Matteo Quattropani

Mixing and cover time on sparse random digraphs

Dottorato in Matematica, Roma Tre (2020).

<http://www.bdim.eu/item?id=tesi_2020_QuattropaniMatteo_1>

Dipartimento di Matematica e Fisica

Mixing and cover time on sparse random digraphs

Tesi di dottorato in matematica
XXXII Ciclo

Candidato:
Matteo Quattropani Firma:________________________

Relatore di tesi:
Prof. Pietro Caputo Firma:________________________

Coordinatore:
Prof. Angelo Felice Lopez Firma:________________________

Anno Accademico 2018-2019
Dipartimento di Matematica e Fisica, Università Roma Tre
Largo San Leonardo Murialdo 1, 00146, Roma

Mixing and cover time
on sparse random digraphs

Matteo Quattropani
Advisor: Prof. P. Caputo

2019-11-17

Contents

CHAPTER 1

INTRODUCTION

This thesis aims at studying geometric properties of large random combinatorial structures. In particular, we will speak about *digraphs*: namely, graphs with directed edges. More specifically, we will be concerned with a slightly more general object, i.e., a so-called *multi-digraph*. We will often use the word *graph* in order to refer to a multi-digraph. The set of multi-digraphs on n vertices is in one-to-one correspondence with the set of square matrices of size n with non-negative integer entries. Given a multi-digraph G on n labeled vertices $[n] = \{1, \dots, n\}$, we associate to it the matrix A which we will refer to as the *adjacency* matrix of G, such that $A(x, y)$ is the number of directed links—or *edges*—connecting vertex x to vertex y. The out-degree of a vertex x is the sum of the entries in the x-th row of A, and will be denoted by d_x^+. Similarly, the in-degree of x, d_x^-, is the sum of the entries in the x-th column of A. We will call $\mathbf{d}^- = (d_x^-)_{x\in[n]}$ and $\mathbf{d}^+ = (d_x^+)_{x\in[n]}$, the *in/out-degree sequences* of G.

1.1 Sparse random digraphs ensembles

The results of this thesis concern what we will call *(di)graph ensembles*. A digraph ensemble is a sequence of sets, $\mathcal{G} \equiv \mathcal{G}_n$, of multi-digraphs over the vertex set $[n]$, equipped with a sequence of probability measures over $\mathcal{G}_n$, which we refer to as $\mathbb{P} \equiv \mathbb{P}_n$. By saying that

$G \equiv G_n$ is a *graph from the ensemble* we mean that G is a random multi-digraph with vertex set $[n]$, sampled from the set $\mathcal{G}$ with the probability measure $\mathbb{P}$. Since we will be interested in the *large volume* asymptotic properties of the graph ensemble—in the sense that we will consider the case in which n grows to infinity—the dependence on n is often omitted in the notation.

The digraphs ensembles we are concerned with in this thesis are parametrised by the degree sequences. More precisely, to every $n \geq 2$ we associate one or more sequences of length n, which we use to define the set $\mathcal{G}_n$ and the probability law $\mathbb{P}_n$.

We first consider the set of multi-digraphs parametrised, for each $n \geq 2$, by the in/out-degree sequences

$$\mathbf{d}^- = (d_x^-)_{x\in[n]}, \qquad \mathbf{d}^+ = (d_x^+)_{x\in[n]}.$$

In other words, the set $\mathcal{G}$ will be given by all the digraphs on the vertex set $[n]$ in which the degree sequences agree with $\mathbf{d}^\pm$. Notice that in order for the sequences to be sensible we need to require that

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

Given that $x \in [n]$ has degrees d_x^- and d_x^+, we can represent this vertex has having d_x^- *heads* and d_x^+ *tails* attached to it. If we call $\mathcal{E}$ the set of all the tails and $\mathcal{F}$ the set of all the heads, we clearly have that $|\mathcal{E}| = |\mathcal{F}| = m$. We then consider ω a uniformly random bijection $\omega : \mathcal{E} \to \mathcal{F}$. We will refer to the random bijection ω with the name *configuration*, we will call $\mathfrak{C} \equiv \mathfrak{C}_n$ the set of configuration and $G(\omega)$ the digraph induced by $\omega \in \mathfrak{C}$. Notice that each configuration ω specifies a multi-digraph G but, if we ignore the labels of heads and tails, a given multi-digraph G can be obtained as the result of different configurations. We will call *Directed Configuration Model* (DCM) the graph ensemble associated to the degree sequences $\mathbf{d}^-$ and $\mathbf{d}^+$ equipped with the probability measure $\mathbb{P}$ induced by the uniformly random choice of the bijection $\omega \in \mathfrak{C}$.

On the other hand, we can fix—for each $n \geq 2$—only the out-degree sequence $\mathbf{d}^+$. We consider the set $\mathcal{G}$ as given by the digraphs in which the number of connections between any two vertices is constrained to be 0 or 1, and such that the out-degree sequence is in agreement with $\mathbf{d}^+$. We will call *Out Configuration Model* (OCM) the digraph ensemble on such a set $\mathcal{G}$, equipped with the uniform measure. We remark that also in this case a random graph G can be sampled similarly to the DCM case. We can equip each node x with d_x^+ tails and pick for every x, independently, a uniformly random injective map ω_x from the set of tails of x to the set of all vertices. For all $x, y \in [n]$, we add a directed edge (x, y) if a tail from x is mapped into y through ω_x. We write $\omega = (\omega_x)_{x\in[n]}$ and, similarly to the DCM, we will call ω a *configuration* and $G(\omega)$ the corresponding induced digraph.

Both the DCM and the OCM are natural generalisations to the directed setting of the well-known *Configuration Model* (CM) for undirected graphs, considered in the seminal paper [11]. In the next sections we will recall the main achievements obtained in the analysis of these two models in the last decades, and we will point out the main tools employed in their study.

1.1.1 The *Directed Configuration Model* (DCM)

The DCM has been studied in [24], with few differences with respect to the definition given above. The authors of [24] study the size of the largest strongly connected component as a function of the degree sequences $\mathbf{d}^-$ and $\mathbf{d}^+$, under full generalities of the latter two.

For what concerns this thesis, instead, we will focus on the *sparse irreducible regime*; namely, we will work under the following assumption.

Assumption 1 *For every* $n \geq 2$

$$\delta := \min\left\{\min_{x\in[n]} d_x^+, \min_{x\in[n]} d_x^-\right\} \geq 2, \qquad \textit{and} \qquad \Delta := \max\left\{\max_{x\in[n]} d_x^+, \max_{x\in[n]} d_x^-\right\} = O(1).$$

The results in [24] imply that, under Assumption 1, the probability to sample a digraph from the DCM($\mathbf{d}^\pm$) ensemble which is strongly connected converges to 1 as $n \to \infty$. In what follows we will use the term *with high probability*, w.h.p., to mean that the probability of a certain event displays such a behaviour.

We will call *diameter* of the graph the maximum directed distance between two different vertices x and y, i.e., the length of the shortest directed path going from x to y. We can rephrase the result in [24] by saying that for sufficiently large values of n, the diameter is finite with probability near to one. We will present later a w.h.p. upper bound on the diameter, that is therefore sufficient to recover the strong connectivity result of [24].

At the core of our study we analyze the statistical properties of in- and out-neighbourhoods in the DCM. By in-neighbourhood of a vertex y of height h we mean the subgraph induced by the vertices which have a path of length at most h to y. On the other hand, the out-neighbourhood of a vertex x of height h is the subgraph induced by the vertices which are reachable by x in at most h steps. We will show that when we consider $h = \varepsilon \log(n)$ for some small $\varepsilon > 0$, in the DCM the sizes of both in- and out-neighbourhoods are "asymptotically equal" for all vertices. More precisely, by making use of *branching approximation* techniques, we show that each in/out-neighbourhood of height h contains $\widetilde{\Theta}(\nu^h)$ vertices, where

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

and the notation $\widetilde{\Theta}$ hides the poly-logarithmic corrections. As a consequence of this, it is not surprising that the diameter of a digraph from the DCM ensemble is typically of size $\log_\nu(n)$.

Theorem 1.1 ([20]) *Consider the DCM$(\mathbf{d}^\pm)$ ensemble, and set* $\mathrm{d}_\star = \log_\nu n$. *For every* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left((1-\varepsilon)\,\mathrm{d}_\star \le \mathrm{diam}(G) \le (1+\varepsilon)\,\mathrm{d}_\star\right) = 1.$$

Moreover, for any $x, y \in [n]$ *with* $x \neq y$

$$\lim_{n\to\infty} \mathbb{P}\left((1-\varepsilon)\,\mathrm{d}_\star \le d(x,y) \le (1+\varepsilon)\,\mathrm{d}_\star\right) = 1.$$

Theorem 1.1 also shows that the diameter is actually the *typical distance* between two vertices. By this we mean that the diameter is not realized by a special couple of vertices but rather, fixed any two vertices x and y, w.h.p. we will have $d(x,y) \sim \log_\nu(n)$.

1.1.2 The *Out Configuration Model* (OCM)

Also in the case of the OCM ensemble, we will work in the *sparse irreducible regime*, i.e., under the following assumption.

Assumption 2 *For every* $n \ge 2$

$$\delta := \min_{x\in[n]} d_x^+ \ge 2, \qquad \textit{and} \qquad \Delta := \max_{x\in[n]} d_x^+ = O(1).$$

Notice that the assumption above does not imply that the random in-degree sequence will enjoy the same bound as the out-degrees.

As far as we know, the Out Configuration Model has been introduced in [36], where the author investigates the case in which all the out-degrees coincide. We will refer to this special ensemble with the name *regular Out Configuration Model*, rOCM(d), where the symbol d stays for $d_x^+ = d$, for every vertex $x \in [n]$.

In the rOCM the out-neighbourhood of most vertices will look like a d-regular directed tree up to some relatively small height h. Nonetheless, if we focus on the in-neighbourhoods, these are well approximated by Galton Watson trees with Poisson offspring distribution of mean d. The first thing to notice is that the latter heuristics suggests that w.h.p. the digraph will be not strongly connected. This fact can be seen clearly by realizing that the probability that any given vertex y has in-degree zero is lower bounded by a constant uniformly in n. The first rigorous result about connectivity in the regular OCM ensemble comes with [36], where it is shown that w.h.p. the digraph has a unique strongly connected component, having size approximately $c_d n$, where the constant c_d is related to the survival probability of a Poisson(d) Galton-Watson tree, i.e.,

$$c_d := \max\{x \in \mathbb{R} \mid 1 - x = e^{-dx}\}.$$

Theorem 1.2 ([36]) *Fix $\varepsilon > 0$ and consider the a regular OCM with out-degree d. Then, called G_0 the largest strongly connected component of G, it holds*

$$\lim_{n\to\infty} \mathbb{P}\left((1-\varepsilon)c_d n \le |G_0| \le (1+\varepsilon)c_d n\right) = 1, \qquad \lim_{n\to\infty} \mathbb{P}\left(\forall x \in [n],\ \max_{y\in G_0} \mathrm{dist}(x,y) < \infty\right) = 1.$$

Henceforth, the intuition of the poissonian approximation is supported by rigorous quantitative results. It is worth noting that the second part of Theorem 1.2 ensures also that the strongly connected component is unique and that it is *globally attracting*, in the sense that w.h.p. every vertex admits a path toward G_0.

On the same line of investigation, in [2], the authors show that the typical value of the diameter is related to the same constant c_d where, of course, here by diameter is meant the diameter of the strongly connected component G_0. The heuristics suggests that among the "poissonian in-neighbourhoods" of the $c_d n$ vertices in G_0, there will be some that are *extremely thin* up to a relatively large height. Clearly, these vertices will be "far" from most of the origins, so that they give rise to a large value for the diameter. Let us remark, as a benchmark, that by Theorem 1.1 the diameter of a typical graph in the d-out-regular DCM ensemble is $\log_d(n)$, independently on the prescribed in-degree sequence. In the case of the rOCM, the presence of such *thin in-neighbourhoods* translates in a typical diameter of $(1+c'_d)\log_d(n)$, where $c'_d > 0$ is defined by

$$c'_d := \frac{\log d}{dc_d - \log d} \ge 0.$$

Theorem 1.3 ([2]) *Consider the rOCM(d) ensemble. For every $\varepsilon > 0$ it holds*

$$\lim_{n\to\infty} \mathbb{P}\left((1+c'_d-\varepsilon)\log_d(n) \le \mathrm{diam}(G) \le (1+c'_d+\varepsilon)\log_d(n)\right) = 1.$$

1.2 Random walks on random digraphs

In this thesis we are mainly concerned with structural properties of typical graphs from the two ensembles which can be better described through the lens of *random walks*. A random walk on a directed graph is one of the easiest examples of Markovian stochastic process. The latter can be thought of as a moving particle sitting on the vertex set of the graph. At each discrete time step the particle chooses uniformly at random one of the out-going links (tails) of the vertex it is currently visiting, and crosses it. It is clear that any claim related to the law of a random walk on a given digraph can be translated in a claim on the geometry of the digraph. Nonetheless, in our case the digraph is a random object, so the law of the randomly moving particle will be a random object itself. In the language of *statistical physics* the latter is an example of a so-called *disordered system*, namely a random dynamical system evolving in a random geometry. The main difference between the time evolution of the random walk on a directed graph with respect to the undirected case is the *irreversibility* of the dynamics. The mathematical theory of *reversible* systems is much richer, rooting its solid basis on the mathematics of self-adjoint operators and spectral theory. Contrarily, *irreversible disordered systems* are less understood, and it is nowadays a challenging task to design suitable setting and techniques that could help in the mathematical understanding of the subject.

1.2.1 The stationary distribution

Much of the theory of finite Markov chains is concerned with the *equilibrium* properties of the process under investigation. Given a transition matrix P, we will call *stationary distribution* any probability (row) vector satisfying $\pi = \pi P$. It is classical that, if P is the the transition matrix of a random walk on the directed graph G, the strong connectedness of G yields the uniqueness of π. In the realm of undirected graphs, it is easy to check that if π is unique then it must be proportional to the degree vector $\mathbf{d}$, independently of the other specifics of G. What makes undirected graphs specials in this sense is not the *reciprocity* of the link structure, but rather the *balancedness* of in- and out-degrees. Indeed, if the graph is directed but Eulerian, i.e., $d_x^- = d_x^+$ for every vertex $x \in [n]$, then $\pi \propto \mathbf{d}^+$ regardless of the other features of the digraph. As soon as an *imbalanced* vertex is present in the graph, the stationary measure has to be though of as a *global* property of the digraph geometry and there are no general procedures to compute the stationary measure π in a *local way*.

In the case of digraphs from the DCM ensemble, the w.h.p. strong connectivity ensures that a typical digraph admits a unique stationary measure for the random walk. Nonetheless, the actual value of the vector π is random. In such a framework two natural questions that arise are: "how does the stationary distribution look like in the large n limit?", "is there any reasonable approximation of π in terms of local observables of the graph?". These questions have been addressed in [13, 14, 18, 20]. In particular in [13] the authors analyze the *bulk* of the stationary measure of the random walk on the DCM. More precisely, they consider the (rescaled) *empirical stationary distribution*, which is simply the random probability measure

$$\frac{1}{n}\sum_{x\in[n]}\delta_{n\pi(x)}(\cdot)$$

where we denote by δ_a the Dirac mass at a. The random measure above can be thought of as the law of the stationary distribution of a uniformly sampled vertex. In [13] it is shown that the empirical stationary distribution is well approximated by a deterministic law in the Wasserstein-1 sense. The crucial idea underlying this result is the analysis of an L^2-bounded martingale which is related to the branching approximation of the in-neighbourhood of a uniformly sampled vertex. The approximating deterministic distribution can be characterized as the unique solution, μ, of a distributional fixed point equation. This equation has been intensively studied for other reasons e.g. in [44, 41, 43], where the author provides detailed results on the absolutely continuity of the solution and presents a precise analysis of its tails.

The case of the OCM ensemble can be carried out analogously, see [18]. In the special case of the rOCM(d) ensemble the result is easier to read and can be, again, explained in terms of Galton-Watson trees. In fact, in that case, the martingale approximating the empirical stationary distribution is the classical martingale associated to a Galton-Watson process of mean-offspring d, i.e., Z_t/d^t, where Z_t is the population of the Galton-Watson process at time t. This fact, in particular, can be read as an alternative proof of the first part of Theorem 1.2. Indeed, the mass at zero of the empirical stationary distribution is exactly the size of fraction of vertices which are not on G_0, so that it follows from the convergence result in [18] that the fraction of vertices in the strongly connected component is asymptotically equal to the probability that a Poisson Galton-Watson tree of mean offspring d survives at infinity.

Unfortunately, the techniques used to study the *typical entry* of the stationary distribution, do not automatically extend to recover results on the *extremal values*. In fact, even if the convergence of the empirical distribution yields a *limit distribution* whose tail behaviour is known, it is not clear how we could "bring this back" to the finite setting. In other words, if for a large finite n we ask questions about the existence of a vertex with stationary distribution smaller than some $f(n) = o(n^{-1})$ or greater than some $g(n) = \omega(n^{-1})$, the answer cannot follow from the analysis of the limiting distribution. Nevertheless, the latter could provide us with a *reasonable guess* about the correct answer. As mentioned above, in [44, 41, 43] the author studies the tails of the limiting law μ, and one could be tempted to use those results in the following, heuristic, way. We can solve with respect to the variable x the equation

$$\mu[0, x(n)) \approx \frac{1}{n}$$

where $\mu[0, x)$ denotes the probability mass associated to the interval $[0, x)$ by the solution of the distributional fixed point equation which approximates the empirical stationary distribution. It turns out that the value predicted by this heuristics is indeed the right answer. In fact, we show rigorously in [20] that the minimal entry of the stationary distribution is of the form $(n \log^C(n))^{-1}$ where the value of the exponent C coincides with the prediction.

In order to state our result avoiding technicalities, let us focus on the special case in which half of the vertices have degrees (δ, Δ) whereas the other half share the degrees (Δ, δ). We will refer to this specific sub-model with the name *binary Directed Configuration Model*, bDCM(δ, Δ). We stress that is this specific case our results are sharper than in the general scenario of Assumption 1. We call

$$\gamma = \frac{\log \Delta}{\log \delta}.$$

The results in [41] imply

$$\log \mu[0, x) \approx x^{-\frac{1}{\gamma - 1}},$$

so that we can expect

$$n\pi_{\min} \overset{\text{w.h.p.}}{\approx} \log^{1-\gamma}(n).$$

The same argument applies to the case of the maximal entry of the stationary measure π. One of the main results of this thesis, presented in [20], shows that the following holds true.

Theorem 1.4 ([20]) *Consider the bDCM(δ, Δ) ensemble. Set $\pi_{\min} = \min_{x \in [n]} \pi(x)$ and $\pi_{\max} = \max_{x \in [n]} \pi(x)$ and call*

$$\gamma = \frac{\log \Delta}{\log \delta} \geq 1.$$

There exists a constant $C \equiv C(\Delta) > 0$ such that

$$\lim_{n \to \infty} \mathbb{P}\left(C^{-1} \log^{1-\gamma}(n) \leq n\pi_{\min} \leq C \log^{1-\gamma}(n)\right) = 1.$$

$$\lim_{n \to \infty} \mathbb{P}\left(C^{-1} \log^{1-\gamma^{-1}}(n) \leq n\pi_{\max} \leq C \log^{1-\gamma^{-1}}(n)\right) = 1.$$

For what concerns the rOCM, it has been shown in [2], as a byproduct of the tools used in analyzing the diameter, that

$$n\pi_{\min} \overset{\text{w.h.p.}}{\approx} n^{-c'_d},$$

where the constant c'_d is the same as in Theorem 1.3. This result can be predicted by a heuristics similar to the one for the DCM.

1.2.2 Cover time

When studying the law of a random walk on a graph, it can be of interest to characterize the latter by means of its *extremal properties*, such as the *maximal hitting time* or the *cover time*. This can be somehow thought of as a line of investigation in the same spirit of the study of the extremal values of the stationary distribution. Indeed, as we will see in a moment, the three quantities are strongly related.

The problem of determining the cover time of a graph is a central one in combinatorics and probability [6, 3, 39, 4, 45, 31, 32, 47, 48]. By definition, the cover time of a Markov chain is the maximum over the starting states of the expected time needed to visit all the states of the chain. In a couple of papers, [31, 32], Feige presents general tight upper and lower bounds for the cover time in the undirected graph setting. In particular, he shows that the cover time can be at most polynomial in the size of the vertex set for *every* undirected graph. Nonetheless, the results do not extend to the directed case. In fact, it is not hard to exhibit examples in which the cover time is exponential in the size of the graph. A first—well known—result on the cover time for a general Markov chain is the so called *Matthews' bound*, [47]. By means of a *coupon-collector argument*, the author shows that the cover time must lie within a logarithmic factor from the *maximal hitting time*, which is defined as the maximum over the ordered couples (x, y) of the expected hitting time of y starting at x.

Passing to the *random graph setting*, in recent years the cover time of random walks on random graphs has been extensively studied [38, 27, 25, 28, 1]. All these works consider undirected graphs, with the notable exception of the paper [28] by Cooper and Frieze, where the authors compute the cover time of directed Erdős-Renyi random graphs in the regime of strong connectivity, that is with a logarithmically diverging average degree. The main difficulty in the directed case is that, in contrast with the undirected setting, the graph's stationary distribution is an unknown random variable.

The techniques developed by Cooper and Frieze—which are the same we adopted in [20], where we study the cover time for the DCM ensemble—are crucially based on a lemma, which they named "First Visit Time Lemma" (FVTL). In our sparse and directed setting, the FVTL claims that the random variable that counts the number of steps needed by the walker to visit a given vertex x after a suitable notion of *mixing time*, has an exponential right tail with rate given by the stationary measure of x. As we mentioned above, having a precise control on the stationary measure is a challenging task. Clearly, a special role is played by the subclass of Eulerian digraphs. We show in [20] that the cover time in this case is of order $Cn\log(n)$, with the constant C depending mostly on the vertices of low degree. More precisely, since in the *sparse irreducible regime* we can have at most a finite number of classes of vertices with different degrees, we show that the constant C depends on the single class of vertices in which a certain trade-off between degree and size is optimal.

Theorem 1.5 ([20]) *Consider an Eulerian DCM($\mathbf{d}^{\pm}$) ensemble. Call $\mathcal{V}_d$ the set of vertices of degree d, and write $\bar{d} = m/n$ for the average degree. Assume*

$$|\mathcal{V}_d| = n^{\alpha_d + o(1)} \tag{1.1}$$

for some constants $\alpha_d \in [0,1]$, for each type d. Then, for every $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left((\beta-\varepsilon)n\log n \le T_{\text{cov}} \le (\beta+\varepsilon)n\log n\right) = 1,$$

where $\beta_n \equiv \beta := \bar{d}\,\max_d \frac{\alpha_d}{d}$.

In the general case, we can exploit the results about the extremal values of the stationary distribution in Theorem 1.4 to find asymptotic estimates of the cover time. Of course, given that our estimates on the minimal entry of π are up to constant, we cannot expect to have a result on the cover time that is able to recover the first order asymptotic. By using the *Cooper and Frieze's recipe* we can prove the following result.

Theorem 1.6 ([20]) *Consider the binary DCM(δ, Δ) ensemble. Then, there exists a constant $C \equiv C(\Delta)$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left(C^{-1} n \log^{\gamma}(n) \le T_{cov} \le C n \log^{\gamma}(n)\right) = 1.$$

1.3 Mixing time: cutoffs and trichotomies

Consider a Markov chain on n states, with transition matrix P and stationary measure π. A classical problem, see e.g. [40], is to determine the time at which—informally—the process "forgets its starting state". More precisely, given a starting distribution λ on $[n]$, and a metric on the space of probability distributions over the n states, one is interested in the time at which the distance between the measure λP and the stationary measure π is below a given threshold. The *total variation distance* between two probabilities μ, ν on $[n]$ is defined as half of the ℓ_1 distance between the vectors μ and ν, namely

$$\|\mu - \nu\|_{\mathrm{TV}} := \frac{1}{2} \sum_{x \in [n]} |\mu(x) - \nu(x)|.$$

For any $\varepsilon \in (0,1)$, the *ε-mixing time* is the first time at which the total variation distance between λP and π is smaller than ε for every starting distribution λ. In formula

$$T^{(n)}_{mix}(\varepsilon) \equiv T_{mix}(\varepsilon) = \inf \left\{ t \geq 1 \;\middle|\; \max_{x \in [n]} \|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} < \varepsilon \right\}.$$

It is easy to show that for every measure λ the quantity $\|\lambda P^t - \pi\|_{\mathrm{TV}}$ is a non-increasing function of t. In the seminal paper [5], the authors formalized for the first time such a concept, presenting an example of a sequence of chains—parametrized by the size of the state space—for which asymptotically, for every $\varepsilon \in (0,1)$,

$$\lim_{n \to \infty} \frac{T_{mix}(\varepsilon)}{T_{mix}(1-\varepsilon)} = 1.$$

Notice that the insensitivity of the leading order of $T_{mix}(\varepsilon)$ to the value of $\varepsilon \in (0,1)$ is equivalent to say that the distance to equilibrium approaches a step function on a certain time scale. This phenomenon is commonly referred to as *cutoff*. The number of examples of chains exhibiting such a behaviour is constantly growing, see e.g. [29, 40] for a review.

In particular, for what concerns this thesis, we restrict to the following setting. Let G be a random graph on n vertices, $P^t(x,\cdot)$ denotes the distribution after t steps of the random walk started at x, and π represents the stationary distribution, which we assume to be unique. We can therefore consider, for every starting point $x \in [n]$, the random sequence

$$\mathcal{D}_x(t) = \|P^t(x,\cdot) - \pi\|_{\mathrm{TV}}, \qquad \forall t \geq 0.$$

If a graph from the ensemble admits a unique stationary distribution w.h.p., we are interested in showing claims of the following form: there exist some deterministic sequence $T_\star \equiv T_\star(n)$ such that for every $\varepsilon > 0$

$$\lim_{n \to \infty} \mathbb{P}\left(\max_{x \in [n]} |\mathcal{D}_x(sT_\star) - \vartheta(s)| \leq \varepsilon \right) = 1\,, \qquad \forall s > 0,\ s \neq 1, \tag{1.2}$$

where ϑ denotes the step function $\vartheta(s) = 1$ if $s \leq 1$ and $\vartheta(s) = 0$ if $s > 1$. We remark that the definition in (1.2) is slightly different than the definition of cutoff given above, since (1.2) requires the uniformity over the starting position. We refer to [46, 9] for similar results in the case of undirected graphs.

The mixing properties for the models we focus on in this thesis have been studied by Bordenave, Caputo and Salez in [13, 14]. In order to present their results in a unified fashion, we need to introduce some quantities. Therefore, we shall adopt the following unified notation. Let us define the *in-degree distribution*

$$\mu_{\text{in}}(x) := \frac{1}{n} \times \begin{cases} d_x^-/\bar{d} & \text{DCM}(\mathbf{d}^\pm) \\ 1 & \text{OCM}(\mathbf{d}^+) \end{cases} \tag{1.3}$$

where we use the notation

$$\bar{d} := \frac{1}{n} \sum_{x \in V} d_x^+ = \frac{m}{n}$$

for the average degree. Let the *average row entropy* H and the associated *entropic time* T_{ENT} be defined by

$$H := \sum_{x \in V} \mu_{\text{in}}(x) \, \log d_x^+, \qquad T_{\text{ENT}} := \frac{\log n}{H}.$$

Note that in the sparse irreducible regime the deterministic quantities H, T_{ENT} satisfy $H = \Theta(1)$ and $T_{\text{ENT}} = \Theta(\log n)$. By employing the same terminology as in (1.2), the main result in [13, 14] is the following.

Theorem 1.7 ([13, 14]) *Consider the DCM($\mathbf{d}^\pm$) ensemble or the OCM($\mathbf{d}^+$) ensemble. For every* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x \in [n]} |\mathcal{D}^x(sT_{\text{ENT}}) - \vartheta(s)| \leq \varepsilon \right) = 1\,, \qquad \forall s > 0,\ s \neq 1.$$

We remark that, in the case of the DCM, by Jensen's inequality the mixing time $T_{\text{ENT}} = \frac{\log n}{H}$ is always larger than the diameter $\mathrm{d}_\star = \log_\nu(n)$,

$$H = \sum_{x=1}^{n} \frac{d_x^-}{m} \log d_x^+ \leq \log \left(\sum_{x=1}^{n} \frac{d_x^-}{m} d_x^+ \right) = \log \nu,$$

with equality if and only if the sequence is out-regular, that is $d_x^+ = d$, for every $x \in [n]$. Thus, the analysis of convergence to stationarity requires investigating the graph on a length scale that may well exceed the diameter. The cutoff phenomenon described in [13, 14] can be read in terms of a law of large number for the weight of the path followed by the random walk on the digraph. We will use the notation $\mathbf{P}_x$ to denote the *quenched law* of the random walk, namely the random probability law that is determined by the realization of the digraph G. Bordenave, Caputo and Salez show that for every $t = \Theta(\log(n))$ w.h.p. the walker will follow a path of weight approximately e^{-Ht}. More explicitly

Theorem 1.8 ([13, 14]) *Consider the DCM($\mathbf{d}^{\pm}$) or the OCM($\mathbf{d}^{+}$) ensemble. For every $\varepsilon, \varepsilon' > 0$, if $t = \Theta(\log(n))$ it holds*

$$\lim_{n\to\infty} \mathbb{P}\left(\min_{x\in[n]} \mathbf{P}_x\left(\prod_{s=0}^{t-1} P(X_s, X_{s+1}) \in \left[e^{-(1+\varepsilon)Ht}, e^{-(1-\varepsilon)Ht}\right]\right) > 1-\varepsilon'\right) = 1.$$

The theorem above implies that if $t < T_{\mathrm{ENT}}$ then the probability distribution $P^t(x,\cdot)$ will be w.h.p. concentrated on a small number of vertices. Hence, one of the main ingredient in proving the *cutoff at the entropic time* is such a law of large numbers. In [19] we extend the result in Theorem 1.8 to the case in which the underlying random digraph is replaced by an independent and identically distributed copy at some $s \le t$. In particular, we show that in such a framework the complete *cutoff phenomenology* is preserved, w.h.p. with respect to the joint sample. More formally, in [19] we prove the following theorem.

Theorem 1.9 ([19]) *Consider the DCM($\mathbf{d}^{\pm}$) or the OCM($\mathbf{d}^{+}$) ensemble. Fix any $C > 0$ and set $t = CT_{\mathrm{ENT}}$ and $s \le t$. Fixed any couple of configurations $\omega, \omega' \in \mathfrak{C}$, we consider the probability distribution*

$$[Q^t_s(x,y)](\omega,\omega') \equiv Q^t_s(x,y) := \sum_{z\in[n]} P^s_\omega(x,z) P^{t-s}_{\omega'}(z,y).$$

Let ω and ω' two uniformly random and independent configurations from DCM($\mathbf{d}^{\pm}$) or from OCM($\mathbf{d}^{+}$) ensemble. Let $\mathbb{P}$ denote the probability law of the joint and independent sample. Then, for every $\varepsilon > 0$

1. *If $C < 1$:*

$$\lim_{n\to\infty} \mathbb{P}\left(\min_{x\in[n]} \|Q^t_s(x,\cdot) - \pi_{\omega'}\|_{TV} > 1-\varepsilon\right) = 1.$$

2. *If $C > 1$ and $t - s = \omega(1)$*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \|Q^t_s(x,\cdot) - \pi_{\omega'}\|_{TV} < \varepsilon\right) = 1.$$

In other words, the latter says that it is the rate at which the probability mass spreads in the graph, rather than the particular geometry, to give rise to the cutoff phenomenon.

Beside the cutoff phenomenon, in the very last years a new form of mixing appeared in the probabilistic literature, see [8, 18, 19, 50]. In [8], the authors show an example of parametric (non-Markovian) stochastic process exhibiting a particular mixing phenomenon. Such a behaviour can be thought of as an interpolation between a *cutoff* and an exponential decay for the total variation distance, depending on the value of the parameter. By the title of the article [8], we refer to this kind of dependency of the mixing time on the parameter of the model with the name of *trichotomy phenomenon*.

In [18, 19] we present three examples of Markovian and non-Markovian processes exhibiting *trichotomy* in their total variation mixing. Each of these examples is related to the particular underlying geometry, which is taken to be a digraph from the DCM or the OCM ensemble.

The following two subsections are devoted to the presentation of the models we investigated and the corresponding results.

1.3.1 Mixing time of the PageRank surfer

Given a directed graph G on n vertices and a parameter $\alpha \in (0,1)$, the PageRank surf on G with damping factor $1-\alpha$ is the Markov chain with state space $[n]$ and transition probabilities given by

$$P_\alpha(x,y) = (1-\alpha)P(x,y) + \frac{\alpha}{n}, \tag{1.4}$$

where P denotes the transition matrix of the simple random walk on G. The interpretation is that of a surfer that at each step, with probability $1-\alpha$ moves as a simple random walk, and with probability α moves to a uniformly random vertex. The surfer reaches eventually a stationary distribution π_α over $[n]$, called the PageRank of G. Since its introduction by Brin and Page in the seminal paper [17], PageRank has played a fundamental role in the ranking functions of all major search engines; see e.g. [30, 35]. A common generalisation is the so-called customised or generalised PageRank, where the uniform resampling is replaced by an arbitrary probability distribution λ over $[n]$, so that (1.4) becomes

$$P_{\alpha,\lambda}(x,y) = (1-\alpha)P(x,y) + \alpha\lambda(y).$$

The resulting stationary distribution $\pi_{\alpha,\lambda}$, characterised by the equation

$$\pi_{\alpha,\lambda}(y) = \sum_{x\in[n]} \pi_{\alpha,\lambda}(x)P_{\alpha,\lambda}(x,y),$$

depends in a nontrivial way on the parameter α and the distribution λ. There have been several investigations of the structural properties of $\pi_{\alpha,\lambda}$; see e.g. [37, 7, 16]; we refer in particular to the recent works [22, 34] for cases where the graph G is a random.

We focus on the dynamical problem of determining the time needed for the surfer to reach the equilibrium distribution $\pi_{\alpha,\lambda}$, namely we study the mixing time of the Markov chain with transition matrix $P_{\alpha,\lambda}$. Even for graphs where the mixing time of the simple random walk is well understood, it is in general not immediate to deduce the influence of the parameter α and of the resampling distribution λ on the speed of convergence to equilibrium.

To obtain explicit asymptotic statements we shall assume that $\alpha \equiv \alpha_n \in (0, 1)$ is a sequence such that $\alpha \to 0$ and that the limit

$$\gamma := \lim_{n\to\infty} \alpha\, T_{\mathrm{ENT}} \in [0, \infty] \tag{1.5}$$

exists, with possibly $\gamma = 0$ or $\gamma = \infty$.

As for the simple random walk, for every vertex $x \in [n]$ we can define the total variation distance from stationarity as the sequence

$$\mathcal{D}^x_{\alpha,\lambda}(t) := \left\| P^t_{\alpha,\lambda}(x, \cdot) - \pi_{\alpha,\lambda} \right\|_{\mathrm{TV}}, \qquad \forall t \geq 0.$$

Let us remark that when $\alpha = 0$ we have, regardless of the probability distribution λ, $\pi_{0,\lambda} \equiv \pi$ and $\mathcal{D}^x_{0,\lambda}(t) \equiv \mathcal{D}_x(t)$.

It is intuitively reasonable to guess that if the parameter α is suitably large compared to the inverse of the mixing time of the graph G, then the time to reach stationarity will be essentially the expected time needed to make the first λ-resampling transition, that is a geometric random variable with parameter α. On the other hand, if α is suitably small compared to the inverse of the mixing time of the graph G, then one should reach stationarity well before the first λ-resampling, so that the speed of convergence to equilibrium will be essentially that of the simple random walk on G. Moreover, one could expect that when α is neither too small nor too large compared to the inverse of the mixing time of the graph G, then some interpolation between the two opposite behaviours should take place. In [18] we substantiate this intuitive picture for a typical graph from the DCM or the OCM ensemble, under the assumption that λ is either not too strongly localized, or that λ is very strongly localized, in the following sense.

On the one hand, the class of *widespread* probability measures is defined as follows.

Definition 1.1 (Widespread measure) *A sequence of probability measures $\lambda \equiv \lambda_n$ on $[n]$ is* widespread *if*

(i) *There exists $\varepsilon > 0$ such that*

$$|\lambda|_\infty = \max_{x\in[n]} \lambda(x) = O(n^{-1/2-\varepsilon}).$$

(ii) *Has a bounded ℓ_2-norm, in the following sense:*

$$n \sum_{j\in[n]} \lambda(j)^2 = O(1).$$

Note that there is no requirement on the minimum of $\lambda(x)$, so that large portions of the set of vertices are allowed to receive zero mass.

On the other hand, we consider also measures which are *strongly localized*.

Definition 1.2 (Strongly localized measure) *A sequence of probability measures $\lambda = \lambda_n$ on $[n]$ is* strongly localized *if it is the convex combination of a finite number of Dirac masses on vertices. Namley, there exists a constant $C > 0$ independent of n, a set $F \subset [n]$ with cardinality $|F| \leq C$ and coefficients $(a_z)_{z\in F}$ such that $a_z \in [0,1]$, $\sum_{z\in F} a_z = 1$, and*

$$\lambda = \sum_{z\in F} a_z\, \delta_z$$

where δ_z is the Dirac mass at vertex z.

The main result of [18] is the following.

Theorem 1.10 ([18]) *Consider the DCM($\mathbf{d}^\pm$) or the OCM($\mathbf{d}^+$) ensemble. Let $\alpha \equiv \alpha(n) \in (0,1)$ be parameters as in (1.5), and let $\lambda \equiv \lambda_n$ be either widespread or strongly localized measures. Then, for every $\varepsilon > 0$ according to the value of γ there are three scenarios:*

(1) *If $\gamma = 0$ then for all $s > 0, s \neq 1$:*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \left| \mathcal{D}^x_{\alpha,\lambda}(s\, T_{\mathrm{ENT}}) - \vartheta(s) \right| < \varepsilon \right) = 1.$$

(2) *If $\gamma \in (0,\infty)$ then for all $s > 0, s \neq 1$:*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \left| \mathcal{D}^x_{\alpha,\lambda}(s/\alpha) - e^{-s}\vartheta(s/\gamma) \right| < \varepsilon \right) = 1.$$

(3) *If $\gamma = \infty$ then for all $s > 0$:*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \left|\mathcal{D}^x_{\alpha,\lambda}(s/\alpha) - e^{-s}\right| < \varepsilon\right) = 1.$$

where ϑ is the step function $\vartheta(s) = 1$ if $s \le 1$ and $\vartheta(s) = 0$ otherwise.

The trichotomy displayed in Theorem 1.10 reflects the competition between two distinct mechanisms of relaxation to equilibrium: the simple random walk dominates in the first scenario, while the λ-resampling dominates in the third; the intermediate scenario interpolates between the two extremes; see Fig. 1.1.

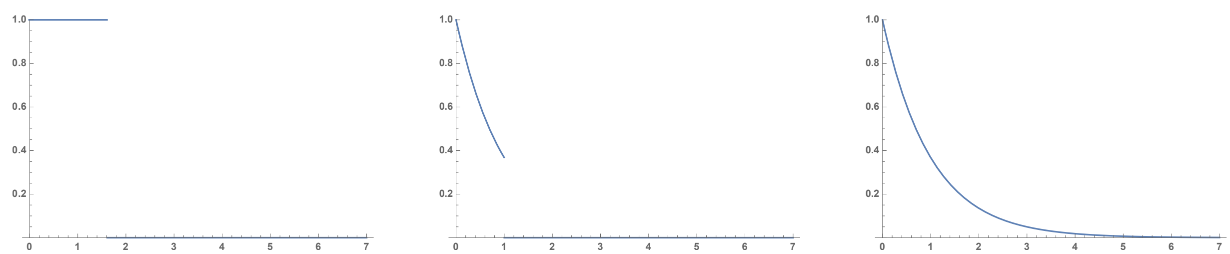

Figure 1.1: On the right a plot of the TV-distance to $\pi_{\alpha,\lambda}$ in the case $\gamma = 0$, on the scale $\Theta(\log(n))$ when $H = \log(5)$. In the middle, the case $\gamma = 1$ and the plot is on the scale $\Theta(\alpha^{-1})$. On the right the case $\gamma = \infty$ on the scale $\Theta(\alpha^{-1})$.

As mentioned above, essentially the same trichotomy was uncovered recently in [8] in a model of random walk on dynamically evolving undirected graphs. In that case, the role of the resampling is played by the underlying reshuffling of the graph edges. It is interesting to observe that, in contrast with the undirected case considered in [8], in our setting the two competing processes may well have very distinct goals, and the overall stationary distribution $\pi_{\alpha,\lambda}$ is the result of a nontrivial balance.

The fact that the above result holds in the two opposite regimes of widespread measures or strongly localized measures may come as a surprise. Indeed, the behaviour of the stationary distribution can be very different in these two cases. As we shall see, some parts of the proof require rather different strategies for the two regimes.

Inspired by the results in [13, 14], in [18] we make use of a key fact to attack the case of a general widespread measure λ, which we used also in the following works [20, 19]. The observation is that, if we start with a widespread distribution λ, then the time needed to reach stationarity for the simple random walk is much smaller than the entropic time T_{ENT}. More precisely we established the following fact.

Theorem 1.11 ([18]) *Consider the DCM($\mathbf{d}^{\pm}$) or the OCM($\mathbf{d}^{+}$) ensemble. If $\lambda \equiv \lambda_n$ is widespread, then for any sequence $t \equiv t_n \to \infty$, and every $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathbb{P}\left(\left\| \lambda P^t - \pi_0 \right\|_{TV} < \varepsilon \right) = 1.$$

In the special case where $\lambda = \mu_{\text{in}}$, and only for DCM ensemble, a similar result was already obtained in [13]. Following the same approach, the proof of Theorem 1.11 is based on the construction of a martingale approximation for the distribution λP^t. The latter rests on the branching approximation mentioned in Section 1.1, which allows one to couple the in-neighbourhood of a uniformly distributed random vertex of G with a marked Galton-Watson tree up to depth $t = o(\log n)$.

1.3.2 Mixing time on regenerating dynamic digraphs

As mentioned above, the emergence of a *trichotomy* in the TV-mixing is due to the coexistence of two competing mixing mechanisms. In the same stream of [8] we study the mixing of a joint chain on the product set of the digraphs from the ensemble and the vertex set. Informally, we consider a simple random walk moving in an evolving graph. At each time step the walker perform the usual step with probability $(1-\alpha)$, while with probability α the walker sticks on the vertex it is currently visiting and a new configuration is sampled. Fixed the digraph ensemble DCM($\mathbf{d}^{\pm}$) or OCM($\mathbf{d}^{+}$), we will refer with the symbol ω to the configuration inducing the digraph $G(\omega)$ from the ensemble under consideration. Moreover, we use the notation $\mathbf{u}(\omega)$ to denote the probability of sampling configuration $\omega \in \mathfrak{C}$, namely $\mathbf{u}(\omega) = |\mathfrak{C}|^{-1}$, and the symbol $\mathbf{P}^{\mathbf{J}}_{\omega,x}$ to denote the law of joint process when the starting state is (ω, x). The joint transition matrix is given by

$$\mathcal{P}_\alpha((\omega,x),(\omega',y)) := (1-\alpha)P_\omega(x,y)\mathbf{1}_\omega(\omega') + \alpha\mathbf{u}(\omega')\mathbf{1}_x(y).$$

Clearly, the process described by the first coordinate is Markovian and mixes exponentially fast on the scale $\Theta(\alpha^{-1})$, yet, what can be said about the joint chain? It is easy to see that the joint process admits a unique stationary distribution, which we will refer to as π^J_α. If we consider the stochastic non-Markovian process obtained by projecting on the second coordinate, what can be said about the asymptotic behaviour of the total variation distance between the evolution of the walk and the marginal distribution of the second coordinate of π^J_α? We answer to these questions in [19] where we show that both the quantities exhibit a trichotomy in dependence of the limit quantity γ, defined in (1.5). The first question we answer concerns the stationary distribution π^J_α. In [19] we show that, regardless on the specific sequence $\alpha \equiv \alpha_n$ satisfying $\alpha \to 0$, a good TV-proxy for $\pi^J_\alpha \equiv \pi^J$ is given by the measure

$$\lambda(\omega, x) := \mathbf{u}(\omega)\pi_\omega(x), \tag{1.6}$$

where we called $\pi_\omega(\cdot)$ the stationary measure of the random walk on ω if it exists unique, and measure μ_{in} otherwise. Called $(\xi_t, X_t)_{t\geq 0}$ the trajectory of the joint Markovian process, we show the result just mentioned by presenting sharp bounds on the total variation distance

$$\mathcal{D}^{\mathrm{J},\alpha}_{\omega,x}(t) = \|\mathbf{P}^{\mathbf{J}}_{\omega,x}(\xi_t = \cdot, X_t = \cdot) - \lambda\|_{\mathrm{TV}}.$$

Theorem 1.12 ([19]) *Consider in DCM($\mathbf{d}^{\pm}$) or the OCM($\mathbf{d}^{+}$) ensemble. Call $\mathbb{P}$ the law of the starting configuration $\xi_0 \equiv \omega$. Then for every sequence $\alpha \equiv \alpha_n \to 0$ and for every $\varepsilon > 0$ it holds the following trichotomy*

1. *If $\gamma = 0$ then for all $s > 0$:*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \left|\mathcal{D}^{\mathrm{J},\alpha}_{\omega,x}(s\alpha^{-1}) - e^{-s}\right| < \varepsilon\right) = 1.$$

2. *If $\gamma = \infty$ then for all $s > 0$:*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \left|\mathcal{D}^{\mathrm{J},\alpha}_{\omega,x}(s\alpha^{-1}) - (1+s)e^{-s}\right| < \varepsilon\right) = 1.$$

3. *If $\gamma \in (0,\infty)$ then for all $s > 0$, $s \neq \gamma$:*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \left|\mathcal{D}^{\mathrm{J},\alpha}_{\omega,x}(s\alpha^{-1}) - \psi_\gamma(s)\right| < \varepsilon\right) = 1.$$

where

$$\psi_\gamma(\beta) = \begin{cases} (1+s)e^{-s} & \textit{if } s < \gamma \\ e^{-s} & \textit{if } s > \gamma \end{cases}.$$

As one can suspect, the result in Theorem 1.9 plays a fundamental role in proving the trichotomy in Theorem 1.12.

