi((x,y))}\left(\tau_\partial = t\right) \tilde{\mathbf{P}}^{\tilde{\mu}}\left(\tilde{X}_{S-t} \in A\right) \\
&- \sum_{t=0}^{S} \sum_{z\in[n]} \mathbf{P}^{\otimes 2}_{(x,y)}\left(\tau_{\text{meet}} = t, X^{\otimes 2}_t = (z,z)\right) \mathbf{P}^{\otimes 2}_{(z,z)}\left(X^{\otimes 2}_{S-t} \in \varphi^{-1}(A)\right) \Bigg| \qquad (2.212) \\
&= \Bigg| \sum_{t=0}^{S} \sum_{z\in[n]} \mathbf{P}^{\otimes 2}_{(x,y)}\left(\tau_{\text{meet}} = t, X^{\otimes 2}_t = (z,z)\right) \\
&\times \left[\tilde{\mathbf{P}}^{\tilde{\mu}}\left(\tilde{X}_{S-t} \in A\right) - \mathbf{P}^{\otimes 2}_{(z,z)}\left(X^{\otimes 2}_{S-t} \in \varphi^{-1}(A)\right)\right] \Bigg|.
\end{aligned}$$

Let us fix a constant $\alpha \in (0,1)$ and partition the above sum over t into two parts centered at αS. By the triangle inequality and (2.210) we get

$$\begin{aligned}
&\max_{(x,y)\in[n]^2} \sup_{A\subset\tilde{V}} \left|\tilde{\mathbf{P}}^{\tilde{\mu}}_{\varphi((x,y))}\left(\tilde{X}_S \in A, \tau_\partial < S\right) - \mathbf{P}^{\otimes 2}_{(x,y)}\left(X^{\otimes 2}_S \in \varphi^{-1}(A), \tau_{\text{meet}} < S\right)\right| \\
&\quad \leq \max_{(x,y)\in[n]^2} \sup_{A\subset\tilde{V}} \Bigg| \sum_{t=0}^{\alpha S} \sum_{z\in[n]} \mathbf{P}^{\otimes 2}_{(x,y)}\left(\tau_{\text{meet}} = t, X^{\otimes 2}_t = (z,z)\right) \times \\
&\qquad\qquad \times \sup_{t\leq \alpha S} \max_{z\in[n]} \left[\tilde{\mathbf{P}}^{\tilde{\mu}}_\partial\left(\tilde{X}_{S-t} \in A\right) - \mathbf{P}^{\otimes 2}_{(z,z)}\left(X^{\otimes 2}_{S-t} \in \varphi^{-1}(A)\right)\right] + \\
&\qquad\qquad + \sum_{t=\alpha S}^{S} \sum_{z\in[n]} \mathbf{P}^{\otimes 2}_{(x,y)}\left(\tau_{\text{meet}} = t, X^{\otimes 2}_t = (z,z)\right) \Bigg| \\
&\quad \leq \max_{(x,y)\in[n]^2} \sup_{A\subset\tilde{V}} \Bigg| \sup_{t\leq\alpha S} \max_{z\in[n]} \left[\tilde{\mathbf{P}}^{\tilde{\mu}}_\partial\left(\tilde{X}_{S-t} \in A\right) - \mathbf{P}^{\otimes 2}_{(z,z)}\left(X^{\otimes 2}_{S-t} \in \varphi^{-1}(A)\right)\right] \\
&\qquad\qquad + \sum_{t=\alpha S}^{S} \sum_{z\in[n]} \mathbf{P}^{\otimes 2}_{(x,y)}\left(\tau_{\text{meet}} = t, X^{\otimes 2}_t = (z,z)\right) \Bigg|,
\end{aligned} \tag{2.213}$$

where we used the rough bounds

$$\left|\tilde{\mathbf{P}}^{\tilde{\mu}}_\partial\left(\tilde{X}_{S-t} \in A\right) - \mathbf{P}^{\otimes 2}_{(z,z)}\left(X^{\otimes 2}_{S-t} \in \varphi^{-1}(A)\right)\right| \leq 1, \qquad t \in [\alpha S, S], z \in [n],$$

and

$$\sum_{t=0}^{\alpha S} \sum_{z\in[n]} \mathbf{P}^{\otimes 2}_{(x,y)} \left(\tau_{\text{meet}} = t, X_t^{\otimes 2} = (z,z)\right) \leq 1\,.$$

Putting together (2.209), (2.212) and (2.213), using the triangular inequality and the fact ,that for any $A \subset \tilde{V}$ it holds that $\tilde{\pi}(A) = \pi^{\otimes 2}(\varphi^{-1}(A))$, we get

$$\begin{aligned}\max_{\boldsymbol{x}\in\tilde{V}} \|\tilde{P}^S(\boldsymbol{x},\cdot) - \tilde{\pi}(\cdot)\|_{\text{TV}} \leq \sup_{t\leq \alpha S} \|\tilde{P}^{S-t}(\partial,\cdot) - \tilde{\pi}\|_{\text{TV}} \\ + \sup_{t\leq\alpha S} \max_{z\in[n]} \| \left(P^{\otimes 2}\right)^{S-t} ((z,z),\cdot) - \pi^{\otimes 2}\|_{\text{TV}} \\ + \max_{(x,y)\in V^2} \sum_{t=\alpha S}^{S} \mathbf{P}^{\otimes 2}_{(x,y)} (\tau_{\text{meet}} = t)\ .\end{aligned} \tag{2.214}$$