Finally, our last result concerns the marginal distribution of the position of the walk, namely the non-Markovian process obtained by projecting the chain $(\xi_t, X_t)_{t\geq 0}$ on the second coordinate. As we will show below, the law of X_t, for t and n suitably large, should

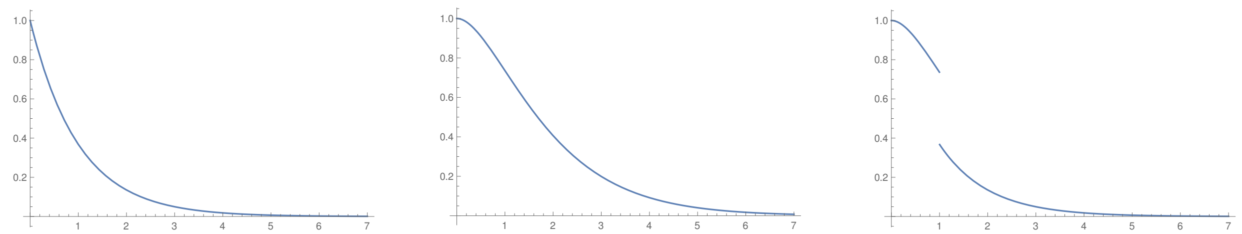

Figure 1.2: The asymptotic behavior on the scale α^{-1} of the quantity $\mathcal{D}^{\mathrm{J},\alpha}_{\sigma,x}(t)$ for a typical starting environment σ and arbitrary $x \in [n]$ in the case $\gamma = 0$ (left), $\gamma = \infty$ (center) and $\gamma \in (0, \infty)$ (right). The transition point in this last scenario is $t = \gamma\alpha^{-1} \sim T_{\mathrm{ENT}}$, and we set $\gamma = 1$.

be well approximated by μ_{in}. The next result quantifies this statement by exhibiting once again a trichotomy. Define

$$\mathcal{D}^{rw,\alpha}_{\omega,x}(t) := \|\mathbf{P}^{\mathbf{J}}_{\omega,x}(X_t = \cdot) - \mu_{\mathrm{in}}\|_{\mathrm{TV}}\,, \qquad q := \mathbb{E}\|\pi_\omega - \mu_{\mathrm{in}}\|_{\mathrm{TV}}.$$

We remark that if the sequences $\mathbf{d}^{\pm}$ are Eulerian then $\pi_\omega = \mu_{\mathrm{in}}$ is stationary. Thus in this case $q = 0$. On the other hand, results from [42, 44] imply that if the sequence is not Eulerian then q is bounded away from zero and one.

Theorem 1.13 ([19]) *Consider the DCM($\mathbf{d}^{\pm}$) or the OCM($\mathbf{d}^{+}$) ensemble. Fix a sequence $\alpha \equiv \alpha_n \to 0$. Then, for all $s > 0$*

$$\limsup_{n\to\infty} \max_{\omega,x} \mathcal{D}^{rw,\alpha}_{\omega,x}(s\alpha^{-1}) \le e^{-s}.$$

Moreover, for every $\varepsilon > 0$ the following trichotomy takes place

1. *If $\gamma = 0$ then for all $s > 0$:*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \left|\mathcal{D}^{rw,\alpha}_{\omega,x}(s\alpha^{-1}) - q\,e^{-s}\right| < \varepsilon\right) = 1.$$

Moreover, if $s \neq 1$ then

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \left|\mathcal{D}^{rw,\alpha}_{\omega,x}(sT_{\mathrm{ENT}}) - \varphi(s)\right| < \varepsilon\right) = 1,$$

where

$$\varphi(s) := \begin{cases} 1 & \text{if } s < 1 \\ q & \text{if } s > 1. \end{cases}$$

2. *If $\gamma = \infty$, then for all $s > 0$:*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \left| \mathcal{D}^{rw,\alpha}_{\omega,x}(s\alpha^{-1}) - e^{-s} \right| < \varepsilon \right) = 1.$$

3. *If $\gamma(0,\infty)$ then for all $s > 0$, $s \neq \gamma$:*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} \left| \mathcal{D}^{rw,\alpha}_{\omega,x}(s\alpha^{-1}) - \varphi(s/\gamma)e^{-s} \right| < \varepsilon \right) = 1.$$

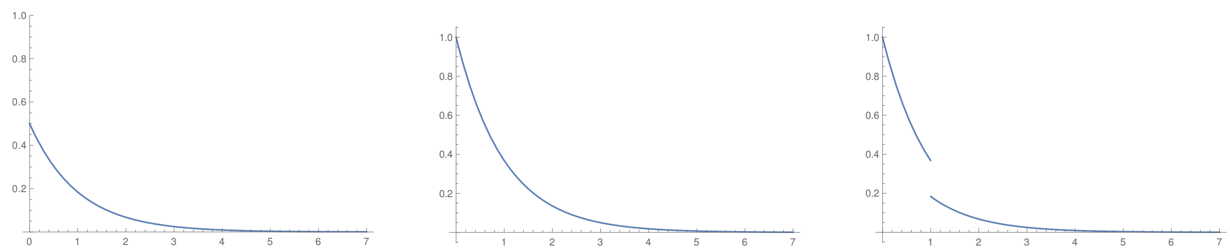

Figure 1.3: The asymptotic behavior on the scale α^{-1} of the quantity $\mathcal{D}^{\alpha}_{\sigma,x}(t)$ for a typical starting environment σ and arbitrary $x \in [n]$ in the case $\gamma = 0$ (left), $\gamma = \infty$ (center) and $\gamma \in (0,\infty)$ (right). The transition point in the latter case is $t = \gamma\alpha^{-1} \sim T_{\mathrm{ENT}}$. In this picture we take $\gamma = 1$ and $q = 1/2$.

1.4 Organization of the thesis

This thesis contains the results I obtained during my PhD, jointly with my advisor Prof. P. Caputo. These have been presented in the preprints [18, 19, 20]. The organization of the rests of the thesis is the following.

In Chapter 2 we introduce the main ingredients for the analysis of the geometry of a digraph from one of the two ensembles. In particular, in Section 2.1 we give a more detailed definition of the two ensembles, and we introduce the basic notation. In Section 2.2 we analyze the property of the branching processes which are the candidate to approximate the in/out-neighbourhoods of a vertex in each of the two models. Moreover, we consider the martingales associated to these processes, highlighting their convergence properties. In Section 2.3 we show how to couple the construction of the neighbourhoods with the respective branching process in order to translate the convergence results in Section 2.2 into asymptotic results on the geometry of the digraph. Finally, in Section 2.4 we give an application of the approximation results above, presenting the proof of Theorem 1.1. The material presented in this chapter is condenses the technical lemmata in [18, 20].

In Chapter 3 we study the mixing behaviour of the PageRank surfer, and present the proof of Theorem 1.10 and Theorem 1.11. The material of this chapter is part of the preprint [18].

Chapter 4 is devoted to the analysis of the model with regenerating underlying graph. The core of the chapter lies on the proofs of Theorem 1.12 and Theorem 1.13, which are presented in Section 4.1 and Section 4.2, respectively. Then, in Section 4.3, we present a proof of Theorem 1.9. The latter is an adaptation of the arguments used in [13, 14] in the proof of Theorem 1.7. This chapter coincide with the work presented in [19].

In Chapter 5 we analyze the extremal values of the stationary distribution in the DCM ensemble. In particular, we start in Section 5.1 by describing a *local approximation* of the stationary distribution at a given vertex x. The rest of the chapter is devoted to the proof of Theorem 1.4. This chapter is taken from the preprint [20].

Finally, in Chapter 6, we make use of the bounds obtained in Chapter 5 and of the strategy developed by Cooper and Frieze to show the result in Theorem 1.6 and Theorem 1.5, which are proved in Section 6.1 and Section 6.2, respectively. This last chapter is part of the preprint [20].

CHAPTER 2

STRUCTURAL PROPERTIES

2.1 Two models of sparse random digraphs

This section is devoted to the study of the structural properties of the DCM and the OCM ensemble. We start with a formal definition of the two ensembles.

2.1.1 Directed configuration model

Let V be a set of n vertices. For simplicity we often write $V = [n]$, with $[n] = \{1, \dots, n\}$. For each n, we are given two finite sequences $\mathbf{d}^+ = (d_x^+)_{x \in [n]}$ and $\mathbf{d}^- = (d_x^-)_{x \in [n]}$ of non negative integers such that

$$m = \sum_{x \in V} d_x^+ = \sum_{x \in V} d_x^- . \tag{2.1}$$

The *directed configuration model* DCM($\mathbf{d}^\pm$), is the distribution of the random graph G obtained as follows: 1) equip each node x with d_x^+ tails and d_x^- heads; 2) pick uniformly at random one of the $m!$ bijective maps from the set of all tails into the set of all heads, call it ω; 3) for all $x, y \in V$, add a directed edge (x, y) every time a tail from x is mapped into a head from y through ω. The resulting graph G may have self-loops and multiple edges, however it is classical that by conditioning on the event that there are no multiple edges and no self-loops one obtains a uniformly random simple digraph with in degree sequence $\mathbf{d}^-$ and out degree sequence $\mathbf{d}^+$.

Structural properties of random graphs obtained in this way have been extensively studied in [24]. Here we shall consider the sparse case corresponding to bounded degree sequences. Moreover, in order to avoid non irreducibility issues, we shall assume that all degrees are at least 2. Thus, throughout this thesis it will always be assumed that

$$\delta_\pm = \min_{x\in[n]} d_x^\pm \geq 2 \qquad \Delta_\pm = \max_{x\in[n]} d_x^\pm = O(1). \tag{2.2}$$

We often use the notation $\Delta = \max_{x\in[n]} d_x^- \vee d_x^+$. Under the first assumption it is known that DCM($\mathbf{d}^\pm$) is strongly connected with high probability. Under the second assumption, it is known that DCM($\mathbf{d}^\pm$) has a uniformly (in n) positive probability of having no self-loops nor multiple edges. In particular, any property that holds with high probability for DCM($\mathbf{d}^\pm$) will also hold with high probability for a uniformly chosen simple graph subject to the constraint that in and out degrees be given by $\mathbf{d}^-$ and $\mathbf{d}^+$ respectively. For what follows it is worth to introduce the quantity

$$\nu = \frac{1}{m}\sum_{y=1}^{n} d_y^- d_y^+. \tag{2.3}$$

Notice that ν can be interpreted in two alternative ways. On the one hand, that is the expected in-degree of the vertex connected to a uniformly random tail. On the other hand, ν is the expected out-degree of a vertex connected to a uniformly random head. Moreover, to discuss some of the forthcoming results it is convenient to introduce the following notation.

Definition 2.1 *We say that a vertex $x \in [n]$ is of type (i,j), and write $x \in \mathcal{V}_{i,j}$, if $(d_x^-, d_x^+) = (i,j)$. We call $\mathcal{C} = \mathcal{C}(\mathbf{d}^\pm)$ the set of all types that are present in the double sequence $\mathbf{d}^\pm$, that is $\mathcal{C} = \{(i,j) : |\mathcal{V}_{i,j}| > 0\}$. The assumption (2.2) implies that the number of distinct types is bounded by a fixed constant C independent of n, that is $|\mathcal{C}| \leq C$. We say that the type (i,j) has linear size, if*

$$\liminf_{n\to\infty} \frac{|\mathcal{V}_{i,j}|}{n} > 0, \tag{2.4}$$

and call $\mathcal{L} \subset \mathcal{C}$ the set of types with linear size.

2.1.2 Out Configuration Model

To define the second model, for each n let $\mathbf{d}^+ = (d_x^+)_{x\in[n]}$ be a finite sequence of non negative integers and define the *out-configuration model* OCM($\mathbf{d}^+$) as the distribution of the random graph G obtained as follows: 1) equip each node x with d_x^+ tails; 2) pick, for every x independently, a uniformly random injective map from the set of tails at x to the set of all vertices V, call it ω_x; 3) for all $x, y \in V$, add a directed edge (x, y) if a tail from x is mapped into y through ω_x. Equivalently, G is the graph whose adjacency matrix is uniformly random in the set of all $n \times n$ matrices with entries 0 or 1 such that every row x sums to d_x^+. Notice that G may have self-loops, but there are no multiple edges in this construction. This is due to the requirement that the maps ω_x be injective. The latter choice is only a matter of convenience, and everything we say below is actually seen to hold as well for the model obtained by dropping that requirement. We write $\omega = (\omega_x)_{x\in[n]}$ for the collection of maps. As before we shall make the assumptions

$$\min_{x\in[n]} d_x^+ \geq 2, \qquad \max_{x\in[n]} d_x^+ = O(1), \tag{2.5}$$

and use the notation $\Delta = \max_{x\in[n]} d_x^+$. We remark that under the above assumptions there can still be vertices with in-degree zero, and therefore in this case G is not necessarily strongly connected.

In what follows, if not differently stated, $G = G(\omega)$ denotes a given realization of either the directed configuration model DCM($\mathbf{d}^\pm$) or the out-configuration model OCM($\mathbf{d}^+$) and all the results to be discussed will hold w.h.p. within these two ensembles.

2.2 Branching approximation

We start with the definition of the relevant branching processes and the associated martingales. These will later be used in a coupling argument to provide an approximate description of the in-neighbourhood of a vertex in our random graphs. Since the constructions differ slightly for the two models DCM($\mathbf{d}^\pm$) or OCM($\mathbf{d}^+$) we will define two distinct random trees $\mathcal{T}^-(\mathbf{d}^\pm)$ and $\mathcal{T}^-(\mathbf{d}^+)$.

2.2.1 Marked Galton-Watson trees

Given $n \in \mathbb{N}$, and a double sequence $\mathbf{d}^\pm$ of degrees satisfying (2.1) and (2.2), for each $i \in [n]$, we define the rooted random marked tree $\mathcal{T}_i^-(\mathbf{d}^\pm)$ recursively with the following rules:

- the root is given the mark i;

- every vertex with mark j has d_j^- children, each of which is given independently the mark $k \in [n]$ with probability d_k^+/m.

On the other hand, given $n \in \mathbb{N}$, and a sequence $\mathbf{d}^+$ of degrees satisfying (2.5), for each $i \in [n]$, the rooted random marked tree $\mathcal{T}_i^-(\mathbf{d}^+)$ is defined by:

- the root is given the mark i;
- regardless of its own mark every vertex has, for each $j \in [n]$ independently with probability d_j^+/n, a child with mark j.

There are several differences between the two trees $\mathcal{T}_i^-(\mathbf{d}^\pm)$ and $\mathcal{T}_i^-(\mathbf{d}^+)$. In the first case the number of children of a given vertex is a deterministic function of the vertex' mark, whereas in the second case it is a random variable D that can be written as

$$D = \sum_{j\in[n]} Y_j\,, \qquad Y_j = \mathrm{Ber}(d_j^+/n), \tag{2.6}$$

where the Y_j are independent Bernoulli random variables with parameters d_j^+/n. In particular, the average degree of any given vertex in $\mathcal{T}_i^-(\mathbf{d}^+)$ is

$$\mathbb{E}[D] = \sum_{j\in[n]} \frac{d_j^+}{n} = \frac{m}{n} =: \bar{d}. \tag{2.7}$$

Since D can be zero, in contrast with the tree $\mathcal{T}_i^-(\mathbf{d}^\pm)$, the tree $\mathcal{T}^-(\mathbf{d}^+)$ is finite with positive probability. However, the two trees share several common features and we shall try to treat the two cases in a unified fashion as much as possible.

We write $\mathbf{o}$ for the root and $\mathbf{x}, \mathbf{y}$ for other vertices of the tree, with the notation $\mathbf{y} \to \mathbf{x}$ if $\mathbf{y}$ is a child of $\mathbf{x}$. Each vertex $\mathbf{x}$ of the tree has a mark, which we denote by $i(\mathbf{x})$. If $\mathcal{I}$ denotes an independent uniformly random $i \in [n]$, and the root is given the mark $i(\mathbf{o}) = \mathcal{I}$, then we write $\mathcal{T}^-(\mathbf{d}^\pm) = \mathcal{T}_{\mathcal{I}}^-(\mathbf{d}^\pm)$ and $\mathcal{T}^-(\mathbf{d}^+) = \mathcal{T}_{\mathcal{I}}^-(\mathbf{d}^+)$. Notice that $\mathcal{T}^-(\mathbf{d}^\pm)$ and $\mathcal{T}^-(\mathbf{d}^+)$ have the same average degree at the root, given by (2.7). We often write $\mathcal{T}^-$ for short if this creates no confusion. For each $t \in \mathbb{N}$ we let $\mathcal{T}^{-,t}$ denote the set of vertices in the generation t of the tree. Each vertex $\mathbf{x} \in \mathcal{T}^{-,t}$ has a unique path $(\mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_1, \mathbf{x}_0)$ connecting it to the root with $\mathbf{x}_t = \mathbf{x}$ and $\mathbf{x}_0 = \mathbf{o}$. To any such $\mathbf{x}$ we associate the weight

$$w(\mathbf{x}) = \prod_{u=1}^{t} \frac{1}{d_{i(\mathbf{x}_u)}^+}. \tag{2.8}$$

The previous construction can be reversed to obtain the following random trees. For each $x \in [n]$, define the rooted random marked tree $\mathcal{T}_x^+(\mathbf{d}^\pm)$ recursively with the following rules:

- the root is given the mark x;
- every vertex with mark y has d_y^+ children, each of which is given independently the mark $z \in [n]$ with probability d_z^-/m.

On the other hand, the rooted random marked tree $\mathcal{T}_x^+(\mathbf{d}^+)$ is defined by:

- the root is given the mark x;
- every vertex with mark y has d_y^+ children, each of which is given independently the mark $z \in [n]$ with probability $1/n$.

To have a unified notation we write $\mathcal{T}_{x,t}^+$ for the first t generations of either $\mathcal{T}_x^+(\mathbf{d}^\pm)$ or $\mathcal{T}_x^+(\mathbf{d}^+)$. Notice that this forward construction has bounded degrees for both models.

2.2.2 Martingale approximation

Given a function $\varphi : [n] \mapsto \mathbb{R}$, we define the process

$$X_t(\varphi) = \sum_{\mathbf{x} \in \mathcal{T}^{-,t}} \varphi(i(\mathbf{x}))w(\mathbf{x}), \qquad X_0(\varphi) = \varphi(i(\mathbf{o})). \tag{2.9}$$

We write $\mathcal{F}_t$ for the σ-algebra generated by the random tree $\mathcal{T}^-$ up to and including generation t.

Lemma 2.1 *Let $\mathcal{T}^-$ be either $\mathcal{T}^-(\mathbf{d}^\pm)$ or $\mathcal{T}^-(\mathbf{d}^+)$, and write $\bar\varphi = \sum_{i=1}^n \varphi(i)$. Then, for all $t \in \mathbb{N}$:*

$$\mathbb{E}[X_t(\varphi)|\mathcal{F}_{t-1}] = X_{t-1}(\bar\varphi\mu_{\text{in}}). \tag{2.10}$$

Proof: If $\mathbf{y} \to \mathbf{x}$, then $w(\mathbf{y}) = w(\mathbf{x})/d^+_{i(\mathbf{y})}$. Therefore,

$$\begin{aligned}\mathbb{E}[X_t(\varphi)|\mathcal{F}_{t-1}] &= \sum_{\mathbf{x} \in \mathcal{T}^{-,t-1}} \mathbb{E}\left[\sum_{\mathbf{y}\to\mathbf{x}} \varphi(i(\mathbf{y}))w(\mathbf{y})|\mathcal{F}_{t-1}\right] \\ &= \sum_{\mathbf{x} \in \mathcal{T}^{-,t-1}} w(\mathbf{x})\,\mathbb{E}\left[\sum_{\mathbf{y}\to\mathbf{x}} \frac{\varphi(i(\mathbf{y}))}{d^+_{i(\mathbf{y})}} \,\Big|\, \mathcal{F}_{t-1}\right].\end{aligned} \tag{2.11}$$

For the tree $\mathcal{T}^-(\mathbf{d}^\pm)$ we have

$$\mathbb{E}\left[\sum_{\mathbf{y}\to\mathbf{x}} \frac{\varphi(i(\mathbf{y}))}{d^+_{i(\mathbf{y})}} \,\Big|\, \mathcal{F}_{t-1}\right] = d^-_{i(\mathbf{x})} \sum_{j=1}^n \frac{d_j^+}{m}\frac{\varphi(j)}{d_j^+} = \bar\varphi\,\mu_{\text{in}}(i(\mathbf{x})). \tag{2.12}$$

For the tree $\mathcal{T}^-(\mathbf{d}^+)$ we have

$$\mathbb{E}\left[\sum_{\mathbf{y}\to\mathbf{x}} \frac{\varphi(i(\mathbf{y}))}{d^+_{i(\mathbf{y})}} \,\Big|\, \mathcal{F}_{t-1}\right] = \sum_{j=1}^n \frac{d_j^+}{n}\frac{\varphi(j)}{d_j^+} = \bar{\varphi}\,\mu_{\text{in}}(i(\mathbf{x})). \tag{2.13}$$

This proves (2.10). □

In particular, when $\varphi = \mu_{\text{in}}$, then

$$\mathbb{E}[X_t(\mu_{\text{in}})|\mathcal{F}_{t-1}] = X_{t-1}(\mu_{\text{in}})\,, \qquad t\in\mathbb{N}.$$

Therefore, $X_t(\mu_{\text{in}})$ is a martingale with respect to the filtration $\mathcal{F}_t$. It is convenient to normalize it and consider instead the martingale defined as

$$M_t = nX_t(\mu_{\text{in}}) = \sum_{\mathbf{x}\in\mathcal{T}^{-,t}} n\mu_{\text{in}}(i(\mathbf{x}))w(\mathbf{x}), \qquad M_0 = n\mu_{\text{in}}(i(\mathbf{o})). \tag{2.14}$$

Notice that $\mathbb{E}[M_t] = \mathbb{E}[M_0] = n\mathbb{E}[\mu_{\text{in}}(\mathcal{I})] = 1$. In the case of the DCM, the following convergence result was already discussed in [13, Proposition 15].

Proposition 2.1 *For every fixed n, as $t\to\infty$ the martingale M_t converges to a limit M_∞, both almost surely and in L^2, and for all $t\in\mathbb{N}$:*

$$\mathbb{E}[(M_t - M_\infty)^2] = C\rho^t \tag{2.15}$$

where the constants ρ, C are given by

$$\rho = \sum_{j=1}^n \mu_{\text{in}}(j)\,\frac{1}{d_j^+}\,, \qquad C = \begin{cases} \frac{n}{m(1-\rho)}\sum_{j=1}^n \frac{(d_j^- - d_j^+)^2}{m d_j^+} & \text{DCM}(\mathbf{d}^\pm) \\ \frac{\rho-1/n}{1-\rho} & \text{OCM}(\mathbf{d}^+) \end{cases} \tag{2.16}$$

Proof: Consider the increments

$$\Delta_t = M_{t+1} - M_t = \sum_{\mathbf{x}\in\mathcal{T}^{-,t}} n\mu_{\text{in}}(i(\mathbf{x}))w(\mathbf{x})\psi(\mathbf{x})\,, \tag{2.17}$$

where

$$\psi(\mathbf{x}) = \sum_{\mathbf{y}\to\mathbf{x}} \frac{\mu_{\text{in}}(i(\mathbf{y}))}{\mu_{\text{in}}(i(\mathbf{x}))d^+_{i(\mathbf{y})}} - 1 = \sum_{\mathbf{y}\to\mathbf{x}}\left[\frac{\mu_{\text{in}}(i(\mathbf{y}))}{\mu_{\text{in}}(i(\mathbf{x}))d^+_{i(\mathbf{y})}} - \frac{1}{d^-_{i(\mathbf{x})}}\right]. \tag{2.18}$$

As in Lemma 2.1 one has $\mathbb{E}[\psi(\mathbf{x})\,|\,\mathcal{F}_t] = 0$. Let us compute $\mathbb{E}[\psi(\mathbf{x})^2|\,\mathcal{F}_t]$. For the tree $\mathcal{T}^-(\mathbf{d}^\pm)$ we have

$$\mathbb{E}[\psi(\mathbf{x})^2|\,\mathcal{F}_t] = d^-_{i(\mathbf{x})}\sum_{j=1}^n \frac{d_j^+}{m}\left(\frac{d_j^-}{d^-_{i(\mathbf{x})}d_j^+} - \frac{1}{d^-_{i(\mathbf{x})}}\right)^2 = \frac{C_1}{d^-_{i(\mathbf{x})}}\,, \tag{2.19}$$

where we use the notation

$$C_1 = \sum_{j=1}^n \frac{(d_j^- - d_j^+)^2}{m d_j^+}.$$

For the tree $\mathcal{T}^-(\mathbf{d}^+)$ we have

$$\begin{aligned}\mathbb{E}[\psi(\mathbf{x})^2\,|\,\mathcal{F}_t] &= \mathbb{E}\left[\left(\sum_{\mathbf{y}\to\mathbf{x}}\frac{1}{d^+_{i(\mathbf{y})}}\right)^2 - 2\sum_{\mathbf{y}\to\mathbf{x}}\frac{1}{d^+_{i(\mathbf{y})}} + 1\right]\\ &= \sum_{j\neq j'}\frac{d_j^+d_{j'}^+}{n^2}\frac{1}{d_j^+d_{j'}^+} + \sum_j \frac{d_j^+}{n}\frac{1}{(d_j^+)^2} - 2\sum_j\frac{d_j^+}{n}\frac{1}{d_j^+} + 1 = \rho - \frac{1}{n}\,,\end{aligned} \tag{2.20}$$

where a is as in (2.16). Since $\mathbb{E}[\psi(\mathbf{x})\psi(\mathbf{x}')\,|\,\mathcal{F}_t] = 0$ for all $\mathbf{x},\mathbf{x}' \in \mathcal{T}^{-,t}$ with $\mathbf{x}\neq\mathbf{x}'$,

$$\mathbb{E}[\Delta_t^2\,|\,\mathcal{F}_t] = \sum_{\mathbf{x}\in\mathcal{T}^{-,t}} n^2\mu_{\text{in}}(i(\mathbf{x}))^2 w(\mathbf{x})^2\mathbb{E}[\psi(\mathbf{x})^2\,|\,\mathcal{F}_t]. \tag{2.21}$$

Therefore, combining (2.19) and (2.20) we have

$$\mathbb{E}[\Delta_t^2\,|\,\mathcal{F}_t] = C(1-\rho)\sum_{\mathbf{x}\in\mathcal{T}^{-,t}} n\mu_{\text{in}}(i(\mathbf{x}))w(\mathbf{x})^2\,, \tag{2.22}$$

where ρ, C are given by (2.16). Furthermore, observe that in both models one has

$$\begin{aligned}\mathbb{E}[\Delta_t^2\,|\,\mathcal{F}_{t-1}] &= \mathbb{E}\left[\mathbb{E}[\Delta_t^2\,|\,\mathcal{F}_t]\,|\,\mathcal{F}_{t-1}\right]\\ &= C(1-\rho)\sum_{\mathbf{x}\in\mathcal{T}^{-,t-1}} n\mu_{\text{in}}(i(\mathbf{x}))w(\mathbf{x})^2\mathbb{E}\left[\sum_{\mathbf{y}\to\mathbf{x}}\frac{\mu_{\text{in}}(i(\mathbf{y}))}{\mu_{\text{in}}(i(\mathbf{x}))(d^+_{i(\mathbf{y})})^2}\,|\,\mathcal{F}_{t-1}\right]\\ &= C(1-\rho)\rho\sum_{\mathbf{x}\in\mathcal{T}^{-,t-1}} n\mu_{\text{in}}(i(\mathbf{x}))w(\mathbf{x})^2 = \rho\,\mathbb{E}[\Delta_{t-1}^2\,|\,\mathcal{F}_{t-1}].\end{aligned} \tag{2.23}$$

Thus, iterating we obtain

$$\mathbb{E}[\Delta_t^2] = \mathbb{E}[\Delta_0^2]\rho^t = C(1-\rho)\mathbb{E}[n\mu_{\text{in}}(\mathcal{I})]\rho^t = C(1-r)\rho^t. \tag{2.24}$$

Since $\rho \leq 1/2$, we see that M_t is a martingale bounded in L^2, and therefore $M_t \to M_\infty$ almost surely and in L^2, for some $M_\infty \in L^2$. Using the orthogonality $\mathbb{E}[\Delta_t \Delta_{t'}] = 0$ for all $t \neq t'$, (2.15) follows by summing (2.24) from t to $+\infty$. $\square$

Remark 2.1 *For each fixed $n \in \mathbb{N}$, one can characterise the random variable M_∞ as the solution to a distributional fixed point equation. For the directed configuration model DCM($\mathbf{d}^\pm$) this is discussed in [13, Lemma 16]. With a similar reasoning, for the out-configuration model OCM($\mathbf{d}^+$) one obtains that*

$$M_\infty \stackrel{d}{=} \sum_{j=1}^{n} \frac{Y_j}{d_j^+} M_{\infty,j}, \tag{2.25}$$

where $\stackrel{d}{=}$ stands for equality of distributions, $M_{\infty,j}$ are i.i.d. copies of M_∞ and Y_j are independent Bernoulli random variables with parameter d_j^+/n.

The next result will be crucial for the analysis of convergence to stationarity when the starting measure is *widespread* in the sense of Definition 1.1, see Section 3.2. Notice that the constant $\gamma(\lambda)$ appearing in the estimate below is bounded uniformly in n if and only if λ is widespread.

Proposition 2.2 *For any probability vector λ, and any $t \in \mathbb{N}$:*

$$\mathbb{E}[(M_t - nX_t(\lambda))^2] \leq \gamma(\lambda)\rho^t\,, \tag{2.26}$$

where $\rho \in (0,1)$ is as in Proposition 2.1 and $\gamma(\lambda)$ is defined as

$$\gamma(\lambda) = \frac{n}{2}\sum_{j=1}^{n}(\lambda(j) - \mu_{\text{in}}(j))^2 \tag{2.27}$$

Proof: Setting $\varphi(j) = n(\mu_{\text{in}}(j) - \lambda(j))$, we write $M_t - nX_t(\lambda) = X_t(\varphi)$. Since $\bar{\varphi} = 0$, Lemma 2.1 shows that $\mathbb{E}[M_t - nX_t(\lambda)|\mathcal{F}_{t-1}] = 0$. We now compute

$$\Gamma_t := \mathbb{E}[(M_{t+1} - nX_{t+1}(\lambda))^2|\mathcal{F}_t].$$

Using $\bar{\varphi} = 0$ one has

$$\begin{aligned}\Gamma_t &= \mathbb{E}[X_{t+1}(\varphi)^2|\mathcal{F}_t] \\ &= \sum_{\mathbf{x}\in\mathcal{T}^{-,t}} w(\mathbf{x})^2\, \mathbb{E}\left[\left(\sum_{\mathbf{y}\to\mathbf{x}} \frac{\varphi(i(\mathbf{y}))}{d^+_{i(\mathbf{y})}}\right)^2 \Big| \mathcal{F}_t\right].\end{aligned} \tag{2.28}$$

For the tree $\mathcal{T}^-(\mathbf{d}^\pm)$ we have

$$\mathbb{E}\left[\left(\sum_{\mathbf{y}\to\mathbf{x}}\frac{\varphi(i(\mathbf{y}))}{d^+_{i(\mathbf{y})}}\right)^2 \,\Big|\, \mathcal{F}_t\right] = d^-_{i(\mathbf{x})}\sum_{j=1}^{n}\frac{d^+_j}{m}\frac{\varphi(j)^2}{(d^+_j)^2} = \mu_{\mathrm{in}}(i(\mathbf{x}))\sum_{j=1}^{n}\frac{\varphi(j)^2}{d^+_j}. \tag{2.29}$$

On the other hand for the tree $\mathcal{T}^-(\mathbf{d}^+)$ we have

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{\mathbf{y}\to\mathbf{x}}\frac{\varphi(i(\mathbf{y}))}{d^+_{i(\mathbf{y})}}\right)^2 \,\Big|\, \mathcal{F}_t\right] &= \sum_{j\neq j'}\frac{\varphi(j)\varphi(j')}{n^2} + \sum_{j}\frac{\varphi(j)^2}{nd^+_j} \\ &= \frac{1}{n}\sum_{j=1}^{n}\frac{\varphi(j)^2}{d^+_j}\left(1-\frac{d^+_j}{n}\right) \end{aligned} \tag{2.30}$$

Summarising, we have shown that

$$\Gamma_t = C(\lambda)\sum_{\mathbf{x}\in\mathcal{T}^{-,t}} n\mu_{\mathrm{in}}(i(\mathbf{x}))w(\mathbf{x})^2\,, \qquad C(\lambda) = \frac{1}{n}\begin{cases}\sum_{j=1}^{n}\frac{\varphi(j)^2}{d^+_j} & \mathrm{DCM}(\mathbf{d}^\pm) \\ \sum_{j=1}^{n}\frac{\varphi(j)^2}{d^+_j}\left(1-\frac{d^+_j}{n}\right) & \mathrm{OCM}(\mathbf{d}^+)\end{cases} \tag{2.31}$$

Thus, the same argument used in (2.23) implies that in both models

$$\mathbb{E}[\Gamma_t \,|\, \mathcal{F}_{t-1}] = \rho\Gamma_{t-1}. \tag{2.32}$$

Therefore,

$$\mathbb{E}[\Gamma_t] = \mathbb{E}[\Gamma_0]\rho^t = C(\lambda)\mathbb{E}[n\mu_{\mathrm{in}}(\mathcal{I})]\rho^t = C(\lambda)\rho^t. \tag{2.33}$$

The desired bound follows from the fact that in both models $C(\lambda) \le \gamma(\lambda)$. □

2.3 Neighbourhoods

The t-in-neighbourhood of a vertex v, denoted $\mathcal{B}^-_x(t)$, is defined as the subgraph of G induced by the set of directed paths of length t in G which terminate at vertex x. In this section we present algorithmic procedures to generate the in/out-neighbourhood if a vertex. Moreover, we observe that for any fixed $x \in [n]$, if t is a small multiple of $\log n$ then with high probability $\mathcal{B}^-_x(t)$ can be coupled to the first t generations of the random trees presented in Section 2.2.

2.3.1 Neighbourhoods in the DCM($\mathbf{d}^{\pm}$)

Each vertex x has d_x^- labeled heads and d_x^+ labeled tails, and we call E_x^- and E_x^+ the sets of heads and tails at x respectively. The uniform bijection ω between heads $E^- = \cup_{x\in[n]} E_x^-$ and tails $E^+ = \cup_{x\in[n]} E_x^+$, viewed as a matching, can be sampled by iterating the following steps until there are no unmatched heads left:

1) pick an unmatched head $f \in E^-$ according to some priority rule;

2) pick an unmatched tail $e \in E^+$ uniformly at random;

3) match f with e, i.e. set $\omega(f) = e$, and call ef the resulting edge.

This gives the desired uniform distribution over matchings $\omega : E^- \mapsto E^+$ regardless of the priority rule chosen at step 1. The digraph G is obtained by adding a directed edge (x, y) whenever $f \in E_y^-$ and $e \in E_x^+$ in step 3 above.

We will use the notation

$$\delta = \min\{\delta_-, \delta_+\}, \qquad \Delta = \max\{\Delta_-, \Delta_+\}. \tag{2.34}$$

For any $h \in \mathbb{N}$, the h-in-neighbourhood of a vertex y, denoted $\mathcal{B}_h^-(y)$, is the digraph defined as the union of all directed paths of length $\ell \leq h$ in G which terminate at vertex y. In the sequel a path is always understood as a sequence of directed edges $(e_1 f_1, \ldots, e_k f_k)$ such that $v_{f_i} = v_{e_{i+1}}$ for all $i = 1, \ldots, k-1$, and we use the notation v_e (resp. v_f) for the vertex x such that $e \in E_x^+$ (resp. $f \in E_x^-$).

To generate the random variable $\mathcal{B}_h^-(y)$, we use the following breadth-first procedure. Start at vertex y and run the sequence of steps described above, by giving priority to those unmatched heads which have minimal distance to vertex y, until this minimal distance exceeds h, at which point the process stops. Similarly, for any $h \in \mathbb{N}$, the h-out-neighbourhood of a vertex x, denoted $\mathcal{B}_h^+(x)$ is defined as the subgraph induced by the set of directed paths of length $\ell \leq h$ which start at vertex x. To generate the random variable $\mathcal{B}_h^+(x)$, we use the same breadth-first procedure described above except that we invert the role of heads and tails. With slight abuse of notation we sometimes write $\mathcal{B}_h^{\pm}(x)$ for the vertex set of $\mathcal{B}_h^{\pm}(x)$. We also warn the reader that to simplify the notation we often avoid taking explicitly the integer part of the various parameters entering our proofs. In particular, whenever we write $\mathcal{B}_h^{\pm}(x)$ it is always understood that $h \in \mathbb{N}$.

Let us now describe a coupling of the in-neighbourhood $\mathcal{B}_h^-(x)$ and the marked tree $\mathcal{T}_{x,h}^-(\mathbf{d}^{\pm})$, where $\mathcal{T}_{x,h}^-(\mathbf{d}^{\pm})$ stands for the marked tree $\mathcal{T}_x^-(\mathbf{d}^{\pm})$ up to generation h; see Section 2.2.1 for the definition of $\mathcal{T}_h^-(\mathbf{d}^{\pm})$. Clearly, step 2 above can be modified by picking e

uniformly at random among all (matched or unmatched) tails and rejecting the proposal if the tail was already matched. The tree can then be generated by iteration of the same sequence of steps with the difference that at step 2 we never reject the proposal and at step 3 we add a new leaf to the current tree, with mark x if $e_+ \in E_x^+$, together with a new set of d_x^- unmatched heads attached to it. Call τ the first time that a uniform random choice among all tails gives $e_+ \in E_x^+$ with x already in the tree. By construction, the in-neighbourhood and the tree coincide up to time τ. At the k-th iteration, the probability of picking a tail with a mark already used is at most $k\Delta/m$, where Δ is the maximum degree. Therefore, by a union bound,

$$\mathbb{P}(\tau \le k) \le \frac{k^2\Delta}{m}. \tag{2.35}$$

Taking $k = \Delta^{t+1}$ steps, we have necessarily uncovered the whole in-neighbourhood $\mathcal{B}_h^-(x)$. Thus, given the symmetry between in and out-neighbourhood, we have proved the following statement.

Lemma 2.2 *The h-in-neighbourhood $\mathcal{B}_h^-(x)$ and the marked tree $\mathcal{T}_{x,h}^-(\mathbf{d}^\pm)$ can be coupled in such a way that*

$$\mathbb{P}\left(\mathcal{B}_h^-(x) \neq \mathcal{T}_{x,h}^-(\mathbf{d}^\pm)\right) \le \frac{\Delta^{2t+3}}{m}. \tag{2.36}$$

The same result holds for the out-neighbourhood $\mathcal{B}_h^+(x)$ and the tree $\mathcal{T}_{x,h}^+(\mathbf{d}^\pm)$ obtained by reversing the role of head and tails in the procedure above.

More generally, in the generation process of the in-neighbourhood, say that a *collision* occurs whenever a tail gets chosen, whose end-point x was already exposed, in the sense that some tail in E_x^+ or head in E_x^- had already been matched. Since less than $2k$ vertices are exposed when the k^{th} tail gets matched, less than $2\Delta k$ of the $m-k+1$ possible choices can result in a collision. Thus, the conditional chance that the k^{th} step causes a collision, given the past, is less than $p_k = \frac{2\Delta k}{m-k+1}$. It follows that the number Z_k of collisions caused by the first k arcs is stochastically dominated by the binomial random variable $\mathrm{Bin}(k, p_k)$. In particular,

$$\mathbb{P}\left(Z_k \ge \ell\right) \le \frac{k^\ell p_k^\ell}{\ell!}, \qquad \ell \in \mathbb{N}. \tag{2.37}$$

The same applies to out-neighbourhoods simply by inverting the role of heads and tails.

For any digraph G, define the tree excess of G as

$$\textsc{tx}(G) = 1 + |E| - |V|,$$

where E is the set of directed edges and V is the set of vertices of G. In particular, $\textsc{tx}(\mathcal{B}_h^{\pm}(x)) = 0$ iff $\mathcal{B}_h^{\pm}(x)$ is a directed tree, and $\textsc{tx}(\mathcal{B}_h^{\pm}(x)) \leq 1$ iff there is at most one collision during the generation of the neighbourhood $\mathcal{B}_h^{\pm}(x)$. Define the events

$$\mathcal{G}_x(h) = \left\{\textsc{tx}(\mathcal{B}_h^{-}(x)) \leq 1 \text{ and } \textsc{tx}(\mathcal{B}_h^{+}(x)) \leq 1\right\}, \qquad \mathcal{G}(h) = \cap_{x\in[n]}\mathcal{G}_x(h). \tag{2.38}$$

Set also

$$\hbar = \frac{1}{5}\log_{\Delta}(n)\,. \tag{2.39}$$

Proposition 2.3 *There exists $\chi > 0$ such that $\mathbb{P}(\mathcal{G}_x(\hbar)) = 1 - O(n^{-1-\chi})$ for any $x \in [n]$. In particular,*

$$\mathbb{P}(\mathcal{G}(\hbar)) = 1 - O(n^{-\chi}). \tag{2.40}$$

Proof: During the generation of $\mathcal{B}_h^{-}(x)$ one creates at most Δ^h edges. It follows from (2.37) with $\ell = 2$ that the probability of the complement of $\mathcal{G}_x(\hbar)$ is $O(n^{-1-\chi})$ for all $x \in [n]$ for some absolute constant $\chi > 0$:

$$\mathbb{P}(\mathcal{G}_x(\hbar)) = 1 - O(n^{-1-\chi}). \tag{2.41}$$

The conclusion follows from the union bound. $\square$

We will need to control the size of the boundary of our neighbourhoods. To this end, we introduce the notation $\partial\mathcal{B}_t^{-}(y)$ for the set of vertices $x \in [n]$ such that $d(x, y) = t$. Similarly, $\partial\mathcal{B}_t^{+}(x)$ is the set of vertices $y \in [n]$ such that $d(x, y) = t$. Clearly, $|\partial\mathcal{B}_t^{\pm}(y)| \leq \Delta^h$ for any $y \in [n]$ and $h \in \mathbb{N}$.

Lemma 2.3 *There exists $\chi > 0$ such that for all $y \in [n]$,*

$$\mathbb{P}\left(|\partial\mathcal{B}_h^{\pm}(y)| \geq \tfrac{1}{2}\delta_{\pm}^h,\ \forall h \in [1, \hbar]\right) = 1 - O(n^{-1-\chi}). \tag{2.42}$$

Proof: By symmetry we may restrict to the case of in-neighbourhoods. By (2.41) it is sufficient to show that $|\partial\mathcal{B}_h^{\pm}(y)| \geq \frac{1}{2}\delta_{\pm}^h$, for all $h \in [1, \hbar]$, if $\mathcal{G}_y(\hbar)$ holds. If the tree excess of the h-in-neighbourhood $\mathcal{B}_h^{-}(y)$ is at most 1 then there is at most one collision in the generation of $\mathcal{B}_h^{-}(y)$. This collision can be of two types:

1. there exists some $1 \leq t \leq h$ and a $v \in \partial\mathcal{B}_t^{-}(y)$ s.t. v has two out-neighbours $w, w' \in \partial\mathcal{B}_{t-1}^{-}(y)$;

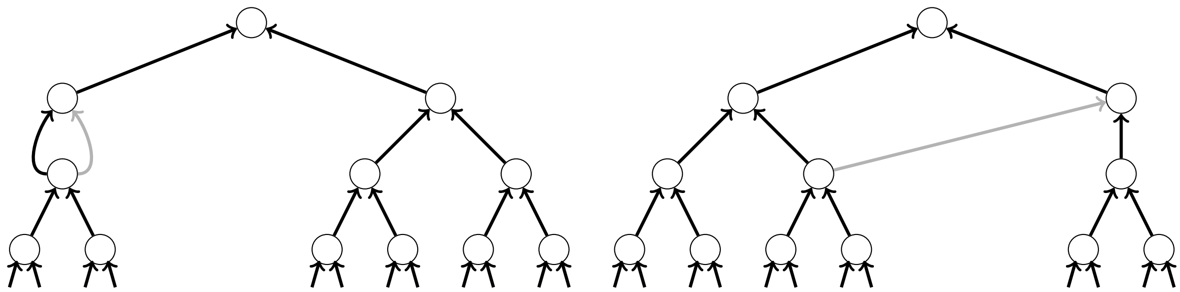

Figure 2.1: The light-coloured arrow represents a collision of type (1a) (left) and a collision of type (1b) (right).

2. there exists some $0 \le t \le h$ and a $v \in \partial\mathcal{B}_t^-(y)$ s.t. v has an in-neighbour w in $\mathcal{B}_t^-(y)$.

The first case can be further divided into two cases: a) $w = w'$, and b) $w \neq w'$; see Fig. 2.1.

In case 1a) we note that the $(h-t)$-in-neighbourhood of v must be a directed tree with at least δ_-^{h-t} elements on its boundary and with no intersection with the $(h-t)$-in-neighbourhoods of other $v' \in \partial\mathcal{B}_t^-(y)$. Moreover, $\mathcal{B}_{t-1}^-(y)$ must be a directed tree with $|\partial\mathcal{B}_{t-1}^-(y)| \ge \delta_-^{t-1}$, and all elements of $\partial\mathcal{B}_{t-1}^-(y)$ except one have disjoint $(h-t+1)$-in-neighbourhoods with δ_-^{h-t+1} elements on their boundary. Therefore

$$|\partial\mathcal{B}_h^-(y)| \ge (\delta_-^{t-1} - 1)\delta_-^{h-t+1} + (\delta_- - 1)\delta_-^{h-t} \ge \frac{1}{2}\delta_-^h.$$

In case 1b) one has that $t \ge 2$, $\mathcal{B}_{t-1}^-(y)$ is a directed tree with $|\partial\mathcal{B}_{t-1}^-(y)| \ge \delta_-^{t-1}$, and for all $z \in \partial\mathcal{B}_t^-(y)$, the $(h-t)$-in-neighbourhoods of z are disjoint directed trees with at least δ_-^{h-t} elements on their boundary. Since $|\partial\mathcal{B}_t^-(y)| \ge \delta_-^t - 1$ it follows that

$$|\partial\mathcal{B}_h^-(y)| \ge (\delta_-^t - 1)\delta_-^{h-t} \ge \frac{1}{2}\delta_-^h.$$

Collisions of type 2 can be further divided into two types: a) $w \in \partial\mathcal{B}_s^-(y)$ with $s < t$ and there is no path from v to w of length $t-s$, or $w \in \partial\mathcal{B}_t^-(y)$ and $w \neq v$, and b) $w \in \partial\mathcal{B}_s^-(y)$ with $s < t$ and there is a path from v to w of length $t-s$, or $w = v$. Note that in contrast with collisions of type 2a), a collision of type 2b) creates a directed cycle within $\mathcal{B}_t^-(y)$; see Fig. 2.2 and Fig. 2.3.

We remark that in either case 2a) or case 2b), $\partial\mathcal{B}_t^-(y)$ has at least δ_-^t elements, and the vertex $v \in \partial\mathcal{B}_t^-(y)$ has at least $\delta_- - 1$ in-neighbours whose $(h-t-1)$-in-neighbourhoods are disjoint directed trees. All other $v' \in \partial\mathcal{B}_t^-(y)$ have $(h-t)$-in-neighbourhoods that are disjoint directed trees. Therefore, in case 2):

$$|\partial\mathcal{B}_h^-(y)| \ge (\delta_-^t - 1)\delta_-^{h-t} + (\delta_- - 1)\delta_-^{h-t-1} \ge \frac{1}{2}\delta_-^h.$$

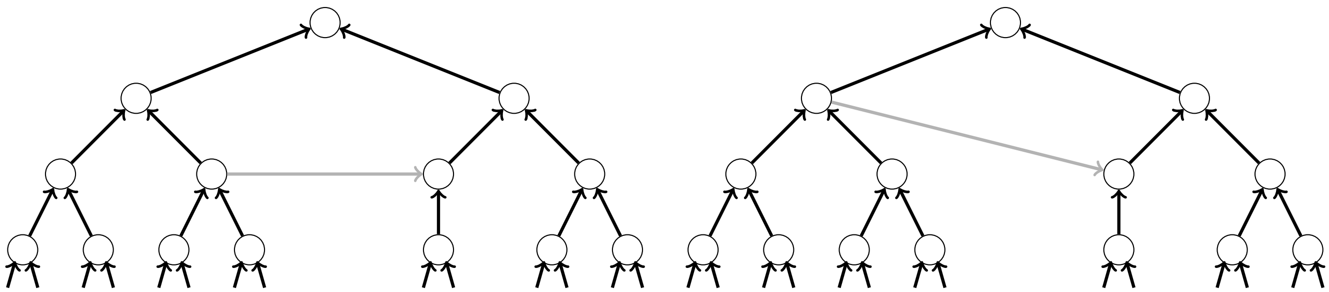

Figure 2.2: Two examples of collision of type $(2a)$.