The second term on the right-hand side of (2.214) goes to zero in probability, thanks to Theorem 2.4.3 and the fact that the total-variation distance is non-increasing in t. As for the third term in (2.214), we have that

$$\begin{aligned}&\max_{(x,y)\in V^2} \sum_{t=\alpha S}^{S} \mathbf{P}^{\otimes 2}_{(x,y)} (\tau_{\text{meet}} = t) \leq \\ &\leq S \sup_{t>\alpha S} \max_{(x,y)\in V^2} \left[\mathbf{P}^{\otimes 2}_{(x,y)} \left(X_t^{\otimes 2} \in \varphi^{-1}(\partial)\right) - \tilde{\pi}(\partial)\right] + S\,\tilde{\pi}(\partial) \\ &\leq S \sup_{t>\alpha S} \max_{(x,y)\in V^2} \| \left(P^{\otimes 2}\right)^t ((x,y),\cdot) - \pi^{\otimes 2}\|_{\text{TV}} + S\,\tilde{\pi}(\partial) = o_{\mathbb{P}}(1)\,,\end{aligned}$$

where the last asymptotic estimate follows from

$$S\,\tilde{\pi}(\partial) = \log^3(n) \sum_{x\in[n]} \pi^2(x) \leq n\,\log^3(n) \left[\max_{y\in[n]} \pi(y)\right]^2\,,$$

and Theorem 2.4.4, while

$$\begin{aligned}&S \sup_{t>\alpha S} \max_{(x,y)\in V^2} \| \left(P^{\otimes 2}\right)^t ((x,y),\cdot) - \pi^{\otimes 2}\|_{\text{TV}} \leq \\ &\leq S \max_{(x,y)\in V^2} \| \left(P^{\otimes 2}\right)^{\alpha S} ((x,y),\cdot) - \pi^{\otimes 2}\|_{\text{TV}} = o_{\mathbb{P}}(1)\,,\end{aligned}$$

thanks to Theorem 2.4.3 and monotonicity of the total-variation distance. In conclusion, for any constant $\alpha \in (0,1)$

$$\max_{\boldsymbol{x}\in\tilde{V}} \|\tilde{P}^T(\boldsymbol{x},\cdot) - \tilde{\pi}(\cdot)\|_{\text{TV}} \leq \sup_{t\leq\alpha S} \|\tilde{P}^{S-t}(\partial,\cdot) - \tilde{\pi}\|_{\text{TV}} + o_{\mathbb{P}}(1)\,. \tag{2.215}$$

Choosing, e.g, $\alpha = \frac{1}{2}$, to conclude the proof of Proposition 2.8.1 we are left to show that

$$\|\tilde{P}^{\frac{S}{2}}(\partial, \cdot) - \tilde{\pi}\|_{\mathrm{TV}} \xrightarrow{\mathbb{P}} 0\,. \tag{2.216}$$

The proof of the latter convergence is provided in next subsection.

Proof of (2.216)

Recall the definition of $\hbar$ and $V_\star$ in (2.55) and (2.57), respectively. Consider the process $(\tilde{X}_t)_{t\geq 0}$ with law $\tilde{\mathbf{P}}^{\tilde{\mu}}$ and, if $\tilde{X}_t \neq \partial$, call X_t and Y_t the projections of the two coordinates of $\tilde{X}_t$. Define

$$\tau_{\mathrm{dev}} := \inf\{t > 0 \,:\, \tilde{X}_t \neq \partial\,, \mathcal{B}^+_{X_t}(\hbar) \cap \mathcal{B}^+_{Y_t}(\hbar) = \emptyset \text{ and } X_t, Y_t \in V_\star\}\,, \tag{2.217}$$

to be the first time such that the walks are visiting vertices the $\hbar$ out-neighborhood of which are not intersecting trees. Moreover, let $\nu_{\mathrm{dev}} \in \mathcal{P}(\tilde{V} \setminus \partial)$ be the distribution of the process $\tilde{X}$ started at ∂, at the occurrence of the stopping time τ_{dev}. Namely,

$$\nu_{\mathrm{dev}}(\boldsymbol{x}) := \tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tilde{X}_{\tau_{\mathrm{dev}}} = \boldsymbol{x}\right)\,, \qquad \boldsymbol{x} \in \tilde{V} \setminus \partial. \tag{2.218}$$

The proof is articulated into three main lemmas, Lemma 2.8.3, Lemma 2.8.4 and Lemma 2.8.5. First, in Lemma 2.8.3 we show that for a typical realization of the graph, the process $\tilde{X}$ started at ∂ is such that the probability that both τ^+_∂ and τ_{dev} are "large" is arbitrarily small. Then, in Lemma 2.8.4, we use that result as a bootstrap to show that the probability that τ_{dev} itself is "large" is arbitrarily small. Finally, in Lemma 2.8.5 (and in Corollary 2.8.6) we show that starting at ν_{dev} the probability to hit ∂ before time S is arbitrarily small for a typical realization of graph. To conclude the proof it suffices to collect the pieces: if both $\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}(\tau_{\mathrm{dev}} < S/4)$ and $\tilde{\mathbf{P}}^{\tilde{\mu}}_{\nu_{\mathrm{dev}}}(\tau_\partial > S)$ are $1 - o_{\mathbb{P}}(1)$, then with the same probability the processes $\tilde{\mathbf{P}}^{\tilde{\mu}}$ and $\mathbf{P}^{\otimes 2}$ can be perfectly coupled within time $S/2$ for a consecutive interval of time having size $S/4$. At this point, the desired result is a consequence of Theorem 2.4.3. We now present the main three lemmas and the respective proofs, and in the last part of this section we spell-out in detail the above-mentioned concluding argument.

Lemma 2.8.3. *Let $h_\star = h_{\star,n}$ be any sequence such that $h_\star \to \infty$ and $h_\star = o(\log(n))$, and call*

$$\tau^+_\partial := \inf\{t \geq 1 \mid \tilde{X}_t = \partial\}\,. \tag{2.219}$$

Then,

$$\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}(\tau_{\mathrm{dev}} \wedge \tau^+_\partial > h_\star) \xrightarrow{\mathbb{P}} 0. \tag{2.220}$$

Proof. Let $V_\star \subset [n]$ as in (2.57) and $\hbar$ as in (2.55). Recall that, thanks to Proposition 2.8.2, $\tilde{\mu}(V_\star) = 1 - o_{\mathbb{P}}(1)$. Hence, fixed any $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0$ we can bound for all n large

enough

$$\begin{aligned}\mathbb{P}\left(\tilde{\mathbf{P}}_{\partial}^{\tilde{\mu}}(\tau_{\mathrm{dev}} \wedge \tau_{\partial}^{+} > h_{\star}) \le \varepsilon_1\right) \\ &\ge \mathbb{P}(\tilde{\mu}(V_{\star}) > 1-\varepsilon_2) - \mathbb{P}\left(\tilde{\mathbf{P}}_{\partial}^{\tilde{\mu}}(\tau_{\mathrm{dev}} \wedge \tau_{\partial}^{+} > h_{\star}) > \varepsilon_1\,,\, \tilde{\mu}(V_{\star}) > 1-\varepsilon_2\right) \\ &\ge 1-\varepsilon_3 - \mathbb{P}\left(\tilde{\mathbf{P}}_{\partial}^{\tilde{\mu}}(\tau_{\mathrm{dev}} \wedge \tau_{\partial}^{+} > h_{\star}) > \varepsilon_1\,,\, \tilde{\mu}(V_{\star}) > 1-\varepsilon_2\right)\,.\end{aligned} \tag{2.221}$$

Consider the (random) subset of states

$$M \coloneqq \left\{\boldsymbol{x} = (x_1, x_2) \in \tilde{V} \setminus \{\partial\} \mid x_1 \in V_{\star}\,,\, x_1 \to x_2 \text{ or } x_2 \in V_{\star}\,,\, x_2 \to x_1\right\}\,, \tag{2.222}$$

and notice that for every $\boldsymbol{x} = (x_1, x_2)$ it always possible to identify a *chasing* and a *escaping* walk. In that case, we write x_c and x_e to avoid confusion on the roles of x_1 and x_2. As already pointed out in (2.211), starting the process $\tilde{X}$ at any $\boldsymbol{x} = (x_1, x_2) \in M$, the first hitting time of ∂ equals the first meeting time of two independent walks started at $(x_1, x_2) \in [n]^2 \setminus \Delta$. We now show that, starting at any $\boldsymbol{x} \in M$, the probability of the event $\{\tau_{\mathrm{dev}} > h_{\star}\} \cap \{\tau_{\mathrm{meet}} > h_{\star}\}$ is $o_{\mathbb{P}}(1)$. Without loss of generality, assume that $\boldsymbol{x} = (x_c, x_e)$. Under the event $\{\tau_{\mathrm{meet}} > s\}$, the only way to realize $\{\tau_{\mathrm{dev}} > s\}$ is to have the chasing walk (the one starting at x_c) following the steps of the escaping one (the one starting at x_e) without reaching it. Hence, calling $A_{x_c}(s)$ the random number of jumps made by the walk starting at x_c before $s > 0$, for all $s \in [0, \hbar)$, $x_c \in V_{\star}$, $x_e \in [n]$ such that $x_c \to x_e$ we have

$$\mathbf{P}_{(x_c,x_e)}^{\otimes 2}(\tau_{\mathrm{dev}} \wedge \tau_{\mathrm{meet}} > s) \le (d_{\min}^{+})^{-s/3} + \mathbf{P}_{(x_c,x_e)}^{\otimes 2}\left(A_{x_c}(s) \le \tfrac{1}{3}s\right)\,. \tag{2.223}$$