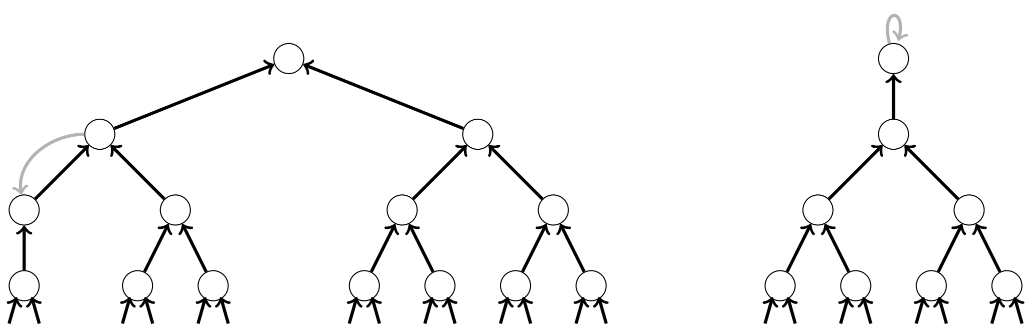

Figure 2.3: Two examples of collision of type (2b).

□

2.3.2 Neighbourhoods in the OCM($\mathbf{d}^+$)

Recall that each vertex x has d_x^+ tails, and call E_x^+ the sets of tails at x. Consider the following *exploration process* of the in-neighbourhood at a fixed vertex x. The process is defined as a triple $(\mathcal{C}_\ell, \mathcal{A}_\ell, \phi_\ell)$ where $\mathcal{C}_\ell, \mathcal{A}_\ell \subset [n]$ are respectively the *completed* set and the *active* set at time ℓ, and $\phi_\ell : [n] \mapsto \mathbb{Z}_+$ is a map such that $\phi_\ell(y) \in \{0, \dots, d_y^+\}$ for each $y \in [n]$, $\ell \in \mathbb{Z}_+$. At time zero we set $\mathcal{C}_0 = \emptyset, \mathcal{A}_0 = \{x\}$, and $\phi_\ell(y) = 0$ for all $y \in [n]$. The ℓ-th iteration of the exploration determines the triple $(\mathcal{C}_\ell, \mathcal{A}_\ell, \phi_\ell)$ by executing the following steps:

1) pick a vertex $v \in \mathcal{A}_{\ell-1}$ according to some priority rule;

2) for each $y = 1, \dots, n$ independently, sample $X_{\ell,y}$ defined as the Bernoulli random variable with parameter
$$p_\ell(y) = \frac{d_y^+ - \phi_{\ell-1}(y)}{n - \ell + 1},$$
and call V_ℓ the set of $y \in [n]$ such that $X_{\ell,y} = 1$, and $W_\ell = (\mathcal{C}_{\ell-1} \cup \mathcal{A}_{\ell-1})^c \cap V_\ell$;

3) define the new triple $(\mathcal{C}_\ell, \mathcal{A}_\ell, \phi_\ell)$ as
$$\mathcal{C}_\ell = \mathcal{C}_{\ell-1} \cup \{v\}\,, \quad \mathcal{A}_\ell = \mathcal{A}_{\ell-1} \setminus \{v\} \cup W_\ell\,, \quad \phi_\ell(y) = \phi_{\ell-1}(y) + \mathbf{1}(y \in V_\ell),\ y = 1, \dots, n.$$

Note that this process stops when $\mathcal{A}_\ell$ becomes empty. Let us call $\tau_\emptyset$ this random time:
$$\tau_\emptyset = \min\{\ell \geq 1 :\ \mathcal{A}_\ell = \emptyset\}. \tag{2.43}$$

For instance, $\tau_\emptyset = 1$ with probability $\prod_{y=1}^n (1 - d_y^+/n)$. We may construct a digraph $G_x(\ell)$ along with the above process by adding the directed edges (y, v) for all $y \in V_\ell$ at step 2. Notice that when the process stops $G_v(\tau_\emptyset)$ is a sample of the subgraph of G induced by all directed paths in G that terminate at x. In particular, if the priority in step 1 is given to v which have minimal distance to x, and if we stop the process as soon as all active vertices have distance to x larger than h in the current graph $G_x(\ell)$, we obtain the in-neighbourhood of x at distance h, namely the digraph $\mathcal{B}_h^-(x)$ for the OCM($\mathbf{d}^+$). More formally, if τ_h denotes the minimal ℓ such that all $v \in \mathcal{A}_\ell$ have distance to x at least $h+1$ in $G_x(\ell)$ then, $\mathcal{B}_h^-(x)$ is given by the subgraph of $G_x(\tau_h \wedge \tau_\emptyset)$ induced by the completed set $\mathcal{C}_{\tau_h \wedge \tau_\emptyset}$, where $a \wedge b$ denotes the minimum of a, b.

Let us now describe a coupling of $\mathcal{B}_h^-(x)$ and the marked tree $\mathcal{T}_{x,h}^-(\mathbf{d}^+)$, where we write $\mathcal{T}_{x,h}^-(\mathbf{d}^+)$ for the marked tree $\mathcal{T}_x^-(\mathbf{d}^+)$ up to generation h; see Section 2.2.1. First, observe that the tree $\mathcal{T}_x^-(\mathbf{d}^+)$ is obtained by iterating the steps above with the difference that at step 2 the probability $p_{\ell,y}$ must be taken always equal to d_ℓ^+/n, and that each $y \in V_\ell$ yields a new child with mark y in the current tree. Let $\mathcal{T}_x^-(\ell)$ denote the tree obtained after ℓ iterations, and let $\Delta = \max_v d_v^+$.

Lemma 2.4 *The random variables $G_x(\ell), \mathcal{T}_x^-(\ell)$ can be coupled in such a way that for every $\ell \in \mathbb{N}$:*

$$\mathbb{P}(G_x(\ell) \neq \mathcal{T}_x^-(\ell)) \leq \frac{\Delta^2 \ell^2}{n - \ell}. \tag{2.44}$$

Proof: Let $E_\ell = \{G_x(\ell) \neq \mathcal{T}_x^-(\ell)\}$. Since at time 0 one has $G_x(0) = \mathcal{T}_x^-(0) = \{x\}$, the event E_ℓ satisfies $E_\ell = \cup_{k=1}^{\ell} E_{k-1}^c \cap E_k$, so that

$$\mathbb{P}\left(G_x(\ell) \neq \mathcal{T}_x^-(\ell)\right) \leq \sum_{k=1}^{\ell} \mathbb{P}(E_{k-1}^c \cap E_k) \tag{2.45}$$

Consider now the k-th iteration, and assume that $G_x(k-1) = \mathcal{T}_x^-(k-1)$. Thus, we may pick the same v in step 1 for both samples. At step 2, let $X_{k,y}$ denote the Bernoulli random variables with parameter $p_k(y)$ used for the sampling of $G_x(k)$ and let $\tilde{X}_{k,y}$ be the Bernoulli random variables with parameter d_y^+/n used for the sampling of $\mathcal{T}_x^-(k)$. For each y independently we may couple $(X_{k,y}, \tilde{X}_{k,y})$ with probability $1 - |p_{k,y} - d_y^+/n|$. Notice that if $G_x(k) \neq \mathcal{T}_x^-(k)$, then either at least one of the pairs $(X_{k,y}, \tilde{X}_{k,y})$ fails to couple, or at least one of the $y \in \mathcal{C}_{\ell-1} \cup \mathcal{A}_{\ell-1}$ has $\tilde{X}_{k,y} = 1$. Thus, on the event E_{k-1}^c, the probability of E_k given the history up to the $(k-1)$-th iteration is bounded above by

$$\begin{aligned}
&\sum_{y=1}^{n} |p_{k,y} - d_y^+/n| + \sum_{y=1}^{n} \frac{d_y^+}{n} \mathbf{1}(y \in \mathcal{C}_{k-1} \cup \mathcal{A}_{k-1}) \\
&\leq \sum_{y=1}^{n} \frac{(k-1)d_y^+}{n(n-k+1)} + \frac{\Delta}{n} |\mathcal{C}_{k-1} \cup \mathcal{A}_{k-1}| \\
&\leq \frac{\Delta}{n-k} \left(k + Z_{k-1}\right),
\end{aligned}$$

where we use that for every $y \in [n]$ one has $(k-1) \geq \phi_{k-1}(y)$, and we write Z_ℓ for the number of edges in the tree $\mathcal{T}_x^-(\ell)$. Thus, letting $\mathcal{F}_\ell$ denote the σ-algebra generated by the two processes up to time ℓ, we have obtained

$$\begin{aligned}\mathbb{P}(E_{k-1}^c \cap E_k) &= \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}(E_{k-1}^c \cap E_k) \,|\, \mathcal{F}_{k-1}\right]\right] \\ &\leq \frac{\Delta}{n-k}\left(k + \mathbb{E}[Z_{k-1}]\right). \end{aligned} \tag{2.46}$$

From (2.7) we deduce $\mathbb{E}[Z_{k-1}] = (k-1)\bar{d} \leq (k-1)\Delta$. Therefore, the estimate (2.44) follows from (2.45) and (2.46). □

The next lemma establishes the coupling estimate for the t-in-neighbourhood $B_{v,t}^-$ and the tree $\mathcal{T}_{v,t}^-(\mathbf{d}^+)$. The estimate could be refined but (2.47) below will be more than sufficient for our purposes.

Lemma 2.5 *The random variables $\mathcal{B}_h^-(x)$ and the tree $\mathcal{T}_{x,h}^-(\mathbf{d}^+)$ can be coupled in such a way that for every $h \leq \frac{\log n}{4\log\Delta}$, for all n large enough:*

$$\mathbb{P}\left(\mathcal{B}_h^-(x) \neq \mathcal{T}_{x,h}^-(\mathbf{d}^+)\right) \leq \frac{\Delta^{3t}(\log n)^4}{n}. \tag{2.47}$$

Proof: Let $|\mathcal{T}_{x,h}^-|$ denote the number of edges in the tree $\mathcal{T}_{x,h}^- = \mathcal{T}_{x,h}^-(\mathbf{d}^+)$. Since at each iteration the number of edges added is stochastically dominated by a binomial random variable with parameters n and Δ/n, one has a large deviation bound for $|\mathcal{T}_{x,h}^-|$ of the form: there exist absolute constants $a, A > 0$ such that

$$\mathbb{P}\left(|\mathcal{T}_{x,h}^-| > s\Delta^h\right) \leq A\,e^{-a\,s}, \qquad s \geq 1. \tag{2.48}$$

The estimate (2.48) can be proved e.g. by repeating the argument in [15, Lemma 23]. Next, observe that if $|\mathcal{T}_{x,h}^-| \leq s\Delta^h$ and $\mathcal{B}_h^-(x) \neq \mathcal{T}_{x,h}^-$, then there must exist $\ell = 1, \ldots, s\Delta^h$ such that $G_x(\ell) \neq \mathcal{T}_x^-(\ell)$. The latter probability can be bounded via Lemma 2.4. Summarizing,

$$\begin{aligned}\mathbb{P}\left(\mathcal{B}_h^-(x) \neq \mathcal{T}_{x,h}^-(\mathbf{d}^+)\right) &\leq \mathbb{P}\left(|\mathcal{T}_{x,h}^-| > s\Delta^h\right) + \mathbb{P}\left(\mathcal{B}_h^-(x) \neq \mathcal{T}_{x,h}^-(\mathbf{d}^+); |\mathcal{T}_{x,h}^-| \leq s\Delta^h\right) \\ &\leq A\,e^{-a\,s} + \sum_{\ell=1}^{s\Delta^h} \mathbb{P}(G_v(\ell) \neq \mathcal{T}_v^-(\ell)) \leq A\,e^{-a\,s} + \frac{s^3\Delta^{3h+2}}{n - s\Delta^h}. \end{aligned} \tag{2.49}$$

The estimate (2.47) follows by taking $s = K\log n$ for some large enough constant K, and by taking n sufficiently large. □

Concerning the out-neighbourhoods, given that the maximum offspring is bounded, the same argument of Lemma 2.2 proves that,

Lemma 2.6 *For any fixed vertex $x \in [n]$, for all $t \in \mathbb{N}$, one has a coupling such that, for some constant $C > 0$ independent of t, n holds*

$$\mathbb{P}\left(\mathcal{B}_t^+(x) \neq \mathcal{T}_{x,t}^+(\mathbf{d}^+)\right) \leq C\,\frac{\Delta^{2t}}{n}.$$

2.4 Diameter and typical distance in the DCM($\mathbf{d}^{\pm}$)

In this section we analyze the diameter of a graph from the DCM($\mathbf{d}^{\pm}$) ensemble. The proof of Theorem 2.1 is a directed version of a classical argument for undirected graphs [12]. It requires controlling the size of in- and out-neighbourhoods of a node, which in turn follows ideas from [2] and [13]. The value $\mathrm{d}_\star = \log_\nu n$ can be interpreted as follows: both the in- and the out-neighbourhood of a node are tree-like with average branching given by ν, so that their boundary at depth h has typically size ν^h, see Lemma 2.7; if the in-neighbourhood of y and the out-neighbourhood of x are exposed up to depth h, one finds that the value $h = \frac{1}{2}\log_\nu(n)$ is critical for the formation of an arc connecting the two neighbourhoods.

Theorem 2.1 *Set $\mathrm{d}_\star = \log_\nu n$. There exists $\varepsilon_n = O\left(\frac{\log\log(n)}{\log(n)}\right)$ such that*

$$\mathbb{P}\left((1-\varepsilon_n)\,\mathrm{d}_\star \le \mathrm{diam}(G) \le (1+\varepsilon_n)\,\mathrm{d}_\star\right) = 1 - o(1). \tag{2.50}$$

Moreover, for any $x, y \in [n]$

$$\mathbb{P}\left((1-\varepsilon_n)\,\mathrm{d}_\star \le d(x,y) \le (1+\varepsilon_n)\,\mathrm{d}_\star\right) = 1 - o(1). \tag{2.51}$$

In particular, Theorem 2.1 shows that w.h.p. the digraph is strongly connected, so there exists a unique stationary distribution π characterized by the equation

$$\pi(x) = \sum_{y=1}^{n} \pi(y) P(y,x)\,, \qquad x \in [n], \tag{2.52}$$

with the normalization $\sum_{x=1}^{n} \pi(x) = 1$.

2.4.1 Controlling the size of the neighbourhoods

We shall need a bound for the size of $\partial\mathcal{B}_h^{\pm}(y)$ for values of h that are larger than $\hbar$. Recall the definition (2.3) of the parameter ν. We use the following notation in the sequel:

$$\ell_0 = 4\log_\delta \log(n), \qquad h_\eta = (1-\eta)\log_\nu(n). \tag{2.53}$$

Lemma 2.7 *For every $\eta \in (0,1)$, there exist constants $c_1, c_2 > 0, \chi > 0$ such that for all $y \in [n]$,*

$$\mathbb{P}\left(\nu^h \log^{-c_1}(n) \le |\partial\mathcal{B}_h^{\pm}(y)| \le \nu^h \log^{c_2}(n)\,,\ \forall h \in [\ell_0, h_\eta]\right) = 1 - O(n^{-1-\chi}). \tag{2.54}$$

Proof: We run the proof for the in-neighbourhood only since the case of the out-neighbourhood is obtained in the same way. We generate $\mathcal{B}_h^-(y)$, $h \in [\ell_0, h_\eta]$ sequentially in a breadth first fashion. After the depth j neighbourhood $\mathcal{B}_j^-(y)$ has been sampled, we call $\mathcal{F}_j$ the set of all heads attached to vertices in $\partial\mathcal{B}_j^-(y)$. Set

$$u = \log^{-7/8}(n).$$

For any $h \geq \ell_0$ define

$$\kappa_h := [\nu(1-u)]^{h-\ell_0} \log^{7/2}(n), \qquad \widehat{\kappa}_h := [\nu(1+u)]^{h-\ell_0} \Delta^{\ell_0}. \tag{2.55}$$

We are going to prove

$$\mathbb{P}\left(\kappa_h \leq |\mathcal{F}_h| \leq \widehat{\kappa}_h\,,\ \forall h \in [\ell_0, h_\eta]\right) = 1 - O(n^{-1-\chi}). \tag{2.56}$$

Notice that, choosing suitable constants $c_1, c_2 > 0$, (2.54) is a consequence of (2.56).

Consider the events

$$A_j = \left\{ |\mathcal{F}_i| \in [\kappa_i, \widehat{\kappa}_i]\ ,\ \forall i \in [\ell_0, j] \right\}. \tag{2.57}$$

Thus, we need to prove $\mathbb{P}(A_h) = 1 - O(n^{-1-\chi})$, for $h = h_\eta$. From Lemma 2.3 and the choice of ℓ_0, it follows that

$$\mathbb{P}(A_{\ell_0}) = 1 - O(n^{-1-\chi}). \tag{2.58}$$

For $h > \ell_0$ we write

$$\mathbb{P}(A_h) = \mathbb{P}(A_{\ell_0}) \prod_{j=\ell_0+1}^{h} \mathbb{P}(A_j | A_{j-1}). \tag{2.59}$$

To estimate $\mathbb{P}(A_j|A_{j-1})$, note that A_{j-1} depends only on the in-neighbourhood $\mathcal{B}_{j-1}^-(y)$, so if σ_{j-1} denotes a realization of $\mathcal{B}_{j-1}^-(y)$ with a slight abuse of notation we write $\sigma_{j-1} \in A_{j-1}$ if A_{j-1} occurs for this given σ_{j-1}. Then

$$\mathbb{P}(A_j|A_{j-1}) = \frac{\sum_{\sigma_{j-1}} \mathbb{P}(\sigma_{j-1})\mathbb{P}(A_j|\sigma_{j-1}) 1_{\sigma_{j-1} \in A_{j-1}}}{\mathbb{P}(A_{j-1})}. \tag{2.60}$$

Therefore, to prove a lower bound on $\mathbb{P}(A_j|A_{j-1})$ it is sufficient to prove a lower bound on $\mathbb{P}(A_j|\sigma_{j-1})$ that is uniform over all $\sigma_{j-1} \in A_{j-1}$.

Suppose we have generated the neighbourhood σ_{j-1} up to depth $j-1$, for a $\sigma_{j-1} \in A_{j-1}$. In some arbitrary order we now generate the matchings of all heads $f \in \mathcal{F}_{j-1}$. We define the random variable $X_f^{(j)}$, $f \in \mathcal{F}_{j-1}$, which, for every f evaluates to the in-degree d_z^- of the vertex z that is matched to f if the vertex z was not yet exposed, and evaluates to zero otherwise. In this way

$$|\mathcal{F}_j| = \sum_{f\in\mathcal{F}_{j-1}} X_f^{(j)}. \tag{2.61}$$

Therefore,

$$\begin{aligned}\mathbb{P}(A_j|\sigma_{j-1}) &= \mathbb{P}\left(\nu(1-u)\kappa_{j-1} \le |\mathcal{F}_j| \le \nu(1+u)\widehat{\kappa}_{j-1} \,|\, \sigma_{j-1}\right) \\ &= 1 - \mathbb{P}\Big(\sum_{f\in\mathcal{F}_{j-1}} X_f^{(j)} < \nu(1-u)\kappa_{j-1} \,|\, \sigma_{j-1}\Big) - \mathbb{P}\Big(\sum_{f\in\mathcal{F}_{j-1}} X_f^{(j)} > \nu(1+u)\widehat{\kappa}_{j-1} \,|\, \sigma_{j-1}\Big).\end{aligned} \tag{2.62}$$

To sample the variables $X_f^{(j)}$, at each step we pick a tail uniformly at random among all unmatched tails and evaluate the in-degree of its end point if it is not yet exposed. Since $\sigma_{j-1} \in A_{j-1}$, at any such step the number of exposed vertices is at most $K = O(n^{1-\eta/2})$. In particular, for any $f \in \mathcal{F}_{j-1}$ and any $d \in [\delta, \Delta]$, $\sigma_{j-1} \in A_{j-1}$:

$$\mathbb{P}\left(X_f^{(j)} = d \,|\, \sigma_{j-1}\right) \ge \frac{\left[\left(\sum_{k=1}^n d_k^+ \mathbf{1}_{d_k^- = d}\right) - \Delta K\right]_+}{m} =: p(d),$$

where $[\cdot]_+$ denotes the positive part. This shows that $X_f^{(j)}$ stochastically dominates the random variable $Y^{(j)}$ and is stochastically dominated by the random variable $\widehat{Y}^{(j)}$, where $Y^{(j)}$ and $\widehat{Y}^{(j)}$ are defined by

$$\begin{gathered}\forall d \in [\delta, \Delta], \qquad \mathbb{P}(Y^{(j)} = d) = \mathbb{P}(\widehat{Y}^{(j)} = d) = p(d) \\ \mathbb{P}\left(\widehat{Y}^{(j)} = \Delta + 1\right) = \mathbb{P}\left(Y^{(j)} = 0\right) = 1 - \sum_{d=\delta}^{\Delta} p(d).\end{gathered}$$

Notice that

$$\mathbb{E}\left(Y^{(j)}\right) = \sum_{d=\delta}^{\Delta} dp(d) \ge \nu - \frac{\Delta^2 K}{m} = \nu - O(n^{-\eta/2}). \tag{2.63}$$

Similarly,

$$\mathbb{E}\left(\widehat{Y}^{(j)}\right) \le \nu + \frac{\Delta^2 K}{m} = \nu + O(n^{-\eta/2}). \tag{2.64}$$

Moreover, letting $Y_i^{(j)}$ and $\widehat{Y}_i^{(j)}$ denote i.i.d. copies of the random variables $Y^{(j)}$ and $\widehat{Y}^{(j)}$ respectively, since $\sigma_{j-1} \in A_{j-1}$, the sum in (2.61) stochastically dominates $\sum_{i=1}^{\kappa_{j-1}} Y_i^{(j)}$, and is stochastically dominated by $\sum_{i=1}^{\widehat{\kappa}_{j-1}} Y_i^{(j)}$. Therefore, $\sum_{f\in\mathcal{F}_{j-1}} X_f^{(j)} < \nu(1-u)\kappa_{j-1}$ implies that

$$\sum_{i=1}^{\kappa_{j-1}} \left[Y_i^{(j)} - \mathbb{E}\left(Y^{(j)}\right)\right] \le -\frac{1}{2}\, u\kappa_{j-1}, \tag{2.65}$$

if n is large enough. Similarly, $\sum_{f\in\mathcal{F}_{j-1}} X_f^{(j)} > \nu(1+u)\widehat{\kappa}_{j-1}$ implies that

$$\sum_{i=1}^{\widehat{\kappa}_{j-1}} \left[\widehat{Y}_i^{(j)} - \mathbb{E}\left(\widehat{Y}^{(j)}\right)\right] \ge \frac{1}{2}\, u\widehat{\kappa}_{j-1}. \tag{2.66}$$

An application of Hoeffding's inequality shows that the probability of the events (2.65) and (2.66) is bounded by $e^{-cu^2\kappa_{j-1}}$ and $e^{-cu^2\widehat{\kappa}_{j-1}}$ respectively, for some absolute constant $c > 0$. Hence, from (2.62) we conclude that for some constant $c > 0$:

$$\mathbb{P}(A_j|\sigma_{j-1}) \ge 1 - e^{-cu^2\kappa_{j-1}} - e^{-cu^2\widehat{\kappa}_{j-1}}.$$

Therefore, using $u^2\widehat{\kappa}_{j-1} \ge u^2\kappa_{j-1} \ge u^2\kappa_0 \ge \log^{3/2}(n)$,

$$\mathbb{P}(A_j|\sigma_{j-1}) = 1 - O(n^{-3}), \tag{2.67}$$

uniformly in $j \in [\ell_0, h_\eta]$ and $\sigma_{j-1} \in A_{j-1}$. By (2.60) the same bound applies to $\mathbb{P}(A_j|A_{j-1})$ and going back to (2.59), for $h = h_\eta$ we have obtained

$$\mathbb{P}(A_h) = 1 - O(n^{-1-\chi}).$$

□

We shall also need the following refinement of Lemma 2.7. Define the events

$$F_y^{\pm} = F_y^{\pm}(\eta, c_1, c_2) = \left\{\nu^h \log^{-c_1}(n) \le |\partial\mathcal{B}_h^{\pm}(y)| \le \nu^h \log^{c_2}(n)\,,\ \forall h \in [\ell_0, h_\eta]\right\}. \tag{2.68}$$

Lemma 2.7 states that

$$\mathbb{P}\left((F_y^{\pm})^c\right) = O(n^{-1-\chi}).$$

Let $\mathcal{G}(\hbar)$ be the event from Proposition 2.3.

Lemma 2.8 *For every $\eta \in (0,1)$, there exist constants $c_1, c_2 > 0, \chi > 0$ such that for all $y \in [n]$,*

$$\mathbb{P}\left((F_y^{\pm})^c; \mathcal{G}(\hbar)\right) = O(n^{-2-\chi}). \tag{2.69}$$

Proof: By symmetry we may prove the inequality for the event F_y^- only. Consider the set $\mathcal{D}_y^-$ of all possible 2-in-neighbourhoods of y compatible with the event $\mathcal{G}(\hbar)$, that is the set of labeled digraphs D such that

$$\mathbb{P}(\mathcal{B}_2^-(y) = D\,;\,\mathcal{G}(\hbar)) > 0. \tag{2.70}$$

Then

$$\mathbb{P}\left((F_y^-)^c;\mathcal{G}(\hbar)\right) \le \sup_{D\in\mathcal{D}_y^-} \mathbb{P}\left((F_y^\pm)^c \,|\, \mathcal{B}_2^-(y) = D\right). \tag{2.71}$$

Thus it is sufficient to prove that

$$\mathbb{P}\left((F_y^\pm)^c \,|\, \mathcal{B}_2^-(y) = D\right) = O(n^{-2-\chi}), \tag{2.72}$$

uniformly in $D \in \mathcal{D}_y^-$. To this end, we may repeat exactly the same argument as in the proof of Lemma 2.7 with the difference that now we condition from the start on the event $\mathcal{B}_2^-(y) = D$ for a fixed $D \in \mathcal{D}_y$. The key observation is that (2.58) can be strenghtened to $O(n^{-2-\chi})$ if we condition on $\mathcal{B}_2^-(y) = D$. That is, for some $\chi > 0$, uniformly in $D \in \mathcal{D}_y$,

$$\mathbb{P}\left(A_{\ell_0} \,|\, \mathcal{B}_2^-(y) = D\right) = 1 - O(n^{-2-\chi}), \tag{2.73}$$

To prove (2.73) notice that if the 2-in-neighbourhood of y is given by $\mathcal{B}_2^-(y) = D \in \mathcal{D}_y$ then the set $\mathcal{F}_2^-(y)$ has at least 4 elements. Therefore, taking a sufficiently large constant C, for the event $|\mathcal{F}_{\ell_0}^-(y)| \ge \delta^{\ell_0}/C$ to fail it is necessary to have at least 3 collisions in the generation of $\mathcal{B}_t^-(y)$, $t \in \{3,\dots,\ell_0\}$. From the estimate (2.37) the probability of this event is bounded by $p_k^3 k^3$ with $k = \Delta^{\ell_0}$, which implies (2.73) if $\chi \in (0,1)$. Once (2.73) is established, the rest of the proof is a repetition of the argument in (2.59)-(2.67). $\square$

2.4.2 Upper bound on the diameter

The upper bound in Theorem 2.1 is reformulated as follows.

Lemma 2.9 *There exist constants $C, \chi > 0$ such that if $\varepsilon_n = \frac{C\log\log(n)}{\log(n)}$,*

$$\mathbb{P}\left(\mathrm{diam}(G) > (1+\varepsilon_n)\,\mathrm{d}_\star\right) = O(n^{-\chi}). \tag{2.74}$$

Proof: From Proposition 2.3 we may restrict to the event $\mathcal{G}(\hbar)$. From the union bound

$$\mathbb{P}\left(\operatorname{diam}(G) > (1+\varepsilon_n)\,\mathrm{d}_\star; \mathcal{G}(\hbar)\right) \le \sum_{x,y\in[n]} \mathbb{P}\left(d(x,y) > (1+\varepsilon_n)\mathrm{d}_\star; \mathcal{G}(\hbar)\right). \tag{2.75}$$

From Lemma 2.8, for all $x, y \in [n]$

$$\mathbb{P}\left(d(x,y) > (1+\varepsilon_n)\mathrm{d}_\star; \mathcal{G}(\hbar)\right) = \mathbb{P}\left(d(x,y) > (1+\varepsilon_n)\mathrm{d}_\star; F_x^+ \cap F_y^-\right) + O(n^{-2-\chi}). \tag{2.76}$$

Fix

$$k = \frac{1+\varepsilon_n}{2} \log_\nu n.$$

Let us use sequential generation to sample first $\mathcal{B}_k^+(x)$ and then $\mathcal{B}_{k-1}^-(y)$. Call σ a realization of these two neighbourhoods. Consider the event

$$U_{x,y} = \{|\partial\mathcal{B}_k^+(x)| \ge \nu^k \log^{-c_1}(n)\,;\ |\partial\mathcal{B}_{k-1}^-(y)| \ge \nu^{k-1}\log^{-c_1}(n)\}.$$

Clearly, $F_x^+ \cap F_y^- \subset U_{x,y}$. Moreover $U_{x,y}$ depends only on σ. Note also that $\{d(x,y) > (1+\varepsilon_n)\mathrm{d}_\star\} \subset E_{x,y}$, where we define the event

$$E_{x,y} = \{\text{There is no path of length} \le 2k-1 \text{ from } x \text{ to } y\}. \tag{2.77}$$

The event $E_{x,y}$ depends only on σ. We say that $\sigma \in U_{x,y} \cap E_{x,y}$ if σ is such that both $E_{x,y}$ and $U_{x,y}$ occur. Thus, we write

$$\begin{aligned}\mathbb{P}\left(d(x,y) > (1+\varepsilon_n)\mathrm{d}_\star; F_x^+ \cap F_y^-\right) &\le \mathbb{P}\left(d(x,y) > (1+\varepsilon_n)\mathrm{d}_\star; U_{x,y}\cap E_{x,y}\right)\\ &\le \sup_{\sigma\in U_{x,y}\cap E_{x,y}} \mathbb{P}\left(d(x,y) > (1+\varepsilon_n)\mathrm{d}_\star \,|\, \sigma\right).\end{aligned} \tag{2.78}$$

Fix a realization $\sigma \in U_{x,y} \cap E_{x,y}$. The event $E_{x,y}$ implies that all vertices on $\partial\mathcal{B}_{k-1}^-(y)$ have all their heads unmatched and the same holds for all the tails of vertices in $\partial\mathcal{B}_k^+(x)$. Call $\mathcal{F}_{k-1}$ the heads attached to vertices in $\partial\mathcal{B}_{k-1}^-(y)$ and $\mathcal{E}_k$ the tails attached to vertices in $\partial\mathcal{B}_k^+(x)$. The event $d(x,y) > (1+\varepsilon_n)\mathrm{d}_\star$ implies that there are no matchings between $\mathcal{F}_{k-1}$ and $\mathcal{E}_k$. The probability of this event is dominated by

$$\left(1 - \frac{|\mathcal{E}_k|}{m}\right)^{|\mathcal{F}_{k-1}|} \le \left(1 - n^{-\frac{1}{2}+\frac{\varepsilon_n}{4}}\right)^{n^{\frac{1}{2}+\frac{\varepsilon_n}{4}}} \le \exp\left(-n^{\varepsilon_n/2}\right),$$

if n is large enough and $\varepsilon_n = C\log\log n/\log n$ with C large enough. Therefore, uniformly in $\sigma \in U_{x,y}\cap E_{x,y}$,

$$\mathbb{P}\left(d(x,y) > (1+\varepsilon_n)\mathrm{d}_\star \,|\, \sigma\right) \le \exp\left(-n^{\varepsilon_n/2}\right) = O(n^{-2-\chi}).$$

Inserting this in (2.75)-(2.76) completes the proof. □

2.4.3 Lower bound on the diameter

We prove the following lower bound on the diameter. Note that Lemma 2.9 and Lemma 2.10 imply Theorem 2.1.

Lemma 2.10 *There exists $C > 0$ such that taking $\varepsilon_n = \frac{C \log\log(n)}{\log(n)}$, for any $x, y \in [n]$,*

$$\mathbb{P}\left(d(x,y) \leq (1-\varepsilon_n)\mathrm{d}_\star\right) = o(1). \tag{2.79}$$

Proof: Define

$$\ell = \frac{1-\varepsilon_n}{2}\log_\nu n.$$

We start by sampling the out-neighbourhood of x up to distance ℓ. Consider the event

$$J_x = \left\{|\mathcal{B}_\ell^+(x)| \leq n^{\frac{1-\varepsilon_n}{2}}\log^{c_2}(n)\right\}.$$

From Lemma 2.7, $\mathbb{P}(J_x) = 1 - O(n^{-1-\chi})$ for suitable constants $c_2, \chi > 0$, and therefore

$$\mathbb{P}(y \in \mathcal{B}_\ell^+(x)) = \mathbb{P}(y \in \mathcal{B}_\ell^+(x); J_x) + O(n^{-1-\chi}). \tag{2.80}$$

If J_x holds, in the generation of $\mathcal{B}_\ell^+(x)$ there are at most $K := n^{\frac{1-\varepsilon_n}{2}}\log^{c_2}(n)$ attempts to include y in $\mathcal{B}_\ell^+(x)$, each with probability at most $d_y^-/(m-K) \leq 2\Delta/m$ of success, so that

$$\mathbb{P}(y \in \mathcal{B}_\ell^+(x); J_x) \leq \frac{2\Delta}{m}K = O(n^{-\frac{1}{2}}). \tag{2.81}$$

Once the out-neighbourhood $\mathcal{B}_\ell^+(x)$ has been generated, if $y \notin \mathcal{B}_\ell^+(x)$, we generate the in-neighbourhood $\mathcal{B}_\ell^-(y)$. If $d(x,y) \leq (1-\varepsilon_n)\mathrm{d}_\star$ then there must be a collision with $\partial\mathcal{B}_\ell^+(x)$, and

$$\mathbb{P}(d(x,y) \leq (1-\varepsilon_n)\mathrm{d}_\star\,;\, y \notin \mathcal{B}_\ell^+(x)) = \mathbb{P}(y \notin \mathcal{B}_\ell^+(x)\,;\, \mathcal{B}_\ell^-(y) \cap \partial\mathcal{B}_\ell^+(x) \neq \emptyset). \tag{2.82}$$

Consider the event

$$J_y = \left\{|\mathcal{B}_\ell^-(y)| < n^{\frac{1-\varepsilon_n}{2}}\log^{c_2}(n)\right\}.$$

From Lemma 2.7 it follows that $\mathbb{P}(J_y) = 1 - O(n^{-1-\chi})$ for suitable constants $c_2, \chi > 0$. If J_x and J_y hold, in the generation of $\mathcal{B}_\ell^-(y)$ there are at most $K = n^{\frac{1-\varepsilon_n}{2}}\log^{c_2}(n)$ attempts to collide with $\partial\mathcal{B}_\ell^+(x)$, each of which with success probability at most $\Delta K/m$, and therefore

$$\mathbb{P}(y \notin \mathcal{B}_\ell^+(x)\,;\, \mathcal{B}_\ell^-(y) \cap \partial\mathcal{B}_\ell^+(x) \neq \emptyset) \leq \frac{\Delta K^2}{m} = O(n^{-\varepsilon_n/2}) = o(1), \tag{2.83}$$

where we take the constant C in the definition of ε_n sufficiently large. In conclusion,

$$\mathbb{P}\left(d(x,y) \leq (1-\varepsilon_n)\mathrm{d}_\star\right) \leq \mathbb{P}\left(y \in \mathcal{B}_\ell^+(x)\right) + \mathbb{P}\left(d(x,y) \leq (1-\varepsilon_n)\mathrm{d}_\star\,;\; y \notin \mathcal{B}_\ell^+(x)\right),$$

and the inequalities (2.80)-(2.83) end the proof. $\square$

CHAPTER 3

MIXING TIME FOR THE PAGERANK SURFER

In this chapter we study the model presented in Section 1.3.1. We start by briefly recalling the definitions given in Chapter 1, then we present the strategy of proof of Theorem 1.10.

Given a directed graph $G = (V, E)$, a parameter $\alpha \in (0, 1)$, and a probability measure λ on V, the *generalised PageRank* surf on G with damping factor $1 - \alpha$ and resampling λ is the Markov chain

$$P_{\alpha,\lambda}(x, y) = (1 - \alpha)P(x, y) + \alpha\lambda(x). \tag{3.1}$$

The resulting stationary distribution $\pi_{\alpha,\lambda}$, characterised by the equation

$$\pi_{\alpha,\lambda}(y) = \sum_{x\in V} \pi_{\alpha,\lambda}(x)P_{\alpha,\lambda}(x, y), \tag{3.2}$$

depends in a nontrivial way on the parameter α and the distribution λ. Starting at a node x the distribution of the PageRank surfer after t steps is $P^t_{\alpha,\lambda}(x, \cdot)$, and the distance to equilibrium is defined by

$$\mathcal{D}^x_{\alpha,\lambda}(t) = \left\| P^t_{\alpha,\lambda}(x, \cdot) - \pi_{\alpha,\lambda} \right\|_{\mathrm{TV}}. \tag{3.3}$$

This defines a non-increasing function of $t \in \mathbb{N}$. It is convenient to extend it to a monotone function of $t \in [0, \infty)$, e.g. by considering the integer part of the argument. Finally, for any $\varepsilon \in (0,1)$, the ε*-mixing time* is defined by

$$T_{\alpha,\lambda}(\varepsilon) = \inf\left\{t \geq 0 : \max_{x\in V} \mathcal{D}^x_{\alpha,\lambda}(t) \leq \varepsilon\right\}. \tag{3.4}$$

When $\alpha = 0$, we write $\mathcal{D}^x_0(t)$ and $T_0(\varepsilon)$ for the corresponding quantities.

As mentioned in Section 1.3.1, we shall assume that $\alpha = \alpha(n) \in (0,1)$ is a sequence such that $\alpha \to 0$ and such that the limit

$$\gamma = \lim_{n\to\infty} \alpha\, T_{\mathrm{ENT}} \in [0,\infty] \tag{3.5}$$

exists, with possibly $\gamma = 0$ or $\gamma = \infty$.

Theorem 3.1 *Let G be a random graph from either the directed configuration model DCM($\mathbf{d}^{\pm}$) or the out-configuration model OCM($\mathbf{d}^{+}$). Let $\alpha = \alpha(n) \in (0,1)$ be parameters as in (3.5), and let $\lambda = \lambda_n$ be either widespread or strongly localized measures. Then, according to the value of γ there are three scenarios:*

(1) *If $\gamma = 0$ then for all $s > 0, s \neq 1$:*

$$\max_{x\in[n]} \left|\mathcal{D}^x_{\alpha,\lambda}(s\,T_{\mathrm{ENT}}) - \vartheta(s)\right| \xrightarrow{\mathbb{P}} 0. \tag{3.6}$$

(2) *If $\gamma \in (0,\infty)$ then for all $s > 0, s \neq 1$:*

$$\max_{x\in[n]} \left|\mathcal{D}^x_{\alpha,\lambda}(s\,T_{\mathrm{ENT}}) - e^{-\gamma s}\vartheta(s)\right| \xrightarrow{\mathbb{P}} 0. \tag{3.7}$$

(3) *If $\gamma = \infty$ then for all $s > 0$:*

$$\max_{x\in[n]} \left|\mathcal{D}^x_{\alpha,\lambda}(s/\alpha) - e^{-s}\right| \xrightarrow{\mathbb{P}} 0. \tag{3.8}$$

where θ is the step function $\theta(s) = 1$ if $s \leq 1$ and $\theta(s) = 0$ otherwise.

In terms of mixing times, Theorem 3.1 implies the following statements.

Corollary 3.1 *In the setting of Theorem 3.1:*

(1) *If $\gamma = 0$ then for all $\varepsilon \in (0,1)$*

$$\frac{T_{\alpha,\lambda}(\varepsilon)}{T_{\mathrm{ENT}}} \xrightarrow{\mathbb{P}} 1, \tag{3.9}$$

(2) If $\gamma \in (0, \infty)$:

$$\frac{T_{\alpha,\lambda}(\varepsilon)}{T_{\mathrm{ENT}}} \xrightarrow{\mathbb{P}} \begin{cases} 1 & \text{if } \varepsilon \in (0, e^{-\gamma}) \\ \frac{1}{\gamma}\log(1/\varepsilon) & \text{if } \varepsilon \in [e^{-\gamma}, 1). \end{cases} \tag{3.10}$$

(3) If $\gamma = \infty$ then for all $\varepsilon \in (0,1)$:

$$\alpha\, T_{\alpha,\lambda}(\varepsilon) \xrightarrow{\mathbb{P}} \log(1/\varepsilon). \tag{3.11}$$

To give some guidelines, below we illustrate the main ideas involved in the proof.

The starting point is the observation that the distance to stationarity $\mathcal{D}^x_{\alpha,\lambda}(t)$ satisfies the following general identity at all times t, for all choices of the parameter α and distribution λ:

$$\left\|P^t_{\alpha,\lambda}(x,\cdot) - \pi_{\alpha,\lambda}\right\|_{\mathrm{TV}} = (1-\alpha)^t \left\|P^t(x,\cdot) - \pi_{\alpha,\lambda}P^t\right\|_{\mathrm{TV}}. \tag{3.12}$$

Here we use the notation $\mu P^t(y) = \sum_{x\in V}\mu(x)P^t(x,y)$ for the distribution at time t of the simple random walk started at a random vertex distributed according to some distribution μ. The relation (3.12) follows from a simple coupling argument; see Proposition 3.3 below. Moreover, the stationary distribution admits the power series expansion

$$\pi_{\alpha,\lambda} = \alpha\sum_{k=0}^{\infty}(1-\alpha)^k \lambda P^k\,, \tag{3.13}$$

see Proposition 3.2 below. A particularly simple special case is when the resampling distribution λ equals the stationary distribution π_0. Indeed, in this case the stationary distribution is the result of a trivial balance and $\pi_{\alpha,\lambda} = \pi_0$, so that (3.12) becomes

$$\mathcal{D}^x_{\alpha,\pi_0}(t) = (1-\alpha)^t \mathcal{D}^x_0(t)\,. \tag{3.14}$$

Therefore, when $\lambda = \pi_0$ the result of Theorem 3.1 is an immediate consequence of Theorem 1.7.

The key observation to attack the case of a general widespread measure λ will be that, if we start with such a distribution λ, then the time needed to reach stationarity for the simple random walk is much smaller than the entropic time T_{ENT}. More precisely we shall establish the following fact.

Lemma 3.1 *Let G be a random graph from either the directed configuration model DCM($\mathbf{d}^{\pm}$) or the out-configuration model OCM($\mathbf{d}^{+}$). If $\lambda = \lambda_n$ is widespread, then for any sequence $t = t(n) \to \infty$,*

$$\left\|\lambda P^t - \pi_0\right\|_{TV} \xrightarrow{\mathbb{P}} 0. \tag{3.15}$$

Once Lemma 3.1 is available, the proof of Theorem 3.1 for the case of widespread measures is not difficult. Indeed, as we shall see, Lemma 3.1 and (3.13) imply that for all widespread measures, for all three scenarios regarding the sequence α, the stationary measures $\pi_{\alpha,\lambda}$ and π_0 become indistinguishable:

$$\|\pi_{\alpha,\lambda} - \pi_0\|_{\mathrm{TV}} \xrightarrow{\mathbb{P}} 0. \tag{3.16}$$

At this point Theorem 3.1 will follow directly from (3.12) and Theorem 1.7.

Let us now turn to the case of strongly localized λ. We will actually reduce the proof to the fully localized case $\lambda = \delta_z$ for some vertex z. In this case, $\lambda P^t = P^t(z, \cdot)$ and therefore (3.15) must fail for all $t = sT_{\mathrm{ENT}}$, with $s \in (0, 1)$ fixed, since by Theorem 1.7 we know that in this case

$$\left\|P^t(z, \cdot) - \pi_0\right\|_{\mathrm{TV}} \xrightarrow{\mathbb{P}} 1. \tag{3.17}$$

The approximation (3.16) can still be expected to hold for the scenario $\alpha T_{\mathrm{ENT}} \to 0$, since in that case the simple random walk has enough time to reach equilibrium between successive visits to the reference vertex z. However, if instead $\alpha T_{\mathrm{ENT}} \to \gamma > 0$, then $\pi_{\alpha,\lambda}$ should be a nontrivial mixture of π_0 and a localized distribution that is singular w.r.t. π_0. We refer to Lemma 3.4 below for the precise version of this statement. A key technical point for the proof of Theorem 3.1 will be the following fact.

Proposition 3.1 *Let G be a random graph from either the directed configuration model DCM($\mathbf{d}^{\pm}$) or the out-configuration model OCM($\mathbf{d}^{+}$). If $\lambda = \delta_z$ for some vertex z, then for fixed $\gamma > 0$, including $\gamma = \infty$, and $s \in (0, \gamma)$, for any sequence $\alpha \to 0$, satisfying $\alpha T_{\mathrm{ENT}} \to \gamma$,*

$$\inf_{x\in[n]} \left\|P^t(x, \cdot) - \pi_{\alpha,\delta_z} P^t\right\|_{TV} \xrightarrow{\mathbb{P}} 1\,, \qquad t = s/\alpha\,. \tag{3.18}$$

3.1 Preliminaries

In this section we collect some simple general facts about the PageRank surf. The statements in the rest of this section do not depend on the graph G where the original walk takes place. Therefore, we fix an arbitrary digraph G with vertex set $V = [n]$, and let P be the transition matrix in of the simple random walk on G. If $d_x^+ = 0$ for some x we may define $P(x, x) = 1$ and $P(x, y) = 0$ for all $y \in V \setminus \{x\}$.