In particular, taking $s \in (h_{\star}, \hbar)$ in the last display and using the fact that

$$\mathbf{P}_{(x_c,x_e)}^{\otimes 2}\left(A_{x_c}(s) \le \tfrac{1}{3}s\right) = \mathbf{P}\left(\mathrm{Bin}\left(s, \tfrac{1}{2}\right) \le \tfrac{1}{3}s\right) = o(1)\,,$$

we conclude that, for every $\varepsilon_1 > 0$ and all n large enough

$$\mathbb{P}\left(\max_{\boldsymbol{x} \in M} \mathbf{P}_{\boldsymbol{x}}^{\otimes 2}(\tau_{\mathrm{dev}} \wedge \tau_{\mathrm{meet}} > s) \le \varepsilon_1\right) = 1\,, \tag{2.224}$$

where we used that $d_{\min}^{+} \ge 2$. Exploiting (2.224), we finally obtain, for all sufficiently

large n,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}(\tau_{\mathrm{dev}} \wedge \tau^{+}_{\partial} > h_\star) > \varepsilon_1\,,\, \tilde{\mu}(V_\star) > 1-\varepsilon_2\right)\\
&\leq \mathbb{P}\left(\sum_{\substack{(x_1,x_2)\in[n]^2:\\ \boldsymbol{x}=(x_1,x_2)\in M}} \frac{\tilde{\mu}(x_c)}{d^{+}_{x_c}} \mathbf{P}^{\otimes 2}_{(x_c,x_e)}(\tau_{\mathrm{dev}} \wedge \tau_{\mathrm{meet}} > h_\star - 1) > \varepsilon_1\,,\, \tilde{\mu}(V_\star) > 1-\varepsilon_2\right) + \varepsilon_2\\
&\leq \mathbb{P}\left(\max_{(x_c,x_e)\text{ s.t. } x_c\in V_\star, x_c\to x_e} \mathbf{P}^{\otimes 2}_{(x_c,x_e)}(\tau_{\mathrm{dev}} \wedge \tau_{\mathrm{meet}} > h_\star - 1) > \varepsilon_1\right) + \varepsilon_2 = \varepsilon_2\,,
\end{aligned} \tag{2.225}$$

and the desired result follows from (2.221) and (2.225), by letting ε_1, ε_2 and ε_3 going to zero. □

Lemma 2.8.4. *Let $h_\star = h_{\star,n}$ be any sequence such that $h_\star \to \infty$ and $h_\star = o(\log(n))$. Then*

$$\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}(\tau_{\mathrm{dev}} > h_\star) \xrightarrow{\mathbb{P}} 0\,. \tag{2.226}$$

Proof. Recall the definition of the set M in (2.222). For any positive r, define

$$\mathcal{R}_r = \bigcup_{\boldsymbol{x}\in M} \bigcap_{\ell=0}^{r-1} \{\tilde{X}_{\tau^{(\ell)}_{\partial}+1} = \boldsymbol{x}\}\,,$$

where we define

$$\tau^{(j)}_{\partial} := \inf\{t > \tau^{(j-1)}_{\partial} \mid \tilde{X}_t \in \partial\}\,, \qquad j \in \mathbb{N}\,.$$

In words, $\mathcal{R}_r$ reads as follows: the first r times in which the process visits ∂ (including time zero), it exits ∂ by reaching (in a single step) some $\boldsymbol{x} = (x_1, x_2) \in M$. Recall that if $\boldsymbol{x} = (x_1, x_2) \in M$ it is always possible to distinguish a *chasing* and an *escaping* starting point between x_1 and x_2, which we refer to as x_c and x_e. Notice that, $\mathbb{P}$ – a.s.,

$$\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}(\mathcal{R}_r) = \tilde{\mu}(V_\star)^r\,. \tag{2.227}$$

Therefore,

$$\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}(\tau_{\mathrm{dev}} > h_\star) \leq (1 - \tilde{\mu}(V_\star)^r) + \tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}(\tau_{\mathrm{dev}} > h_\star\,,\, \mathcal{R}_r)\,. \tag{2.228}$$

Notice that under $\mathcal{R}_r$ the probability that ∂ is visited more than ℓ times before τ_{dev} is exponentially small in ℓ. Indeed, for any $\ell < r$,

$$\begin{aligned}
\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tau^{(\ell)}_{\partial} < \tau_{\mathrm{dev}},\, \mathcal{R}_r\right) = \tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tau^{(1)}_{\partial} < \tau_{\mathrm{dev}},\, \mathcal{R}_2\right)^{\ell} &\leq \tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tau_{\mathrm{dev}} > 2,\, \mathcal{R}_2\right)^{\ell}\\
&\leq \left(\frac{1}{2}\left(\frac{1}{d^{+}_{\min}} + 1\right)\right)^{\ell} \leq \left(\frac{3}{4}\right)^{\ell}\,,
\end{aligned} \tag{2.229}$$

as, under $\mathcal{R}_r$, the event $\tau_{\mathrm{dev}} > 2$ starting at ∂ requires that: at the step right after the exit from ∂, either the chasing walk reaches the escaping one, or the escaping one moves forward in the tree. Moreover, by definition of the process $\tilde{X}$, $\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}(\tilde{X}_1 = \partial, \mathcal{R}_2) = 0$, since in order to see a transition $\partial \to \partial$ in the $\tilde{\mu}$-reset case one has to sample a vertex having a self-loop, hence not a member of $V_\star$.