3.1.1 The stationary distribution $\pi_{\alpha,\lambda}$

Proposition 3.2 *For any $\alpha \in (0,1)$, any probability vector λ, let $P_{\alpha,\lambda}$ be defined by (3.1). There exists a unique probability vector $\pi_{\alpha,\lambda}$ satisfying $\pi_{\alpha,\lambda}P_{\alpha,\lambda} = \pi_{\alpha,\lambda}$. Moreover, $\pi_{\alpha,\lambda}$ is given by*

$$\pi_{\alpha,\lambda} = \alpha \sum_{k=0}^{\infty} (1-\alpha)^k \lambda P^k. \tag{3.19}$$

Proof: The equation $\pi_{\alpha,\lambda}P_{\alpha,\lambda} = \pi_{\alpha,\lambda}$ is equivalent to

$$\pi_{\alpha,\lambda}(\mathbf{1} - (1-\alpha)P) = \alpha\lambda.$$

Since P is a stochastic matrix, the matrix $\mathbf{1} - (1-\alpha)P$ is strictly diagonally dominant, and therefore invertible. Then (3.19) follows by expanding the expression $\pi_{\alpha,\lambda} = \alpha\lambda(\mathbf{1} - (1-\alpha)P)^{-1}$. $\square$

In particular, (3.19) and the triangle inequality imply that for any other probability vector μ:

$$\|\pi_{\alpha,\lambda} - \mu\|_{\mathrm{TV}} \leq \alpha \sum_{k=0}^{\infty} (1-\alpha)^k \|\lambda P^k - \mu\|_{\mathrm{TV}}. \tag{3.20}$$

3.1.2 Walk vs. teleport

A trajectory of the PageRank surf can be sampled as follows. At each time unit independently, we flip a α-biased coin: if heads (with probability α) then the surfer is teleported to a new vertex, chosen according to λ; if tails (with probability $1-\alpha$) then the surfer walks one step according to the transition matrix P. The probability associated to this construction will be denoted by $\mathbb{P}$. If τ_α denotes the first time the surfer is teleported, then for all $t \in \mathbb{N}$:

$$\mathbb{P}(\tau_\alpha > t) = (1-\alpha)^t. \tag{3.21}$$

Proposition 3.3 *For any $\alpha \in (0,1)$, any probability vector λ, and all $t \in \mathbb{N}$, $x \in [n]$:*

$$\|P^t_{\alpha,\lambda}(x,\cdot) - \pi_{\alpha,\lambda}\|_{TV} = (1-\alpha)^t \|P^t(x,\cdot) - \pi_{\alpha,\lambda}P^t\|_{TV}. \tag{3.22}$$

Proof: We use the construction introduced above, and write X_t^x for the position of the surfer at time t with initial vertex x. By using the same sample of the teleporting distribution λ we couple two trajectories X_t^x, X_t^z in such a way that $X_t^x = X_t^z$, for all $t \geq \tau_\alpha$. Therefore, letting $\mathbb{E}$ denote the expectation with respect to this coupling:

$$\begin{aligned} P^t_{\alpha,\lambda}(x,y) - P^t_{\alpha,\lambda}(z,y) &= \mathbb{E}\left[\mathbf{1}(X_t^x = y) - \mathbf{1}(X_t^z = y)\right] \\ &= \mathbb{E}\left[\mathbf{1}(X_t^x = y) - \mathbf{1}(X_t^z = y); \tau_\alpha > t\right]. \end{aligned} \tag{3.23}$$

Moreover,

$$\mathbb{E}\left[\mathbf{1}(X_t^x = y); \tau_\alpha > t\right] = \mathbb{P}(\tau_\alpha > t)\mathbb{P}(X_t = y | X_0 = x,\ \tau_\alpha > t) = \mathbb{P}(\tau_\alpha > t)P^t(x,y). \tag{3.24}$$

Therefore,

$$P^t_{\alpha,\lambda}(x,y) - P^t_{\alpha,\lambda}(z,y) = \mathbb{P}(\tau_\alpha > t)(P^t(x,y) - P^t(z,y)). \tag{3.25}$$

Multiplying by $\pi_{\alpha,\lambda}(z)$, summing over z, and using (3.21) one obtains

$$P^t_{\alpha,\lambda}(x,y) - \pi_{\alpha,\lambda}(y) = (1-\alpha)^t \left(P^t(x,y) - [\pi_{\alpha,\lambda}P^t](y)\right). \tag{3.26}$$

It follows that

$$\begin{aligned} \|P^t_{\alpha,\lambda}(x,\cdot) - \pi_{\alpha,\lambda}\|_{\mathrm{TV}} &= \frac{1}{2}\sum_{y\in V} |P^t_{\alpha,\lambda}(x,y) - \pi_{\alpha,\lambda}(y)| \\ &= (1-\alpha)^t \frac{1}{2}\sum_{y\in V} \left|P^t(x,y) - [\pi_{\alpha,\lambda}P^t](y)\right| \\ &= (1-\alpha)^t \|P^t(x,\cdot) - \pi_{\alpha,\lambda}P^t\|_{\mathrm{TV}}. \end{aligned} \tag{3.27}$$

□

Since the total variation distance is always bounded above by 1, Proposition 3.3 implies the upper bound

$$\mathcal{D}^x_{\alpha,\lambda}(t) = \|P^t_{\alpha,\lambda}(x,\cdot) - \pi_{\alpha,\lambda}\|_{\mathrm{TV}} \leq (1-\alpha)^t. \tag{3.28}$$

The latter, in turn, gives the following upper bound on the mixing time.

Corollary 3.2 *For any $\alpha \in (0,1)$, any probability vector λ, and all $\varepsilon \in (0,1)$, the ε-mixing time* (3.4) *satisfies*

$$T_{\alpha,\lambda}(\varepsilon) \leq \frac{1}{\alpha}\log(1/\varepsilon). \tag{3.29}$$

A further immediate consequence of Proposition 3.3 is that if λ is stationary for P, then the distance to equilibrium $\mathcal{D}^x_{\alpha,\lambda}(t)$ takes a simple form.

Corollary 3.3 *For any $\alpha \in (0,1)$, for all $x \in V$ and all $t \in \mathbb{N}$, if π_0 is a probability vector such that $\pi_0 P = \pi_0$, then taking $\lambda = \pi_0$,*

$$\mathcal{D}^x_{\alpha,\pi_0}(t) = (1-\alpha)^t \, \|P^t(x,\cdot) - \pi_0\|_{TV}. \tag{3.30}$$

Proof: From Proposition 3.2 it follows that $\pi_{\alpha,\pi_0} = \pi_0$, and therefore $\pi_{\alpha,\pi_0} P^t = \pi_0$ for all t. □

Finally, another useful consequence of Proposition 3.3 is that it allows us to control the distance $\mathcal{D}^x_{\alpha,\lambda}(t)$ in terms of the distance $\mathcal{D}^x_{\alpha,\pi_0}(t)$, for some stationary π_0 as in Corollary 3.3, by means of the distance between $\pi_{\alpha,\lambda}$ and π_0.

Corollary 3.4 *For any $\alpha \in (0,1)$, all $t \in \mathbb{N}$, any probability vector λ, if π_0 is such that $\pi_0 P = \pi_0$,*

$$\max_{x\in V} \left|\mathcal{D}^x_{\alpha,\lambda}(t) - \mathcal{D}^x_{\alpha,\pi_0}(t)\right| \le \|\pi_{\alpha,\lambda} - \pi_0\|_{TV}. \tag{3.31}$$

Proof: From the triangle inequality and the fact that $\|\mu P^t - \nu P^t\|_{\mathrm{TV}}$ is monotone in t for all distributions μ, ν, one has

$$\left|\|P^t(x,\cdot) - \pi_{\alpha,\lambda} P^t\|_{\mathrm{TV}} - \|P^t(x,\cdot) - \pi_0\|_{\mathrm{TV}}\right| \le \|\pi_{\alpha,\lambda} - \pi_0\|_{\mathrm{TV}}. \tag{3.32}$$

The conclusion then follows from Proposition 3.3 and Corollary 3.3. □

3.2 Mixing from widespread measures

This section is devoted to the proof of Lemma 3.1[1]. Recall that in both models DCM($\mathbf{d}^{\pm}$) and OCM($\mathbf{d}^{+}$) one has w.h.p. a unique stationary distribution for the simple random walk on G, which we denote π_0. The starting point is a result that follows directly from [13, 14], which allows us to replace the unknown distribution π_0 with a local approximation.

Proposition 3.4 *For any fixed $\varepsilon > 0$, taking $h = \varepsilon T_{\mathrm{ENT}}$, as $n \to \infty$ both models satisfy*

$$\left\|\mu_{\mathrm{in}} P^h - \pi_0\right\|_{TV} \xrightarrow{\mathbb{P}} 0. \tag{3.33}$$

[1] That is Theorem 1.11 in Chapter 1.

Proof: From [13, Section 6] for the DCM and from [13, Section 4] for the OCM, it follows that there exists $\varepsilon_0 > 0$ such that for all fixed $\varepsilon \in (0, \varepsilon_0)$, setting $h = \varepsilon T_{\rm ENT}$, and $t = (1 + \varepsilon/2)T_{\rm ENT}$ one has

$$\max_{x\in[n]} \left\|\mu_{\rm in}P^h - P^t(x,\cdot)\right\|_{\rm TV} \xrightarrow{\mathbb{P}} 0. \tag{3.34}$$

To prove (3.33), note that $\left\|\mu_{\rm in}P^h - \pi_0\right\|_{\rm TV}$ is monotone in h, and therefore it is not restrictive to assume that $\varepsilon \in (0, \varepsilon_0)$. Thus (3.33) is a consequence of (3.34) and the simple inequality, valid for any t:

$$\left\|\mu_{\rm in}P^h - \pi_0\right\|_{\rm TV} = \left\|\mu_{\rm in}P^h - \pi_0 P^t\right\|_{\rm TV} \le \max_{x\in[n]} \left\|\mu_{\rm in}P^h - P^t(x,\cdot)\right\|_{\rm TV}. \tag{3.35}$$

□

To prove Lemma 3.1, by monotonicity of $\|\lambda P^t - \pi_0\|_{\rm TV}$ as a function of t, we may restrict to sequences $t = t(n) \to \infty$ with $t = o(\log n)$. Thus, taking advantage of Proposition 3.4, the conclusion of Lemma 3.1 is a consequence of the following result.

Proposition 3.5 *There exists $\varepsilon > 0$ such that if $h = \varepsilon T_{\rm ENT}$, then for any $t = t(n) \to \infty$ with $t = o(\log n)$, for any widespread measure λ:*

$$\left\|\lambda P^t - \mu_{\rm in}P^h\right\|_{TV} \xrightarrow{\mathbb{P}} 0. \tag{3.36}$$

Proof: The proof is divided into two steps. We first show that the random variable appearing in (3.36) is concentrated around its average:

$$\left\|\lambda P^t - \mu_{\rm in}P^h\right\|_{\rm TV} - \mathbb{E}\left[\left\|\lambda P^t - \mu_{\rm in}P^h\right\|_{\rm TV}\right] \xrightarrow{\mathbb{P}} 0. \tag{3.37}$$

Define the random variable

$$Z(\omega) = \left\|\lambda P^t_\omega - \mu_{\rm in}P^h_\omega\right\|_{\rm TV}, \tag{3.38}$$

where ω denotes the map that defines the configuration model as in Section 2.1, and P_ω is the corresponding random walk transition matrix. We now compare the value of Z at two different configurations ω, ω'. In the case of the DCM we assume that ω, ω' differ by a swap of two coordinates, namely that there exist tails e_0^+, e_1^+ and heads e_0^-, e_1^- such that $\omega(e) = \omega'(e)$ for all tails $e \neq e_0^+, e_1^+$ and $\omega(e_i^+) = e_i^-$, $\omega'(e_i^+) = e_{1-i}^-$, $i = 0, 1$. In the case of the OCM we take $\omega = (\omega_x)_{x\in[n]}$ and $\omega' = (\omega'_x)_{x\in[n]}$ such that $\omega_x = \omega'_x$ for all $x \neq x_0$ and $\omega_{x_0} \neq \omega'_{x_0}$ for some fixed $x_0 \in [n]$. Let us show that in either case one has

$$|Z(\omega) - Z(\omega')| \le \Delta|\lambda|_\infty \max_{v\in[n]} |\mathcal{B}_t^{-,\omega}(v)| + \Delta|\mu_{\rm in}|_\infty \max_{v\in[n]} |\mathcal{B}_h^{-,\omega}(v)|. \tag{3.39}$$

where $|\lambda|_\infty = \max_{z\in[n]} \lambda(z)$ and $|\mu_{\text{in}}|_\infty = \max_{z\in[n]} \mu_{\text{in}}(z)$. By the triangle inequality

$$\begin{aligned} |Z(\omega) - Z(\omega')| &\le Z_1(\omega,\omega') + Z_2(\omega,\omega')\,, \\ Z_1(\omega,\omega') = \left\|\lambda P^t_\omega - \lambda P^t_{\omega'}\right\|_{\text{TV}}\,, &\quad Z_2(\omega,\omega') = \left\|\mu_{\text{in}} P^h_\omega - \mu_{\text{in}} P^h_{\omega'}\right\|_{\text{TV}}\,. \end{aligned} \tag{3.40}$$

Let $X_t^{\lambda,\omega}$ denote the position at time t of the random walk which starts with distribution λ, so that $\lambda P^t_\omega(y) = \mathbb{P}(X_t^{\lambda,\omega} = y)$. One may couple exactly $(X_s^{\lambda,\omega})_{s\ge 0}$ and $(X_t^{\lambda,\omega'})_{s\ge 0}$ until the first time τ that the trajectory $(X_0^{\lambda,\omega},\dots,X_t^{\lambda,\omega})$ passes through one of the edges in the symmetric difference of ω and ω'. Therefore, for all ω,ω' as above

$$Z_1(\omega,\omega') \le \mathbb{P}(\tau \le t).$$

Notice that for both models there are at most Δ edges in ω that are not in ω', and that for each such edge, say (u,v), the probability that the trajectory $(X_0^{\lambda,\omega},\dots,X_t^{\lambda,\omega})$ passes through (u,v) is always bounded above by $\lambda(\mathcal{B}_t^{-,\omega}(v))$, where $\mathcal{B}_t^{-,\omega}(v)$ denotes the t-in-neighbourhood of v in the configuration ω. Therefore,

$$Z_1(\omega,\omega') \le \Delta\lambda(\mathcal{B}_t^{-,\omega}(v)) \le \Delta|\lambda|_\infty \max_{x\in[n]} |\mathcal{B}_t^{-,\omega}(x)|.$$

The same argument shows that

$$Z_2(\omega,\omega') \le \Delta\mu_{\text{in}}(\mathcal{B}_h^{-,\omega}(v)) \le \Delta|\mu_{\text{in}}|_\infty \max_{x\in[n]} |\mathcal{B}_h^{-,\omega}(x)|.$$

This proves (3.39) for both models. We now apply this bound to obtain the desired concentration inequality. We start with the DCM, where one has deterministic upper bounds on the size of the in-neighbourhoods, which makes the argument simpler.

In the case of the DCM we observe that $\max_{x\in[n]} |\mathcal{B}_t^{-,\omega}(x)| \le \Delta^{t+1}$ for all $t\in\mathbb{N}$, and all ω. Therefore (3.39) implies

$$|Z(\omega) - Z(\omega')| \le |\lambda|_\infty \Delta^{t+2} + \frac{\Delta^{h+3}}{m}\,. \tag{3.41}$$

By an application of the Azuma–Hoeffding inequality (see e.g. [49, Section 3.2]), (3.41) implies that under the uniform measure over ω, the random variable $\omega \mapsto Z(\omega)$ satisfies the following concentration inequality for all $\eta > 0$:

$$\mathbb{P}\left(|Z - \mathbb{E}[Z]| \ge \eta\right) \le 2\exp\left(-\frac{\eta^2}{2mb^2}\right), \tag{3.42}$$

where $b = |\lambda|_\infty \Delta^{t+2} + \frac{\Delta^{h+3}}{m}$. We may now let $n \to \infty$. If $h = \varepsilon T_{\mathrm{ENT}}$, with $\varepsilon > 0$ fixed but small enough, we have $\Delta^{2h}/m \to 0$, and using $t = o(\log n)$ and the assumption $|\lambda|_\infty = O(n^{-1/2-\delta})$ for some $\delta > 0$, we have $m\Delta^{2t}|\lambda|_\infty^2 = O(n^{-\delta/2})$. In conclusion $mb^2 \to 0$, as desired. This concludes the proof of (3.37) for the DCM.

In the case of the OCM, consider the event

$$E = \left\{ \omega : \max_{x \in [n]} |\mathcal{B}_s^{-,\omega}(x)| \le \Delta^s (\log n)^2, \ \forall s = 1, \dots h \right\}. \tag{3.43}$$

The union bound implies

$$\mathbb{P}(E^c) \le n \sum_{s=1}^{h} \mathbb{P}\left(|\mathcal{B}_s^-(1)| > \Delta^s (\log n)^2 \right). \tag{3.44}$$

To estimate the probability in the right hand side of (3.44) we return to the coupling used in Lemma 2.4 and Lemma 2.5. Instead of coupling the in-neighbourhood $\mathcal{B}_s^-(1)$ with the tree $\mathcal{T}_1^{-,s}$, we use the slightly enlarged tree $\widehat{\mathcal{T}}_1^{-,s}$ obtained by replacing the random variables $\tilde{X}_{k,y}$ in Lemma 2.4 by new variables $\widehat{X}_{k,y}$ defined as Bernoulli with parameter $d_y^+/(n - \sqrt{n})$. The point is that the new tree $\widehat{\mathcal{T}}_1^{-,s}$ obtained in this way is such that, as long as $|\widehat{\mathcal{T}}_1^{-,s}| \le \sqrt{n}$ we also have $|\mathcal{B}_s^-(1)| \le |\widehat{\mathcal{T}}_1^{-,s}|$. Indeed, as long as the number of edges added does not exceed $\sqrt{n}$ then $p_{k,y} \le d_y^+/(n - \sqrt{n})$ for all y. Therefore,

$$\mathbb{P}\left(|\mathcal{B}_s^-(1)| > \Delta^s (\log n)^2 \right) \le \mathbb{P}\left(|\widehat{\mathcal{T}}_1^{-,s}| > \sqrt{n} \right) + \mathbb{P}\left(|\widehat{\mathcal{T}}_1^{-,s}| > \Delta^s (\log n)^2 \right). \tag{3.45}$$

Now we use the bound (2.48) adapted to the enlarged tree $\widehat{\mathcal{T}}_1^{-,s}$. Since at each iteration the number of edges added to the tree is stochastically dominated by a binomial random variable with parameters n and $\Delta/(n-\sqrt{n}) \le 2\Delta/n$, one has a large deviation bound as in (2.48), with possibly different constants. In conclusion, for some new constants $A, a > 0$, for all $s \le h \le \frac{\log n}{3 \log \Delta}$:

$$\mathbb{P}\left(|\mathcal{B}_s^-(1)| > \Delta^s (\log n)^2 \right) \le A\, e^{-a(\log n)^2}. \tag{3.46}$$

Therefore (3.44) implies

$$\mathbb{P}(E^c) \le nhA\, e^{-a(\log n)^2}. \tag{3.47}$$

From the bound in (3.39) we see that

$$|Z(\omega) - Z(\omega')| \le |\lambda|_\infty \Delta^t (\log n)^2 + \frac{\Delta^h (\log n)^2}{n}, \qquad \omega, \omega' \in E. \tag{3.48}$$

We can now apply the following modified version of the standard Azuma–Hoeffding inequality from [23]. Setting $b = |\lambda|_\infty \Delta^t(\log n)^2 + \frac{1}{n}\,\Delta^h(\log n)^2$ and $p = \mathbb{P}(E^c)$, [23, Theorem 2.1] implies

$$\mathbb{P}\left(|Z - \mathbb{E}[Z \,|\, E]| \geq \eta\right) \leq 2\left(p + \exp\left(-\frac{2(\eta - npb)^2}{nb^2}\right)\right), \tag{3.49}$$

where $\mathbb{E}[Z \,|\, E]$ denotes the expected value of Z conditioned on the event E. Notice that if $h = \varepsilon T_{\text{ENT}}$, with $\varepsilon > 0$ small enough, as before we have $nb^2 \to 0$. Moreover, $npb \to 0$ by (3.47). Finally, since $|Z| \leq 1$ we may estimate $|\mathbb{E}[Z \,|\, E] - \mathbb{E}[Z]| \leq p/(1-p) \to 0$. This concludes the proof of the concentration estimate (3.37) for the OCM.

The second step of the proof of Proposition 3.5 is to show that

$$\lim_{n\to\infty} \mathbb{E}\left[\left\|\lambda P^t - \mu_{\text{in}} P^h\right\|_{\text{TV}}\right] = 0. \tag{3.50}$$

Observe that

$$\begin{aligned}\mathbb{E}\left[\left\|\lambda P^t - \mu_{\text{in}} P^h\right\|_{\text{TV}}\right] &= \frac{1}{2}\sum_{j\in[n]} \mathbb{E}\left[\left|\lambda P^t(j) - \mu_{\text{in}} P^h(j)\right|\right] \\ &\leq \frac{1}{2}\mathbb{E}\left[\left|n\lambda P^t(\mathcal{I}) - n\mu_{\text{in}} P^t(\mathcal{I})\right|\right] + \frac{1}{2}\mathbb{E}\left[\left|n\mu_{\text{in}} P^t(\mathcal{I}) - n\mu_{\text{in}} P^h(\mathcal{I})\right|\right],\end{aligned} \tag{3.51}$$

where $\mathcal{I}$ denotes an independent uniformly random vertex in $[n]$ and the expectation $\mathbb{E}$ is understood to include the expectation over $\mathcal{I}$ as well. Consider the first term above. We are going to use Lemma 2.2 for the DCM and Lemma 2.5 for the OCM. Notice that since these estimates apply to any fixed vertex v, they apply just as well if the vertex v is taken to be uniformly random in $[n]$, i.e. if $v = \mathcal{I}$ as it is the case here. In particular, since $t = o(\log n)$, as $n \to \infty$,

$$\mathbb{P}\left(\mathcal{B}_t^-(\mathcal{I}) \neq \mathcal{T}_t^-\right) \to 0, \tag{3.52}$$

where we use the unified notation $\mathcal{T}_t^-$ for the first t generations of the tree $\mathcal{T}_\mathcal{I}^-$ in either DCM or OCM. Next, note that by definition, if $\mathcal{B}_t^-(\mathcal{I}) = \mathcal{T}_t^-$, then

$$n\lambda P^t(\mathcal{I}) - n\mu_{\text{in}} P^t(\mathcal{I}) = nX_t(\lambda) - M_t\,,$$

where we use the notation from (2.9) and (2.14). Therefore,

$$\mathbb{E}\left[\left|n\lambda P^t(\mathcal{I}) - n\mu_{\text{in}} P^t(\mathcal{I})\right|\right] \leq \mathbb{P}\left(\mathcal{B}_t^-(\mathcal{I}) \neq \mathcal{T}_t^-\right) + \mathbb{E}\left[|M_t - nX_t(\lambda)|\right]$$

Using Schwarz' inequality and Proposition 2.2 it follows that

$$\mathbb{E}\left[|M_t - nX_t(\lambda)|\right]^2 \leq \gamma(\lambda)\rho^t.$$

Since $t = t(n) \to \infty$ as $n \to \infty$ and $\rho \in (0,1)$, using (3.52) we conclude that

$$\lim_{n\to\infty} \mathbb{E}\left[\left|n\lambda P^t(\mathcal{I}) - n\mu_{\text{in}} P^t(\mathcal{I})\right|\right] = 0,$$

for all widespread measure λ. This settles the convergence of the first term in (3.51). To handle the second term, reasoning as above we obtain

$$\mathbb{E}\left[\left|n\mu_{\text{in}} P^t(\mathcal{I}) - n\mu_{\text{in}} P^h(\mathcal{I})\right|\right] \le \mathbb{P}\left(\mathcal{B}_h^-(\mathcal{I}) \ne \mathcal{T}_h^-\right) + \mathbb{E}\left[|M_t - M_h|\right]$$

If $h \le \frac{1}{5}\log_\Delta n$, Lemma 2.2 and Lemma 2.5 imply that both models satisfy

$$\mathbb{P}\left(\mathcal{B}_h^-(\mathcal{I}) \ne \mathcal{T}_h^-\right) \to 0. \tag{3.53}$$

Moreover, Schwarz' inequality and Proposition 2.1 imply

$$\begin{aligned}\mathbb{E}\left[|M_t - M_h|\right]^2 &\le \mathbb{E}[(M_t - M_h)^2] \\ &\le \mathbb{E}[(M_t - M_\infty)^2] = C\rho^t.\end{aligned}$$

Since the constant C is bounded, letting $n \to \infty$ concludes the proof. □

3.3 Fully localized case

The goal of this section is to prove Proposition 3.1. Recall that we have $\lambda = \delta_z$ for a fixed vertex z and we are assuming $\alpha T_{\text{ENT}} \to \gamma > 0$. We start with the simpler case $\gamma = +\infty$.

Proof of Proposition 3.1: the case $\gamma = +\infty$

Let $\mathcal{B}_t^+(z)$ denote the t-out-neighbourhood of a vertex z, that is the subgraph of G induced by the set of directed paths of length t in G which start at vertex z. Since $\gamma = +\infty$, there exists a sequence $u = u(n)$ such that $u = u(n) = o(\log n)$ and $\alpha u \to +\infty$. Next, set $t = s/\alpha$ for some fixed $s \in (0,\infty)$, and notice that for all z, the measure π_{α,δ_z} satisfies

$$\lim_{n\to\infty} \pi_{\alpha,\delta_z} P^t(\mathcal{B}_u^+(z)) = 1, \tag{3.54}$$

with probability 1. Indeed, from Proposition 3.2, we see that

$$\pi_{\alpha,\delta_z} P^t(\mathcal{B}_u^+(z)) = \alpha \sum_{k=0}^{\infty} (1-\alpha)^k P^{k+t}(z, \mathcal{B}_u^+(z))$$
$$\geq \alpha \sum_{k=0}^{u/2} (1-\alpha)^k = 1 - (1-\alpha)^{u/2} \to 1, \tag{3.55}$$

where we have used the obvious fact that $P^{k+t}(z, \mathcal{B}_u^+(z)) = 1$ for all k such that $k + t \leq u$, and that $u/2 + t \leq u$ for n large enough. Thus, the proof of Proposition 3.1 would be achieved if we could show that for all x:

$$P^t\big(x, \mathcal{B}_u^+(z)\big) \xrightarrow{\mathbb{P}} 0. \tag{3.56}$$

However, we need this estimate to hold uniformly in $x \in [n]$ while the above statement cannot hold if e.g. $x = z$, since $P^t(z, \mathcal{B}_u^+(z)) = 1$. Thus, we shall actually prove the following slightly different statement.

Lemma 3.2 *Suppose $\alpha T_{\mathrm{ENT}} \to \infty$. Fix $z \in [n]$, $t = s/\alpha$, $s \in (0, \infty)$, and a sequence $u = o(\log n)$ such that $\alpha u \to \infty$. For each $\eta > 0$, with high probability: for each x there exists a set $F_x \subset \mathcal{B}_u^+(z)$ such that*

$$\max_{x\in[n]} P^t\big(x, \mathcal{B}_u^+(z) \setminus F_x\big) \leq \eta, \qquad \min_{x\in[n]} \pi_{\alpha,\delta_z} P^t(\mathcal{B}_u^+(z) \setminus F_x) \geq 1 - \eta. \tag{3.57}$$

Notice that Lemma 3.2 implies Proposition 3.1 in the case $\gamma = +\infty$. The proof of Lemma 3.2 is broken into several steps. We first recall that by Section 2.2.1, $\mathcal{B}_u^+(z)$ can be coupled to a tree in both models. Next, we will introduce the notion of *gates* from a vertex x to the set $\mathcal{B}_u^+(z)$, which will be the basis of the construction of the sets F_x needed for the proof of (3.57).

From Lemma 2.6, for our fixed vertex z, and for our choice of the sequence u, the event $E_z = \{\mathcal{B}_{2u}^+(z) = \mathcal{T}_{z,2u}^+\}$ satisfies

$$\lim_{n\to\infty} \mathbb{P}(E_z) = 1. \tag{3.58}$$

We will need the following definition.

Definition 3.1 *For $x \in [n]$, $t \in \mathbb{N}$ and a directed subgraph $\Gamma \subset G$ with vertex set $V(\Gamma)$, the set of* t-gates *from x to Γ, denoted $\mathcal{G}_t(x, \Gamma)$ is defined as the subset of vertices $y \in \mathcal{B}_t^+(x) \cap V(\Gamma)$ such that: either $x = y \in V(\Gamma)$, or $y \in \mathcal{B}_t^+(x) \setminus \{x\}$ and there exists a directed path $(x_0, x_1, \ldots, x_s)$ in G with $1 \leq s \leq t$ edges, such that $x_0 = x$, $x_{s-1} \notin V(\Gamma)$ and $x_s = y \in V(\Gamma)$.*

Lemma 3.3 *For any fixed vertex z, if $t = o(u)$ and $u = o(\log n)$, then*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{x\in[n]} |\mathcal{G}_t(x, \mathcal{B}_u^+(z))| \geq 3\right) = 0. \tag{3.59}$$

Proof: Fix $x \in [n]$. Recall the generation algorithm from Section 2.3.1. If we invert the role of heads vs. tails and we use the priority rule given by the minimal distance to vertex z, and we stop when the minimal distance exceeds u, the algorithm generates the u-out-neighbourhood $\mathcal{B}_u^+(z)$. Once the digraph $\mathcal{B}_u^+(z)$ has been generated we turn to the generation of the t-in-neighbourhood $B_{x,t}^+$. Here we use the same algorithm, without erasing the pairs already matched in the generation of $\mathcal{B}_u^+(z)$. By definition of gate, if $y \in B_{x,t}^+\setminus\{x\}$ is a gate, there is a time in the generation of $B_{x,t}^+$ where an arc (w, y) is formed between a tail $e_+ \in E_w^+$ with vertex $w \notin \mathcal{B}_u^+(z)$ and a head $e_- \in E_y^-$ with vertex $y \in \mathcal{B}_u^+(z)$. Given the realization of $\mathcal{B}_u^+(z)$, at the k-th step of the generation of $B_{x,t}^+$, conditioned on the previous history of the generation process, the probability of forming such an arc is at most

$$\frac{\Delta|\mathcal{B}_u^+(z)|}{m - |\mathcal{B}_u^+(z)| - k} \leq \frac{\Delta^{u+2}}{m - \Delta^{u+1} - \Delta^{t+1}}, \tag{3.60}$$

where we use the fact that $|\mathcal{B}_u^+(z)| \leq \Delta^{u+1}$ and $k \leq \Delta^{t+1}$. From our assumptions on t, u, one has that this probability is less that $p := n^{-1+\delta}$, for any fixed $\delta > 0$. Thus, the number of such arcs is stochastically dominated by the binomial random variable X with parameters Δ^{t+1} and p. It follows that the number of gates is stochastically dominated by $1 + X$, where the 1 takes into account the possibility that x itself is a gate, and therefore for any $K \in \mathbb{N}$:

$$\mathbb{P}\left(|\mathcal{G}_t(x, \mathcal{B}_u^+(z))| \geq K\right) \leq \mathbb{P}(X + 1 \geq K) \leq (p\Delta^{t+1})^{K-1}. \tag{3.61}$$

Since $t = o(\log n)$ we obtain that $(p\Delta^{t+1})^{K-1}$ is $o(1/n)$ already with $K = 3$. A union bound over $x \in [n]$ implies (3.59). □

We turn to the application of the notion of gates to our problem. Let us write Q_x for the law of the trajectory $(X_0, X_1, \dots)$ of the random walk started at $X_0 = x$ in the graph G, so that

$$P^t(x, A) = Q_x(X_t \in A), \tag{3.62}$$

Notice that if $B_{x,t}^+$ is a tree, then $X_0 = x$ and $X_t \in \mathcal{B}_u^+(z)$ imply that the trajectory passes through a t-gate from x to $\mathcal{B}_u^+(z)$ and that for each gate y the first visit to y must occur, if ever, at a deterministic time t_y, which is the height of y in $\mathcal{T}^+_{x,t}$. Therefore, if $\mathcal{B}^+_{2u}(z)$ is also a tree, then the measure

$$A \mapsto Q_x(X_t \in \mathcal{B}^+_u(z) \cap A)\,, \tag{3.63}$$

is supported on sets L_{z,s_y}, where $s_y = h_y + t - t_y$ and we use the notation $L_{z,s}$ for the set of vertices at distance s from z in the tree $\mathcal{B}^+_u(z)$, and h_y is such that $y \in L_{z,h_y}$. Note that here we are using the fact that once the walk enters the set $\mathcal{B}^+_u(z) \subset \mathcal{B}^+_{2u}(z)$ then it cannot exit $\mathcal{B}^+_u(z)$ and come back to it within time $t \le u$ since $\mathcal{B}^+_{2u}(z)$ is a directed tree. Thus, if $B^+_{x,t}$ and $\mathcal{B}^+_{2u}(z)$ are both trees, and we define the set

$$A_{x,t} = \cup_{y \in \mathcal{G}_t(x,\mathcal{B}^+_u(z))} L_{z,s_y}\,, \tag{3.64}$$

then

$$P^t\big(x, \mathcal{B}^+_u(z) \setminus A_{x,t}\big) = 0. \tag{3.65}$$

Proof of Lemma 3.2

For $h \in \mathbb{N}$, let $V_{*,h} \subset V$ denote the subset of vertices $x \in V$ such that $\mathcal{B}^+_h(x)$ is a tree. As in [13, Propositon 6] one shows that for both models, with high probability:

$$\max_{x \in V} Q_x(X_\ell \in V \setminus V_{*,h}) \le 2^{-\ell}, \qquad \ell \le h \le \frac{\log n}{10 \log \Delta}\,. \tag{3.66}$$

From (3.58) and Lemma 3.3, we may assume that $\mathcal{B}^+_{2u}(z)$ is a tree and that $|\mathcal{G}_t(x,\mathcal{B}^+_{2u}(z))| \le 2$ for all $x \in [n]$. For every $x \in V_{*,t}$ let $A_{x,t}$ denote the sets defined in (3.64). For any $\eta > 0$, take ℓ such that $2^{-\ell} < \eta$. For every $x \in [n]$, define

$$F_x = \cup_{w \in V_{*,t-\ell} \cap \mathcal{B}^+_\ell(x)} A_{w,t-\ell}. \tag{3.67}$$

From (3.65) and (3.66) it follows that with high probability

$$\max_{x \in [n]} P^t\big(x, \mathcal{B}^+_u(z) \setminus F_x\big) \le 2^{-\ell} + \max_{w \in V_{*,t-\ell} \cap \mathcal{B}^+_\ell(x)} P^{t-\ell}\big(w, \mathcal{B}^+_u(z) \setminus A_{w,t-\ell}\big) = 2^{-\ell}. \tag{3.68}$$

Note that F_x is the union of at most $2\Delta^{\ell+1}$ levels of the form $L_{z,s}$. Since $\mathcal{B}^+_{2u}(z)$ is a tree, for any fixed set $L_{z,s} \subset \mathcal{B}^+_u(z)$, $P^{k+t}(z, L_{z,s})$ can be nonzero for only one index $k \le u/2$. Therefore, reasoning as in (3.55) and using $\alpha \to 0$,

$$\max_{x \in [n]} \pi_{\alpha,\delta_z} P^t(F_x) \le 2\Delta^{\ell+1}\alpha + (1-\alpha)^{u/2} \le \eta, \tag{3.69}$$

as soon as n is large enough. This ends the proof of Lemma 3.2. $\square$

Proof of Proposition 3.1: the case $\gamma \in (0, \infty)$

We start by approximating the measure $\pi_{\alpha,\delta_z} P^t$ by a convex combination of π_0 and a more localized probability vector μ. Afterwards we recall some key facts about the random walk that were already established in [13, 14]. Later we combine these ingredients with a strategy similar to that employed in the proof of Lemma 3.2 above to finish the proof.

Lemma 3.4 *Fix $\gamma \in (0, \infty)$ and $s \in (0, \gamma)$. Then there exists a constant $C > 0$ such that for all $\eta > 0$ small enough, for $\alpha T_{\mathrm{ENT}} \to \gamma$ and $t = s/\alpha$, with high probability:*

$$\|\pi_{\alpha,\delta_z} P^t - a\mu - (1-a)\pi_0\|_{TV} \leq C\eta, \tag{3.70}$$

where $a \in (0,1)$ and the probability vector $\mu = \mu_{z,\eta}$ are given by

$$a = \sum_{k=0}^{(1-\eta)T_{\mathrm{ENT}}-t} \alpha(1-\alpha)^k, \qquad \mu = \frac{1}{a} \sum_{k=0}^{(1-\eta)T_{\mathrm{ENT}}-t} \alpha(1-\alpha)^k P^{k+t}(z, \cdot). \tag{3.71}$$

Proof: For any $a < b$, define the probability vector

$$\nu_{a,b} = \frac{1}{Z_{a,b}} \sum_{k=aT_{\mathrm{ENT}}}^{bT_{\mathrm{ENT}}-1} \alpha(1-\alpha)^k P^{k+t}(z, \cdot), \qquad Z_{a,b} = \sum_{k=aT_{\mathrm{ENT}}}^{bT_{\mathrm{ENT}}-1} \alpha(1-\alpha)^k. \tag{3.72}$$

Since $t = s/\alpha$, $s \in (0, \gamma)$, and $\alpha T_{\mathrm{ENT}} \to \gamma$ we may equivalently set $t = \kappa T_{\mathrm{ENT}}$, $\kappa \in (0,1)$. Using Proposition 3.2, for any $0 < \eta < (1-\kappa)$ we write

$$\pi_{\alpha,\delta_z} P^t = Z_{0,1-\eta-\kappa}\, \nu_{0,1-\eta-\kappa} + Z_{1-\eta-\kappa,1+\eta-\kappa}\, \nu_{1-\eta-\kappa,1+\eta-\kappa} + Z_{1+\eta-\kappa,\infty}\, \nu_{1+\eta-\kappa,\infty} \tag{3.73}$$

Using $\alpha T_{\mathrm{ENT}} \to \gamma \in (0, \infty)$, by Riemann integration it follows that for all n large enough

$$Z_{1-\eta-\kappa,1+\eta-\kappa} \leq \sum_{k=(1-\eta-\kappa)T_{\mathrm{ENT}}}^{(1+\eta-\kappa)T_{\mathrm{ENT}}} \alpha e^{-k\alpha} \leq C\eta, \tag{3.74}$$

for some constant $C > 0$. Next, using Theorem 1.7, w.h.p.

$$\sup_{k \geq (1+\eta-\kappa)T_{\mathrm{ENT}}} \|P^{k+t}(z, \cdot) - \pi_0\|_{\mathrm{TV}} \leq \eta. \tag{3.75}$$

It follows that w.h.p.

$$\|\nu_{1+\eta-\kappa,\infty} - \pi_0\|_{\mathrm{TV}} \leq \eta. \tag{3.76}$$

Taking a and μ as in (3.71) and adjusting the value of the constant C concludes the proof. $\square$

Let us recall the following key facts established in [13, Section 6] for the DCM and from [14, Section 4] for the OCM. For every z, and every $\eta > 0$, there exists a directed tree $\mathcal{T}_z(\eta)$ rooted at z such that the trajectory $(X_0, \dots, X_u)$ of the random walk started at z satisfies with high probability $(X_0, \dots, X_u) \subset \mathcal{T}_z(\eta)$ for all $u \leq (1-\eta)T_{\mathrm{ENT}}$. Here the notation $(X_0, \dots, X_u) \subset \mathcal{T}_z(\eta)$ means that the directed path $(X_0, \dots, X_u)$ is a subgraph of the directed tree. More precisely, fix

$$h = c \log n, \tag{3.77}$$

where $c = \eta^2/(4 \log \Delta)$, and recall the set $V_{*,h}$ introduced in the proof of Lemma 3.2. Then, with the notation Q_x used in (3.62), one has[2] : for all $\eta T_{\mathrm{ENT}} \leq u \leq (1-\eta)T_{\mathrm{ENT}}$

$$\max_{z \in V_{*,h}} Q_z\left((X_0, \dots, X_u) \subset \mathcal{T}_z(\eta)\right) \xrightarrow{\mathbb{P}} 1, \tag{3.78}$$

see [14, Lemma 11]. Moreover, the number of vertices in the tree $\mathcal{T}_z(\eta)$ satisfies

$$|\mathcal{T}_z(\eta)| \leq n^{1-\eta^2/2}, \tag{3.79}$$

see [14, Section 4.1]. The bound (3.79) can be used to establish another crucial fact, namely that for a walk that is out of $\mathcal{T}_z(\eta)$ at some time, then it is very unlikely to get to $\mathcal{T}_z(\eta)$ at some later time. The precise statement we need is as follows.

Lemma 3.5 *For any fixed $z \in V_{*,h}$, all $\eta T_{\mathrm{ENT}} \leq u \leq (1-\eta)T_{\mathrm{ENT}}$, with $u \geq h$:*

$$\max_{x \in [n]} Q_x\left((X_h, \dots, X_u) \not\subset \mathcal{T}_z(\eta) \text{ and } X_u \in \mathcal{T}_z(\eta)\right) \xrightarrow{\mathbb{P}} 0. \tag{3.80}$$

The proof of Lemma 3.5 follows the strategy introduced in [14, Section 2.4]. We omit the details and refer the interested reader to [14, Lemma 12] for the proof of a very similar statement.

The analogue of Lemma 3.2 in our present setting reads as follows.

[2] In [14] the notation $\mathcal{T}_z((1-\eta)T_{\mathrm{ENT}})$ is used instead of $\mathcal{T}_z(\eta)$, and the sequence h is defined as $h = c \log n$, with a constant c different from our choice $c = \eta^2/(4 \log \Delta)$. However, it can be checked that the choice of c is not important for the result (3.78), and here it is convenient to work with $c = \eta^2/(4 \log \Delta)$ for reasons that will be clear below.

Lemma 3.6 *Suppose $\alpha T_{\mathrm{ENT}} \to \gamma \in (0,\infty)$. Fix $z \in [n]$ and take $t = \kappa T_{\mathrm{ENT}}$ with some $\kappa \in (0,1)$. For each $\eta > 0$, with high probability: for each x there exists a set $F_x \subset \mathcal{T}_z(\eta)$ such that*

$$\max_{x\in[n]} P^t(x, \mathcal{T}_z(\eta) \setminus F_x) \le \eta, \qquad \min_{x\in[n]} \mu(\mathcal{T}_z(\eta) \setminus F_x) \ge 1-\eta, \tag{3.81}$$

where μ is the measure from Lemma 3.4.

Proof: Suppose A is a subset of vertices in $\mathcal{T}_z(\eta)$. If $\eta > 0$ is small enough, then $\eta T_{\mathrm{ENT}} \le t \le (1-\eta)T_{\mathrm{ENT}}$ and $t \ge h$. Thus, using Lemma 3.5, w.h.p.:

$$\begin{aligned} P^t(x,A) &= Q_x(X_t \in A) \\ &= Q_x\left((X_h,\dots,X_t) \not\subset \mathcal{T}_z(\eta),\ X_t \in A\right) + Q_x((X_h,\dots,X_t) \subset \mathcal{T}_z(\eta),\ X_t \in A) \\ &\le \eta/2 + Q_x((X_h,\dots,X_t) \subset \mathcal{T}_z(\eta),\ X_t \in A). \end{aligned} \tag{3.82}$$

Consider the measure

$$A \mapsto \bar{Q}_x(A) := Q_x((X_h,\dots,X_t) \subset \mathcal{T}_z(\eta),\ X_t \in A)\,. \tag{3.83}$$

A trajectory started at $X_0 = x$ that satisfies $(X_h,\dots,X_t) \subset \mathcal{T}_z(\eta)$ must have entered the set $\mathcal{T}_z(\eta)$ at a vertex $y \in \mathcal{B}_h^+(x)$ that is also an h-gate from x to $\mathcal{T}_z(\eta)$; see Definition 3.1. In particular, if $x \in V_{*,h}$, i.e. if $\mathcal{B}_h^+(x)$ is a tree, then an h-gate y from x to $\mathcal{T}_z(\eta)$ has a deterministic first visit time t_y for the walk started at x and such that $(X_h,\dots,X_t) \subset \mathcal{T}_z(\eta)$. It follows that $\bar{Q}_x$ is supported on on sets L_{z,s_y}, where $s_y = h_y + t - t_y$, where $L_{z,s}$ is the set of vertices at distance s from z in the tree $\mathcal{T}_z(\eta)$, and h_y is such that $y \in L_{z,h_y}$. Thus, if $\mathcal{B}_h^+(x)$ is a tree, we define the set

$$A_{x,t} = \cup_{y\in\mathcal{G}_h(x,\mathcal{T}_z(\eta))} L_{z,s_y}\,. \tag{3.84}$$

With these definitions one has

$$\bar{Q}_x\left(\mathcal{T}_z(\eta) \setminus A_{x,t}\right) = 0\,, \qquad x \in V_{*,h}. \tag{3.85}$$

Next, take $\ell \in \mathbb{N}$ such that $2^{-\ell} \le \eta/2$. For every $x \in [n]$, define

$$F_x = \cup_{w\in V_{*,h}\cap\mathcal{B}_\ell^+(x)} A_{w,t-\ell}. \tag{3.86}$$

As in (3.68), using (3.82), we obtain that with high probability

$$\max_{x\in[n]} P^t(x, \mathcal{T}_z(\eta) \setminus F_x) \le 2^{-\ell} + \max_{w\in V_{*,h}\cap\mathcal{B}_\ell^+(x)} P^{t-\ell}(w, \mathcal{T}_z(\eta) \setminus A_{w,t-\ell}) \le \eta. \tag{3.87}$$

This gives the desired bound in (3.81) for the measure $P^t(x,\cdot)$. Let us now turn to a proof of the bound in (3.81) on the measure μ. We first give an estimate on the number of gates $\mathcal{G}_h(x,\mathcal{T}_z(\eta))$. From the argument in the proof of Lemma 3.3, we know that $|\mathcal{G}_h(x,\mathcal{T}_z(\eta))|$ is stochastically dominated by 1 plus a binomial with parameters Δ^{h+1} and p given by

$$\frac{\Delta|\mathcal{T}_z(\eta)|}{m-|\mathcal{T}_z(\eta)|-\Delta^{h+1}} \le \Delta\, n^{-\frac{1}{2}\eta^2}, \tag{3.88}$$

where we have used (3.79). Therefore for any $K \in \mathbb{N}$:

$$\mathbb{P}\left(|\mathcal{G}_h(x,\mathcal{T}_z(\eta))| \ge K\right) \le (\Delta^{h+2} n^{-\frac{1}{2}\eta^2})^{K-1}. \tag{3.89}$$

Since $h = c\log n$ with $c = \eta^2/(4\log\Delta)$, we see that (3.89) is less than $\Delta^{2K} n^{-K\eta^2/4}$. Thus, if $K = K_\eta$ is larger than say $5\eta^{-2}$, this probability is $o(1/n)$ and then a union bound shows that with high probability

$$\max_{x\in[n]} |\mathcal{G}_h(x,\mathcal{T}_z(\eta))| \le K_\eta. \tag{3.90}$$

Thus from now on we assume that for all x one has $|\mathcal{G}_h(x,\mathcal{T}_z(\eta))| \le K_\eta$. Moreover, we may assume that $z \in V_{*,h}$. Indeed, for fixed z this holds w.h.p. by Lemma 2.2 for the DCM and by Lemma 2.5 for the OCM. Next, observe that the measure μ satisfies w.h.p.

$$\mu(\mathcal{T}_z(\eta)) \ge 1-\eta/2. \tag{3.91}$$

This follows from the expression (3.71) for μ and the fact (3.78) that the walk started at $z \in V_{*,h}$ stays in $\mathcal{T}_z(\eta)$ up to time $(1-\eta)T_{\mathrm{ENT}}$ with large probability. On the other hand,

$$\max_{x\in[n]} \mu(F_x) \le \alpha\Delta^{\ell+1}K_\eta \tag{3.92}$$

where we use the fact that there are at most $\Delta^{\ell+1}$ vertices in $\mathcal{B}^+_\ell(x)$, that there are at most K_η levels of the form $L_{z,s}$ in a set $A_{w,t-\ell}$ and that each of these levels contributes at most α to the mass of μ. For each $\eta > 0$, using $\alpha \to 0$ the above bound is less than $\eta/2$ as soon as n is large enough. Therefore,

$$\mu(\mathcal{T}_z(\eta) \setminus F_x) \ge 1-\eta, \tag{3.93}$$

which completes the proof of Lemma 3.6. $\square$

To finish the proof of Proposition 3.1, remark that by Theorem 1.7, since $t = \kappa T_{\mathrm{ENT}}$, $\kappa \in (0,1)$, we know that for all $\eta > 0$ with high probability: for all x there exists a set B_x such that

$$\pi_0(B_x) \geq 1 - \eta, \qquad P^t(x, B_x) \leq \eta. \tag{3.94}$$

Therefore, taking $B'_x = B_x \cup (\mathcal{T}_z(\eta) \setminus F_x)$ and using Lemma 3.6 we obtain

$$a\mu(B'_x) + (1-a)\pi_0(B'_x) \geq 1 - \eta, \qquad P^t(x, B'_x) \leq 2\eta, \tag{3.95}$$

with a and μ from Lemma 3.4. In conclusion, if C is the constant in Lemma 3.4, we have proved that for all $\eta > 0$, with high probability:

$$\begin{aligned} \left\| P^t(x,\cdot) - \pi_{\alpha,\delta_z} P^t \right\|_{\mathrm{TV}} &\geq \left\| P^t(x,\cdot) - a\mu - (1-a)\pi_0 \right\|_{\mathrm{TV}} - C\eta \\ &\geq \left(a\mu(B'_x) + (1-a)\pi_0(B'_x) - P^t(x, B'_x) \right) - C\eta \geq 1 - (C+3)\eta. \end{aligned} \tag{3.96}$$

Since $\eta > 0$ is arbitrary this completes the proof of Proposition 3.1. $\square$

3.4 Proof of the trichotomy

In this section we show how to prove Theorem 3.1 from the facts established above. Thus, G is a random graph from either the directed configuration model DCM($\mathbf{d}^{\pm}$) or the out-configuration model OCM($\mathbf{d}^{+}$), where the degree sequences satisfy the assumptions (2.2) and (2.5) respectively, and π_0 denotes the (w.h.p.) unique stationary distribution for the simple random walk on G.