Therefore, for every $\ell < r$,

$$\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tau_{\mathrm{dev}} > h_\star, \mathcal{R}_r\right) \le \tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tau^{(\ell)}_{\partial} > \tau_{\mathrm{dev}} > h_\star, \mathcal{R}_r\right) + \left(\frac{3}{4}\right)^{\ell} . \tag{2.230}$$

Notice that, under the event $\{\tau^{(\ell)}_{\partial} > \tau_{\mathrm{dev}} > h_\star, \mathcal{R}_r\}$ it must exist some $j \le \ell$ such that the j-th excursion from ∂ to ∂ lasted at least $h_\star/\ell$ steps and τ_{dev} did not realize in such an interval of time. Hence, thanks to Lemma 2.8.3, if $\ell = o(h_\star)$,

$$\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tau^{(\ell)}_{\partial} > \tau_{\mathrm{dev}} > h_\star, \mathcal{R}_r\right) \le \tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tau_{\mathrm{dev}} \wedge \tau^{+}_{\partial} \ge \frac{h_\star}{\ell}\right) = o_{\mathbb{P}}(1) . \tag{2.231}$$

Plugging (2.230) and (2.230) into (2.228), choosing, e.g., $\ell = \sqrt{h_\star}$, we finally obtain

$$\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}(\tau_{\mathrm{dev}} > h_\star) \le \left(\frac{3}{4}\right)^{\sqrt{h_\star}} + o_{\mathbb{P}}(1) , \tag{2.232}$$

from which the desired result follows. □

Lemma 2.8.5. *Consider the (random) set*

$$D := \{(x, y) \in V_\star^2 \mid \mathcal{B}^+_x(\hbar) \cap \mathcal{B}^+_y(\hbar) = \emptyset\} .$$

For all n large enough it holds

$$\max_{(x,y)} \mathbb{P}\left(\mathbb{1}_{(x,y)\in D}\mathbf{P}^{\otimes 2}_{(x,y)}(\tau_{\mathrm{meet}} < S) > \varepsilon\right) \le n^{-3} , \tag{2.233}$$

where S is defined in (2.199)*.*

Corollary 2.8.6. *It holds*

$$\max_{\boldsymbol{x}\in \mathrm{supp}(\nu_{\mathrm{dev}})} \tilde{\mathbf{P}}^{\tilde{\mu}}_{\boldsymbol{x}}(\tau_{\partial} < S) \xrightarrow{\mathbb{P}} 0 , \tag{2.234}$$

where $\nu_{\mathrm{dev}} \in \mathcal{P}(\tilde{V} \setminus \{\partial\})$ is defined in (2.218)*.*

Proof of Corollary 2.8.6. The result follows by coupling $\tilde{\mathbf{P}}$ and $\mathbf{P}^{\otimes 2}$ up to τ_∂, using Lemma 2.8.5, a union bound on $(x, y) \in D$, and the fact that the support of ν_{dev} is contained in D by definition. □

Proof of Lemma 2.8.5. Fix some constant $\kappa \in \mathbb{N}$. By the generalized Markov inequality we have

$$\max_{(x,y)} \mathbb{P}(\mathbb{1}_{(x,y)\in D}\mathbf{P}^{\otimes 2}_{(x,y)}(\tau_{\text{meet}} < S) > \varepsilon) \leq \frac{\max_{(x,y)} \mathbb{E}[\mathbb{1}_{(x,y)\in D}\mathbf{P}^{\otimes 2}_{(x,y)}(\tau_{\text{meet}} < S)^{\kappa}]}{\varepsilon^{\kappa}} . \tag{2.235}$$

Call $\sigma = \sigma_{x,y}$ a realization of the complete out-neighborhood of x and y up to length $\hbar$, Notice that the event $(x, y) \in D$ is σ-measurable. With a slight abuse of notation we say that $\sigma \in D$ if $(x, y) \in D$ under σ. Then, it is enough to show that there exists some constant $\kappa \in \mathbb{N}$ such that

$$\max_{(x,y)} \max_{\sigma\in D} \mathbb{E}[\mathbf{P}^{\otimes 2}_{(x,y)}(\tau_{\text{meet}} < S)^{\kappa} \mid \sigma] \leq n^{-3} , \tag{2.236}$$

for all n large enough. To this aim, we use another (multiple) *annealing argument*. In particular, let the law $\mathbb{P}^{\otimes 2,\,\kappa\text{-an}|\sigma}_{(x,y)}$ refer to the non-Markovian process introduced in Remark 2.6.2 with initial environment σ. In such a probability space we consider the events

$$\mathcal{E}_i = \{\exists t \leq S \text{ s.t. } W^{(i,1)}_t = W^{(i,2)}_t\} , \qquad i \leq \kappa , \tag{2.237}$$

and we notice that

$$\max_{(x,y)} \max_{\sigma\in D} \mathbb{E}[\mathbf{P}^{\otimes 2}_{(x,y)}(\tau_{\text{meet}} < S)^{\kappa} \mid \sigma] = \max_{(x,y)} \max_{\sigma\in D} \mathbb{P}^{\otimes 2,\,\kappa\text{-an}|\sigma}_{(x,y)}(\cap_{i\leq\kappa}\mathcal{E}_i) . \tag{2.238}$$