Scenario 1, all λ's

We begin with scenario 1, namely when $\alpha T_{\mathrm{ENT}} \to 0$. In this case, we actually prove a result that holds regardless of the resampling distribution λ. In particular, it will imply the first part of the trichotomy in Theorem 3.1 for arbitrary λ.

Proposition 3.6 *For any sequence α such that $\alpha T_{\mathrm{ENT}} \to 0$,*

$$\sup_{\lambda} \left\| \pi_{\alpha,\lambda} - \pi_0 \right\|_{TV} \xrightarrow{\mathbb{P}} 0, \tag{3.97}$$

where the supremum is over all possible probability vectors λ on $[n]$.

Proof: We need to show that, uniformly in λ, for any $\delta > 0$,

$$\|\pi_{\alpha,\lambda} - \pi_0\| \le \delta\,, \qquad w.h.p. \tag{3.98}$$

The upper bound (3.20) shows that for all $t \in \mathbb{N}$:

$$\|\pi_{\alpha,\lambda} - \pi_0\|_{\mathrm{TV}} \le (1 - (1-\alpha)^t) + \sum_{k>t} \alpha(1-\alpha)^k \|\lambda P^k - \pi_0\|_{\mathrm{TV}}. \tag{3.99}$$

Take $t = sT_{\mathrm{ENT}}$, with some fixed $s > 1$, and observe that by Theorem 1.7 we know that for all $k > t$, for all λ:

$$\begin{aligned} \|\lambda P^k - \pi_0\|_{\mathrm{TV}} &\le \|\lambda P^{sT_{\mathrm{ENT}}} - \pi_0\|_{\mathrm{TV}} \\ &\le \max_{x \in V} \|P^{sT_{\mathrm{ENT}}}(x,\cdot) - \pi_0\|_{\mathrm{TV}} \le \delta/2\,, \qquad w.h.p. \end{aligned} \tag{3.100}$$

In particular, using $\alpha t \to 0$:

$$\sup_{\lambda} \|\pi_{\alpha,\lambda} - \pi_0\|_{\mathrm{TV}} \le (1 - (1-\alpha)^t) + \delta/2 \le \delta\,, \qquad w.h.p. \tag{3.101}$$

□

The claim (3.6), for arbitrary λ, is thus a consequence of Corollary 3.3, Corollary 3.4 and Theorem 1.7.

Scenario 2 and 3, widespread λ's

We first show that if λ is widespread, then Lemma 3.1 ensures that $\|\pi_{\alpha,\lambda} - \pi_0\|_{\mathrm{TV}} \to 0$ in probability.

Lemma 3.7 *If $\lambda = \lambda_n$ is widespread, then for any sequence $\alpha = \alpha(n) \to 0$,*

$$\|\pi_{\alpha,\lambda} - \pi_0\|_{TV} \xrightarrow{\mathbb{P}} 0. \tag{3.102}$$

Proof: Take $t = t(n) \to \infty$ such that $\alpha t \to 0$. From Lemma 3.1 we know that

$$\left\|\lambda P^t - \pi_0\right\|_{\mathrm{TV}} \xrightarrow{\mathbb{P}} 0. \tag{3.103}$$

As in (3.99), from the upper bound (3.20) we obtain:

$$\|\pi_{\alpha,\lambda} - \pi_0\|_{\mathrm{TV}} \le (1 - (1-\alpha)^t) + \|\lambda P^t - \pi_0\|_{\mathrm{TV}}. \tag{3.104}$$

Using (3.103) and $\alpha t \to 0$ we conclude the proof. □

Using the approximation (3.102), claims (3.7) and (3.8) of Theorem 3.1, for arbitrary widespread λ, are a consequence of Corollary 3.3, Corollary 3.4 and Theorem 1.7.

Scenario 3, strongly localized λ's

When $\alpha T_{\mathrm{ENT}} \to +\infty$, $t = s/\alpha$ for some fixed $s \in (0, \infty)$, Proposition 3.1 shows that

$$\inf_{x\in[n]} \left\|P^t(x,\cdot) - \pi_{\alpha,\lambda}P^t\right\|_{\mathrm{TV}} \xrightarrow{\mathbb{P}} 1\,, \tag{3.105}$$

whenever $\lambda = \delta_z$ for some vertex z. Let us show that (3.105) actually holds for all strongly localized λ. From Proposition 3.1 we know that for every $\eta > 0$ and $z \in [n]$, w.h.p. for all x there is a set $B^\eta_{x,z}$ such that

$$P^t(x, B^\eta_{x,z}) \le \eta\,, \qquad \pi_{\alpha,\delta_z}P^t(B^\eta_{x,z}) \ge 1-\eta\,. \tag{3.106}$$

If $\lambda = \sum_{z\in F} a_z\delta_z$, then by Proposition 3.2 $\pi_{\alpha,\lambda} = \sum_{z\in F} a_z\pi_{\alpha,\delta_z}$ and therefore, taking $B^\eta_x = \cup_{z\in F}B^\eta_{x,z}$:

$$P^t(x, B^\eta_x) \le |F|\eta\,, \qquad \pi_{\alpha,\lambda}P^t(B^\eta_x) \ge \sum_{z\in F} a_z\pi_{\alpha,\delta_z}P^t(B^\eta_{x,z}) \ge 1-\eta\,. \tag{3.107}$$

It follows that for all $\eta > 0$ w.h.p.

$$\inf_{x\in[n]} \left\|P^t(x,\cdot) - \pi_{\alpha,\lambda}P^t\right\|_{\mathrm{TV}} \ge 1 - (|F|+1)\eta. \tag{3.108}$$

Since η is arbitrarily small, this implies (3.105).

Once we have (3.105), from Proposition 3.3 and the upper bound (3.28) we obtain:

$$\max_{x\in[n]} \left|\left\|P^t_{\alpha,\lambda}(x,\cdot) - \pi_{\alpha,\lambda}\right\|_{\mathrm{TV}} - (1-\alpha)^t\right| \xrightarrow{\mathbb{P}} 0\,. \tag{3.109}$$

Equivalently,

$$\max_{x\in[n]} \left|\mathcal{D}^x_{\alpha,\lambda}(s/\alpha) - e^{-s}\right| \xrightarrow{\mathbb{P}} 0\,. \tag{3.110}$$

Scenario 2, strongly localized λ's

Here $\alpha T_{\mathrm{ENT}} \to \gamma \in (0,\infty)$. We take $t = u/\alpha$, with fixed $u \in (0,\infty)$. We consider separately the case $u \in (\gamma, \infty)$ and the case $u \in (0,\gamma)$.

Suppose first $u \in (\gamma,\infty)$. By Proposition 3.3 and the triangle inequality

$$\begin{aligned}\mathcal{D}^x_{\alpha,\lambda}(u/\alpha) &\le \left\|P^t(x,\cdot) - \pi_{\alpha,\lambda}P^t\right\|_{\mathrm{TV}} \\ &\le \left\|P^t(x,\cdot) - \pi_0\right\|_{\mathrm{TV}} + \max_{y\in V}\left\|\pi_0 - P^t(y,\cdot)\right\|_{\mathrm{TV}}\,.\end{aligned} \tag{3.111}$$

If $u \in (\gamma, \infty)$, then for some $\varepsilon > 0$ we have $t \geq (1+\varepsilon)T_{\mathrm{ENT}}$. Therefore, by Theorem 1.7 it follows that

$$\max_{x\in[n]} \mathcal{D}^x_{\alpha,\lambda}(u/\alpha) \xrightarrow{\mathbb{P}} 0\,, \qquad u \in (\gamma, \infty). \tag{3.112}$$

On the other hand, suppose that $u \in (0, \gamma)$. Here we can apply Proposition 3.1 and the argument in (3.107) above to obtain

$$\inf_{x\in[n]} \left\| P^t(x,\cdot) - \pi_{\alpha,\lambda} P^t \right\|_{\mathrm{TV}} \xrightarrow{\mathbb{P}} 1\,. \tag{3.113}$$

Therefore, by Proposition 3.3 and the upper bound (3.28),

$$\max_{x\in[n]} \left| \left\| P^t_{\alpha,\lambda}(x,\cdot) - \pi_{\alpha,\lambda} \right\|_{\mathrm{TV}} - (1-\alpha)^t \right| \xrightarrow{\mathbb{P}} 0\,. \tag{3.114}$$

Equivalently,

$$\max_{x\in[n]} \left| \mathcal{D}^x_{\alpha,\lambda}(u/\alpha) - e^{-u} \right| \xrightarrow{\mathbb{P}} 0\,, \qquad u \in (0, \gamma). \tag{3.115}$$

Combining (3.111) and (3.115), we have proved (3.7) for all strongly localized λ.

CHAPTER 4

MIXING TIMES ON REGENERATING DYNAMIC DIGRAPHS

In this chapter we obtain a precise description of the mixing times for the random walk on dynamic digraphs undergoing a particularly simple evolution, namely for digraphs with given degree sequences that are fully regenerated at independent geometrically distributed random time intervals.

We shall consider both families of directed graphs: DCM and OCM.

In what follows $\sigma \in \mathfrak{C}$ denotes a given realisation of either the directed configuration model DCM($\mathbf{d}^{\pm}$) or the out-configuration model OCM($\mathbf{d}^{+}$), and we write $G(\sigma)$ for the corresponding realisation of the digraph. We will treat both models on an equal footing as much as possible, and when we need to distinguish between them we often refer to these as *model 1* and *model 2* respectively. As usual, for a fixed configuration $\sigma \in \mathfrak{C}$, we consider the transition matrix

$$P_\sigma(x,y) = \frac{\#(\sigma; x \to y)}{d_x^+}, \qquad x, y \in [n], \tag{4.1}$$

where $\#(\sigma; x \to y)$ denotes the number of directed edges from x to y in $G(\sigma)$. We use the notation $\mathbf{P}_x^\sigma(\cdot)$ for the law of the trajectory $(X_0, X_1, \dots)$ when $X_0 = x$, so that in particular, for any $x, y \in [n]$, and $t \in \mathbb{N}$:

$$\mathbf{P}_x^\sigma(X_t = y) = P_\sigma^t(x, y). \tag{4.2}$$

We now introduce the joint evolution of the digraph and the random walk. Given $\alpha \in (0,1)$, we consider the Markov chain with state space $\mathfrak{C} \times [n]$ and with transition matrix

$$\mathcal{P}_\alpha((\sigma, x), (\eta, y)) = (1-\alpha) P_\sigma(x, y) \mathbf{1}_\sigma(\eta) + \alpha\, \mathbf{u}(\eta) \mathbf{1}_x(y), \tag{4.3}$$

where $\mathbf{1}_a(b)$ stands for 1 if $a = b$ and 0 otherwise and $\mathbf{u}(\eta) = |\mathfrak{C}|^{-1}$ denotes the uniform distribution over the set $\mathfrak{C}$. In words, at each time $t \in \mathbb{N}$ independently, we sample a Bernoulli(α) random variable J_t; if $J_t = 1$ we pick a uniformly random $\eta \in \mathfrak{C}$ and move from the current state (σ, x) to the new state (η, x), while if $J_t = 0$ we move to the new state (σ, y) where y is chosen uniformly at random among the out-neighbours of x in the digraph $G(\sigma)$. We write (ξ_t, X_t) for the trajectory of the Markov chain and write $\mathbf{P}^{\mathbf{J}}_{\sigma,x}(\cdot)$ for its law when started at $\xi_0 = \sigma$ and $X_0 = x$. It is not hard to check that $(\xi_t, X_t)_{t\geq 0}$ is an irreducible and aperiodic Markov chain and therefore it admits a unique stationary distribution π^{J}. A consequence of our results, see Remark 4.1 below, is that π^{J} is well approximated in total variation distance by the probability measure ν on $\mathfrak{C} \times [n]$ defined by

$$\nu(\sigma, x) = \mathbf{u}(\sigma) \pi_\sigma(x). \tag{4.4}$$

We know that π_σ is uniquely defined for all σ in a set $\Omega_n \subset \mathfrak{C}$ with $\mathbf{u}(\Omega_n) \to 1$ as $n \to \infty$. To extend ν to all $\mathfrak{C} \times [n]$ we may define e.g. $\pi_\sigma = \mu_{\mathrm{in}}$ for $\sigma \in \mathfrak{C} \setminus \Omega_n$. We define

$$\mathcal{D}^{\mathrm{J},\alpha}_{\sigma,x}(t) = \|\mathbf{P}^{\mathbf{J}}_{\sigma,x}(\xi_t = \cdot, X_t = \cdot) - \nu\|_{\mathrm{TV}}. \tag{4.5}$$

For each $t \in \mathbb{N}$, the quantity $\mathcal{D}^{\mathrm{J},\alpha}_{\sigma,x}(t)$ is regarded as a random variable with respect to the uniform choice of the configuration $\sigma \in \mathfrak{C}$. Moreover, we extend $\mathcal{D}^{\mathrm{J},\alpha}_{\sigma,x}(\cdot)$ to all positive reals by taking the integer part of the argument.

Theorem 4.1 *Fix a sequence $\alpha = \alpha_n$ such that $\alpha_n \to 0$ as $n \to \infty$. Then, for all $\beta > 0$*

$$\limsup_{n\to\infty} \max_{\sigma \in \mathfrak{C}, x \in [n]} \mathcal{D}^{\mathrm{J},\alpha}_{\sigma,x}(\beta\alpha^{-1}) \leq (1+\beta) e^{-\beta}. \tag{4.6}$$

Next, assume that

$$\gamma = \lim_{n\to\infty} \alpha\, T_{\mathrm{ENT}} \in [0, \infty]. \tag{4.7}$$

Then, according to the value of γ there are three scenarios:

1. *If $\gamma = 0$ then for all $\beta > 0$:*

$$\max_{x\in[n]} \left| \mathcal{D}^{\mathrm{J},\alpha}_{\sigma,x}(\beta\alpha^{-1}) - e^{-\beta} \right| \xrightarrow{\mathbb{P}} 0. \tag{4.8}$$

2. *If $\gamma = \infty$ then for all $\beta > 0$:*

$$\max_{x \in [n]} \left| \mathcal{D}^{\mathrm{J},\alpha}_{\sigma,x}(\beta\alpha^{-1}) - (1+\beta)e^{-\beta} \right| \xrightarrow{\mathbb{P}} 0. \tag{4.9}$$

3. *If $\gamma \in (0,\infty)$ then for all $\beta > 0$, $\beta \neq \gamma$:*

$$\max_{x \in [n]} \left| \mathcal{D}^{\mathrm{J},\alpha}_{\sigma,x}(\beta\alpha^{-1}) - \psi_\gamma(\beta) \right| \xrightarrow{\mathbb{P}} 0. \tag{4.10}$$

where

$$\psi_\gamma(\beta) = \begin{cases} (1+\beta)e^{-\beta} & \text{if } \beta < \gamma \\ e^{-\beta} & \text{if } \beta > \gamma \end{cases}.$$

The trichotomy displayed in Theorem 4.1 can be interpreted as follows; see also Fig. 1.2. On the time scale α^{-1} the regeneration times, that is the $t \in \mathbb{N}$ such that $J_t = 1$, converge to a Poisson process of intensity 1. Then $e^{-\beta}$ and $(1+\beta)e^{-\beta}$ represent the probability of having no regeneration and at most one regeneration up to time $\beta\alpha^{-1}$ respectively. Thus Theorem 4.1 essentially says that when the walk is far from being mixed within the current digraph then two regenerations are necessary and sufficient for a complete loss of memory of the initial state, whereas if the walk has already mixed within the current digraph then all it is required to reach stationarity is one regeneration.

Remark 4.1 *From Theorem 4.1 it follows that*

$$\lim_{n\to\infty} \|\nu - \pi^{\mathrm{J}}\|_{TV} = 0, \tag{4.11}$$

which in turn implies that all statements in Theorem 4.1 holds with ν replaced by π^{J}. Indeed, to prove (4.11) observe that for any $t \in \mathbb{N}$

$$\|\nu - \pi^{\mathrm{J}}\|_{TV} \le \max_{\sigma \in \mathfrak{C}, x \in [n]} \mathcal{D}^{\mathrm{J},\alpha}_{\sigma,x}(t).$$

Taking $t = \beta\alpha^{-1}$, (4.6) implies that $\limsup_{n\to\infty} \|\nu - \pi^{\mathrm{J}}\|_{TV} \le (1+\beta)e^{-\beta}$, and letting $\beta \to \infty$ we obtain (4.11).

The proof of Theorem 4.1 will be crucially based on Theorem 1.7 and Lemma 3.1. The dynamic setting however requires an important extension of these results that can be formulated as follows. For any $(\sigma, \eta) \in \mathfrak{C} \times \mathfrak{C}$ and integers $0 \le s \le t$, define

$$Q^{s,t}_{\sigma,\eta}(x,y) = \sum_{z \in [n]} P^s_\sigma(x,z) P^{t-s}_\eta(z,y). \tag{4.12}$$

Theorem 4.2 (Cutoff on double digraphs) *Fix $\beta > 0$ and set $t = \beta T_{\mathrm{ENT}}$ and $0 \le s \le t$. Let σ and η be two independent uniformly random configurations in $\mathfrak{C}$, and let $\mathbb{P}$ denote the associated probability. Then for fixed $\beta > 0$:*

1. *If $\beta < 1$:*
$$\min_{x\in[n]} \|Q^{s,t}_{\sigma,\eta}(x,\cdot) - \pi_\eta\|_{TV} \xrightarrow{\mathbb{P}} 1.$$

2. *If $\beta > 1$ and $t - s \to \infty$ as $n \to \infty$:*
$$\max_{x\in[n]} \|Q^{s,t}_{\sigma,\eta}(x,\cdot) - \pi_\eta\|_{TV} \xrightarrow{\mathbb{P}} 0.$$

Another key ingredient for the proof of Theorem 4.1 is the control of the annealed walk. By this we mean the law
$$\mathbf{P}^{\mathbf{an}}_x(\cdot) = \sum_{\eta\in\mathfrak{C}} \mathbf{u}(\eta)\mathbf{P}^{\eta}_x(\cdot), \tag{4.13}$$
where $\mathbf{P}^{\eta}_x$ is defined before (4.2).

Lemma 4.1 *The annealed law satisfies*
$$\lim_{n\to\infty} \sup_{x\in[n],t\ge 1} \|\mathbf{P}^{\mathbf{an}}_x(X_t = \cdot) - \mu_{\mathrm{in}}\|_{TV} = 0. \tag{4.14}$$

Once Lemma 4.1 and Theorem 4.2 are available, we shall obtain Theorem 4.1 by a decomposition of the law at time t according to the location of the regeneration times; see Section 4.1.

Finally, our last main result concerns the marginal distribution of the position of the walk, namely the non-Markovian process obtained by projecting the chain $(\xi_t, X_t)_{t\ge 0}$ on the second coordinate. According to Theorem 4.1 and Lemma 4.1 the law of X_t, for t and n suitably large, should be well approximated by μ_{in}. The next result quantifies this statement by exhibiting once again a trichotomy. Define
$$\mathcal{D}^{\alpha}_{\sigma,x}(t) := \|\mathbf{P}^{\mathbf{J}}_{\sigma,x}(X_t = \cdot) - \mu_{\mathrm{in}}\|_{\mathrm{TV}}\,, \qquad q := \mathbb{E}\|\pi_\sigma - \mu_{\mathrm{in}}\|_{\mathrm{TV}}. \tag{4.15}$$

We remark that if the sequences $\mathbf{d}^{\pm}$ are *eulerian*, that is $d^+_x = d^-_x$ for all $x \in [n]$, then $\pi_\sigma = \mu_{\mathrm{in}}$ is stationary for all $\sigma \in \mathfrak{C}$. Thus in this case $q = 0$. On the other hand, results from [42, 44] imply that if the sequence is not eulerian then q is bounded away from 0 and 1; see [13, Theorem 4] and Remark 2.1 for more details.

Theorem 4.3 *Fix a sequence $\alpha = \alpha_n$ such that $\alpha_n \to 0$ as $n \to \infty$. Then, for all $\beta > 0$*

$$\limsup_{n\to\infty} \max_{\sigma\in\mathfrak{C}, x\in[n]} \mathcal{D}^{\alpha}_{\sigma,x}(\beta\alpha^{-1}) \le e^{-\beta}. \tag{4.16}$$

Next, assume that

$$\gamma = \lim_{n\to\infty} \alpha\, T_{\mathrm{ENT}} \in [0,\infty]. \tag{4.17}$$

Then, according to the value of γ there are three scenarios:

1. *If $\gamma = 0$ then for all $\beta > 0$:*

$$\max_{x\in[n]} \left|\mathcal{D}^{\alpha}_{\sigma,x}(\beta\alpha^{-1}) - q\, e^{-\beta}\right| \xrightarrow{\mathbb{P}} 0. \tag{4.18}$$

Moreover, if $\beta \neq 1$ then

$$\max_{x\in[n]} \left|\mathcal{D}^{\alpha}_{\sigma,x}(\beta T_{\mathrm{ENT}}) - \varphi(\beta)\right| \xrightarrow{\mathbb{P}} 0, \tag{4.19}$$

where

$$\varphi(\beta) := \begin{cases} 1 & \text{if } \beta < 1 \\ q & \text{if } \beta > 1. \end{cases} \tag{4.20}$$

2. *If $\gamma = \infty$, then for all $\beta > 0$:*

$$\max_{x\in[n]} \left|\mathcal{D}^{\alpha}_{\sigma,x}(\beta\alpha^{-1}) - e^{-\beta}\right| \xrightarrow{\mathbb{P}} 0. \tag{4.21}$$

3. *If $\gamma(0,\infty)$ then for all $\beta > 0$, $\beta \neq \gamma$:*

$$\max_{x\in[n]} \left|\mathcal{D}^{\alpha}_{\sigma,x}(\beta\alpha^{-1}) - \varphi(\beta/\gamma)e^{-\beta}\right| \xrightarrow{\mathbb{P}} 0. \tag{4.22}$$

The above results can be roughly interpreted as follows. If we follow only the position of the particle then after the first regeneration time the walk has the annealed law, and by Lemma 4.1 this is given by μ_{in}. Thus, a complete loss of memory of the initial state with relaxation to the limiting state μ_{in} occurs essentially at the time of the first regeneration of the digraph. On the other hand, if no regeneration occurs then a partial loss of memory occurs at time T_{ENT} because of the static mixing cutoff phenomenon, and this is quantified by the drop by a factor q in total variation. The competition between these two effects explains the above triad; see Fig. 1.3.

4.1 Trichotomy for the joint process

We start with the proof of Lemma 4.1, and then prove Theorem 4.1 assuming the validity of Theorem 4.2. The proof of the latter is given in Section 4.3 below.

Proof of Lemma 4.1

We divide the proof in two cases: $t \le 2T_{\mathrm{ENT}}$, and $t > 2T_{\mathrm{ENT}}$. If $t \le 2T_{\mathrm{ENT}}$, in particular one has $t = O(\log n)$, and we will see in Lemma 5.8 that

$$\mathbf{P}^{\mathbf{an}}_x(X_t = y) = \mu_{\mathrm{in}}(y)(1+o(1)), \qquad \mathbf{P}^{\mathbf{an}}_x(X_t = x) = O\left(n^{-1}\log(n)\right), \tag{4.23}$$

for $t = O(\log n)$, uniformly in $x, y \in [n]$. The proof of (4.23) is carried out in detail in Lemma 5.8 for the DCM only, but the very same arguments imply the validity of the statements for the OCM as well. The estimates in (4.23) are enough to conclude that uniformly in $x \in [n]$:

$$\|\mathbf{P}^{\mathbf{an}}_x(X_t = \cdot) - \mu_{\mathrm{in}}\|_{\mathrm{TV}} = \frac{1}{2}\sum_{y\in[n]} |\mathbf{P}^{\mathbf{an}}_x(X_t = y) - \mu_{\mathrm{in}}(y)| = o(1)\,, \qquad t = O(\log n). \tag{4.24}$$

We now turn to the case $t > 2T_{\mathrm{ENT}}$. By the triangle inequality we have

$$\|\mathbf{P}^{\mathbf{an}}_x(X_t = \cdot) - \mu_{\mathrm{in}}\|_{\mathrm{TV}} \le \mathbb{E}\|P^t_\sigma(x,\cdot) - \pi_\sigma\|_{\mathrm{TV}} + \|\mathbb{E}\pi_\sigma - \mu_{\mathrm{in}}\|_{\mathrm{TV}}. \tag{4.25}$$

Concerning the first term on the right hand side, we use Theorem 1.7 to obtain

$$\mathbb{E}\|P^t_\sigma(x,\cdot) - \pi_\sigma\|_{\mathrm{TV}} = o(1)\,, \tag{4.26}$$

uniformly in $x \in [n]$, and $t > 2T_{\mathrm{ENT}}$. The second term on the right hand side of (4.25) can be bounded by a combination of the arguments in (4.24) and (4.26). Indeed, using again the triangle inequality and setting $s = \lfloor 2T_{\mathrm{ENT}} \rfloor$:

$$\|\mathbb{E}\pi_\sigma - \mu_{\mathrm{in}}\|_{\mathrm{TV}} \le \mathbb{E}\|P^s_\sigma(x,\cdot) - \pi_\sigma\|_{\mathrm{TV}} + \|\mathbf{P}^{\mathbf{an}}_x(X_s = \cdot) - \mu_{\mathrm{in}}\|_{\mathrm{TV}} = o(1). \tag{4.27}$$

This ends the proof of Lemma 4.1. □

Proof of Theorem 4.1

For every $(\eta, y) \in \mathfrak{C} \times [n]$ define

$$\mu_t^{\sigma,x}(\eta, y) = \mathbf{P}^{\mathbf{J}}_{\sigma,x}(\xi_t = \eta, X_t = y).$$

Recall that J_s, $s \in \mathbb{N}$ are i.i.d. Bernoulli(α) random variables indicating the occurrence of the regeneration event. For each $t \geq 1$, consider the random variable $\tau = \tau(t)$ defined by

$$\tau = \mathbf{1}\left(\exists s \in \{1, \dots, t\} :\ J_s = 1\right) \sup\{s \leq t \mid J_s = 1\}. \tag{4.28}$$

We may write

$$\begin{aligned}
\mu_t^{\sigma,x}(\eta, y) &= \sum_{s=0}^{t} \mathbf{P}^{\mathbf{J}}(\tau = s) \mathbf{P}^{\mathbf{J}}_{\sigma,x}\left((\xi_t, X_t) = (\eta, y) \mid \tau = s\right) \\
&= (1-\alpha)^t \mathbf{1}_\sigma(\eta) P_\sigma^t(x, y) + \sum_{s=1}^{t} \alpha(1-\alpha)^{t-s} \sum_{z\in[n]} \sum_{\xi\in\mathfrak{C}} \mathbf{u}(\eta) \mu_{s-1}^{\sigma,x}(\xi, z) P_\eta^{t-s}(z, y).
\end{aligned}$$

Since $\mu_{s-1}^{\sigma,x}(\xi, z)$ admits the same decomposition we obtain the expansion:

$$\mu_t^{\sigma,x}(\eta, y) = A_t^{\sigma,x}(\eta, y) + B_t^{\sigma,x}(\eta, y) + C_t^{\sigma,x}(\eta, y),$$

where

$$A_t^{\sigma,x}(\eta, y) = (1-\alpha)^t \mathbf{1}_\sigma(\eta) P_\sigma^t(x, y),$$

$$B_t^{\sigma,x}(\eta, y) = \alpha(1-\alpha)^{t-1} \sum_{s=1}^{t} \sum_{z\in[n]} \mathbf{u}(\eta) P_\sigma^{s-1}(x, z) P_\eta^{t-s}(z, y),$$

$$C_t^{\sigma,x}(\eta, y) = \sum_{s=1}^{t} \sum_{r=1}^{s-1} \alpha^2 (1-\alpha)^{t-1-r} \sum_{v,z\in[n]} \sum_{\xi,\omega\in\mathfrak{C}} \mathbf{u}(\eta)\mathbf{u}(\xi) \mu_{r-1}^{\sigma,x}(\omega, v) P_\xi^{s-1-r}(v, z) P_\eta^{t-s}(z, y),$$

Notice that for every choice of $W = \nu, B_t^{\sigma,x}, C_t^{\sigma,x}$, for any fixed choice of $\sigma \in \mathfrak{C}$ one has

$$\sum_{\eta:\,\eta\neq\sigma} \sum_{y\in[n]} W(\eta, y) = \sum_{\eta\in\mathfrak{C}} \sum_{y\in[n]} W(\eta, y) + O\left(|\mathfrak{C}|^{-1}\right).$$

Therefore,

$$
\begin{aligned}
2\|\mu_t^{\sigma,x}-\nu\|_{\mathrm{TV}} &= \sum_{\eta\in\mathfrak{C}}\sum_{y\in[n]} |\mu_t^{\sigma,x}(\eta,y)-\nu(\eta,y)| \\
&= \sum_{y\in[n]} |\mu_t^{\sigma,x}(\sigma,y)-\nu(\sigma,y)| + \sum_{\eta\neq\sigma}\sum_{y\in[n]} |\mu_t^{\sigma,x}(\eta,y)-\nu(\eta,y)| \\
&= (1-\alpha)^t + \sum_{\eta\in\mathfrak{C}}\sum_{y\in[n]} |B_t^{\sigma,x}(\eta,y)+C_t^{\sigma,x}(\eta,y)-\mathbf{u}(\eta)\pi_\eta(y)| + o(1). \qquad (4.29)
\end{aligned}
$$

We may rewrite $C_t^{\sigma,x}(\eta,y)=\chi\mathbf{u}(\eta)\widehat{C}_t^{\sigma,x}(\eta,y)$, where

$$
\chi = 1-(1-\alpha)^t-\alpha t(1-\alpha)^{t-1} = \alpha^2\sum_{s=1}^{t}\sum_{r=1}^{s-1}(1-\alpha)^{t-r-1},
$$

$$
\widehat{C}_t^{\sigma,x}(\eta,y) = \frac{1}{\chi}\alpha^2\sum_{s=1}^{t}\sum_{r=1}^{s-1}(1-\alpha)^{t-r-1}\sum_{z\in[n]}\sum_{v\in[n]}\mu_{r-1}^{\sigma,x}(v)\mathbf{P}_v^{\mathbf{an}}\left(X_{s-1-r}=z\right)P_\eta^{t-s}(z,y),
$$

and we use the notation $\mu_{r-1}^{\sigma,x}(v):=\sum_{\omega\in\mathfrak{C}}\mu_{r-1}^{\sigma,x}(\omega,v)$. Notice that $\widehat{C}_t^{\sigma,x}(\eta,\cdot)$ is a probability on $[n]$. Define also the probability λ_η by

$$
\lambda_\eta(y) = \frac{1}{\chi}\alpha^2\sum_{s=1}^{t}\sum_{r=1}^{s-1}(1-\alpha)^{t-r-1}\mu_{\mathrm{in}}P_\eta^{t-s}(y).
$$

Lemma 4.1 implies that uniformly in $\eta\in\mathfrak{C}$:

$$
\|\widehat{C}_t^{\sigma,x}(\eta,\cdot)-\pi_\eta\|_{\mathrm{TV}} = \|\lambda_\eta-\pi_\eta\|_{\mathrm{TV}} + o(1). \qquad (4.30)
$$

Moreover, Lemma 3.1 implies that whenever $t-s\to\infty$:

$$
\sum_\eta \mathbf{u}(\eta)\|\mu_{\mathrm{in}}P_\eta^{t-s}-\pi_\eta\|_{\mathrm{TV}} = o(1). \qquad (4.31)
$$

Since $\alpha\to 0$, (4.31) implies

$$
\sum_\eta \mathbf{u}(\eta)\|\lambda_\eta-\pi_\eta\|_{\mathrm{TV}} = o(1). \qquad (4.32)
$$

Inserting (4.30) and (4.32) in (4.29) we obtain

$$
2\|\mu_t^{\sigma,x}-\nu\|_{\mathrm{TV}} = (1-\alpha)^t + \sum_{\eta\in\mathfrak{C}}\sum_{y\in[n]}|B_t^{\sigma,x}(\eta,y)+(1-\chi)\mathbf{u}(\eta)\pi_\eta(y)| + o(1). \qquad (4.33)
$$

Let us now take $t = \beta\alpha^{-1}$, for some fixed constant $\beta > 0$. Since $\alpha \to 0$ we have $1 - \chi \to e^{-\beta}(1+\beta)$ and

$$2\|\mu_t^{\sigma,x} - \nu\|_{\text{TV}} = e^{-\beta} + \sum_{\eta\in\mathfrak{C}} \mathbf{u}(\eta)\psi_t(\eta) + o(1), \tag{4.34}$$

where we define

$$\psi_t(\eta) = \sum_{y\in[n]} \left|\beta e^{-\beta}\widehat{B}_t^{\sigma,x}(\eta,y) - e^{-\beta}(1+\beta)\pi_\eta(y)\right|, \tag{4.35}$$

with $\widehat{B}_t^{\sigma,x}(\eta,\cdot)$ the probability on $[n]$ defined by

$$\widehat{B}_t^{\sigma,x}(\eta,y) = \frac{1}{t}\sum_{s=1}^{t}\sum_{z\in[n]} P_\sigma^{s-1}(x,z)P_\eta^{t-s}(z,y).$$

We start by noting that

$$\psi_t(\eta) \le 2\beta e^{-\beta}\|\widehat{B}_t^{\sigma,x}(\eta,\cdot) - \pi_\eta\|_{\text{TV}} + e^{-\beta}.$$

In particular, (4.34) shows that uniformly in $(\sigma,x) \in \mathfrak{C}\times[n]$

$$\|\mu_t^{\sigma,x} - \nu\|_{\text{TV}} \le (1+\beta)e^{-\beta} + o(1),$$

which proves (4.6). At this point we split the analysis in four cases.

Scenario 1: $\alpha T_{\text{ENT}} \to \gamma = \infty$

In this case we notice that $t = o(\log n)$. Therefore, for every $\sigma \in \mathfrak{C}$, $x \in [n]$ there must exist a set $\mathcal{I} \subset [n]$ such that for all $s \le t$:

$$P_\sigma^{s-1}(x,\mathcal{I}) = 1, \qquad |\mathcal{I}| \le \Delta^t = n^{o(1)}.$$

Moreover, for every $\eta \in \mathfrak{C}$ and for every $z \in [n]$ there exists a set $\mathcal{J}_z \in [n]$ such that for every $s \le t$

$$P_\eta^{t-s}(z,\mathcal{J}_z) = 1, \qquad |\mathcal{J}_z| = n^{o(1)}.$$

Therefore, setting $\mathcal{J} = \cup_{z\in\mathcal{I}}\mathcal{J}_z$,

$$|\mathcal{J}| = n^{o(1)}, \qquad \widehat{B}_t^{\sigma,x}(\eta,\mathcal{J}) = 1 - o(1).$$

Moreover, w.h.p. with respect to η one has $\pi_\eta(\mathcal{J}) = o(1)$. Indeed, we know that for some constant $C > 0$, $\sum_{x\in[n]} \pi(x)^2 \le Cn^{-1}$ by Lemma 4.3 below, and for any $U \subset [n]$, Cauchy-Schwarz implies

$$\pi_\eta(U)^2 \le |U| \sum_{x\in[n]} \pi(x)^2 \le C|U|n^{-1}. \tag{4.36}$$

It follows that w.h.p.

$$\psi_t(\eta) = 2\beta e^{-\beta} + e^{-\beta} + o(1).$$

In conclusion, (4.34) implies

$$\|\mu_t^{\sigma,x} - \nu\|_{\mathrm{TV}} = (1+\beta)e^{-\beta} + o(1),$$

which proves (4.9). Note that because of the uniform average over $\eta \in \mathfrak{C}$ the convergence in (4.9) actually holds uniformly in $\sigma \in \mathfrak{C}$ rather than in $\mathbb{P}$-probability as stated.

Scenario 2: $\alpha T_{\mathrm{ENT}} \to \gamma = 0$.

In this case it possible to find a sequence $\upsilon = \upsilon(n) = o(1)$ that vanishes sufficiently slowly that

$$\upsilon t = \upsilon\beta\alpha^{-1} = \omega(T_{\mathrm{ENT}}).$$

If $\widehat{E}_t^{\sigma,x}(\eta,\cdot)$ denotes the probability on $[n]$

$$\widehat{E}_t^{\sigma,x}(\eta,y) = \frac{1}{(1-2\upsilon)t} \sum_{s=\upsilon t}^{(1-\upsilon)t} \sum_{z\in[n]} P_\sigma^{s-1}(x,z) P_\eta^{t-s}(z,y),$$

then

$$\|\widehat{B}_t^{\sigma,x}(\eta,\cdot) - \widehat{E}_t^{\sigma,x}(\eta,\cdot)\|_{\mathrm{TV}} = O(\upsilon).$$

Let us write

$$\sum_{z\in[n]} P_\sigma^{s-1}(x,z) P_\eta^{t-s}(z,\cdot) =: \lambda P_\eta^{t-s}(\cdot),$$

and notice that

$$\|\lambda P_\eta^{t-s} - \pi_\eta\|_{\mathrm{TV}} \le \max_{x\in[n]} \|P_\eta^{t-s}(x,\cdot) - \pi_\eta\|_{\mathrm{TV}}.$$

Since $t - s = \omega(T_{\mathrm{ENT}})$, from Theorem 1.7 we conclude that w.h.p. with respect to η:

$$\|\widehat{E}_t^{\sigma,x}(\eta,\cdot) - \pi_\eta\|_{\mathrm{TV}} = o(1).$$

Therefore, w.h.p.

$$\|\widehat{B}_t^{\sigma,x}(\eta,\cdot) - \pi_\eta\|_{\mathrm{TV}} = o(1).$$

By the triangular inequality and (4.34),

$$\|\mu_t^{\sigma,x} - \nu\|_{\mathrm{TV}} = e^{-\beta} + o(1).$$

This proves (4.8). As in the previous case, it is worth noting that the convergence in (4.8) actually holds uniformly in $\sigma \in \mathfrak{C}$ rather than in $\mathbb{P}$-probability.

Scenario 3i: $\alpha T_{\mathrm{ENT}} \to \gamma \in (0,\infty)$ **and** $\beta < \gamma$

We want to control $\psi_t(\eta)$ as defined in (4.35). If $\beta < \gamma$ then $t = (1-\epsilon)T_{\mathrm{ENT}}$ for some $\epsilon \in (0,1)$. We argue that w.h.p. with respect to the independent pair (σ,η),

$$\|\widehat{B}_t^{\sigma,x}(\eta,\cdot) - \pi_\eta\|_{\mathrm{TV}} = 1 - o(1). \tag{4.37}$$

Call $Y_{s,t}$ the set of $y \in [n]$ such that

$$\sum_{z\in[n]} P_\sigma^{s-1}(x,z)P_\eta^{t-s}(z,y) \geq e^{-(1+\epsilon)Ht}$$

Summing over $y \in Y_{s,t}$, we must have

$$|Y_{s,t}| \leq e^{(1+\varepsilon)Ht} = n^{1-\epsilon^2}$$

Lemma 4.5 below implies in particular that

$$\sum_{y\in Y_{s,t}} \sum_{z\in[n]} P_\sigma^{s-1}(x,z)P_\eta^{t-s}(z,y) = 1 - o(1).$$

Setting $Y = \cup_{s=1}^t Y_{s,t}$ and noticing that $|Y| \leq T_{\mathrm{ENT}}\, n^{1-\epsilon^2} = o(n)$, we have

$$\sum_{y\in Y} \widehat{B}_t^{\sigma,x}(\eta,y) \geq \frac{1}{t}\sum_{s=1}^{t} \sum_{y\in Y_{s,t}} \sum_{z\in[n]} P_\sigma^{s-1}(x,z)P_\eta^{t-s}(z,y) = 1 - o(1).$$

Since $|Y| = o(n)$, $\widehat{B}_t^{\sigma,x}(\eta,\cdot)$ is w.h.p. asymptotically singular with respect to π_η; see the argument in (4.36). This proves (4.37). Inserting this in (4.34)-(4.35), it follows that w.h.p. with respect to $\sigma \in \mathfrak{C}$:

$$\|\mu_t^{\sigma,x} - \nu\|_{\mathrm{TV}} = (1+\beta)e^{-\beta} + o(1).$$

Scenario 3ii: $\alpha T_{\mathrm{ENT}} \to \gamma \in (0,\infty)$ **and** $\beta > \gamma$**.**

By definition there must exist some $\epsilon > 0$ such that

$$t = \beta\alpha^{-1} = \frac{\beta}{H\gamma}\log(n) > (1+\epsilon)T_{\mathrm{ENT}}.$$

For every $\upsilon \in (0, \epsilon/2)$, at the price of an additive error $O(\upsilon)$ in total variation, we can replace $\widehat{B}_t^{\sigma,x}(\eta,\cdot)$ by the probability $\widehat{B}_1(\cdot)$ defined as

$$\widehat{B}_1(y) = \frac{1}{(1-2\upsilon)t}\sum_{s=\upsilon t}^{(1-\upsilon)t}\sum_{z\in[n]} P_\sigma^{s-1}(x,z)P_\eta^{t-s}(z,y).$$

Since $t > (1+\epsilon)T_{\mathrm{ENT}}$ and $t - s \to \infty$, we can use Theorem 4.2 to obtain that w.h.p. with respect to the independent pair (σ,η),

$$\|\widehat{B}_1 - \pi_\eta\|_{\mathrm{TV}} = o(1) \tag{4.38}$$

From (4.35),

$$\psi_t(\eta) = e^{-\beta}\sum_{y\in[n]}|\beta\widehat{B}_1(y) - (1+\beta)\pi_\eta(y)| + O(\upsilon) = e^{-\beta} + O(\upsilon) + o(1).$$

Since υ is arbitrarily small, from (4.34) we obtain that w.h.p. with respect to $\sigma \in \mathfrak{C}$:

$$\|\mu_t^{\sigma,x} - \nu\|_{\mathrm{TV}} = e^{-\beta} + o(1).$$

□

4.2 Trichotomy for the random walk

Here we prove Theorem 4.3. The main observation can be stated as follows.

Proposition 4.1 *Let $\tau = \tau(t)$ denote the random variable in (4.28). Then, uniformly in $t \geq 2$:*

$$\lim_{n\to\infty}\max_{\sigma,x}\|\mathbf{P}^{\mathbf{J}}_{\sigma,x}(X_t = \cdot \mid 1 \leq \tau < t) - \mu_{\mathrm{in}}\|_{TV} = 0$$

Proof: Observe that

$$\mathbf{P}^{\mathbf{J}}_{\sigma,x}\left(\tau \in \{0,t\}\right) = 1 - \mathbf{P}^{\mathbf{J}}_{\sigma,x}\left(1 \le \tau < t\right) = (1-\alpha)^t + \alpha(1-\alpha)^{t-1} = (1-\alpha)^{t-1}. \tag{4.39}$$

Moreover, if $1 \le s < t$,

$$\mathbf{P}^{\mathbf{J}}_{\sigma,x}\left(X_t = y;\ \tau = s\right) = \alpha(1-\alpha)^{t-s} \sum_{\eta \in \mathfrak{C}} \mathbf{u}(\eta)\lambda P^{t-s}_{\eta}(y)$$

where λ is the probability measure

$$\lambda(z) = \mathbf{P}^{\mathbf{J}}_{\sigma,x}(X_s = z \mid \tau = s).$$

We then compute the conditional probability

$$\begin{aligned}
\mathbf{P}^{\mathbf{J}}_{\sigma,x}\left(X_t = y \mid 1 \le \tau < t\right) =& \frac{1}{1-(1-\alpha)^{t-1}} \sum_{s=1}^{t-1} \mathbf{P}^{\mathbf{J}}_{\sigma,x}\left(X_t = y;\ \tau = s\right) &(4.40)\\
=& \frac{1}{1-(1-\alpha)^{t-1}} \sum_{s=1}^{t-1} \alpha(1-\alpha)^{t-s} \sum_{\eta \in \mathfrak{C}} \mathbf{u}(\eta)\lambda P^{t-s}_{\eta}(y). &(4.41)
\end{aligned}$$

Now we can rely on the uniform bound of Lemma 4.1 to conclude

$$\begin{aligned}
&\|\mathbf{P}^{\mathbf{J}}_{\sigma,x}\left(X_t = \cdot \mid 1 \le \tau < t\right) - \mu_{\mathrm{in}}\|_{\mathrm{TV}} \\
&\quad \le \sum_{s=1}^{t-1} \frac{\alpha(1-\alpha)^{t-s}}{1-(1-\alpha)^{t-1}} \max_{x \in [n]} \|\mathbf{P}^{\mathbf{an}}_{x}(X_{t-s} = \cdot) - \mu_{\mathrm{in}}\|_{\mathrm{TV}} = o(1).
\end{aligned}$$

□

Corollary 4.1 *Uniformly in $t \ge 1$:*

$$\|\mathbf{P}^{\mathbf{J}}_{\sigma,x}(X_t = \cdot) - \mu_{\mathrm{in}}\|_{TV} = (1-\alpha)^t \|P^t_{\sigma}(x,\cdot) - \mu_{\mathrm{in}}\|_{TV} + o(1).$$

Proof: Note that

$$\mathbf{P}^{\mathbf{J}}_{\sigma,x}(X_t = \cdot \mid \tau = 0) = P^t_{\sigma}(x,\cdot)\,, \qquad \mathbf{P}^{\mathbf{J}}_{\sigma,x}(X_t = \cdot \mid \tau = t) = P^{t-1}_{\sigma}(x,\cdot).$$

Using Proposition 4.1, and $\alpha \to 0$, the triangle inequality shows that

$$
\begin{aligned}
&\|\mathbf{P}^{\mathbf{J}}_{\sigma,x}(X_t = \cdot) - \mu_{\text{in}}\|_{\text{TV}} \\
&\quad = \|(1-\alpha)^{t-1}\mathbf{P}^{\mathbf{J}}_{\sigma,x}(X_t = \cdot \mid \tau \in \{0,t\}) + \mathbf{P}^{\mathbf{J}}_{\sigma,x}\,(1 \le \tau < t)\,\mathbf{P}^{\mathbf{J}}_{\sigma,x}(X_t = \cdot \mid 1 \le \tau < t) - \mu_{\text{in}}\|_{\text{TV}} \\
&\quad = (1-\alpha)^{t-1}\|\mathbf{P}^{\mathbf{J}}_{\sigma,x}(X_t = \cdot \mid \tau \in \{0,t\}) - \mu_{\text{in}}\|_{\text{TV}} + o(1) \\
&\quad = (1-\alpha)^{t-1}\|(1-\alpha)P^t_\sigma(x,\cdot) + \alpha P^{t-1}_\sigma(x,\cdot) - \mu_{\text{in}}\|_{\text{TV}} + o(1) \\
&\quad = (1-\alpha)^t\|P^t_\sigma(x,\cdot) - \mu_{\text{in}}\|_{\text{TV}} + o(1).
\end{aligned}
$$

□

Therefore, to prove Theorem 4.3 it is sufficient to prove the next lemma.