We now show that there exists some $\delta > 0$ such that

$$\max_{j\leq\kappa} \max_{(x,y)} \max_{\sigma\in D} \mathbb{P}^{\otimes 2,\,\kappa\text{-an}|\sigma}_{(x,y)}(\mathcal{E}_j \mid \cap_{i<j}\mathcal{E}_i) \leq n^{-\delta} , \tag{2.239}$$

for all n large enough. Then, from (2.238) and (2.239) follows that

$$\max_{(x,y)} \max_{\sigma\in D} \mathbb{E}[\mathbf{P}^{\otimes 2}_{(x,y)}(\tau_{\text{meet}} < S)^{\kappa} \mid \sigma] = \max_{(x,y)} \max_{\sigma\in D} \prod_{j=1}^{\kappa} \mathbb{P}^{\otimes 2,\,\kappa\text{-an}|\sigma}_{(x,y)}(\mathcal{E}_j \mid \cap_{i<j}\mathcal{E}_i) \leq n^{-\kappa\,\delta} . \tag{2.240}$$

and the desired result follows by taking κ large enough, e.g., such that $\kappa\,\delta > 3$.

We are left with proving (2.239). Fix any $(x, y) \in [n]^2 \setminus \Delta$, and fix any $\sigma = \sigma_{x,y}$ such that $\sigma \in D$. Then, when $i = 1$ and $t = 0$ we have a subgraph of G consisting of two out-going trees of length $\hbar$, rooted at x and y, that do not intersect. We let the first couple of walks evolve for time S and bound from above the probability of the event $\mathcal{E}_1$. Notice that, by the fact that $\sigma \in D$ we have

$$W^{(i,1)}_s \neq W^{(i,2)}_s \qquad \forall s < \hbar,\ i \leq \kappa \qquad \mathbb{P}^{\otimes 2,\,\kappa\text{-an}|\sigma}_{(x,y)} - \text{a.s.} . \tag{2.241}$$

In order for the event $\mathcal{E}_1$ to occur, one of the two walks has to exit its σ-tree and hit one of the vertices that are already discovered, that is, that have at least a head or a

tailed matched. Since there are at most S steps available and, uniformly on the past, the chance that such a hitting takes place at a fixed time $t \le S$ is bounded by

$$w_1 := \frac{d^-_{\max}[2(d^+_{\max})^{2\hbar} + 2S]}{m - 2(d^+_{\max})^{\hbar} - 2S}\,, \tag{2.242}$$

as at any time t, for each of the two walks, there are at most $2(d^+_{\max})^{2\hbar} + 2S$ vertices that have been explored already. We conclude that

$$\mathbb{P}^{\otimes 2,\,\kappa\text{-an}|\sigma}_{(x,y)}(\mathcal{E}_1) \le S\, w_1\,. \tag{2.243}$$

Now condition on an arbitrary realization of the paths of the first $i-1$ couples of walks in which $\cap_{j<i}\mathcal{E}_j$ is realized. Notice that, for $z = x, y$, within the leafs of $\mathcal{B}^+_z(\hbar)$ there are at most $(2S)^{i-1} \le (2\log(n))^{3i}$ having at least a matched tail. Hence, the probability that the walk that starts at z exits $\mathcal{B}^+_z(\hbar)$ at a leaf that has already been discovered can be bounded (uniformly in (x,y) on σ, and on the behavior of the previous couples of walks) for all n large enough, by

$$\bar{w} := (d^+_{\min})^{-\hbar}(2\log(n))^{3\kappa} \le \frac{1}{2}n^{-2/5}\,, \tag{2.244}$$

thanks to the definition of $\hbar$ in (2.55) and the fact that $d^+_{\min} \ge 2$. Indeed, being $\mathcal{B}^+_z(\hbar)$ a tree, the probability that the i-th walk ends up in specific leaf is given by the probability to follow the unique path from z to that leaf.

If the walk that start at z exists $\mathcal{B}^+_z(\hbar)$ from a leaf that has no matched tails, then the argument before (2.242) applies, and this time the probability of a meeting time before S can be bounded by

$$w_i := \frac{d^-_{\max}[2(d^+_{\max})^{2\hbar} + 2iS]}{m - 2(d^+_{\max})^{\hbar} - 2iS} \le n^{-2/5}\,, \qquad i \le \kappa\,, \tag{2.245}$$

for all n large enough. In conclusion, for all n large enough,

$$\max_{j\le\kappa}\max_{(x,y)}\max_{\sigma\in D}\mathbb{P}^{\otimes 2,\,\kappa\text{-an}|\sigma}_{(x,y)}(\mathcal{E}_j \mid \cap_{i<j}\mathcal{E}_i) \le \bar{w} + S\, w_\kappa \le 2n^{-2/5}\,, \tag{2.246}$$

and (2.239) readily follows. □

We are now in shape to prove (2.216), and in turn conclude the proof of Proposition 2.8.1.