Lemma 4.2 *If $t = \beta T_{\text{ENT}}$ then for any fixed $\beta > 0$, $\beta \neq 1$:*

$$\max_{x\in[n]} \left| \|P^t_\sigma(x,\cdot) - \mu_{\text{in}}\|_{TV} - \varphi(\beta) \right| \xrightarrow{\mathbb{P}} 0, \tag{4.42}$$

where $\varphi(\beta) = 1$ if $\beta < 1$ and $\varphi(\beta) = q$ for $\beta > 1$, and q is defined in (4.15).

Proof: From Theorem 1.7 it is sufficient to show that if $t = \beta T_{\text{ENT}}$ with $\beta < 1$, then for any $\varepsilon > 0$ w.h.p.

$$\min_{x\in[n]} \left\|P^t_\eta(x,\cdot) - \mu_{\text{in}}\right\|_{\text{TV}} \ge 1 - \varepsilon, \tag{4.43}$$

and that

$$q - \|\pi_\eta - \mu_{\text{in}}\|_{\text{TV}} \xrightarrow{\mathbb{P}} 0. \tag{4.44}$$

The concentration (4.44) has been already proved in [13, Lemma 17] (see also the proof of Proposition 3.5). Concerning the estimate (4.43), we can use Lemma 4.5 below to show that if $t = \beta T_{\text{ENT}}$ with $\beta < 1$ then there exists a set $U_x \subset [n]$ with $|U_x| = o(n)$ such that w.h.p.

$$P^t_\eta(x, U_x) \ge 1 - o(1).$$

Since $\mu_{\text{in}}(U_x) = o(1)$, this ends the proof. □

4.3 Cutoff in changing environment

This section is devoted to the proof of Theorem 4.2. In the proof we will make a crucial use of the fact that the stationary distribution of a digraph from the DCM/OCM ensemble is a widespread measure w.h.p.. This result is presented in the following subsection.

4.3.1 The stationary measure is widespread

Lemma 4.3 *There exists a constant $C \equiv C(\Delta) > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left(n \sum_{z\in[n]} \pi(z)^2 \le C \right) = 1.$$

Proof: Call $Z = n\sum_{z\in[n]} \pi(z)^2$. Let $t = \log^3(n)$ and consider the event

$$\mathcal{D} = \left\{ \max_{x,z\in[n]} |\pi(z) - P^t(x,z)| = o(n^{-3}) \right\}. \tag{4.45}$$

A simple consequence of Theorem 1.7 (see Section 5.3) is that $\mathbb{P}(\mathcal{D}) = 1 - o(1)$. Therefore,

$$\mathbb{P}(Z > C) \le \mathbb{P}(Z > C; \mathcal{D}) + o(1).$$

By Markov's inequality

$$\mathbb{P}\left(Z > C;\ \mathcal{D}\right) \le \frac{\mathbb{E}[Z^K \mathbf{1}_{\mathcal{D}}]}{C^K}, \qquad \forall K \ge 1.$$

Therefore, it is sufficient to show that $\mathbb{E}[Z^K \mathbf{1}_{\mathcal{D}}] \leq (C/2)^K$ for some $K = \omega(1)$. Choose for example $K = \log(n)$,

$$\mathbb{E}[Z^K \mathbf{1}_{\mathcal{D}}] \leq n^K \mathbb{E}\left[\left(\sum_{z\in[n]}\sum_{x\in[n]}\sum_{y\in[n]} \frac{1}{n^2}\left(P^t(x,z)+o(n^{-3})\right)\left(P^t(y,z)+o(n^{-3})\right)\right)^K\right] \tag{4.46}$$

$$\leq n^K \left(\mathbb{E}\left[\left(o(n^{-1}) + \sum_{z\in[n]}\sum_{x\in[n]}\sum_{y\in[n]} \frac{1}{n^2}P^t(x,z)P^t(y,z)\right)^K\right]\right) \tag{4.47}$$

$$\leq (2n)^K \left(\mathbb{E}\left[\left(\sum_{z\in[n]}\sum_{x\in[n]}\sum_{y\in[n]} \frac{1}{n^2}P^t(x,z)P^t(y,z)\right)^K\right]\right) + o(1) \tag{4.48}$$

$$= (2n)^K \mathbf{P}^{\mathbf{an}}_{unif}\left(X^{(\ell)}_t = Y^{(\ell)}_t,\ \forall \ell \leq K\right) + o(1) \tag{4.49}$$

where $\mathbf{P}^{\mathbf{an}}_{unif}$ denotes the law of the $2K$ annealed walks $(X^{(k)}_s, Y^{(k)}_s)_{s\leq t}$ for $k \leq K$, each starting at a uniform and vertex independently on the others. The $2K$ walks are independent conditionally on the environment, and the average is both over the walks and the environment. For an explicit construction, we can generate recursively the walks and the environment, letting the trajectories reveal the configuration σ, the ℓ-th trajectory living in the environment discovered by the previous $\ell-1$ trajectories; see Section 5.3 for a more detailed argument. Therefore, it is sufficient to show that it is possible to find a constant $C > 0$ such that for every sufficiently large n

$$\mathbf{P}^{\mathbf{an}}_{unif}\left(X^{(\ell)}_t = Y^{(\ell)}_t,\ \forall \ell \leq K\right) \leq \left(\frac{C}{4n}\right)^K. \tag{4.50}$$

We define the events

$$B_k = \bigcap_{\ell\leq k}\{X^{(\ell)}_t = Y^{(\ell)}_t\}, \tag{4.51}$$

and call A_k the set of vertices which have at least one tail/head revealed by the trajectories $\left(X^{(\ell)}, Y^{(\ell)}\right)_{\ell\leq k}$. We call Ξ_k a realization of the trajectories $(X^{(\ell)}_s, Y^{(\ell)}_s)_{s\leq t,\ell\leq k}$ satisfying B_k, and compute the conditional probability

$$\mathbf{P}^{\mathbf{an}}_{unif}(B_{k+1} \mid \Xi_k) = \sum_{z\in[n]}\sum_{x\in[n]}\sum_{y\in[n]} \frac{1}{n^2}\mathbf{P}^{\mathbf{an}}_{x,y}\left(X^{(k+1)}_t = Y^{(k+1)}_t = z \mid \Xi_k\right). \tag{4.52}$$

We start by showing that if x, y, z are distinct and in $[n] \setminus A_k$ then, uniformly in Ξ_k,

$$\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)} = Y_t^{(k+1)} = z \,|\Xi_k\right) = O\left(\frac{1}{n^2}\right). \tag{4.53}$$

Consider the event $\mathcal{E}_k$ that the trajectory $X^{(k)}$ has no collision with itself nor with the environment previously discovered by $X^{(1)}, Y^{(1)} \dots, X^{(k-1)}, Y^{(k-1)}$, while $\mathcal{Y}_k$ will denote the event that the walk $X^{(k)}$ does not discover y. At any given time, any given walk has probability $O(1/n)$ of hitting a given vertex by generating a fresh new edge. Thus, by a union bound, the event $\mathcal{E}_k^c \cup \mathcal{Y}_k^c$ has probability $O(Kt^2/n)$ uniformly in $k \le K$. We write

$$\begin{aligned}
\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)} = Y_t^{(k+1)} = z \mid \Xi_k\right) =& \mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)} = Y_t^{(k+1)} = z; \mathcal{E}_{k+1}; \mathcal{Y}_{k+1} \mid \Xi_k\right) + \\
&+\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)} = Y_t^{(k+1)} = z; \mathcal{E}_{k+1}^c; \mathcal{Y}_{k+1} \mid \Xi_k\right) + \\
&+\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)} = Y_t^{(k+1)} = z; \mathcal{E}_{k+1}; \mathcal{Y}_{k+1}^c \mid \Xi_k\right) + \\
&+\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)} = Y_t^{(k+1)} = z; \mathcal{E}_{k+1}^c; \mathcal{Y}_{k+1}^c \mid \Xi_k\right).
\end{aligned}$$

Consider first the case where $X^{(k+1)}$, before arriving in z at time t, passes through y and visits at least once an already discovered vertex. The probability of visiting $y \in A_k^c$ and $z \in A_k^c$ can be bounded by $O(t^2/n^2)$, and the probability of visiting a discovered vertex cannot exceed $O(Kt^2/n)$, uniformly in $k \le K$. Thus,

$$\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)} = z; \mathcal{E}_{k+1}^c; \mathcal{Y}_{k+1}^c \mid \Xi_k\right) = O(t/n)O(t/n)O(Kt^2/n) = o(n^{-2}).$$

Similarly,

$$\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)} = Y_t^{(k+1)} = z; \mathcal{E}_{k+1}^c; \mathcal{Y}_{k+1} \mid \Xi_k\right) = O(Kt^2/n)O(t/n)O(Kt^2/n) = o(n^{-2}).$$

Indeed, the walk $X^{(k+1)}$ must visit z and also one of the previously discovered vertices, which has probability $O(Kt^2/n) \times O(t/n)$. Then, in order for the walk $Y^{(k+1)}$ to arrive in z at t it is necessary to visit a vertex that was already discovered (e.g., z itself). The latter event has probability $O\left(Kt^2/n\right)$.

Notice that under $\mathcal{E}_{k+1} \cap \mathcal{Y}_{k+1}$, in order to realize the event $X_t^{(k+1)} = z, Y_t^{(k+1)} = z$ there must be a time $s \le t$ such that $Y^{(k+1)}$ collides at time s with the trajectory of $X^{(k+1)}$, then $Y^{(k+1)}$ stays on this trajectory for $t - s$ units of time, and then finally hits z at time t. On the event $\mathcal{E}_{k+1}$ the probability of $X_t^{(k+1)} = z$ is bounded by $\frac{d_z^-}{m}(1+o(1))$, and the event that $Y^{(k+1)}$ spends h units of time in the path $X^{(k+1)}$ is at most 2^{-h}. Therefore,

$$\mathbf{P}_{x,y}^{\mathbf{an}}\left(X_t^{(k+1)} = Y_t^{(k+1)} = z; \mathcal{E}_{k+1}; \mathcal{Y}_{k+1} \mid \Xi_k\right) \leq \frac{d_z^-}{m}(1+o(1)) \cdot \frac{\Delta}{m}\sum_{h=1}^{t} 2^{-h} \leq \frac{\Delta^2}{n^2}. \tag{4.54}$$

Under $X_0 \neq y$ and The event $\mathcal{E}_{k+1} \cap \{X_t^{(k+1)} = z\} \cap \mathcal{Y}_{k+1}^c$ has probability $O(t/n) \times O(1/n)$. Under this event, when the walk $Y^{(k+1)}$ starts at y the revealed in-neighborhood of z consist of a unique path of length t from x to z and y is a vertex in this path. Since $y \neq x$, to achieve $Y^{(k+1)} = z$ it is necessary that $Y^{(k+1)}$ exits and re-enters the path. This requires hitting the path by creating a fresh edge, which has probability $O(Kt^2/n)$. Hence,

$$\mathbf{P}_{x,y}^{\mathbf{an}}\left(X_t^{(k+1)} = Y_t^{(k+1)} = z; \mathcal{E}_{k+1}; \mathcal{Y}_{k+1}^c \mid \Xi_k\right) = O(t/n)O(1/n)O(Kt^2/n) = o(n^{-2}).$$

In conclusion, we have proved (4.53). In particular,

$$\sum_{z \in A_k^c}\sum_{x \in A_k^c \setminus z}\sum_{y \in A_k^c \setminus z,x} \frac{1}{n^2}\mathbf{P}_{x,y}^{\mathbf{an}}\left(X_t^{(k+1)} = Y_t^{(k+1)} = z \mid \Xi_k\right) \leq n^3\frac{1}{n^2}\frac{\Delta^2}{n^2} = \frac{\Delta^2}{n}. \tag{4.55}$$

We now deal with the probability

$$\mathbf{P}_{x,y}^{\mathbf{an}}\left(X_t^{(k+1)} = Y_t^{(k+1)} \mid \Xi_k\right),$$

when $x \in A_k$ and $y \in A_k^c$ or viceversa. By symmetry we can restrict to the former case. We observe that

$$\mathbf{P}_{x,y}^{\mathbf{an}}\left(X_t^{(k+1)} = Y_t^{(k+1)} \mid \Xi_k\right) = O\left(\frac{Kt^2}{n}\right). \tag{4.56}$$

Indeed,

$$\mathbf{P}_{x,y}^{\mathbf{an}}\left(X_t^{(k+1)} = Y_t^{(k+1)}; \mathcal{Y}_{k+1} \mid \Xi_k\right) = O\left(\frac{Kt^2}{n}\right), \tag{4.57}$$

since the latter event requires that the walk $Y_t^{(k+1)}$ visits a vertex that has been already discovered by $X^{(1)}, Y^{(1)}, \ldots, X^{(k)}, Y^{(k)}, X^{(k+1)}$. On the other hand

$$\mathbf{P}_{x,y}^{\mathbf{an}}\left(\mathcal{Y}_{k+1}^c \mid \Xi_k\right) = O\left(\frac{t}{n}\right). \tag{4.58}$$

Hence, using $|A_k| \leq Kt$,

$$\sum_{x\in A_k}\sum_{y\in A_k^c}\frac{1}{n^2}\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)}=Y_t^{(k+1)}\mid \Xi_k\right)\leq \frac{Kt\cdot n}{n^2}\times O\left(\frac{Kt^2}{n}\right)=o(n^{-1}). \tag{4.59}$$

We are left with considering two cases: $x \in A_k$ and $y \in A_k$ (and $z \in [n]$ arbitrary), and the case $z \in A_k$ while $x, y \in A_k^c$. The probability of these two events are easy to bound. Indeed, $|A_k| \leq Kt$ implies

$$\sum_{x\in A_k}\sum_{y\in A_k}\frac{1}{n^2}\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)}=Y_t^{(k+1)}\mid \Xi_k\right)\leq \frac{K^2t^2}{n^2}=o(n^{-1}). \tag{4.60}$$

Finally, if $z \in A_k$ and $x, y \in A_k^c$ then

$$\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)}=Y_t^{(k+1)}\in A_k;\ \mathcal{Y}_{k+1}\mid \Xi_k\right)=O\left(\frac{K^2t^4}{n^2}\right),$$

since both the walks have to discover the cluster A_k in order to visit z. On the other hand

$$\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)}=Y_t^{(k+1)}\in A_k;\ \mathcal{Y}^c_{k+1}\mid \Xi_k\right)=O\left(\frac{Kt^2}{n}\cdot\frac{t}{n}\right),$$

since the $X^{(k+1)}$ needs to visit both y and the cluster A_k. Hence

$$\sum_{x\in A_k^c}\sum_{y\in A_k^c}\frac{1}{n^2}\mathbf{P}^{\mathbf{an}}_{x,y}\left(X_t^{(k+1)}=Y_t^{(k+1)}\in A_k\mid \Xi_k\right)=O\left(\frac{K^2t^4}{n^2}\right)=o(n^{-1}). \tag{4.61}$$

Therefore, putting together the bounds (4.55), (4.59), (4.60) and (4.61), and reccaling (4.52), we showed that

$$\mathbf{P}^{\mathbf{an}}_{unif}(B_{k+1}\mid \Xi_k)\leq \frac{\Delta^2}{n}+o\left(n^{-1}\right)\leq\frac{2\Delta^2}{n}.$$

By the uniformity in $k \leq K$ and in Ξ_k of the previous argument, we conclude that

$$\mathbf{P}^{\mathbf{an}}_{unif}\left(B_K\right)=\mathbf{P}^{\mathbf{an}}_{unif}\left(B_1\right)\prod_{k=1}^{K-1}\mathbf{P}^{\mathbf{an}}_{unif}\left(B_{k+1}\mid B_k\right)\leq\left(\frac{2\Delta^2}{n}\right)^K.$$

Therefore it is sufficient to choose $C = 8\Delta^2$ to conclude that (4.50) holds.

□

Lemma 4.4 *We have*

$$\lim_{n\to\infty}\mathbb{P}\left(\max_{z\in[n]}\pi(z)\leq\frac{\log^8(n)}{n}\right)=1. \tag{4.62}$$

Proof: For the DCM ensemble we may refer to Section 5.4 for a much more precise result, where 8 is replaced by an optimal constant $a \in [0,1]$. We give here an alternative proof of the weaker bound (4.62) that holds for the OCM as well. We show that if $t = \log^3(n)$, then uniformly in $z \in [n]$

$$\mathbb{P}\left(\sum_{x\in[n]} \frac{1}{n} P^t(x,z) \geq \frac{\log^8(n)}{2n}\right) = o(n^{-1}). \tag{4.63}$$

By the union bound, and the fact that the event $\mathcal{D}$ in (4.45) occurs w.h.p., (4.63) is sufficient to prove (4.62). Define

$$W := \sum_{x\in[n]} \frac{1}{n} P^t(x,z).$$

By Markov inequality, for every $K \geq 1$

$$\mathbb{P}\left(W \geq \frac{\log^8(n)}{2n}\right) \leq \frac{2^K n^K}{\log^{8K}(n)} \mathbb{E}\left[W^K\right].$$

As in the proof of Lemma 4.3, the term in the right hand side of the latter display can be read in terms of the annealed walks. In conclusion, to prove (4.63) it is sufficient to show that for $K = \log(n)$

$$\mathbb{E}\left[W^K\right] = \mathbf{P}^{\mathbf{an}}_{unif}\left(X^{(k)}_t = z,\ \forall k \leq K\right) \leq \left(\frac{C\log^7(n)}{n}\right)^K, \tag{4.64}$$

for some constant $C > 0$, where the $(X^{(k)}_s)_{s\leq t}$, for $k \in \{1,\dots,K\}$, are K independent walks conditionally on the environment and the average is both over the walks and the environment. Reasoning as in the proof of Lemma 4.3, similarly to (4.51) we call

$$B_k = \bigcap_{\ell \leq k} \{X^{(\ell)}_t = z\}.$$

The proof is completed by observing that uniformly in $k \leq K$,

$$\mathbf{P}^{\mathbf{an}}_{unif}(B_{k+1} \mid B_k) = O\left(\frac{Kt^2}{n}\right) = O\left(\frac{\log^7(n)}{n}\right),$$

which is sufficient to prove (4.64). The above estimate simply follows by observing that $X^{(k+1)}_t = z$ implies that $X^{(k+1)}$ hits at some time $s \in [0,t]$ for the first time a vertex already discovered by the walks $X^{(\ell)}$, $\ell \leq k$. $\square$

Lemma 4.3 and Lemma 4.4 provide the result mentioned at the beginning of the section.

Corollary 4.2 *W.h.p. π_σ is a widespread measure.*

4.3.2 Proof of the cutoff

We now turn to the proof of Theorem 4.2. Let σ, η two independent uniformly random configurations in $\mathfrak{C}$. In this section we will assume that $t \in \Theta(T_{\mathrm{ENT}})$ and $s \leq t$. Let $\mathbf{Q}_x^{\sigma,\eta} \equiv \mathbf{Q}_{x,s,t}^{\sigma,\eta}$ denote the quenched law of the walker that starts at $X_0 = x$, goes for s steps trough σ and then, starting at X_s, goes for $t-s$ steps trough η. We will represent the probability distribution on $[n]$ for the position of the walker at time t by the symbol

$$Q_s^t(x,y) = \mathbf{Q}_x^{\sigma,\eta}(X_t = y).$$

Definition 4.1 *We define* path of length t *an arbitrary sequence of vertices* $\mathfrak{p} = (v_0, \dots, v_t)$. *We call weight of the path* $\mathbf{w}(\mathfrak{p})$ *the product*

$$\mathbf{w}(\mathfrak{p}) = \prod_{j=0}^{s-1} P_\sigma(v_j, v_{j+1}) \prod_{i=s}^{t-1} P_\eta(v_i, v_{i+1}).$$

Lemma 4.5 *If* $s \in [0,t]$ *and* $t = \Theta(\log n)$, *for every* $\varepsilon \in (0,1)$

$$\min_{x \in [n]} \mathbf{Q}_x^{\sigma,\eta} \left(\mathbf{w}(X_0, X_1, \dots, X_t) \in \left[e^{-(1+\varepsilon)Ht}, e^{-(1-\varepsilon)Ht} \right] \right) \xrightarrow{\mathbb{P}} 1.$$

Proof: The case $s = 0$ is exactly [13, Proposition 8] for the DCM and [14, Theorem 4] for the OCM. Since $\mathbf{w}$ is a product, if $s, t-s = \Theta(\log n)$ the claim follows by these results in [13, 14] and the independence of σ and η. It remains to consider the case $s = o(\log n)$ and the case $t - s = o(\log n)$. If $s = o(\log n)$ then w.h.p. any path of length s has weight in the window $[\Delta^{-s}, 2\delta^{-s}] = e^{-o(t)}$, see Chapter 2. Hence the result follows again by the case $s = 0$. The same argument works also in the case $t - s = o(\log n)$. □

Proof of the Lower Bound of Theorem 4.2.

Let $t = (1-\varepsilon)T_{\mathrm{ENT}}$ for some $\varepsilon > 0$. Fix any $x \in [n]$ and call U_x the set of vertices y such that $Q_s^t(x,y) > e^{-(1+\varepsilon)Ht} = n^{-1+\varepsilon^2}$. By Lemma 4.5, uniformly in x, $|U_x| = o(n)$ and $Q_s^t(x, U_x) = 1 - o(1)$. Since π_η is widespread by Corollary 4.2, from the argument in (4.36) we have $\pi_\eta(U_x) = o(1)$. Hence, w.h.p. for every $x \in [n]$ the measure $Q_s^t(x, \cdot)$ is asymptotically singular with respect to π_η. This concludes the proof of the lower bound of Theorem 4.2. □

We now turn to proving the upper bound in Theorem 4.2, which is more involved. In fact, an adaptation of the arguments of [13, 14] is not straightforward in this case. Below we present the details in the case of the DCM. The proof for the OCM is very similar.

Remark 4.2 *For what concerns the upper bound, we can restrict to the case* $s, t-s = \Theta(\log(n))$ *because of the following argument. If* $s = o(\log(n))$ *then the upper bound in Theorem 4.2 holds as a trivial consequence of Theorem 1.7. In fact, if* $s = o(\log(n))$ *and* $t = (1+\varepsilon)T_{\rm ENT}$ *for some* $\varepsilon > 0$*, we have* $t - s \geq (1+\varepsilon/2)T_{\rm ENT}$*. Hence,*

$$\|Q_s^t(x,\cdot) - \pi_\eta\|_{TV} \leq \max_{x\in[n]} \|P_\eta^{t-s}(x,\cdot) - \pi_\eta\|_{TV} \xrightarrow{\mathbb{P}} 0.$$

If $t - s = o(\log(n))$ *then* $s \geq (1+\varepsilon/2)T_{\rm ENT}$*. Therefore*

$$\begin{aligned}\|Q_s^t(x,\cdot) - \pi_\eta\|_{TV} \leq & \|Q_s^t(x,\cdot) - \pi_\sigma P_\eta^{t-s}\|_{TV} + \|\pi_\sigma P_\eta^{t-s} - \pi_\eta\|_{TV} \\ \leq & \|P_\sigma^s(x,\cdot) - \pi_\sigma\|_{TV} + \|\pi_\sigma P_\eta^{t-s} - \pi_\eta\|_{TV} \xrightarrow{\mathbb{P}} 0.\end{aligned}$$

where we used that $\|P_\sigma^s(x,\cdot) - \pi_\sigma\|_{TV} \xrightarrow{\mathbb{P}} 0$ *by Theorem 1.7 and by Corollary 4.2* π_σ *is widespread. Then the result follows again by Theorem 1.7.*

In what follow we will assume $t, s, t-s = \Theta(\log(n))$ and $t = (1+\upsilon)T_{\rm ENT}$ for some sufficiently small $\upsilon > 0$. The general case $t = \beta T_{\rm ENT}$ for any $\beta > 1$ follows by monotonicity of the TV-distance with respect to t. Call

$$\hbar := \frac{1}{5}\log_\Delta(n), \qquad h := \hbar \wedge \frac{t-s}{2} = \Theta(T_{\rm ENT}), \qquad r := t - s - h = \Theta(T_{\rm ENT}),$$

and notice that there exists some $\epsilon \in (0,1)$ such that

$$r + s \leq (1-\epsilon)T_{\rm ENT}.$$

Strategy of proof

The general strategy of proof is the same as in [13, 14]. We recall here the main steps and then give the details of its implementation in our general setting. We can replace π_η by $\mu_{\rm in}P^h$ since we know by Lemma 3.1 that w.h.p.

$$\|\mu_{\rm in}P_\eta^h - \pi_\eta\|_{\rm TV} = o(1). \tag{4.65}$$

We will focus on a particular set of starting states. Fixed the realization σ we call $S_\star^\sigma$ the set of vertices for which the out-neighborhood in σ is a tree up to height h. By [13, Proposition 6] (or [14, Lemma 9]) w.h.p. with respect to the sampling of the configuration σ, most of the vertices are in $S^\sigma_\star$, and the quenched probability that the walk is out of the set $S^\omega_\star$ vanishes exponentially fast in time. More precisely, fixed $\ell = \log\log n < s$

$$\max_{x\in[n]} \mathbf{Q}^{\sigma,\eta}_x \left(X_\ell \notin S^\sigma_\star\right) \le 2^{-\ell}.$$

Therefore, by the triangle inequality

$$\max_{x\in[n]} \|Q^t_s(x,\cdot) - \mu_{\rm in}P^h_\eta\|_{\rm TV} \le \max_{x\in[n]} \mathbf{Q}^{\sigma,\eta}_x(X_\ell \notin S^\sigma_\star) + \max_{x\in S^\sigma_\star} \|Q^{t-\ell}_{s-\ell}(x,\cdot) - \mu_{\rm in}P^h_\eta\|_{\rm TV}. \tag{4.66}$$

Thus, in order to show the uniform upper bound in Theorem 4.2 it is sufficient to show an upper bound that holds uniformly in the random set $S^\sigma_\star$.

We will define a set of *nice paths* for the trajectory of the walk. For every couple of vertices x, y we let $\mathcal{N}_{x,y}$ denote the set of *nice paths* from x to y of length t. Consequently, we define

$$\bar{Q}^t_s(x,y) := \sum_{\mathfrak{p}\in\mathcal{N}_{x,y}} \mathbf{w}(\mathfrak{p})$$

the probability to go from x to y along a nice path. Notice that for arbitrarily small constant $\varepsilon > 0$

$$\begin{aligned} \|\mu_{\rm in}P^h_\eta - Q^t_s(x,\cdot)\|_{\rm TV} &= \sum_{y\in[n]} \left[\mu_{\rm in}P^h_\eta(y) - Q^t_s(x,y)\right]_+ \\ &\le \sum_{y\in[n]} \left[\mu_{\rm in}P^h_\eta(y)(1+\varepsilon) + \frac{\varepsilon}{n} - \bar{Q}^t_s(x,y)\right]_+ . \end{aligned} \tag{4.67}$$

Therefore, if we can show that

$$\mu_{\rm in}P^h_\eta(y)(1+\varepsilon) + \frac{\varepsilon}{n} \ge \bar{Q}^t_s(x,y), \tag{4.68}$$

then the positive part in (4.67) can be neglected, and summing over $y \in [n]$ one obtains

$$\begin{aligned} \|\mu_{\rm in}P^h_\eta - Q^t_s(x,\cdot)\| &\le \sum_{y\in[n]} \Big((1+\varepsilon)\mu_{\rm in}P^h_\eta(y) + \frac{\varepsilon}{n} - \bar{Q}^t_s(x,y)\Big) \\ &= \mathbf{Q}^{\sigma,\eta}_x\Big((X_0,\dots,X_t) \notin \bigcup_{y\in[n]} \mathcal{N}_{x,y}\Big) + 2\varepsilon. \end{aligned} \tag{4.69}$$

At this point we are left with showing that the probability of following a path that is not nice is arbitrarily small uniformly in the starting point $x \in S_\star^\sigma$, namely

$$\max_{x \in S_\star^\sigma} \mathbf{Q}_x^{\sigma,\eta}\Big((X_0, \dots, X_t) \notin \bigcup_{y \in [n]} \mathcal{N}_{x,y}\Big) < \varepsilon \qquad \text{w.h.p.} \tag{4.70}$$

We first introduce the notation required to define the set of nice paths. Then we will present a proof of the validity of (4.70) and (4.68).

We will start by constructing the subgraph $\mathcal{G}_x^\sigma(s)$ of σ spanned by the paths of length at most s, starting at x, and with weight greater $e^{-(1+\upsilon)Hs}$. We will construct $\mathcal{G}_x^\sigma(s)$ together with a spanning tree $\mathcal{T}_x^\sigma(s)$ of $\mathcal{G}_x^\sigma(s)$ in the following way.

Definition 4.2 ***Construction of*** $\mathcal{G}_x^\sigma(s)$ ***and*** $\mathcal{T}_x^\sigma(s)$**.**

- *Call* $\mathcal{G}_x^\sigma[0]$ *the empty graph on* $\{x\}$ *and* $E_1^\sigma = E_x^+$.
- *To every* $e_1 \in E_1^\sigma$ *it is associated the weight* $\widehat{\mathbf{w}}_\sigma(e_1) := (d_x^+)^{-1}$.
- *For every* $\ell \geq 1$
  1. *Choose a tail* $e_\ell \in E_\ell^\sigma$ *with maximal weight and reveal* $\sigma(e_\ell) = f_\ell$.
  2. *Add the edge* (e_ℓ, f_ℓ) *to* $\mathcal{G}_x^\sigma[\ell - 1]$ *and call the resulting graph* $\mathcal{G}_x^\sigma[\ell]$.
  3. *Call the edge* (e_ℓ, f_ℓ) open *if* $v(f_\ell) \notin \mathcal{G}_x^\sigma[\ell - 1]$.
  4. *Call* $\mathcal{T}_x^\sigma[\ell]$ *the* open *subgraph of* $\mathcal{G}_x^\sigma[\ell]$.
  5. *If* $v(f_\ell) \notin \mathcal{G}_x^\sigma[\ell - 1]$, *then we associate to any* $e' \in E_{v(f_\ell)}^+$ *the weight* $\widehat{\mathbf{w}}_\sigma(e') := \widehat{\mathbf{w}}_\sigma(e_\ell)(d_{v(f_\ell)}^+)^{-1}$, *and if*

     $$\widehat{\mathbf{w}}_\sigma(e') \geq e^{-(1+\upsilon)Hs} =: \widehat{\mathbf{w}}_{\min,\sigma},$$

     then let $E_{\ell+1}^\sigma = E_\ell^\sigma \setminus \{e_\ell\} \cup E_{v(f_\ell)}^+$. *Otherwise, set* $E_{\ell+1}^\sigma = E_\ell^\sigma \setminus \{e_\ell\}$.
  6. *Remove from* $E_{\ell+1}^\sigma$ *the tails* e' *such that the vertex* $v(e')$ *is at distance greater than* s *from* x *in* $\mathcal{T}_x^\sigma[\ell]$.
- *Iterate the procedure up to the random time* κ_σ *at which* $E_{\kappa_\sigma}^\sigma = \emptyset$, *and call*

  $$\mathcal{T}_x^\sigma(s) := \mathcal{T}_x^\sigma[\kappa_\sigma], \qquad \mathcal{G}^\sigma(s) := \mathcal{G}_x^\sigma[\kappa_\sigma].$$

The definition given above of the subgraphs $\mathcal{T}_x^\sigma(s)$ and $\mathcal{G}_x^\sigma(s)$ coincides with that given in [13, 14]. It was shown in [13, 14] that the random walk on the static environment

σ, starting at $x \in S_\star^\sigma$ and of length s will stay on the tree $\mathcal{T}_x^\sigma(s)$ w.h.p.. Hence, in the double environment case, the walk will be w.h.p. in one of the leaves of $\mathcal{T}_x^\sigma(s)$ at time s. Call $\mathcal{L}_s^{x,\sigma}$ the set of leaves of $\mathcal{T}_x^\sigma(s)$. We will now construct the subgraph $\mathcal{G}^\eta(r)$ of all the paths in η which start at some $z \in \mathcal{L}_s^{x,\sigma}$, have length r, and cumulative weight larger than $e^{-(1+v/2)H(t-h)}$. Similarly to the construction in Definition 4.2, together with $\mathcal{G}^\eta(r)$ we are going to construct a collection of rooted directed trees $\mathcal{T}_z^\eta(r)$, each rooted at $z \in \mathcal{L}_s^{x,\sigma}$, with depth r. The forest

$$\mathcal{W}^\eta(r) := \bigcup_{z \in \mathcal{L}_s^{x,\sigma}} \mathcal{T}_z^\eta(r)$$

seen as a collection of edges, will be our candidate for the support of the second part of the walk.

Definition 4.3 ***Construction of*** $\mathcal{G}^\eta(r)$ ***and*** $\mathcal{W}^\eta(r)$**.**

- *Call* $\mathcal{G}^\eta[0]$ *the empty graph on* $\mathcal{L}_s^{x,\sigma}$ *and call* $E_1^\eta = \cup_{z \in \mathcal{L}_s^{x,\sigma}} E_z^+$.
- *To every* $e_1 \in E_1^\eta$ *we associate the weight*

$$\widehat{\mathbf{w}}(e_1) := \widehat{\mathbf{w}}_\sigma(e_1),$$

of the unique path in $\mathcal{T}_x^\sigma(s)$ *joining* x *to* $v(e_1)$ *times the inverse of the out degree of* $v(e_1)$.

- *For every* $\ell \geq 1$
  1. *Choose a tail in* $e_\ell \in E_\ell^\eta$ *with maximal weight and reveal* $\eta(e_\ell) = f_\ell$.
  2. *Add the edge* (e_ℓ, f_ℓ) *to* $\mathcal{G}^\eta[\ell-1]$ *and call the resulting graph* $\mathcal{G}^\eta[\ell]$.
  3. *Call the edge* (e_ℓ, f_ℓ) open *if* $v(f_\ell) \notin \mathcal{G}^\eta[\ell-1]$.
  4. *Call* $\mathcal{W}^\eta[\ell]$ *the* open *subgraph of* $\mathcal{G}^\eta[\ell]$.
  5. *If* $v(f_\ell) \notin \mathcal{G}^\eta[\ell-1]$, *then we associate to any* $e' \in E_{v(f_\ell)}^+$ *the weight* $\widehat{\mathbf{w}}(e') := \widehat{\mathbf{w}}(e_\ell)(d_{v(f_\ell)}^+)^{-1}$, *and if*

$$\widehat{\mathbf{w}}(e') \geq e^{-(1+v/2)H(t-h)} =: \widehat{\mathbf{w}}_{\min},$$

     then let $E_{\ell+1}^\eta = E_\ell^\eta \setminus \{e_\ell\} \cup E_{v(f_\ell)}^+$. *Otherwise, set* $E_{\ell+1}^\eta = E_\ell^\eta \setminus \{e_\ell\}$.
  6. *Remove from* $E_{\ell+1}^\eta$ *the tails* e' *such that the vertex* $v(e')$ *is at distance greater than* r *from the corresponding root in* $\mathcal{W}^\eta[\ell]$.
- *Iterate the procedure up to the random time* κ'_η *at which* $E_{\kappa'_\eta}^\eta = \emptyset$, *and call*

$$\mathcal{W}^\eta(r) := \mathcal{W}^\eta[\kappa'_\eta] = \cup_{z \in \mathcal{L}_s^{\sigma,x}} \mathcal{T}_z^\eta[\kappa'_\eta], \qquad \mathcal{G}^\eta(r) := \mathcal{G}^\eta[\kappa'_\eta].$$

We know by [13, 14] that the random number of edges revealed by the construction in Definition 4.2, κ_σ, is a.s. $o(n)$. We need an analogous result for the quantity κ'_η in Definition 4.3.

Lemma 4.6 *For any $\sigma \in \mathfrak{C}$ and $x \in [n]$, for all $\eta \in \mathfrak{C}$*

$$\widehat{\mathbf{w}}(e_\ell) \leq \frac{r}{\ell}, \qquad \forall \ell < \kappa'_\eta.$$

In particular, recalling that $t - h = r + s \leq (1-\epsilon)T_{\mathrm{ENT}}$, by choosing $\upsilon < \epsilon$

$$\kappa'_\eta = O\left(\log(n) n^{(1+\upsilon/2)(1-\epsilon)}\right) = O(n^{1-\epsilon^2}).$$

Proof: For each $\ell < \kappa'_\eta$ we consider the forest

$$\widetilde{\mathcal{W}}^\eta[\ell] := \cup_{z \in \mathcal{L}_s^{x,\sigma}} \widetilde{\mathcal{T}}_z^\eta[\ell]$$

constructed as in Definition 4.3, but if an edge $(e_{\ell'}, f_{\ell'})$ for some $\ell' \leq \ell$ is not open, we attach a fictitious leaf (with no future children) to $e_{\ell'}$, to which we assign the weight $\widehat{\mathbf{w}}(e_{\ell'})$. This construction ensures that for every ℓ both the graph $\mathcal{G}^\eta[\ell]$ and the forest $\widetilde{\mathcal{W}}^\eta[\ell]$ have exactly ℓ edges. Call F_ℓ the set of leaves of $\widetilde{\mathcal{W}}^\eta[\ell]$. By construction, for all $v \in F_\ell$ there is a unique $z \in \mathcal{L}_s^{x,\sigma}$ and a unique path $\mathfrak{p}(v) : z \to v$ of length at most r in $\widetilde{W}^\eta[\ell]$. The weight of such a path is given by $\widehat{\mathbf{w}}(e_v)$ where e_v is any tail in E_u^+ if (u, v) is the last edge in the path $\mathfrak{p}(v)$. It follows that

$$\sum_{z \in \mathcal{L}_s^{x,\sigma}} \sum_{v \in F_\ell} \sum_{\mathfrak{p}: z \to v} \widehat{\mathbf{w}}(e_v) \leq 1.$$

Since all $v \in F_\ell$ are such that $\widehat{\mathbf{w}}(e_v) \geq \widehat{\mathbf{w}}(e_\ell)$, we obtain

$$|F_\ell| \widehat{\mathbf{w}}(e_\ell) \leq 1.$$

By the absence of cycles in $\widetilde{\mathcal{W}}^\eta[\ell]$ we also have that

$$\ell \leq r |F_\ell|.$$

In conclusion

$$\widehat{\mathbf{w}}(e_\ell) \leq \frac{1}{|F_\ell|} \leq \frac{r}{\ell}.$$

If we replace $\ell = \kappa'_\eta - 1$ we get

$$\kappa'_\eta - 1 \le \frac{r}{\widehat{\mathbf{w}}(e_{\kappa'_\eta - 1})} \le \frac{r}{\widehat{\mathbf{w}}_{\min}}.$$

□

We are now in shape to define the set of nice paths.

Definition 4.4 *We call* nice *a path* $\mathfrak{p} = (v_0, \ldots, v_t)$ *s.t.*

1. $\mathbf{w}(\mathfrak{p}) \le e^{-(1-\upsilon/2)Ht} = n^{-1-\upsilon/2-\upsilon^2/2} = O(n^{-1-\upsilon/3})$.
2. $\mathfrak{p}$ *belongs to* $\mathcal{T}^\sigma_x(s)$ *up to time* s.
3. $\mathfrak{p}$ *belongs to* $\mathcal{T}^\eta_z(r)$ *from time* s *to time* $t - h = r + s$ *for some* $z \in \mathcal{L}^{x,\sigma}_s$.
4. $(v_{t-h}, \ldots, v_t)$ *is the unique path of length at most* h *in the configuration* η *from* v_{t-h} *to* v_t.

We now focus on proving (4.70), which will be a consequence of the law of large numbers in Lemma 4.5 together with the forthcoming Lemma 4.7. The latter shows, via a martingale argument, that w.h.p. the walk will pass along the edges of one of the trees of the forest $\mathcal{W}^\eta(r)$ for the time steps $s, \ldots, s + r$.

Fix $\sigma \in \mathfrak{C}$ and $x \in [n]$. The set of leaves $\mathcal{L}^{\sigma,x}_s$ is then determined, and we call $\mathbb{P}_{\sigma,x}$ the law of the process defined in Definition 4.3. Consider the the σ-field $(\mathcal{S}_\ell)_{\ell\ge 0}$ generated by the first ℓ steps of the construction described in Definition 4.3. Call $(M_\ell)_{\ell\ge 0}$ the stochastic process adapted to $(\mathcal{S}_\ell)_{\ell\ge 0}$ defined recursively by $M_0 = 0$ and

$$M_{\ell+1} = M_\ell + \mathbf{1}_{\ell+1<\kappa'_\eta}\mathbf{1}_{v(f_{\ell+1})\in\mathcal{G}^\eta[\ell]}\widehat{\mathbf{w}}(e_{\ell+1}).$$

Lemma 4.7 *For every* $\varepsilon > 0$

$$\mathbb{P}_{\sigma,x}\left(M_{\kappa'_\eta} \le \varepsilon\right) = 1 - o(n^{-1}).$$

Proof: We follow [13, 14], where a very similar statement was proved for the walk on a single environment. We compute the first two conditional moments of the increment $M_{\ell+1} - M_\ell$:

$$\mathbb{E}\left[M_{\ell+1} - M_\ell \mid \mathcal{S}_\ell\right] \le \mathbf{1}_{\ell+1<\kappa'_\eta}\frac{\widehat{\mathbf{w}}(e_{\ell+1})\Delta|\mathcal{G}^\eta[\ell]|}{m-\ell}, \tag{4.71}$$

$$\mathbb{E}\left[(M_{\ell+1} - M_\ell)^2 \mid \mathcal{S}_\ell\right] \le \mathbf{1}_{\ell+1<\kappa'_\eta}\frac{\widehat{\mathbf{w}}(e_{\ell+1})^2\Delta|\mathcal{G}^\eta[\ell]|}{m-\ell}. \tag{4.72}$$

Fix any $\bar{\ell} = \Theta(\log(n))$ and observe that since $|\mathcal{G}^\eta[\ell]| \le \ell$, $\widehat{\mathbf{w}}(e_\ell) \le \frac{r}{\ell}$, we have

$$\widehat{\mathbf{w}}(e_{\ell+1})\Delta|\mathcal{G}^\eta[\ell]| = O(\log(n)), \quad \sum_{\ell \ge \bar{\ell}} \widehat{\mathbf{w}}(e_\ell) = O(\log^2(n)).$$

Set

$$a := \sum_{\ell \ge \bar{\ell}} \mathbb{E}\left[M_{\ell+1} - M_\ell \mid \mathcal{S}_\ell\right] = O\left(\log(n) n^{-\epsilon^2}\right) = o(1),$$
$$b := \sum_{\ell \ge \bar{\ell}} \mathbb{E}\left[(M_{\ell+1} - M_\ell)^2 \mid \mathcal{S}_\ell\right] = O\left(\log^3(n) n^{-1}\right).$$

Fix any $\varepsilon \in (0,1)$ and define

$$Z_{\ell+1} = \frac{4}{\varepsilon}\left(M_{\ell+1} - M_\ell - \mathbb{E}\left[M_{\ell+1} - M_\ell \mid \mathcal{S}_\ell\right]\right).$$

We observe that $\mathbb{E}[Z_{\ell+1}|\mathcal{S}_\ell] = 0$ and that $|Z_{\ell+1}| \le 1$, since if $\ell \ge \bar{\ell} = \omega(1)$, then $\mathbf{w}(e_{\ell+1}) \to 0$, and in particular $M_{\ell+1} - M_\ell \le \frac{\varepsilon}{4}$. We now focus on the martingale

$$W_u = \sum_{\ell = \bar{\ell}+1}^{u} Z_\ell, \qquad \forall u > \bar{\ell}.$$

We notice that

$$W_{\kappa'_\eta} = \frac{4}{\varepsilon}\left(M_{\kappa'_\eta} - M_{\bar{\ell}} - a\right) \qquad \text{and} \qquad b' := \sum_{\ell \ge \bar{\ell}} \mathrm{Var}(Z_\ell \mid \mathcal{S}_\ell) \le \frac{16}{\varepsilon^2} b.$$

A martingale version of Bennett's inequality introduced in [33, Theorem 1.6] ensures that, for $c > 0$,

$$\mathbb{P}_{\sigma,x}\left(\exists u \ge \bar{\ell} \text{ s.t. } W_u \ge c\right) \le e^c \left(\frac{b'}{c+b'}\right)^{c+b'}.$$

In particular

$$\mathbb{P}_{\sigma,x}\left(M_{\kappa'_\eta} - M_{\bar{\ell}} \ge \varepsilon\right) = \mathbb{P}_{\sigma,x}\left(\frac{\varepsilon}{4} W_{\kappa'_\eta} + a \ge \varepsilon\right)$$
$$\le \mathbb{P}_{\sigma,x}\left(\frac{\varepsilon}{4} W_{\kappa'_\eta} \ge \frac{\varepsilon}{2}\right) = \mathbb{P}_{\sigma,x}\left(W_{\kappa'_\eta} \ge 2\right) = o(n^{-1}).$$

We are left to show that for every $\varepsilon > 0$ it holds

$$\mathbb{P}_{\sigma,x}\left(M_{\bar{\ell}} \le \varepsilon\right) = 1 - o(n^{-1}).$$

It is easy to check that the probability of having 2 or more edges that are not open in the construction of the first ℓ edges of η is $o(n^{-1})$. Moreover, it trivially holds that for every $\ell \geq 0$, $\widehat{\mathbf{w}}(e_\ell) \leq 2^{-s}$. Hence,

$$\mathbb{P}_{\sigma,x}(M_{\bar{\ell}} \geq 2^{-s+1}) = o(n^{-1}),$$

which is clearly enough to derive the desired conclusions. □

At this point, the proof of (4.70) is just a collection of the results obtained so far.