Conclusion of the Proof of (2.216). Recall that the function φ in (2.208) and notice that φ is injective on $\mathrm{supp}(\nu_{\mathrm{dev}})$, hence there is no ambiguity in identifying ν_{dev} with

its lifting in $[n]^2$. We can rewrite the first term on the right-hand side of (2.214) as follows

$$\begin{aligned}
&\|\tilde{P}^{\frac{S}{2}}(\partial,\cdot)-\tilde{\pi}\|_{\mathrm{TV}} = \\
&= \sup_{A\subset\tilde{V}}\left|\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tilde{X}_{S/2}\in A,\tau_{\mathrm{dev}}>\frac{S}{4}\right)+\tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tilde{X}_{S/2}\in A,\tau_{\mathrm{dev}}\le\frac{S}{4}\right)-\tilde{\pi}(A)\right| \\
&\le \tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tau_{\mathrm{dev}}>\frac{S}{4}\right)+\sup_{\ell\le\frac{S}{4}}\sup_{A\subset\tilde{V}}\left|\tilde{\mathbf{P}}^{\tilde{\mu}}_{\nu_{\mathrm{dev}}}\left(\tilde{X}_{\frac{S}{2}-\ell}\in A\right)-\tilde{\pi}(A)\right| \\
&= \tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tau_{\mathrm{dev}}>\frac{S}{4}\right)+\left\|\nu_{\mathrm{dev}}\tilde{P}^{\frac{S}{4}}-\tilde{\pi}\right\|_{\mathrm{TV}} \\
&\le \tilde{\mathbf{P}}^{\tilde{\mu}}_{\partial}\left(\tau_{\mathrm{dev}}>\sqrt{\log(n)}\right)+\left\|\nu_{\mathrm{dev}}\left(P^{\otimes 2}\right)^{\frac{S}{4}}-\pi^{\otimes 2}\right\|_{\mathrm{TV}} \\
&+\left\|\nu_{\mathrm{dev}}\left(P^{\otimes 2}\right)^{\frac{S}{4}}-\nu_{\mathrm{dev}}\tilde{P}^{\frac{S}{4}}\right\|_{\mathrm{TV}},
\end{aligned}$$

where the last step follows by the triangular inequality.

The first two terms on the right-hand side of the last display are both $o_{\mathbb{P}}(1)$, thanks to Lemma 2.8.4 and Theorem 2.4.3, respectively. Moreover, the last TV-distance appearing on the right-hand side can be bounded by a coupling argument: using the same source of randomness to generate the initial position according to ν_{dev}, and the two processes $P^{\otimes 2}$ and $\tilde{P}$ up to time τ_{meet} (and then letting the chain evolving independently), and call $\mathbf{Q}$ such a (random) coupled probability law. In particular, consider the random variable $\mathbf{Q}(\tau_{\mathrm{fail}}\le s)$, that correspond to the (random) probability that the coupling fails before s, and therefore, by the definition of the TV-distance in terms of optimal coupling, it follows that for every $s\ge 0$,

$$\mathbf{Q}(\tau_{\mathrm{fail}}\le s)\ge\left\|\nu_{\mathrm{dev}}\left(P^{\otimes 2}\right)^{s}-\nu_{\mathrm{dev}}\tilde{P}^{s}\right\|_{\mathrm{TV}},\qquad \mathbb{P}-\text{a.s.}.$$

Thanks to Corollary 2.8.6 we have,

$$\mathbf{Q}(\tau_{\mathrm{fail}}\le S/4)\le\mathbf{Q}(\tau_{\mathrm{fail}}\le S)\le\max_{\boldsymbol{x}\in\mathrm{supp}(\nu_{\mathrm{dev}})}\tilde{\mathbf{P}}^{\tilde{\mu}}_{\boldsymbol{x}}(\tau_{\partial}\le S)=o_{\mathbb{P}}(1),\tag{2.247}$$

from which (2.216) follows. □

§2.9 Conclusions and open problems

With this work we provide the first-order asymptotic of the expected meeting, consensus and coalescence time on a typical sparse random digraphs from the DCM ensemble. This result adds to the list of examples, presented in the Introduction, for which such a precise result can be obtained. Moreover, to the best of our knowledge,

it is the only degree-inhomogeneous class of graphs in which such results has been obtained so far. We conclude the chapter with a short list of open problems and with an outlook on possible future research.

As pointed out in Section 2.5.1, in this chapter we did not aim at understanding the most general conditions on the degree sequence up to which our technique could be stretched. Nevertheless, it is natural to ask this question. More importantly, it looks clear from our results that a change in the scaling of the consensus time should be attained as soon as the in-degree sequence does not admit a bounded second moment. Unfortunately, the understanding of the random walk on the DCM with such extreme heavy-tailed degrees is still completely open, and some new ideas are needed to carry out this investigation.

The question of determining the preconstant for the consensus time of the undirected configuration model remains open. The representation proposed in [HLYZ22, Lemma 6.12] looks promising, but an explicit characterization is still missing. In this respect, as mentioned in Section 2.3.1, non-rigorous results in [SAR08] seems to suggest a phenomenology similar to that in the Eulerian case. In fact, in view of the mean field conditions and the First Visit Time Lemma, it is not hard to realize that the value of ϑ in Eulerian case should provide a lower bound for its undirected counterpart. Indeed, being the stationary distribution unaltered, the only thing to check is that the quantity $R_T(\partial)$ in Proposition 2.5.6 should be larger in the latter case, due to the backtracking feature of the random walk on sparse undirected graphs.