Proposition 4.2 *For every $\varepsilon > 0$:*

$$\lim_{n\to\infty} \mathbb{P}\Big(\min_{x \in S_\star^\sigma} \mathbf{Q}_x^{\sigma,\eta}\Big((X_0,\dots,X_t) \in \cup_{y\in[n]} \mathcal{N}_{x,y}\Big) > 1-\varepsilon \Big) = 1.$$

Proof: We check the conditions in Definition 4.4 one by one:

1. follows from Lemma 4.5;
2. see [14, Proposition 13] and [13, Proposition 10];
3. The third requirement in Definition 4.4 follows from Lemma 4.7. Indeed, $M_{\kappa'_\eta}$ is greater or equal than the probability that a random walk starting at x in σ and visiting $\mathcal{L}_s^{\sigma,x}$ at time s, exits the forest $\mathcal{W}^\eta(r)$ in the time interval $s, t-h$;
4. Notice that in order to satisfy the fourth requirement of Definition 4.4 it is sufficient that $v_{t-h} \in S_\star^\eta$. Therefore we obtain the desired conclusion by noticing that $\max_{z\in[n]} \mathbf{P}_z^\eta(X_r \notin S_\star^\eta) \xrightarrow{\mathbb{P}} 0$, see [13, Proposition 6] and [14, Lemma 9].

□

We are now left with showing the validity of (4.68), which concludes the proof of the upper bound. Such a result is achieved by the following lemma, which is based on the constructions in Definitions 4.2 and 4.3 and on a variation of Azuma's inequality.

Lemma 4.8 *For every $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathbb{P}\left(\forall x, y \in [n],\ \bar{Q}_s^t(x,y) \leq (1+\varepsilon)\mu_{\rm in} P_\eta^h(y) + \tfrac{\varepsilon}{n}\right) = 1.$$

Proof: Fix $x, y \in [n]$. Construct, in the order, σ, $\mathcal{G}^\eta(r)$ as in Definition 4.3 and the in-neighborhood of y up to distance $h-1$. The latter can be constructed in the usual *breadth first* way, see Chapter 2. Let $\mathcal{S}$ denote the σ-field generated by this construction. Clearly, in the construction of the in-neighborhood of y we cannot reveal more than $\Delta^h = o(n)$

edges. Therefore, by Lemma 4.6 at most $o(n)$ edges of η have been revealed up to this point. Let $\mathcal{E}_{\mathcal{W}}$ denote the tails of the leaves of $\mathcal{W}^{\eta}(r)$ and call $\mathcal{F}_y$ the set of heads of the vertices v in the boundary of the in-neighborhood of y such that there is a unique path of length at most $h-1$ to y in the configuration η. Associate to each head $f \in \mathcal{F}_y$ the quantity

$$\widehat{\mathbf{w}}'(f) = P_{\eta}^{h-1}(v(f), y).$$

At this point we notice that by definition of nice paths,

$$\bar{Q}_s^t(x,y) = \sum_{e \in \mathcal{E}_{\mathcal{W}}} \widehat{\mathbf{w}}(e) \sum_{f \in \mathcal{F}_y} \widehat{\mathbf{w}}'(f) \mathbf{1}_{\widehat{\mathbf{w}}(e)\widehat{\mathbf{w}}'(f) \le n^{-1-\upsilon/3}} \mathbf{1}_{\eta(e)=f}.$$

We remark that

$$\sum_{f \in \mathcal{F}_y} \widehat{\mathbf{w}}'(f) \le \mu_{\mathrm{in}} P_{\eta}^{h}(y), \qquad \sum_{e \in \mathcal{E}_{\mathcal{W}}} \widehat{\mathbf{w}}(e) \le 1.$$

Since a matching $\eta(e) = f$ of $e \in \mathcal{E}_{\mathcal{W}}$ and $f \in \mathcal{F}_y$ can only occur after the generation of $\sigma, \mathcal{G}^{\eta}(r), \mathcal{F}_y$,

$$\mathbb{E}[\bar{Q}_s^t(x,y) \mid \mathcal{S}] \le \mu_{\mathrm{in}} P_{\eta}^{h}(y),$$

where we use $\mathbb{E}[\mathbf{1}_{\eta(e)=f} \mid \mathcal{S}] = \frac{1}{m}(1+o(1))$. We rewrite $Z := \bar{Q}_s^t(x,y) = \sum_{e \in \mathcal{E}_{\mathcal{W}}} c(e, \eta(e))$, where

$$c(e,f) = \widehat{\mathbf{w}}(e)\widehat{\mathbf{w}}'(f)\mathbf{1}_{\widehat{\mathbf{w}}(e)\widehat{\mathbf{w}}'(f) \le n^{-1-\upsilon/3}} \le n^{-1-\upsilon/3} =: \|c\|_{\infty}.$$

We can now invoke the concentration inequality (see [21, Proposition 1.1] and [13, Section 6.2])

$$\mathbb{P}\left(Z - \mathbb{E}[Z|\mathcal{S}] > a \mid \mathcal{S}\right) \le \exp\left(-\frac{a^2}{2\|c\|_{\infty}(2\mathbb{E}[Z|\mathcal{S}] + a)}\right)$$

and by choosing $a := \frac{\varepsilon}{2}\mathbb{E}[Z|\mathcal{S}] + \frac{\varepsilon}{n}$ we obtain a probability bounded by $o(n^{-2})$ for every fixed choice of x, y. Taking a union bound we conclude the desired result. $\square$

CHAPTER 5

EXTREMAL VALUES OF THE STATIONARY DISTRIBUTION

This chapter is devoted to the study of the stationary distribution of the simple random walk on a graph G from the DCM($\mathbf{d}^{\pm}$) ensemble. As remarked in Chapter 1, our focus in on give sharp bounds on the extremal values of the stationary distribution, which we will refer to as $\pi_{\min}$ and $\pi_{\max}$. The main results are reported in Theorems 5.1 and 5.2. To discuss them, it is convenient to introduce the following notation.

Definition 5.1 *Recalling the Definition 2.1, we define the parameters*

$$\gamma_0 := \frac{\log \Delta_+}{\log \delta_-}\,, \qquad \gamma_1 := \max_{(k,\ell)\in\mathcal{L}} \frac{\log \ell}{\log k}\,, \qquad \kappa_1 := \min_{(k,\ell)\in\mathcal{L}} \frac{\log \ell}{\log k}\,, \qquad \kappa_0 := \frac{\log \delta_+}{\log \Delta_-}. \tag{5.1}$$

Theorem 5.1 *Set $\pi_{\min} = \min_{x\in[n]} \pi(x)$. There exists a constant $C > 0$ such that*

$$\mathbb{P}\left(C^{-1}\log^{1-\gamma_0}(n) \le n\pi_{\min} \le C\,\log^{1-\gamma_1}(n)\right) = 1 - o(1). \tag{5.2}$$

Moreover, there exists $\beta > 0$ such that

$$\mathbb{P}\Big(\exists S \subset [n],\ |S| \ge n^{\beta}\,,\ n\max_{y\in S}\pi(y) \le C\log^{1-\gamma_1}(n)\Big) = 1 - o(1). \tag{5.3}$$

Remark 5.1 *Notice that* $\gamma_0 \geq \gamma_1 \geq 1$. *If the sequences* $\mathbf{d}^{\pm}$ *are such that* $(\delta_-, \Delta_+) \in \mathcal{L}$, *then* $\gamma_0 = \gamma_1 =: \gamma$, *so in these cases* Theorem 5.1 *implies that*

$$\pi_{\min} \asymp \frac{1}{n} \log^{1-\gamma}(n) \qquad \text{w.h.p.} \tag{5.4}$$

In all other cases, the estimate (5.2) *can be strengthened by replacing* γ_0 *with* γ_0' *where*

$$\gamma_0' := \frac{\log \Delta_+'}{\log \delta_-'}, \qquad \Delta_+' := \max\{\ell : (k,\ell) \in \mathcal{L}_0\}, \quad \delta_-' := \min\{k : (k,\ell) \in \mathcal{L}_0\}, \tag{5.5}$$

and $\mathcal{L}_0 \subset \mathcal{C}$ *is defined as the set of* $(k,\ell) \in \mathcal{C}$ *such that*

$$\limsup_{n\to\infty} \frac{|\mathcal{V}_{k,\ell}|}{n^{1-a}} = +\infty, \qquad \forall a > 0. \tag{5.6}$$

We refer to Remark 5.3 *below for additional details on this improvement.*

Concerning the maximal values of π we establish the following estimates.

Theorem 5.2 *Set* $\pi_{\max} = \max_{x\in[n]} \pi(x)$. *There exists a constant* $C > 0$ *such that*

$$\mathbb{P}\left(C^{-1}\log^{1-\kappa_1}(n) \leq n\pi_{\max} \leq \log^{1-\kappa_0}(n)\right) = 1 - o(1). \tag{5.7}$$

Moreover, there exists $\beta > 0$ *such that*

$$\mathbb{P}\Big(\exists S \subset [n],\ |S| \geq n^{\beta},\ n \min_{y\in S} \pi(y) \geq C^{-1}\log^{1-\kappa_1}(n)\Big) = 1 - o(1). \tag{5.8}$$

Remark 5.2 *Notice that* $\kappa_0 \leq \kappa_1 \leq 1$. *If the sequences* $\mathbf{d}^{\pm}$ *are such that* $(\Delta_-, \delta_+) \in \mathcal{L}$, *then* $\kappa_0 = \kappa_1 =: \kappa$, *and in these cases* Theorem 5.2 *implies*

$$\pi_{\max} \asymp \frac{1}{n} \log^{1-\kappa}(n) \qquad \text{w.h.p.} \tag{5.9}$$

In analogy with Remark 5.1, *if* $(\Delta_-, \delta_+) \notin \mathcal{L}$, *then* (5.7) *can be improved by replacing* κ_0 *with* κ_0' *where*

$$\kappa_0' := \frac{\log \delta_+'}{\log \Delta_-'}, \qquad \delta_+' := \min\{\ell : (k,\ell) \in \mathcal{L}_0\}, \quad \Delta_-' := \max\{k : (k,\ell) \in \mathcal{L}_0\}, \tag{5.10}$$

5.1 The local approximation

We saw in Chapter 2 that the stationary distribution at a typical vertex y admits an approximation in terms of the in-neighbourhood of y at a distance that is much smaller than the mixing time. More precisely, we saw in Lemma 3.1 that for any sequence $t_n \to \infty$

$$\|\pi - \mu_{\text{in}} P^{t_n}\|_{\text{TV}} \xrightarrow{\mathbb{P}} 0. \tag{5.11}$$

While these facts are very useful to study the typical values of π, they give very poor information on its extremal values $\pi_{\min}$ and $\pi_{\max}$, and to prove Theorem 5.1 and Theorem 5.2 we need a stronger control of the local approximation of the stationary distribution.

A key role in our analysis is played by the quantity $\Gamma_h(y)$ defined as follows. Consider the set $\partial\mathcal{B}_h^-(y)$ of all vertices $z \in [n]$ such that $d(z, y) = h$, and define

$$\Gamma_h(y) := \sum_{z \in \partial\mathcal{B}_h^-(y)} d_z^- \, P^h(z, y). \tag{5.12}$$

The definitions (5.12) is such that for any $y \in [n]$ and $h \in \mathbb{N}$

$$\Gamma_h(y) \le m \, \mu_{\text{in}} P^h(y). \tag{5.13}$$

If $\mathcal{B}_h^-(y)$ is a tree, then (5.13) is an equality. In any case, $\Gamma_h(y)$ satisfies the following rough inequalities.

Lemma 5.1 *With high probability, for all $y \in [n]$, for all $h \in [1, \hbar]$:*

$$\left(\frac{\delta_-}{\Delta_+}\right)^h \le \Gamma_h(y) \le 2\Delta_- \left(\frac{\Delta_-}{\delta_+}\right)^h. \tag{5.14}$$

Proof: From Proposition 2.3 we may assume that the event $\mathcal{G}(\hbar)$ holds. From Lemma 2.3 we know that $\frac{1}{2}\delta_-^h \le |\partial\mathcal{B}_h^-(y)| \le \Delta_-^h$. Thus it suffices to show that for any $z \in \partial\mathcal{B}_h^-(y)$, $h \in [1, \hbar]$:

$$\Delta_+^{-h} \le P^h(z, y) \le 2\delta_+^{-h}. \tag{5.15}$$

The bounds in (5.15) follow from the observation that any path of length h from z to y has weight at least Δ_+^{-h} and at most δ_+^{-h}, and that there is at least one and at most two such paths if $z \in \partial\mathcal{B}_h^-(y)$ and $\mathcal{G}(\hbar)$ holds. The latter fact can be seen with the same argument used in the proof of Lemma 2.3. With reference to that proof: in case 1) there are at most two paths from z to y, see Fig. 2.1; in case 2) there is only one path from z to y; see Fig. 2.2

and Fig. 2.3. □

Roughly speaking, in what follows the extremal values of π will be controlled by approximating $\pi(y)$ in terms of $\Gamma_h(y)$ for values of h of order $\log\log n$, for every node y. The next two results allow us to control $\Gamma_h(y)$ in terms of $\Gamma_{h_0}(y)$ for all $h \in [h_0, \hbar]$ where h_0 is of order $\log\log n$.

Lemma 5.2 *There exist constants $c > 0$ and $C > 0$ such that:*

$$\mathbb{P}\left(\forall y \in [n],\ \forall h \in [h_0, \hbar]\,,\ \Gamma_h(y) \geq c\log^{1-\gamma_0}(n)\right) = 1 - o(1), \tag{5.16}$$

where γ_0 is the constant from Theorem 5.1 and $h_0 := \log_{\delta_-}\log(n) + C$.

Proof: From Lemma 2.3 we may assume that $|\partial\mathcal{B}^-_{h_0}(y)| \geq \frac{1}{2}\delta_-^{h_0} =: R$ for all $y \in [n]$, where h_0 is as in the statement above with C to be fixed later. Once we have the in-neighbourhood $\mathcal{B}^-_{h_0}(y)$ we proceed with the generation of the $(h - h_0)$-in-neghborhoods of all $z \in \partial\mathcal{B}^-_{h_0}(y)$. Consider the first R elements of $\partial\mathcal{B}^-_{h_0}(y)$, and order them as $(z_1, \ldots, z_R)$ in some arbitrary way. We sample sequentially $\mathcal{B}^-_{h-h_0}(z_1)$, then $\mathcal{B}^-_{h-h_0}(z_2)$, and so on. We want to couple the random variables $Z_i := \mathcal{B}^-_{h-h_0}(z_i)$, $i = 1, \ldots, R$ with a sequence of independent rooted directed random trees W_i, $i = 1, \ldots, R$, defined as follows. The tree W_i is defined as the first $h - h_0$ generations of the marked random tree $\mathcal{T}_i$ produced by the following instructions:

- the root is given the mark z_i;
- every vertex with mark j has d_j^- children, each of which is given independently the mark $k \in [n]$ with probability d_k^+/m.

Consider the generation of the i-th variable Z_i. This is achieved by the breadth-first sequential procedure, where at each step a head is matched with a tail chosen uniformly at random from all unmatched tails; see Section 2.3.1. If instead we pick the tail uniformly at random from all possible tails, then we need to reject the outcome if the chosen tail belongs to the set of tails that have been already matched. Since the total number of tails matched at any step of this generation is at most $K := \Delta^{\hbar} = O(n^{1/5})$, it follows that the probability of a rejection is bounded by $p := K/m = O(n^{-4/5})$. Let us now consider the event of a collision, that is when the chosen tail belongs to a vertex that has already been exposed during the previous steps, including the generation of $\mathcal{B}^-_{h_0}(y)$ and of the Z_j, $j \leq i$. Notice that the total number of exposed vertices is at most K and therefore the probability of a collision is bounded by $p' = \Delta K/m = O(n^{-4/5})$. Since the generation of Z_i requires at most K matchings, we see that conditionally on the past, a Z_i with no rejections and no collisions is created with probability uniformly bounded from below by $1 - q$, where

$q = O(n^{-3/5})$. We say that Z_i is *bad* if its generation produced a rejection or a collision. Once the Z_i's have been sampled we define a set $\mathcal{I}$ such that $i \in \mathcal{I}$ if and only if either Z_i is bad or there is a bad Z_j such that the generation of Z_j produced a collision with a vertex from Z_i. With this notation, $W_i = Z_i$ for all $i \notin \mathcal{I}$ and

$$\Gamma_h(y) \geq \Delta_+^{-h_0} \sum_{i \notin \mathcal{I}} \Gamma_{h-h_0}(z_i). \tag{5.17}$$

The above construction shows that the cardinality of the set $\mathcal{I}$ is stochastically dominated by twice the binomial $\mathrm{Bin}(R, q)$. Therefore,

$$\mathbb{P}(|\mathcal{I}| \geq 10) \leq \mathbb{P}(\mathrm{Bin}(R, q) \geq 5) \leq (Rq)^5 = o(n^{-2}). \tag{5.18}$$

On the other hand, notice that for all $i \notin \mathcal{I}$:

$$\Gamma_{h-h_0}(z_i) = M^i_{h-h_0}, \tag{5.19}$$

where M^i_t, $t \in \mathbb{N}$, is defined as follows. Let $\mathcal{T}_{t,i}$ denote the set of vertices forming generation t of the tree $\mathcal{T}_i$ rooted at z_i, and for $x \in \mathcal{T}_{t,i}$, write

$$\mathbf{w}(x) := \mathbf{w}\left(x \mapsto z_i; \mathcal{T}_i\right) = \prod_{u=1}^{t} \frac{1}{d^+_{x_u}}, \tag{5.20}$$

for the weight of the path $(x_t = x, x_{t-1}, \ldots, x_1, x_0 = z_i)$ from x to z_i along $\mathcal{T}_i$. Then M^i_t is defined by

$$M^i_t = \sum_{x \in \mathcal{T}_{t,i}} d^-_x \mathbf{w}(x), \qquad M^i_0 = d^-_{z_i}. \tag{5.21}$$

It is not hard to check (see Section 2.2) that for fixed n, $(M^i_t)_{t \geq 0}$ is a martingale with

$$\mathbb{E}[M^i_t] = M^i_0 = d^-_{z_i}.$$

In particular, by truncating at a sufficiently large constant $C_1 > 0$ one has $M^i_{h-h_0} \geq X_i$, where

$$X_i := \min\{M^i_{h-h_0}, C_1\}$$

are independent random variables with $0 \leq X_i \leq C_1$ and $\mathbb{E}[X_i] \geq 1$ for all i. Therefore, Hoeffding's inequality gives, for any $k \in \mathbb{N}$:

$$\mathbb{P}\Big(\sum_{i=1}^{k} M^i_{h-h_0} \leq k/2\Big) \leq e^{-c_1 k}, \tag{5.22}$$

where $c_1 > 0$ is a suitable constant.

Divide the integers $\{1, \dots, R\}$ into 10 disjoint intervals $I_1, \dots, I_{10}$, each containing $R/10$ elements. If $|\mathcal{I}| < 10$ then there must be one of the intervals, say I_{j_*}, such that $I_{j_*} \cap \mathcal{I} = \emptyset$. It follows that if $|\mathcal{I}| < 10$, then

$$\sum_{i \notin \mathcal{I}} \Gamma_{h-h_0}(z_i) \geq \sum_{i \in I_{j^*}} M^i_{h-h_0} \geq \min_{\ell=1,\dots,10} \sum_{i \in I_\ell} M^i_{h-h_0}. \tag{5.23}$$

Using (5.18), and (5.22)-(5.23) we conclude that, for a suitable constant $c_2 > 0$:

$$\begin{aligned}\mathbb{P}\Big(\sum_{i \notin \mathcal{I}} \Gamma_{h-h_0}(z_i) \leq c_2 R\Big) &\leq \mathbb{P}\Big(\min_{\ell=1,\dots,10} \sum_{i \in I_\ell} M^i_{h-h_0} \leq c_2 R\Big) + \mathbb{P}(|\mathcal{I}| \geq 10) \\ &\leq 10 \exp\left(-c_1 R/10\right) + o(n^{-2}). \end{aligned} \tag{5.24}$$

Since $R = \frac{1}{2}\delta_-^{h_0} = \frac{1}{2}\delta_-^{C} \log n$, the probability in (5.24) is $o(n^{-2})$ if C is large enough. From (5.17), on the event $\sum_{i \notin \mathcal{I}} \Gamma_{h-h_0}(z_i) > c_2 R$ one has

$$\Gamma_h(y) \geq \tfrac{1}{2} c_2 \delta_-^{h_0} \Delta_+^{-h_0} = c \log^{1-\gamma_0}(n), \tag{5.25}$$

where $c = \frac{1}{2}c_2(\delta_-/\Delta_+)^C$. Thus the event (5.25) has probability $1 - o(n^{-2})$, and the desired conclusion follows by taking a union bound over $y \in [n]$ and $h \in [h_0, \hbar]$. $\square$

Lemma 5.3 *There exists a constant $K > 0$ such that for all $\varepsilon > 0$, with high probability:*

$$\max_{y \in [n]} \max_{h \in [h_1, \hbar]} \left| \frac{\Gamma_h(y)}{\Gamma_{h_1}(y)} - 1 \right| \leq \varepsilon, \tag{5.26}$$

where $h_1 := K \log\log(n)$.

Proof: For any $h \geq h_1$, let σ_h denote a realization of the in-neighbourhood $\mathcal{B}_h^-(y)$, obtained with the usual breadth-first sequential generation. From Proposition 2.3 we may assume that the tree excess of $\mathcal{B}_h^-(y)$ is at most 1, as long as $h \leq \hbar$. Call $\mathcal{E}_{tot,h}, \mathcal{F}_{tot,h}$ the set of unmatched tails and unmatched heads, respectively, after the generation of σ_h. Let also $\mathcal{E}_h \subset \mathcal{E}_{tot,h}$ denote the set of unmatched tails belonging to vertices not yet exposed, and let $\mathcal{F}_h$ be the subset of heads attached to $\partial\mathcal{B}_h^-(y)$. By construction, all heads attached to $\partial\mathcal{B}_h^-(y)$ must be unmatched at this stage so that $\mathcal{F}_h \subset \mathcal{F}_{tot,h}$. Moreover,

$$\Gamma_h(y) = \sum_{f \in \mathcal{F}_h} P^h(v_f, y), \tag{5.27}$$

where v_f denotes the vertex to which the head f belongs. To compute Γ_{h+1} given σ_h we let $\omega : \mathcal{E}_{tot,h} \mapsto \mathcal{F}_{tot,h}$ denote a uniform random matching of $\mathcal{E}_{tot,h}$ and $\mathcal{F}_{tot,h}$, and notice that a vertex z is in $\partial\mathcal{B}^-_{h+1}(y)$ if and only if z is revealed by matching one of the heads $f \in \mathcal{F}_h$ with one of the tails $e \in \mathcal{E}_h$. Therefore,

$$\begin{aligned}\Gamma_{h+1}(y) &= \sum_{e\in\mathcal{E}_h} \frac{d_e^-}{d_e^+} \sum_{f\in\mathcal{F}_h} P^h(v_f, y)\mathbf{1}_{\omega(e)=f} \\ &= \sum_{e\in\mathcal{E}_{tot,h}} c(e,\omega(e)),\end{aligned} \tag{5.28}$$

where we use the notation $d_e^\pm$ for the degrees of the vertex to which the tail e belongs, and the function c is defined by

$$c(e,f) = \frac{d_e^-}{d_e^+} P^h(v_f,y)\mathbf{1}_{e\in\mathcal{E}_h, f\in\mathcal{F}_h}. \tag{5.29}$$

Since σ_h is such that $\texttt{TX}(\mathcal{B}^-_h(y)) \le 1$, we may estimate $P^h(v_f,y)$ as in (5.15), so that

$$\|c\|_\infty = \max_{e,f} c(e,f) \le 2\Delta\, \delta^{-h-1}. \tag{5.30}$$

We now use a version of Bernstein's inequality proved by Chatterjee ([21, Proposition 1.1]) which applies to any function of a uniform random matching of the form (5.28). It follows that for any fixed σ_h, for any $s > 0$:

$$\mathbb{P}\left(|\Gamma_{h+1}(y) - \mathbb{E}\left[\Gamma_{h+1}(y)\,|\,\sigma_h\right]| \ge s \,|\, \sigma_h\right) \le 2\exp\left(-\frac{s^2}{2\|c\|_\infty \left(2\mathbb{E}\left[\Gamma_{h+1}(y)\,|\,\sigma_h\right] + s\right)}\right). \tag{5.31}$$

Taking $s = a\mathbb{E}\left[\Gamma_{h+1}(y)\,|\,\sigma_h\right]$, $a \in (0,1)$, one has

$$\mathbb{P}\left(|\Gamma_{h+1}(y) - \mathbb{E}\left[\Gamma_{h+1}(y)\,|\,\sigma_h\right]| \ge s \,|\, \sigma_h\right) \le 2\exp\left(-\frac{a^2\mathbb{E}\left[\Gamma_{h+1}(y)\,|\,\sigma_h\right]}{6\|c\|_\infty}\right). \tag{5.32}$$

Since the probability of the event $\omega(e) = f$ conditioned on σ_h is $\frac{1}{|\mathcal{E}_{tot,h}|} = \frac{1}{m}(1 + O(\Delta^h/m))$, we have

$$\begin{aligned}\mathbb{E}\left[\Gamma_{h+1}(y)\,|\,\sigma_h\right] &= \frac{1}{|\mathcal{E}_{tot,h}|}\sum_{e\in\mathcal{E}_h}\frac{d_e^-}{d_e^+}\Gamma_h(y) \\ &= \frac{1}{m}\left(1 + O(\Delta^h/m)\right)\left(m - \sum_{e\notin\mathcal{E}_h}\frac{d_e^-}{d_e^+}\right)\Gamma_h(y) \\ &= \left(1 + O(\Delta^h/m)\right)\Gamma_h(y) = \left(1 + O(n^{-1/2})\right)\Gamma_h(y),\end{aligned} \tag{5.33}$$

for all $h \in [h_1, \hbar]$, where we use the fact that the sum over all tails e (matched or unmatched) of d_e^-/d_e^+ equals m. In particular, from Lemma 5.2 it follows that for some constant $c > 0$:

$$\mathbb{E}\left[\Gamma_{h+1}(y) \mid \sigma_h\right] \geq c \log^{-\gamma_0+1}(n), \tag{5.34}$$

and therefore, using (5.30), one finds

$$\|c\|_\infty^{-1} \mathbb{E}\left[\Gamma_{h+1}(y) \mid \sigma_h\right] \geq \log^6(n), \tag{5.35}$$

for all $h \geq h_1$, if the constant K in the definition of h_1 is large enough. From (5.32), (5.33) and (5.35) it follows that, letting

$$\mathcal{A} := \left\{|\Gamma_{h+1}(y) - \Gamma_h(y)| \leq a\Gamma_h(y)\,,\ \forall h \in [h_1, \hbar]\right\},$$

with $a := \log^{-2}(n)$, then

$$\mathbb{P}\left(\mathcal{A}\right) = 1 - o(1). \tag{5.36}$$

Moreover, on the event $\mathcal{A}$, for all $h \in [h_1, \hbar]$:

$$|\Gamma_h(y) - \Gamma_{h_1}(y)| \leq \sum_{j=h_1}^{h-1} |\Gamma_{j+1}(y) - \Gamma_j(y)| \leq \varepsilon \Gamma_{h_1}(y).$$

□

5.2 Lower bound on $\pi_{\min}$

If for some $t \in \mathbb{N}$ and $a > 0$ one has $P^t(x, y) \geq a$ for all $x, y \in [n]$, then

$$\pi(z) = \sum_{x=1}^{n} \pi(x) P^t(x, z) \geq a, \tag{5.37}$$

and therefore $\pi_{\min} \geq a$. We will prove the lower bound on $P^t(x, y)$ by choosing t of the form $t = (1+\varepsilon)T_{\mathrm{ENT}}$, for some small enough $\varepsilon > 0$. More precisely, fix a constant $\eta > 0$, set $\eta' = 3\eta \frac{H}{\log \delta}$, and define

$$t_\star = h_x + h_y + 1\,, \quad h_x = (1-\eta)T_{\mathrm{ENT}}\,, \quad h_y = \eta' T_{\mathrm{ENT}}. \tag{5.38}$$

Note that $\eta' \geq 3\eta$ and thus $t_\star = t_\star(\eta) \geq (1+2\eta)T_{\mathrm{ENT}}$.

Lemma 5.4 *There exists $\eta_0 > 0$ such that for all $\eta \in (0, \eta_0)$:*

$$\mathbb{P}\left(\forall x, y \in [n],\ \ P^{t_\star+1}(x,y) \geq \tfrac{c}{n}\,\Gamma_{h_y}(y)\right) = 1 - o(1), \tag{5.39}$$

for some constant $c = c(\eta, \Delta) > 0$.

From (5.37) and Lemma 5.4 it follows that w.h.p. for all y

$$\pi(y) \geq \tfrac{c}{n}\,\Gamma_{h_y}(y). \tag{5.40}$$

Lemma 5.2 thus implies, for some new constant $c > 0$

$$\mathbb{P}\left(\pi_{\min} \geq \tfrac{c}{n}\log^{1-\gamma_0}(n)\right) = 1 - o(1), \tag{5.41}$$

which settles the lower bound in Theorem 5.1.

To prove Lemma 5.4 we will restrict to a subset of *nice* paths from x to y. This will allow us to obtain a concentration result for the probability to reach y from x in $t_\star$ steps.

5.2.1 A concentration result for nice paths

The definition of the nice paths follows a construction introduced in [13], which we now recall. In contrast with [13] however, here we need a lower bound on $P^{t_\star}(x,y)$ and thus the argument is somewhat different.

Following [13, Section 6.2] and [14, Section 4.1], we introduce the rooted directed tree $\mathcal{T}(x)$, namely the subgraph of the h_x-out-neighbourhood of x defined by the following process: initially all tails and heads are unmatched and $\mathcal{T}(x)$ is identified with its root, x; throughout the process, we let $\partial_+\mathcal{T}(x)$ (resp. $\partial_-\mathcal{T}(x)$) denote the set of unmatched tails (resp. heads) whose endpoint belongs to $\mathcal{T}(x)$; the height $\mathbf{h}(e)$ of a tail $e \in \partial_+\mathcal{T}(x)$ is defined as 1 plus the number of edges in the unique path in $\mathcal{T}(x)$ from x to the endpoint of e; the weight of $e \in \partial_+\mathcal{T}(x)$ is defined as

$$\mathbf{w}(e) = \prod_{i=0}^{\mathbf{h}(e)-1} \frac{1}{d^+_{x_i}}\,, \tag{5.42}$$

where $(x = x_0, x_1, \ldots, x_{\mathbf{h}(e)-1})$ denotes the path in $\mathcal{T}(x)$ from x to the endpoint of e; we then iterate the following steps:

- a tail $e \in \partial_+\mathcal{T}(x)$ is selected with maximal weight among all $e \in \partial_+\mathcal{T}(x)$ with $\mathbf{h}(e) \leq h_x - 1$ and $\mathbf{w}(e) \geq \mathbf{w}_{min} := n^{-1+\eta^2}$ (using an arbitrary ordering of the tails to break ties);

- e is matched to a uniformly chosen unmatched head f, forming the edge ef;
- if f was not in $\partial_- \mathcal{T}(x)$, then its endpoint and the edge ef are added to $\mathcal{T}(x)$.

The process stops when there are no tails $e \in \partial_+ \mathcal{T}(x)$ with height $\mathbf{h}(e) \le h_x - 1$ and weight $\mathbf{w}(e) \ge \mathbf{w}_{min}$. Note that $\mathcal{T}(x)$ remains a directed tree at each step. The final value of $\mathcal{T}(x)$ represents the desired directed tree. After the generation of the tree $\mathcal{T}(x)$ a total number κ of edges has been revealed, some of which may not belong to $\mathcal{T}(x)$. As in [14, Lemma 7], it is not difficult to see that when exploring the out-neighbourhood of x in this way the random variable κ is deterministically bounded as

$$\kappa \le n^{1-\frac{\eta^2}{2}}. \tag{5.43}$$

At this stage, let us call $\mathcal{E}^*(x)$ the set of unmatched tails $e \in \partial_+ \mathcal{T}(x)$ such that $\mathbf{h}(e) = h_x$.

Definition 5.2 *A path $\mathbf{p} = (x_0 = x, x_1, \ldots, x_{t_\star} = y)$ of length $t_\star$ starting at x and ending at y is called* nice *if it satisfies:*

1. *The first h_x steps of $\mathbf{p}$ are contained in $\mathcal{T}(x)$, and satisfy*

$$\prod_{i=0}^{h_x} \frac{1}{d_{x_i}^+} \le n^{2\eta - 1};$$

2. *$x_{h_x+1} \in \partial \mathcal{B}_{h_y}^-(y)$.*

To obtain a useful expression for the probability of going from x to y along a nice path, we need to generate $\mathcal{B}_{h_y}^-(y)$, the h_y-in-neighbourhood of y. To this end, assume that κ edges in the h_x-out-neighbourhood of x have been already sampled according to the procedure described above, and then sample $\mathcal{B}_{h_y}^-(y)$ according to the sequential generation described in Section 2.3.1. Some of the matchings producing $\mathcal{B}_{h_y}^-(y)$ may have already been revealed during the previous stage. In any case, this second stage creates an additional random number τ of edges, satisfying the crude bound $\tau \le \Delta^{h_y+1}$. We call $\mathcal{F}_{tot}$ the set of unmatched heads, and $\mathcal{E}_{tot}$ the set of unmatched tails after the sampling of these $\kappa + \tau$ edges. Consider the set $\mathcal{F}^0 := \mathcal{F}_{h_y} \cap \mathcal{F}_{tot}$, where $\mathcal{F}_{h_y}$ denotes the set of all heads (matched or unmatched) attached to vertices in $\partial \mathcal{B}_{h_y}^-(y)$. Moreover, call $\mathcal{E}^0 := \mathcal{E}^*(x) \cap \mathcal{E}_{tot}$ the subset of unmatched tails which are attached to vertices at height h_x in $\mathcal{T}(x)$. Finally, complete the generation of the digraph by matching the $m - \kappa - \tau$ unmatched tails $\mathcal{E}_{tot}$ to the $m - \kappa - \tau$ unmatched heads $\mathcal{F}_{tot}$ using a uniformly random bijection $\omega : \mathcal{E}_{tot} \mapsto \mathcal{F}_{tot}$. For any $f \in \mathcal{F}_{h_y}$ we introduce the notation

$$\mathbf{w}(f) := P^{h_y}(v_f, y), \tag{5.44}$$

where v_f denotes the vertex $v \in \partial\mathcal{B}^-_{h_y}(y)$ such that $f \in E^-_v$. With the notation introduced above, the probability to go from x to y in $t_\star$ steps following a nice path can now be written as

$$P_{0,t_\star}(x,y) := \sum_{e\in\mathcal{E}^0}\sum_{f\in\mathcal{F}^0} \mathbf{w}(e)\mathbf{w}(f)\mathbf{1}_{\omega(e)=f}\mathbf{1}_{\mathbf{w}(e)\le n^{2\eta-1}}. \tag{5.45}$$

Note that, conditionally on the construction of the first $\kappa+\tau$ edges described above, each Bernoulli random variable $\mathbf{1}_{\omega(e)=f}$ appearing in the above sum has probability of success at least $1/m$. In particular, if σ denotes a fixed realization of the $\kappa+\tau$ edges, then

$$\mathbb{E}\left[P_{0,t_\star}(x,y)\,|\,\sigma\right] \ge \frac{1}{m}\,A_{x,y}(\sigma)B_{x,y}(\sigma)\,, \tag{5.46}$$

where

$$A_{x,y}(\sigma) := \sum_{e\in\mathcal{E}^0}\mathbf{1}_{\mathbf{w}(e)\le n^{2\eta-1}}\mathbf{w}(e)\,,\quad B_{x,y}(\sigma) := \sum_{f\in\mathcal{F}^0}\mathbf{w}(f). \tag{5.47}$$

Moreover, the probability of $\omega(e)=f$ for any fixed $e\in\mathcal{E}^0$, $f\in\mathcal{F}^0$ is at most $1/(m-\kappa-\tau)$, so that

$$\mathbb{E}\left[P_{0,t_\star}(x,y)\,|\,\sigma\right] \le \frac{(1+o(1))}{m}\,A_{x,y}(\sigma)B_{x,y}(\sigma) \le \frac{(1+o(1))}{m}\,\Gamma_{h_y}(y), \tag{5.48}$$

where we use $A_{x,y}\le 1$ and $B_{x,y}\le\Gamma_{h_y}(y)$. Consider the event

$$\mathcal{Y}_{x,y} = \left\{\sigma:\ A_{x,y}(\sigma)\ge\tfrac{1}{2}\,,\ B_{x,y}(\sigma)\ge\log^{-\gamma_0}(n)\,,\ \mathrm{TX}(\mathcal{B}^-_{h_y}(y))\le 1\right\}, \tag{5.49}$$

where the exponent $-\gamma_0$ is chosen for convenience only and any exponent $-c$ with $c>\gamma_0-1$ would be as good.

Lemma 5.5 *There exists $\eta_0>0$ such that for all $\eta\in(0,\eta_0)$, for any $\sigma\in\mathcal{Y}_{x,y}$, any $a\in(0,1)$:*

$$\mathbb{P}\left(\left|P_{0,t_\star}(x,y)-\mathbb{E}\left[P_{0,t_\star}(x,y)\,|\,\sigma\right]\right|\ge a\mathbb{E}\left[P_{0,t_\star}(x,y)\,|\,\sigma\right]\,\middle|\,\sigma\right)\le 2\exp\left(-a^2n^{\eta/2}\right) \tag{5.50}$$

Proof: Conditioned on σ, $P_{0,t_\star}(x,y)$ is a function of the uniform random permutation $\omega:\mathcal{E}_{tot}\mapsto\mathcal{F}_{tot}$,

$$P_{0,t_\star}(x,y) = \sum_{e\in\mathcal{E}_{tot}} c(e,\omega(e))\,,\quad c(e,f)=\mathbf{w}(e)\mathbf{w}(f)\mathbf{1}_{\mathbf{w}(e)\le n^{2\eta-1}}\mathbf{1}_{e\in\mathcal{E}^0,f\in\mathcal{F}^0}. \tag{5.51}$$

Since we are assuming $\mathrm{TX}(\mathcal{B}^-_{h_y}(y))\le 1$, we can use (5.15) to estimate $\mathbf{w}(f)\le 2\delta^{-h_y}=n^{-3\eta}$ for any $f\in\mathcal{F}^0$. Therefore

$$\|c\|_\infty = \max_{e,f} c(e,f)\le 2n^{-1-\eta}. \tag{5.52}$$

As in Lemma 5.3, and as in [13], we use Chatterjee's concentration inequality for uniform random matchings [21, Proposition 1.1] to obtain for any $s > 0$:

$$\mathbb{P}\left(\left|P_{0,t_\star}(x,y) - \mathbb{E}\left[P_{0,t_\star}(x,y)\,|\,\sigma\right]\right| \geq s\,|\,\sigma\right) \leq 2\exp\left(-\frac{s^2}{2\,\|c\|_\infty\left(2\mathbb{E}\left[P_{0,t_\star}(x,y)\,|\,\sigma\right]+s\right)}\right). \tag{5.53}$$

Taking $s = a\mathbb{E}\left[P_{0,t_\star}(x,y)\,|\,\sigma\right]$, $a \in (0,1)$, one has

$$\mathbb{P}\left(\left|P_{0,t_\star}(x,y) - \mathbb{E}\left[P_{0,t_\star}(x,y)\,|\,\sigma\right]\right| \geq s\,|\,\sigma\right) \leq 2\exp\left(-\frac{a^2\mathbb{E}\left[P_{0,t_\star}(x,y)\,|\,\sigma\right]}{6\,\|c\|_\infty}\right). \tag{5.54}$$

Using (5.46), (5.49), and (5.52) one concludes that (5.50) holds for all $\sigma \in \mathcal{Y}_{x,y}$ and for all n large enough. $\square$

Proof of Lemma 5.4

Let V_* denote the set of all $z \in [n]$ such that $\mathcal{B}^+_{\hbar}(z)$ is a directed tree. As observed in [13, Proposition 6], it is an immediate consequence of Proposition 2.3 that with high probability, for all $x \in [n]$:

$$P(x, V_*) = \sum_{z \in V_*} P(x,z) \geq \tfrac{1}{2}. \tag{5.55}$$

Therefore,

$$P^{t_\star+1}(x,y) \geq \tfrac{1}{2}\min_{x \in V_*} P^{t_\star}(x,y). \tag{5.56}$$

Since $P^{t_\star}(x,y) \geq P_{0,t_\star}(x,y)$ it is sufficient to prove

$$\mathbb{P}\left(\forall x \in V_*, \forall y \in [n],\;\; P_{0,t_\star}(x,y) \geq \tfrac{c}{n}\,\Gamma_{h_y}(y)\right) = 1 - o(1), \tag{5.57}$$

for some constant $c = c(\eta, \Delta) > 0$. The proof of (5.57) is based on Lemma 5.5 and the following estimates which allow us to make sure the events $\mathcal{Y}_{x,y}$ in Lemma 5.5 have large probability.

Lemma 5.6 *The event $\mathcal{A}_1 = \{\forall x \in V_*, \forall y \in [n] : A_{x,y} \geq \frac{1}{2}\}$ has probability*

$$\mathbb{P}\left(\mathcal{A}_1\right) = 1 - o(1)\,.$$

Proof: Let us first note that the event $\widehat{\mathcal{A}}_1 = \{\forall x \in V_* : \sum_{e\in\mathcal{E}^*(x)} \mathbf{w}(e)\mathbf{1}_{\mathbf{w}(e)\leq n^{2\eta-1}} \geq 0.9\}$ satisfies

$$\mathbb{P}\left(\widehat{\mathcal{A}}_1\right) = 1 - o(1).$$

Indeed, this fact is a consequence of [13, 14], which established that for any $\varepsilon > 0$, with high probability

$$\min_{x\in V_*} \sum_{e\in\mathcal{E}^*(x)} \mathbf{w}(e)\mathbf{1}_{\mathbf{w}(e)\leq n^{2\eta-1}} \geq 1 - \varepsilon, \tag{5.58}$$

see e.g. [14, Theorem 4 and Lemma 11]. Thus, it remains to show that replacing $\mathcal{E}^*(x)$ with $\mathcal{E}^0$ does not alter much the sum. Suppose the κ edges generating $\mathcal{T}(x)$ have been revealed and then sample the τ edges generating the neighbourhood $\mathcal{B}^-_{h_y}(y)$. Let K denote the number of collisions between $\mathcal{T}(x)$ and $\mathcal{B}^-_{h_y}(y)$. There are at most $N := \Delta^{h_y} = n^{3\eta \log\Delta/\log\delta}$ attempts each with success probability at most $p := \kappa/(m-\kappa)$. Thus K is stochastically dominated by a binomial $\mathrm{Bin}(N,p)$, and therefore by Hoeffding's inequality

$$\mathbb{P}(K > Np + N) \leq \exp\left(-2N\right) \leq \exp\left(-n^{3\eta}\right).$$

Thus by a union bound we may assume that all x, y are such that the corresponding collision count K satisfies $K \leq Np + N \leq 2N$. Therefore, on the event $\widehat{\mathcal{A}}_1$

$$\sum_{e\in\mathcal{E}^0} \mathbf{w}(e)\mathbf{1}_{\mathbf{w}(e)\leq n^{2\eta-1}} \geq 0.9 - 2N\, n^{2\eta-1} \geq \frac{1}{2},$$

if η is small enough. □

Lemma 5.7 *Fix a constant $c > 0$ and consider the event $\mathcal{A}_2 = \{\forall x, y \in [n] : B_{x,y} \geq c\,\Gamma_{h_y}(y)\}$. If $c > 0$ is small enough*

$$\mathbb{P}\left(\mathcal{A}_2\right) = 1 - o(1)\,.$$

Proof: By definition, $\sum_{f\in\mathcal{F}_{h_y}} \mathbf{w}(f) = \Gamma_{h_y}(y)$. Thus, we need to show that if we replace $\mathcal{F}^0$ by $\mathcal{F}_{h_y}$ the sum defining $B_{x,y}$ is still comparable to $\Gamma_{h_y}(y)$. For any constant $T > 0$, for each $z \in \partial\mathcal{B}^-_{h_y-T}(y)$, let V_z denote the set of $w \in \partial\mathcal{B}^-_{h_y}(y)$ such that $d(w,z) = T$. Notice that if the event $\mathcal{G}(\hbar)$ from Proposition 2.3 holds then for each $z \in \partial\mathcal{B}^-_{h_y-T}(y)$ one has $|V_z| \geq \frac{1}{2}\delta^T$. Consider the generation of the $\kappa + \tau$ edges as above, and call a vertex $z \in \partial\mathcal{B}^-_{h_y-T}(y)$ *bad* if all heads attached to V_z are matched, or equivalently if none of these heads is in $\mathcal{F}_{tot}$. Given a $z \in \partial\mathcal{B}^-_{h_y-T}(y)$, we want to estimate the probability that it is bad. To this end, we use the same construction given in Section 5.2.1 but this time we first generate the in-neighbourhood $\mathcal{B}^-_{h_y}(y)$ and then the tree $\mathcal{T}(x)$. Let K denote the number of collisions between $\mathcal{T}(x)$ and the set V_z. Notice that $|V_z| \le \Delta^T$ and that $|\mathcal{T}(x)| \le n^{1-\eta^2/2}$, so that K is stochastically dominated by the binomial $\mathrm{Bin}(N,p)$ where $N = n^{1-\eta^2/2}$ and $p = \Delta^{T+1}/n$. Therefore,

$$\mathbb{P}\left(K > \tfrac{1}{2}\delta^T\right) \le (Np)^{\frac{1}{2}\delta^T} \le \left(\Delta^{T+1} n^{-\eta^2/2}\right)^{\frac{1}{2}\delta^T}.$$

Since $|V_z| \ge \frac{1}{2}\delta^T$, if z is bad then $K > \frac{1}{2}\delta^T$ and thus the probability of the event that z is bad is at most $O(n^{-\delta^T\eta^2/4})$. The probability that there exists a bad $z \in \partial\mathcal{B}^-_{h_y-T}(y)$ is then bounded by $O(\Delta^{h_y} n^{-\delta^T\eta^2/4})$. In conclusion, if $T = T(\eta)$ is a large enough constant, we can ensure that for any $y \in [n]$ the probability that there exists a bad $z \in \partial\mathcal{B}^-_{h_y-T}(y)$ is $o(n^{-2})$, and therefore, by a union bound, with high probability there are no bad $z \in \partial\mathcal{B}^-_{h_y-T}(y)$, for all $x, y \in [n]$. On this event, for all z we may select one vertex $w \in V_z$ with at least one head $f \in \mathcal{F}^0$ attached to it. Notice that $\mathbf{w}(f) \ge \Delta^{-T-1} P^{h_y-T}(z,y)$. Therefore, assuming that there are no bad $z \in \partial\mathcal{B}^-_{h_y-T}(y)$:

$$\begin{aligned} B_{x,y}(\sigma) &= \sum_{f\in\mathcal{F}^0} \mathbf{w}(f) \\ &\ge \Delta^{-T} \sum_{z\in\partial\mathcal{B}^-_{h_y-T}(y)} P^{h_y-T}(z,y) \ge \Delta^{-T-1}\Gamma_{h_y-T}(y). \end{aligned}$$

From Lemma 5.3 we may finish with the estimate $\Gamma_{h_y-T}(y) \ge \frac{1}{2}\Gamma_{h_y}(y)$. □

We can now conclude the proof of (5.57). Consider the event

$$\mathcal{A} = \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{G}(\hbar). \tag{5.59}$$

For any $s > 0$,

$$\mathbb{P}\left(\forall x, y \in [n],\ \ P_{0,t_\star}(x,y) \ge \tfrac{s}{n}\Gamma_{h_y}(y)\right) \ge \mathbb{P}(\mathcal{A}) - \sum_{x,y\in[n]} \mathbb{P}\left(P_{0,t_\star}(x,y) < \tfrac{s}{n}\Gamma_{h_y}(y); \mathcal{A}\right). \tag{5.60}$$

From Lemma 5.6, Lemma 5.7, and Proposition 2.3 it follows that $\mathbb{P}(\mathcal{A}) = 1 - o(1)$. Let $\mathcal{W}_{x,y}$ denote the event

$$\mathbb{E}\left[P_{0,t_\star}(x,y) \,|\, \sigma\right] \ge \tfrac{c}{2m}\Gamma_{h_y}(y), \tag{5.61}$$

where c is the constant from Lemma 5.7. From Lemma 5.2 we infer that

$$\mathcal{A} \subset \mathcal{W}_{x,y} \cap \mathcal{Y}_{x,y},$$

for all x, y, and for all n large enough. Therefore,

$$\mathbb{P}\left(P_{0,t_\star}(x,y) < \tfrac{s}{n}\,\Gamma_{h_y}(y); \mathcal{A}\right) \leq \sup_{\sigma\in\mathcal{W}_{x,y}\cap\mathcal{Y}_{x,y}} \mathbb{P}\left(P_{0,t_\star}(x,y) < \tfrac{s}{n}\,\Gamma_{h_y}(y) \,|\, \sigma\right). \tag{5.62}$$

Taking $s > 0$ a small enough constant and using (5.48) and (5.61), we see that $P_{0,t_\star}(x,y) < \frac{s}{n}\,\Gamma_{h_y}(y)$ implies

$$|P_{0,t_\star}(x,y) - \mathbb{E}\left[P_{0,t_\star}(x,y) \,|\, \sigma\right]| \geq a\,\mathbb{E}\left[P_{0,t_\star}(x,y) \,|\, \sigma\right],$$

for some constant $a > 0$, and therefore from Lemma 5.5

$$\sup_{\sigma\in\mathcal{W}_{x,y}\cap\mathcal{Y}_{x,y}} \mathbb{P}\left(P_{0,t_\star}(x,y) < \tfrac{s}{n}\,\Gamma_{h_y}(y) \,|\, \sigma\right) = o(n^{-2}). \tag{5.63}$$

The bounds (5.60) and (5.63) end the proof of (5.57). This ends the proof of Lemma 5.4.