§2.10 Appendix

In this Appendix we give an alternative proof of the convergence of $\mathbb{E}[R^\mu]$.

Before to present our result on the behavior of $\mathbb{E}[R^\mu]$ it is convenient to recall the following classical identity.

Lemma 2.10.1. *For all* $\rho \in (0, 1/2]$

$$\sum_{t\geq 1} 2^{-2t+1}\, C_{t-1}\, \rho^{t-1} = \frac{1-\sqrt{1-\rho}}{\rho}\,. \tag{2.248}$$

Proof. For all $\rho \in (0, 1/2]$, it holds that

$$\sum_{t\geq 1} 2^{-2t+1}\, C_{t-1}\, \rho^{t-1} = \frac{1}{2}\sum_{t\geq 0} C_t \left(\frac{\rho}{4}\right)^t = G(\rho/4) = \frac{1-\sqrt{1-\rho}}{\rho}\,,$$

where

$$G(x) = \frac{1-\sqrt{1-4x}}{2\,x}, \quad x \in (0,1),$$

is the generating function of the Catalan numbers, (see, e.g., [GKP94, Ch. 5.4] or [FL03, Eq. 24]). □

At this point we have essentially all the technical ingredients needed to compute $\mathbb{E}[R^\mu]$.

Lemma 2.10.2. *In the same setting of Lemma 2.6.6, defined* $\Phi : [0,1/2]^2 \to \mathbb{R}$ *as*

$$\Phi(\rho, q) := \frac{\rho}{\rho - q\left(1 - \sqrt{1-\rho}\right)} - 1\,, \tag{2.249}$$

it holds

$$\lim_{n\to\infty} \sup_{\mu\in\mathcal{M}(q)} |\mathbb{E}[R^\mu] - (1 + \Phi(\rho, q))| = 0\,, \tag{2.250}$$

where $\rho = \rho_n$ *is defined as in* (2.21).

Proof. Call $\hat{R}^\mu(s,t)$ the expected number of "reset" of the forest process, i.e., of the number of meetings of the red and the blue flag, in the interval $[s,t)$. We claim that it is enough to show that

$$\lim_{n\to\infty} \hat{R}^\mu(0, T+1) = 1 + \Phi(\rho, q)\,. \tag{2.251}$$

Indeed, thanks to Proposition 2.6.4,

$$\mathbb{E}[R^\mu] = \hat{R}^\mu(0, T+1) + O(Tn^{-\varepsilon/3}) \sim \hat{R}^\mu(0, T+1)\,, \tag{2.252}$$

where we also used that $\mathbb{E}[R^\mu] \le T$, $T = o(n^{\varepsilon/3})$ and $\hat{R}^\mu(0, T+1) \ge 1$. Recalling that $\hat{\tau}_\partial^{(i)}$ denotes the time of the i-th reset, we notice that process satisfies the renewal property

$$\hat{\mathbb{P}}(\hat{\tau}_\partial^{(i+1)} = \infty \mid \mathcal{F}_i) = 1 - \sum_{t\ge 1} \hat{\mathbb{P}}(\hat{\tau}_\partial^{(1)} = 2t)\,, \qquad \forall i \ge 1\,, \tag{2.253}$$

where $\mathcal{F}_i$ is the sigma-field generated by the history of the forest process up to time $\hat{\tau}_\partial^{(i)}$. Therefore,

$$\hat{R}^\mu(0,\infty) = \frac{1}{1 - \sum_{t\ge1} \hat{\mathbb{P}}(\hat{\tau}_\partial^{(1)} = 2t)} = 1 + \Phi(\rho, q)\,, \tag{2.254}$$

where the last equality follows from Lemmas 2.6.6 and 2.10.1. We claim that, there exists some universal constant $C > 0$ such that for all $t, s \ge 1$

$$\hat{R}^\mu(t, t+s) \le s\left(2^{-t/3} + \hat{\mathbb{P}}\big(\mathrm{Bin}(t,1/2) < \tfrac{t}{3}\big)\right) \le s\left(2^{-t/3} + e^{-Ct^2}\right). \tag{2.255}$$

The second inequality in (2.255) is a trivial consequence of the Chernoff bound. To see the validity of the first inequality, observe that the number of resets in $(t, t+s)$ is

at most s almost surely and, in order to be different from zero, it must be the case that the chasing walk never deviates from the escaping walk up to time t. At each step the chasing walk is selected with probability $1/2$ and, given that it did not deviate yet, it deviates with probability at least $1/2$, regardless of μ. Hence, or the chasing walk is selected less than $t/3$ times or, otherwise, the chance that it never deviates up to time t is at most $2^{-t/3}$; thus (2.255) follows. Therefore, for n sufficiently large,

$$\hat{R}^{\mu}(T,\infty) = \sum_{k\geq 1} \hat{R}(kT,(k+1)T) \leq \sum_{k\geq 1} T2^{-kT/4} = O\big(T2^{-T/4}\big) = o(1)\,, \qquad (2.256)$$

and (2.251) follows from (2.254) and (2.256). □