Remark 5.3 *Let us show that if the type (δ_-, Δ_+) is not in the set of linear types $\mathcal{L}$ one can improve the lower bound on $\pi_{\min}$ as mentioned in Remark 5.1. The proof given above shows that it is sufficient to replace γ_0 by γ_0' in Lemma 5.2, where γ_0' is defined by (5.5). To this end, for any $\varepsilon > 0$, let $\mathcal{L}_\varepsilon$ denote the set of types $(k,\ell) \in \mathcal{C}$ such that*

$$\limsup_{n\to\infty} \frac{|\mathcal{V}_{k,\ell}|}{n^{1-\varepsilon}} = +\infty\,, \tag{5.64}$$

where $\mathcal{V}_{k,\ell}$ denotes the set of vertices of type (k,ℓ), and define

$$\gamma_\varepsilon' := \frac{\log \Delta_{\varepsilon,+}'}{\log \delta_{\varepsilon,-}'}\,, \qquad \Delta_{\varepsilon,+}' := \max\{\ell:\ (k,\ell)\in\mathcal{L}_\varepsilon\}\,, \quad \delta_{\varepsilon,-}' := \min\{k:\ (k,\ell)\in\mathcal{L}_\varepsilon\}. \tag{5.65}$$

The main observation is that if $(k,\ell)\notin\mathcal{L}_\varepsilon$, then w.h.p. there are at most a finite number of vertices of type (k,ℓ) in all in-neighbourhoods $\mathcal{B}_{h_0}^-(y)$, $y\in[n]$, for any $h_0 = O(\log\log n)$. Indeed, for a fixed $y\in[n]$ the number of $v\in\mathcal{V}_{k,\ell}\cap\mathcal{B}_{h_0}^-(y)$ is stochastically dominated by the binomial $\mathrm{Bin}\left(\Delta^{h_0}, n^{-\varepsilon/2}\right)$, and therefore if $K = K(\varepsilon)$ is a sufficiently large constant then the probability of having more than K such vertices is bounded by $(\Delta^{h_0}n^{-\varepsilon/2})^K = o(n^{-1})$. Taking a union bound over $y\in[n]$ shows that w.h.p. all $\mathcal{B}_{h_0}^-(y)$, $y\in[n]$ have at most K vertices with type (k,ℓ). Then we may repeat the argument of Lemma 5.2 with this constraint, to obtain that for all $\varepsilon>0$, w.h.p. $\Gamma_{h_y}(y) \geq c(\varepsilon)\log^{1-\gamma_\varepsilon'}(n)$. Since the number of types is finite one concludes that if ε is small enough then $\gamma_0' = \gamma_\varepsilon'$ and the desired conclusion follows.

5.3 Upper bound on $\pi_{\min}$

In this section we prove the upper bound on $\pi_{\min}$ given in (5.3). We first show that we can replace $\pi(y)$ in (5.3) by a more convenient quantity. Define the distances

$$d(s) = \max_{x\in[n]} \|P^s(x,\cdot) - \pi\|_{\mathrm{TV}}\,, \quad \bar{d}(s) = \max_{x,y\in[n]} \|P^s(x,\cdot) - P^s(y,\cdot)\|_{\mathrm{TV}}. \tag{5.66}$$

It is standard that, for all $k, s \in \mathbb{N}$,

$$d(ks) \le \bar{d}(ks) \le \bar{d}(s)^k \le 2^k d(s)^k, \tag{5.67}$$

see e.g. [40]. In particular, defining

$$\lambda_t(y) = \frac{1}{n}\sum_{x\in[n]} P^t(x,y)\,, \tag{5.68}$$

for any $k \in \mathbb{N}$, setting $t = 2kT_{\mathrm{ENT}}$, one has

$$\max_{y\in[n]} |\lambda_t(y) - \pi(y)| \le d(2kT_{\mathrm{ENT}}) \le 2^k d(2T_{\mathrm{ENT}})^k. \tag{5.69}$$

From Theorem 1.7 we know that w.h.p. $d(2T_{\mathrm{ENT}}) \le \frac{1}{2e}$ so that the right hand side above is at most e^{-k}. If $k = \Theta(\log^2(n))$ we can safely replace $\pi(y)$ with $\lambda_t(y)$ in (5.3). Thus, it suffices to prove the following statement.

Lemma 5.8 *For some constants $\beta > 0$, $C > 0$, and for any $t = t_n = \Theta(\log^3(n))$:*

$$\mathbb{P}\Big(\exists S \subset [n],\ |S| \ge n^{\beta}\,,\ n\max_{y\in S}\lambda_t(y) \le C\,\log^{1-\gamma_1}(n)\Big) = 1 - o(1). \tag{5.70}$$

Proof: Let $(\delta_*, \Delta_*) \in \mathcal{L}$ denote the type realizing the maximum in the definition of γ_1; see (5.1). Let $V_* = \mathcal{V}_{\delta_*,\Delta_*}$ denote the set of vertices of this type, and let $\alpha_* \in (0,1)$ be a constant such that $|V_*| \ge \alpha_* n$, for all n large enough. Let us fix a constant $\beta_1 \in (0, \frac{1}{4})$. This will be related to the constant β, but we shall not look for the optimal exponent β in the statement (5.70). Consider the first $N_1 := n^{\beta_1}$ vertices in the set V_*, and call them $y_1, \dots, y_{N_1}$. Next, generate sequentially the in-neighbourhoods $\mathcal{B}^-_{h_0}(y_i)$, $i = 1, \dots, N_1$, where

$$h_0 = \log_{\delta_*}\log n - C_0, \tag{5.71}$$

for some constant C_0 to be fixed later. As in the proof of Lemma 5.2 we couple the $\mathcal{B}^-_{h_0}(y_i)$ with independent random trees Y_i rooted at y_i. For each $\mathcal{B}^-_{h_0}(y_i)$ the probability of failing to equal Y_i, conditionally on the previous generations, is uniformly bounded above by $p := N_1 \Delta^{2h_0}/m$. Let $\mathcal{A}$ denote the event that all $\mathcal{B}^-_{h_0}(y_i)$ are successfully coupled to the Y_i's and that they have no intersections. Therefore,

$$\mathbb{P}(\mathcal{A}) \geq 1 - O(N_1 p) \geq 1 - O(n^{3\beta_1 - 1}) = 1 - o(1). \tag{5.72}$$

Consider now a single random tree Y_1. We say that Y_1 is *unlucky* if all labels of the vertices in the tree are of type (δ_*, Δ_*). The probability that Y_1 is unlucky is at least

$$q = \left(\frac{\alpha_* n \Delta_*}{m}\right)^{\delta_*^{h_0}} \geq n^{-\eta},$$

where $\eta = \delta_*^{-C_0} \log(\Delta/2\alpha_*)$ if C_0 is the constant in (5.71). We choose C_0 so large that $0 < \eta \leq \beta_1/4$. Call S_1 the set of $y \in \{y_1, \ldots, y_{N_1}\}$ such that Y_i is unlucky. Since the Y_i are i.i.d. the probability that $|S_1| < n^{\beta_1/2}$ is bounded by the probability that $\mathrm{Bin}(N_1, q) < n^{\beta_1/2}$, which by Hoeffding's inequality is at most

$$\exp\left(-n^{\beta_1/3}\right) \tag{5.73}$$

Fix a realization σ of the in-neighbourhoods $\mathcal{B}^-_{h_0}(y_i)$, $i = 1, \ldots, N_1$. Say that y_i is unlucky if all vertices in $\mathcal{B}^-_{h_0}(y_i)$ are of type (δ_*, Δ_*). Thanks to (5.72) we may assume that $\sigma \in \mathcal{A}$, i.e. $\mathcal{B}^-_{h_0}(y_i) = Y_i$ for all i so that the set of unlucky y_i coincides with S_1, and thanks to (5.73) we may also assume that σ is such that $|S_1| \geq \bar{N} := n^{\beta_1/2}$. We call $\mathcal{A}' \subset \mathcal{A}$ the set of all $\sigma \in \mathcal{A}$ satisfying the latter requirement. Let $\bar{S}$ denote the first $\bar{N}$ elements in S_1. We are going to show that uniformly in $\sigma \in \mathcal{A}'$, for a sufficiently large constant $C > 0$, any $t = \Theta(\log^3(n))$,

$$\mathbb{P}\Big(\sum_{y \in \bar{S}} \lambda_t(y) > \tfrac{C\bar{N}}{2n} \log^{1-\gamma_1}(n) \,\Big|\, \sigma\Big) = o(1). \tag{5.74}$$

Notice that (5.74) says that, conditionally on a fixed $\sigma \in \mathcal{A}'$, with high probability

$$\sum_{y \in \bar{S}} \lambda_t(y) \leq \tfrac{C\bar{N}}{2n} \log^{1-\gamma_1}(n),$$

which implies that there are at most $\bar{N}/2$ vertices $y \in \bar{S}$ with the property that $\lambda_t(y) > \frac{C}{n} \log^{1-\gamma_1}(n)$. Summarizing, the above arguments and (5.74) allow one to conclude the unconditional statement that with high probability there are at least $\frac{1}{2} n^{\beta_1/2}$ vertices $y \in [n]$ such that

$$\lambda_t(y) \leq \tfrac{C}{n} \log^{1-\gamma_1}(n),$$

which implies the desired claim (5.70), taking e.g. $\beta = \beta_1/3$.

To prove (5.74), consider the sum

$$\mathcal{X} = \sum_{y \in \bar{S}} \lambda_t(y).$$

We first establish that, uniformly in $\sigma \in \mathcal{A}'$, for any $t = \Theta(\log^3(n))$,

$$\mathbb{E}\left(\mathcal{X} \mid \sigma\right) = (1 + o(1)) \frac{\delta_*}{m} \bar{N} \Delta_*^{-h_0} \delta_*^{h_0}. \tag{5.75}$$

If y is unlucky then $P^{h_0}(z, y) = \Delta_*^{-h_0}$ for any $z \in \partial \mathcal{B}_{h_0}^-(y)$. Hence, for any $y \in \bar{S}$:

$$\lambda_t(y) = \frac{\Delta_*^{-h_0}}{n} \sum_{x \in [n]} \sum_{z \in \partial \mathcal{B}_{h_0}^-(y)} P^{t-h_0}(x, z) = \Delta_*^{-h_0} \sum_{z \in \partial \mathcal{B}_{h_0}^-(y)} \lambda_{t-h_0}(z).$$

Since $|\partial \mathcal{B}_{h_0}^-(y)| = \delta_*^{h_0}$, and since all $z \in \partial \mathcal{B}_{h_0}^-(y)$ have the same in-degree $d_z^- = \delta_*$, using symmetry the proof of (5.75) is reduced to showing that for any $z \in \partial \mathcal{B}_{h_0}^-(y)$, $t = \Theta(\log^3 n)$,

$$\mathbb{E}\left(\lambda_t(z) \mid \sigma\right) = (1 + o(1)) \frac{d_z^-}{m}. \tag{5.76}$$

To compute the expected value in (5.76) we use the so called *annealed* process. Namely, observe that

$$\mathbb{E}\left(\lambda_t(z) \mid \sigma\right) = \frac{1}{n} \sum_{x \in [n]} \mathbb{E}\left(P^t(x, z) \mid \sigma\right) = \frac{1}{n} \sum_{x \in [n]} \mathbb{P}_x^{a,\sigma}\left(X_t = z\right), \tag{5.77}$$

where X_t is the annealed walk with initial environment σ, and initial position x, and $\mathbb{P}_x^{a,\sigma}$ denotes its law. This process can be described as follows. At time 0 the environment consists of the edges from σ alone, and $X_0 = x$; at every step, given the current environment and position, the walker picks a uniformly random tail e from its current position, if it is still unmatched then it picks a uniformly random unmatched head f, the edge ef is added to the environment and the position is moved to the vertex of f, while if e is already matched then the position is moved to the vertex of the head to which e was matched. Let us show that uniformly in $x \neq z \in \partial \mathcal{B}_{h_0}^-(y)$, uniformly in $\sigma \in \mathcal{A}'$:

$$\mathbb{P}_x^{a,\sigma}\left(X_t = z\right) = (1 + o(1)) \frac{d_z^-}{m}. \tag{5.78}$$

Say that a collision occurs if the walk lands on a vertex that was already visited by using a freshly matched edge. At each time step the probability of a collision is at most $O(t/m)$, and therefore the probability of more than one collision in the first t steps is at most $O(t^4/m^2) = o(m^{-1})$. Thus we may assume that there is at most one cycle in the path of the walk up to time t. There are two cases to consider: 1) there is no cycle in the path up to time t or there is one cycle that does not pass through the vertex z; 2) there is a cycle and it passes through z. In case 1) since $X_t = z$ the walker must necessarily pick one of the heads of z at the very last step. Since all heads of z are unmatched by construction, and since the total number of unmatched heads at that time is at least $m - n^{\beta_1}\Delta^{h_0} - t = (1-o(1))m$, this event has probability $(1+o(1))d_z^-/m$. In case 2) since $x \neq z$ we argue that in order to have a cycle that passes through z, the walk has to visit z at some time before t, which is an event of probability $O(t/m)$, and then must hit back the previous part of the path, which is an event of probability $O(t^2/m)$. This shows that we can upper bound the probability of scenario 2) by $O(t^3/m^2) = o(m^{-1})$. This concludes the proof of (5.78). Next, observe that if $x = z$, then the previous argument gives $\mathbb{P}_z^{a,\sigma}(X_t = z) = O(t/m)$ which is a bound on the probability that the walk hits again z at some point within time t. In conclusion, (5.77) and (5.78) imply (5.76) which establishes (5.75).

Let us now show that

$$\mathbb{E}\left(\mathcal{X}^2 \,|\, \sigma\right) = (1+o(1))\mathbb{E}\left(\mathcal{X} \,|\, \sigma\right)^2. \tag{5.79}$$

Once we have (5.79) we can conclude (5.74) by using Chebyshev's inequality together with (5.75) and the fact that $\delta_*^{h_0}\Delta_*^{-h_0} \leq C_2 \log^{1-\gamma_1}(n)$ for some constant $C_2 > 0$. We write

$$\mathbb{E}\left(\mathcal{X}^2 \,|\, \sigma\right) = \sum_{y,y' \in \bar{S}} \Delta_*^{-2h_0} \frac{1}{n^2} \sum_{x,x' \in [n]} \sum_{z \in \partial\mathcal{B}_{h_0}^-(y)} \sum_{z' \in \partial\mathcal{B}_{h_0}^-(y')} \mathbb{P}_{x,x'}^{a,\sigma}(X_{t-h_0} = z, X'_{t-h_0} = z'), \tag{5.80}$$

where $\mathbb{P}_{x,x'}^{a,\sigma}$ is the law of two trajectories (X_s, X'_s), $s = 0, \dots, t$, that can be sampled as follows. Let X be sampled up to time t according to the previously described annealed measure $\mathbb{P}_x^{a,\sigma}$, call σ' the environment obtained by adding to σ all the edges discovered during the sampling of X and then sample X' up to time t independently, according to $\mathbb{P}_{x'}^{a,\sigma'}$.

Let also $\mathbb{P}_{\mathrm{u}}^{a,\sigma}$ be defined by

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma} = \frac{1}{n^2} \sum_{x,x' \in [n]} \mathbb{P}_{x,x'}^{a,\sigma}.$$

Thus, under $\mathbb{P}_{\mathrm{u}}^{a,\sigma}$ the two trajectories have independent uniformly distributed starting points x, x'. With this notation we write

$$\mathbb{E}\left(\mathcal{X}^2 \,|\, \sigma\right) = \sum_{y,y' \in \bar{S}} \Delta_*^{-2h_0} \sum_{z \in \partial\mathcal{B}_{h_0}^-(y)} \sum_{z' \in \partial\mathcal{B}_{h_0}^-(y')} \mathbb{P}_{\mathrm{u}}^{a,\sigma}(X_{t-h_0} = z, X'_{t-h_0} = z'). \tag{5.81}$$

Let us show that if $z \neq z'$, $t = \Theta(\log^3(n))$:

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma}(X_t = z, X'_t = z') = (1 + o(1))\frac{d_z^- d_{z'}^-}{m^2}. \tag{5.82}$$

Indeed, let A be the event that the first trajectory hits z at time t and visits z' at some time before that. Then reasoning as in (5.78) the event A has probability $O(t/m^2)$. Given any realization X of the first trajectory satisfying this event, the probability of $X'_t = z'$ is at most the probability of colliding with the trajectory X within time t, which is $O(t/m)$. On the other hand, if the first trajectory hits z at time t and does visit z' at any time before that, then the conditional probability of $X'_t = z$, as in (5.78) is given by $(1 + o(1))d_{z'}^-/m$. This proves (5.82) when $z \neq z'$.

If $z = z'$, $t = \Theta(\log^3(n))$, let us show that

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma}(X_t = z, X'_t = z) = O(1/m^2). \tag{5.83}$$

Consider the event A that the first trajectory X has at most one collision. The complementary event A^c has probability at most $O(t^4/m^2)$. If A^c occurs, then the conditional probability of $X'_t = z$ is at most the probability that X' collides with the first trajectory at some time $s \leq t$, that is $O(t/m)$. Hence,

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma}(X_t = z, X'_t = z; A^c) = O(t^5/m^3) = O(1/m^2). \tag{5.84}$$

To prove (5.83), notice that to realize $X'_t = z$ there must be a time $s = 0, \ldots, t$ such that X' collides with the first trajectory X at time s, then X' stays in the digraph D_1 defined by the first trajectory for the remaining $t - s$ units of time, and X' hits z at time t. On the event A the probability of spending h units of time in D_1 is at most $2\delta^{-h}$, and for any $h \in [0, t]$ there are at most $h + 1$ points x which have a path of length h from x to z in D_1. Therefore

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma}(X_t = z, X'_t = z; A) \leq (1 + o(1))\frac{d_z^-}{m}\sum_{h=0}^{t}\frac{2(h+1)}{m}\,2\delta^{-h} = O(1/m^2). \tag{5.85}$$

Hence, (5.83) follows from (5.84) and (5.85).

In conclusion, using (5.82) and (5.83) in (5.81), and recalling (5.75), we have obtained (5.79). $\square$

5.4 Upper bound on $\pi_{\max}$

As in Section 5.3 we start by replacing $\pi(y)$ with $\lambda_t(y) = \frac{1}{n}\sum_x P^t(x,y)$. In (5.69) we have seen that if $t = 2kT_{\mathrm{ENT}}$, then w.h.p.

$$\max_{y\in[n]} |\lambda_t(y) - \pi(y)| \le e^{-k}. \tag{5.86}$$

Thus, using a union bound over $y \in [n]$, the upper bound in Theorem 5.2 follows from the next statement.

Lemma 5.9 *There exists $C > 0$ such that for any $t = t_n = \Theta(\log^3(n))$, uniformly in $y \in [n]$*

$$\mathbb{P}\left(\lambda_t(y) \ge \tfrac{C}{n}\log^{1-\kappa_0}(n)\right) = o(n^{-1}). \tag{5.87}$$

Proof: Fix

$$h_0 = \log_{\Delta_-}\log n,$$

and call σ a realization of the in-neighbourhood $\mathcal{B}^-_{h_0}(y)$. Clearly,

$$\lambda_{t+h_0}(y) = \sum_{z\in\mathcal{B}^-_{h_0}(y)} \lambda_t(z)P^{h_0}(z,y).$$

From (5.15), under the event $\mathcal{G}_y(\hbar)$ from Proposition 2.3, we have $P^{h_0}(z,y) \le 2\delta_+^{-h_0} = 2\log^{-\kappa_0}(n)$ for every $z \in \mathcal{B}^-_{h_0}(y)$. Define

$$\mathcal{X} := \sum_{z\in\mathcal{B}^-_{h_0}(y)} \lambda_t(z) = \lambda_t(\mathcal{B}^-_{h_0}(y)).$$

Then it is sufficient to prove that for some constant C, uniformly in σ and $y \in [n]$:

$$\mathbb{P}\left(\mathcal{X} > \tfrac{C}{n}\log n\,;\ \mathcal{G}_y(\hbar)\ \mid\ \sigma\right) = o(n^{-1}). \tag{5.88}$$

By Markov's inequality, for any $K \in \mathbb{N}$ and any constant $C > 0$:

$$\mathbb{P}\left(\mathcal{X} > \tfrac{C}{n}\log(n); \mathcal{G}_y(\hbar)\,|\,\sigma\right) \le \frac{\mathbb{E}\left[\mathcal{X}^K; \mathcal{G}_y(\hbar)\,|\,\sigma\right]}{\left(\frac{C}{n}\log n\right)^K}. \tag{5.89}$$

We fix $K = \log n$, and claim that there exists an absolute constant $C_1 > 0$ such that

$$\mathbb{E}\left[\mathcal{X}^K; \mathcal{G}_y(\hbar)\,|\,\sigma\right] \le \left(\tfrac{C_1}{n}\log n\right)^K. \tag{5.90}$$

The desired estimate (5.88) follows from (5.90) and (5.89) by taking C large enough.

We compute the K-th moment $\mathbb{E}\left[\mathcal{X}^K; \mathcal{G}_y(\hbar) \,|\, \sigma\right]$ by using the annealed process as in (5.80). This time we have K trajectories instead of 2:

$$\begin{aligned}\mathbb{E}\left[\mathcal{X}^K; \mathcal{G}_y(\hbar) \,|\, \sigma\right] &= \frac{1}{n^K}\sum_{x_1,\dots,x_K} \mathbb{E}\left[P^t(x_1, \mathcal{B}^-_{h_0}(y)) \cdots P^t(x_K, \mathcal{B}^-_{h_0}(y))\,;\, \mathcal{G}_y(\hbar) \,|\, \sigma\right] \\ &= \frac{1}{n^K}\sum_{x_1,\dots,x_K} \mathbb{P}^{a,\sigma}_{x_1,\dots,x_K}\left(X^{(1)}_t \in \mathcal{B}^-_{h_0}(y), \dots, X^{(K)}_t \in \mathcal{B}^-_{h_0}(y)\,;\, \mathcal{G}_y(\hbar)\right), \end{aligned} \tag{5.91}$$

where $X^{(j)} := \{X^{(j)}_s, s \in [0,t]\}$, $j = 1,\dots,K$ denote K annealed walks each with initial point x_j, and $\mathbb{P}^{a,\sigma}_{x_1,\dots,x_K}$ denotes the joint law of the trajectories $X^{(j)}$, $j = 1,\dots,K$, and the environment, defined as follows. Start with the environment σ, and then run the first random walk $X^{(1)}$ up to time t as described after (5.77). After that run the walk $X^{(2)}$ up to time t with initial environment given by the union of edges from σ and the first trajectory, as described in (5.80). Proceed recursively until all trajectories up to time t have been sampled. This produces a new environment, namely the digraph given by the union of σ and all the K trajectories. At this stage there are still many unmatched heads and tails, and we complete the environment by using a uniformly random matching of the unmatched heads and tails. This defines the coupling $\mathbb{P}^{a,\sigma}_{x_1,\dots,x_K}$ between the environment and K independent walks in that environment, which justifies the expression in (5.91). It is convenient to introduce the notation

$$\mathbb{P}^{a,\sigma}_{\mathrm{u}} = \frac{1}{n^K}\sum_{x_1,\dots,x_K} \mathbb{P}^{a,\sigma}_{x_1,\dots,x_K},$$

for the annealed law of the K trajectories such that independently each trajectory starts at a uniformly random point $X^{(j)}_0 = x_j$. Let $D_0 = \sigma$ and let D_ℓ, for $\ell = 1,\dots,K$, denote the digraph defined by the union of $\sigma = \mathcal{B}^-_{h_0}(y)$ with the first ℓ paths

$$\{X^{(j)}_s, 0 \le s \le t\}, \quad j = 1,\dots,\ell.$$

Call $D_\ell(\hbar)$ the subgraph of D_ℓ consisting of all directed paths in D_ℓ ending at y with length at most $\hbar$. We define $\mathcal{G}^\ell_y(\hbar)$ as the event $\textsc{tx}(D_\ell(\hbar)) \le 1$. Notice that if the final environment has to satisfy $\mathcal{G}_y(\hbar)$, then necessarily for every ℓ the digraph D_ℓ must satisfy $\mathcal{G}^\ell_y(\hbar)$. Therefore,

$$\mathbb{E}\left[\mathcal{X}^K; \mathcal{G}_y(\hbar) \,|\, \sigma\right] \le \mathbb{P}^{a,\sigma}_{\mathrm{u}}\left(X^{(1)}_t \in \mathcal{B}^-_{h_0}(y), \dots, X^{(K)}_t \in \mathcal{B}^-_{h_0}(y)\,;\, \mathcal{G}^K_y(\hbar)\right). \tag{5.92}$$

Define

$$\mathcal{W}_\ell = \sum_{x \in V(D_\ell)} [d_x^-(D_\ell) - 1]_+, \tag{5.93}$$

where $V(D_\ell)$ denotes the vertex set of D_ℓ and $d_x^-(D_\ell)$ is the in-degree of x in the digraph D_ℓ. Define also the (ℓ, s) cluster $\mathcal{C}_\ell^s$ as the digraph given by the union of $D_{\ell-1}$ and the truncated path $\{X_u^{(\ell)}, 0 \le u \le s\}$. We say that the ℓ-th trajectory $X^{(\ell)}$ has a *collision* at time $s \ge 1$ if the edge $(X_{s-1}^{(\ell)}, X_s^{(\ell)}) \notin \mathcal{C}_\ell^{s-1}$ and $X_s^{(\ell)} \in \mathcal{C}_\ell^{s-1}$. We say that a collision occurs at time zero if $X_0^{(\ell)} \in D_{\ell-1}$. Notice that at least

$$\sum_{x \notin \mathcal{B}_{h_0}^-(y)} [d_x^-(D_\ell) - 1]_+$$

collisions must have occurred after the generation of the first ℓ trajectories.

Let $\mathcal{Q}_\ell$ denote the total number of collisions after the generation of the first ℓ trajectories. Since $|\mathcal{B}_{h_0}^-(y)| \le \Delta \log n$ one must have

$$\mathcal{W}_\ell \le \Delta \log n + \mathcal{Q}_\ell. \tag{5.94}$$

Notice that the probability of a collision at any given time by any given trajectory is bounded above by $p := 2\Delta(Kt + \Delta_-^{h_0})/m = O(\log^4(n)/n)$ and therefore $\mathcal{Q}_\ell$ is stochastically dominated by the binomial $\mathrm{Bin}(Kt, p)$. In particular, for any $k \in \mathbb{N}$:

$$\mathbb{P}\left(\mathcal{Q}_K \ge k\right) \le (Ktp)^k \le C_2^k \frac{\log^{8k}(n)}{n^k}, \tag{5.95}$$

for some constant $C_2 > 0$. If $A > 0$ is a large enough constant, then

$$\mathbb{P}\left(\mathcal{Q}_K \ge A \log n\right) \le e^{-\frac{A}{2}\log^2(n)}. \tag{5.96}$$

If $A \ge 2$ then (5.96) is smaller than the right hand side of (5.90) with e.g. $C_1 = 1$, and therefore from now on we may restrict to proving the upper bound

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_t^{(1)} \in \mathcal{B}_{h_0}^-(y), \dots, X_t^{(K)} \in \mathcal{B}_{h_0}^-(y)\,;\, \mathcal{Q}_K \le A \log n\,;\, \mathcal{G}_y^K(\hbar)\right) \le \left(\tfrac{C_1}{n}\log n\right)^K, \tag{5.97}$$

for some constant $C_1 = C_1(A) > 0$. To prove (5.97), define the events

$$B_\ell = \{X_t^{(1)} \in \mathcal{B}_{h_0}^-(y), \dots, X_t^{(\ell)} \in \mathcal{B}_{h_0}^-(y)\,;\, \mathcal{Q}_\ell \le A \log n\,;\, \mathcal{G}_y^\ell(\hbar)\}, \tag{5.98}$$

for $\ell = 1, \ldots, K$. Since $B_{\ell+1} \subset B_\ell$, the left hand side in (5.97) is equal to

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(B_1\right)\prod_{\ell=2}^{K}\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(B_\ell \,|\, B_{\ell-1}\right) \tag{5.99}$$

Thus, it is sufficient to show that for some constant C_1:

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(B_\ell \,|\, B_{\ell-1}\right) \le \tfrac{C_1}{n}\log n\,, \tag{5.100}$$

for all $\ell = 1, \ldots, K$, where it is understood that $\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(B_1 \,|\, B_0\right) = \mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(B_1\right).$

Let us partition the event $\{X_t^{(\ell)} \in \mathcal{B}_{h_0}^-(y)\}$ by specifying the last time in which the walk $X^{(\ell)}$ enters the neighbourhood $\mathcal{B}_{h_0}^-(y)$. Unless the walk starts in $\mathcal{B}_{h_0}^-(y)$, at that time it must enter from $\partial\mathcal{B}_{h_0}^-(y)$. Since the tree excess of $\mathcal{B}_{h_0}^-(y)$ is at most 1, once the walker is in $\mathcal{B}_{h_0}^-(y)$, we can bound the chance that it remains in $\mathcal{B}_{h_0}^-(y)$ for k steps by $2\delta_+^{-k}$. Therefore,

$$\begin{aligned}
\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(B_\ell \,|\, B_{\ell-1}\right) &\le \mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_t^{(\ell)} \in \mathcal{B}_{h_0}^-(y) \,|\, B_{\ell-1}\right) \\
&\le 2\delta_+^{-t}\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_0^{(\ell)} \in \mathcal{B}_{h_0}^-(y) \,|\, B_{\ell-1}\right) + \sum_{j=1}^{t} 2\delta_+^{-(t-j)}\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y) \,|\, B_{\ell-1}\right) \\
&\le 2t\delta_+^{-t/2} + \sum_{j=t/2+1}^{t} 2\delta^{-(t-j)}\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y) \,|\, B_{\ell-1}\right)
\end{aligned}$$

Since $t = \Theta(\log^3(n))$, it is enough to show

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y) \,|\, B_{\ell-1}\right) \le \tfrac{C_1}{n}\log n, \tag{5.101}$$

uniformly in $j \in (t/2, t)$ and $1 \le \ell \le K$.

Let $\mathcal{H}_0^\ell$ denote the event that the ℓ-th walk makes its first visit to the digraph $D_{\ell-1}$ at the very last time j, when it enters $\partial\mathcal{B}_{h_0}^-(y)$. Uniformly in the trajectories of the first $\ell - 1$ walks, at any time there are at most $\Delta_-|\partial\mathcal{B}_{h_0}^-(y)| \le \Delta_-^{h_0+1} = \Delta_- \log n$ unmatched heads attached to $\partial\mathcal{B}_{h_0}^-(y)$, and therefore

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y)\,;\, \mathcal{H}_0^\ell \,|\, B_{\ell-1}\right) = O(|\partial\mathcal{B}_{h_0}^-(y)|/m) \le \tfrac{C_1}{n}\log n. \tag{5.102}$$

Let $\mathcal{H}_2^\ell$ denote the event that the ℓ-th walk makes a first visit to $D_{\ell-1}$ at some time $s_1 < j$, then at some time $s_2 > s_1$ it exits $D_{\ell-1}$, and then at a later time $s_3 \le j$ enters again the digraph $D_{\ell-1}$. Since each time the walk is outside $D_{\ell-1}$ the probability of entering $D_{\ell-1}$ at the next step is $O(Kt/m)$, it follows that

$$\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y)\,;\, \mathcal{H}_2^\ell \,|\, B_{\ell-1}\right) = O(K^2t^4/m^2) \leq \tfrac{C_1}{n}\log n. \tag{5.103}$$

It remains to consider the case where the ℓ-th walk enters only once the digraph $D_{\ell-1}$ at some time $s \leq j-1$, and then stays in $D_{\ell-1}$ for the remaining $j-s$ units of time. Calling $\mathcal{H}_{1,s}^\ell$ this event, and summing over all possible values of s, we need to show that

$$\sum_{s=0}^{j-1}\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y)\,;\, \mathcal{H}_{1,s}^\ell \,|\, B_{\ell-1}\right) \leq \tfrac{C_1}{n}\log n. \tag{5.104}$$

We divide the sum in two parts: $s \in [0, j-\hbar+h_0]$ and $s \in (j-\hbar+h_0, j)$. For the first part, note that the walk must spend at least $\hbar - h_0 \geq \hbar/2$ units of time in $D_{\ell-1}(\hbar)$, which has probability at most $2\delta_+^{-\hbar/2} = O(n^{-\varepsilon})$ for some constant $\varepsilon > 0$, because of the condition $\mathcal{G}_y^{\ell-1}(\hbar)$ included in the event $B_{\ell-1}$. Since the probability of hitting $D_{\ell-1}$ at time s is $O(Kt/m)$ we obtain

$$\sum_{s=0}^{j-\hbar+h_0}\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y)\,;\, \mathcal{H}_{1,s}^\ell \,|\, B_{\ell-1}\right) = O(Kt^2n^{-\varepsilon}/m) \leq \tfrac{C_1}{n}\log n. \tag{5.105}$$

To estimate the sum over $s \in (j-\hbar+h_0, j)$, notice that the walk has to enter $D_{\ell-1}$ by hitting a point $z \in D_{\ell-1}$ at time s such that there exists a path of length $h = j-s$ from z to $\partial\mathcal{B}_{h_0}^-(y)$ within the digraph $D_{\ell-1}$. Call L_h the set of such points in $D_{\ell-1}$. Hitting this set at any given time s coming from outside the digraph $D_{\ell-1}$ has probability at most $2\Delta|L_h|/m$, and the path followed once it has entered $D_{\ell-1}$ is necessarily in $D_{\ell-1}(\hbar)$ and therefore has weight at most $2\delta_+^{-h}$. Then,

$$\sum_{s=j-\hbar+h_0+1}^{j-1}\mathbb{P}_{\mathrm{u}}^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y)\,;\, \mathcal{H}_{1,s}^\ell \,|\, B_{\ell-1}\right) \leq \sum_{h=1}^{\hbar-h_0-1}\frac{2\Delta|L_h|}{m}2\delta_+^{-h}, \tag{5.106}$$

Let $A_h \subset L_h$ denote the set of points exactly at distance h from $\partial\mathcal{B}^-_{h_0}(y)$ in $D_{\ell-1}$. We have

$$\begin{aligned}
|A_h| &\le \sum_{x\in A_{h-1}} d^-_x(D_{\ell-1}) \\
&\le |A_{h-1}| + \sum_{x\in A_{h-1}} [d^-_x(D_{\ell-1})-1]_+ \\
&\le |A_{h-2}| + \sum_{x\in A_{h-1}\cup A_{h-2}} [d^-_x(D_{\ell-1})-1]_+ \\
&\le \cdots \le |A_0| + \sum_{x\in A_0\cup\ldots\cup A_{h-1}} [d^-_x(D_{\ell-1})-1]_+ \\
&\le |\partial\mathcal{B}^-_{h_0}(y)| + \mathcal{W}_{\ell-1}.
\end{aligned}$$

Since $h \le \hbar = O(\log n)$ and $|\partial\mathcal{B}^-_{h_0}(y)| \le \log n$, using (5.94) we have obtained

$$|A_h| \le C_2 \log n + \mathcal{Q}_{\ell-1}. \tag{5.107}$$

On the event $B_{\ell-1}$ we know that $\mathcal{Q}_{\ell-1} \le A\log n$, and therefore $|A_h| \le C_3 \log n$ for some absolute constant $C_3 > 0$. In conclusion, for all $h \in (0, \hbar - h_0)$

$$|L_h| \le \sum_{\ell=0}^{h} |A_\ell| \le C_3 h \log n. \tag{5.108}$$

Inserting this estimate in (5.106),

$$\sum_{s=j-\hbar+1}^{j-1} \mathbb{P}^{a,\sigma}_{\mathrm{u}}\left(X^{(\ell)}_j \in \partial\mathcal{B}^-_{h_0}(y)\,;\,\mathcal{H}^\ell_{1,s}\,|\,B_{\ell-1}\right) \le \tfrac{C_4}{n}\log n. \tag{5.109}$$

Combining (5.105) and (5.109) we have proved (5.104) for a suitable constant C_1. $\square$

5.5 Lower bound on $\pi_{\max}$

Lemma 5.10 *There exist constants $\varepsilon, c > 0$ such that*

$$\mathbb{P}\Big(\exists S\subset[n],\ |S|\ge n^\varepsilon\,,\ n\min_{y\in S}\pi(y) \ge c\log^{1-\kappa_1}(n)\Big) = 1-o(1). \tag{5.110}$$

Proof: We argue as in the first part of the proof of Lemma 5.8. Namely, let $(\Delta_*, \delta_*) \in \mathcal{L}$ denote the type realizing the minimum in the definition of κ_1; see (5.1). Let $V_* = \mathcal{V}_{\Delta_*,\delta_*}$ denote the set of vertices of this type, and let $\alpha_* \in (0,1)$ be a constant such that $|V_*| \geq \alpha_* n$, for all n large enough. Fix a constant $\beta_1 \in (0, \frac{1}{4})$ and call $y_1, \dots, y_{N_1}$ the first $N_1 := n^{\beta_1}$ vertices in the set V_*. Then sample the in-neighbourhoods $\mathcal{B}^-_{h_0}(y_i)$ where

$$h_0 = \log_{\Delta_*} \log n - C, \tag{5.111}$$

and call σ a realization of all these neighbourhoods. As in the proof of Lemma 5.8, we may assume that all $\mathcal{B}^-_{h_0}(y_i)$ are successfully coupled with i.i.d. random trees Y_i. Next define a y_i *lucky* if $\mathcal{B}^-_{h_0}(y_i)$ has all its vertices of type (Δ_*, δ_*). Then, if C in (5.111) is large enough we may assume that at least $n^{\beta_1/2}$ vertices y_i are lucky; see (5.73). As before, we call $\mathcal{A}'$ the set of σ realizing these constraints. Given a realization $\sigma \in \mathcal{A}'$, and some $\varepsilon \in (0, \beta_1/2)$ we fix the first n^ε lucky vertices $y_{*,i}$, $i = 1, \dots, n^\varepsilon$. Since $\mathbb{P}(\mathcal{A}') = 1 - o(1)$, letting $S = \{y_{*,i}, i = 1, \dots, n^\varepsilon\}$, it is sufficient to prove that for some constant $c > 0$

$$\max_{\sigma \in \mathcal{A}'} \mathbb{P}\left(\min_{i=1,\dots,n^\varepsilon} n\pi(y_{*,i}) < c \log^{1-\kappa_1}(n) \,|\, \sigma \right) = o(1). \tag{5.112}$$

To prove (5.112) we first observe that by (5.40) and Lemma 5.3 it is sufficient to prove the same estimate with $n\pi(y_{*,i})$ replaced by $\Gamma_{h_1}(y_{*,i})$, where $h_1 = K \log\log n$ for some large but fixed constant K. Therefore, by using symmetry and a union bound it suffices to show

$$\max_{\sigma \in \mathcal{A}'} \mathbb{P}\left(\Gamma_{h_1}(y_*) < c \log^{1-\kappa_1}(n) \,|\, \sigma \right) \leq n^{-2\varepsilon}, \tag{5.113}$$

where $y_* = y_{*,1}$ is the first lucky vertex. By definition of lucky vertex, $\partial\mathcal{B}^-_{h_0}(y_*)$ has exactly $\Delta_*^{h_0}$ elements. For each $z \in \partial\mathcal{B}^-_{h_0}(y_*)$ we sample the in-neighbourhood $\mathcal{B}^-_{h_1-h_0}(z)$. The same argument of the proof of Lemma 5.2 shows that the probability that all these neighbourhoods are successfully coupled to i.i.d. random directed trees is at least $1 - O(\Delta^{2h_1}/n)$. On this event we have

$$\Gamma_{h_1}(y_*) = \delta_*^{-h_0} \sum_{i=1}^{\Delta_*^{h_0}} X_i, \tag{5.114}$$

where $X_i = M^i_{h_1-h_0}$ is defined by (5.21). Then (5.22) shows that

$$\mathbb{P}\left(\Gamma_{h_1}(y_*) < \tfrac{1}{2} \Delta_*^{h_0} \delta_*^{-h_0} \right) \leq \exp\left(-c_1 \Delta_*^{h_0} \right), \tag{5.115}$$

for some constant $c_1 > 0$. Since $\Delta_*^{h_0} = \Delta_*^{-C} \log n$ and $\Delta_*^{h_0} \delta_*^{-h_0} = (\delta_*/\Delta_*)^C \log^{1-\kappa_1}(n)$, this shows that

$$\max_{\sigma \in \mathcal{A}'} \mathbb{P}\left(\Gamma_{h_1}(y_*) < c_2 \log^{1-\kappa_1}(n) \,|\, \sigma \right) \leq n^{-2\varepsilon}, \tag{5.116}$$

for some new constant $c_2 > 0$ and for $\varepsilon = c_1 \Delta_*^{-C}/4$. This ends the proof of (5.113). $\square$

CHAPTER 6

COVER TIME

In this chapter we present our results concerning the cover time of the simple random walk on the DCM($\mathbf{d}^{\pm}$). Let X_t, $t = 0, 1, 2, \ldots$, denote the simple random walk on the digraph G. Consider the hitting times

$$H_y = \inf\{t \geq 0 : \, X_t = y\}, \qquad \tau_{\text{cov}} = \max_{y \in [n]} H_y. \tag{6.1}$$

The cover time $T_{\text{cov}} = T_{\text{cov}}(G)$ is defined by

$$T_{\text{cov}} = \max_{x \in [n]} \mathbf{E}_x[\tau_{\text{cov}}], \tag{6.2}$$

where $\mathbf{E}_x$ denotes the expectation with respect to the law of the random walk (X_t) with initial point $X_0 = x$ in a fixed realization of the digraph G.

Theorem 6.1 *Let γ_0, γ_1 be as in Definition 5.1. There exists a constant $C > 0$ such that*

$$\mathbb{P}\left(C^{-1} n \log^{\gamma_1}(n) \leq T_{\text{cov}} \leq C\, n \log^{\gamma_0}(n)\right) = 1 - o(1). \tag{6.3}$$

Remark 6.1 *For sequences $\mathbf{d}^{\pm}$ such that $(\delta_-, \Delta_+) \in \mathcal{L}$ one has $\gamma_0 = \gamma_1 = \gamma$ and Theorem 6.1 implies*

$$T_{\text{cov}} \asymp n \log^{\gamma}(n), \qquad \text{w.h.p.} \tag{6.4}$$

As in Remark 5.1, if $(\delta_-, \Delta_+) \notin \mathcal{L}$, then Theorem 6.1 can be strengthened by replacing γ_0 with the constant γ_0' defined in (5.5).

Finally, we observe that when the sequences $\mathbf{d}^{\pm}$ are Eulerian, that is $d_x^+ = d_x^-$ for all $x \in [n]$, then the estimates in Theorem 6.1 can be refined considerably, and one obtains results that are at the same level of precision of those already established in the case of random undirected graphs [27, 1].

Theorem 6.2 *Suppose $d_x^- = d_x^+ = d_x$ for every $x \in [n]$. Call $\mathcal{V}_d$ the set of vertices of degree d, and write $\bar{d} = m/n$ for the average degree. Assume*

$$|\mathcal{V}_d| = n^{\alpha_d + o(1)} \tag{6.5}$$

for some constants $\alpha_d \in [0,1]$, for each type d. Then,

$$T_{\mathrm{cov}} = (\beta + o(1))\, n \log n\,, \qquad \text{w.h.p.} \tag{6.6}$$

where $\beta := \bar{d} \max_d \frac{\alpha_d}{d}$.

In particular, if all present types have linear size then $\alpha_d \in \{0,1\}$ for all d and (6.6) holds with $\beta = \bar{d}/\delta$, where δ is the minimum degree. In any case it is not difficult to see that $\beta \geq 1$, since $\bar{d}$ is determined only by types with linear size. For some general bounds on cover times of Eulerian graphs we refer to [10].

6.1 Bounds on the cover time

In this section we show how the control on the extremal values of the stationary distribution obtained in previous sections can be turned into the bounds on the cover time presented in Theorem 6.1. To this end we exploit the full strength of the strategy developed by Cooper and Frieze [27, 25, 26, 28, 1].

6.1.1 The key lemma

Given a digraph G, write X_t for the position of the random walk at time t and write $\mathbf{P}_x$ for the law of $\{X_t, t \geq 0\}$ with initial value $X_0 = x$. In particular, $\mathbf{P}_x(X_t = y) = P^t(x,y)$ denotes the transition probability. Fix a time $T > 0$ and define the event that the walk does not visit y in the time interval $[T,t]$, for $t > T$:

$$\mathcal{A}_y^T(t) = \{X_s \neq y,\ \forall s \in [T,t]\}. \tag{6.7}$$

Moreover, define the generating function

$$R_y^T(z) = \sum_{t=0}^{T} z^t \, \mathbf{P}_y(X_t = y), \qquad z \in \mathbb{C}. \tag{6.8}$$

Thus, $R_y^T(1) \geq 1$ is the expected number of returns to y within time T, if started at y. The following statement is proved in [26], see also [28, Lemma 3].

Lemma 6.1 (First Visit Time Lemma) *Assume that $G = G_n$ is a sequence of digraphs with vertex set $[n]$ and stationary distribution $\pi = \pi_n$, and let $T = T_n$ be a sequence of times such that*

(i) $\max_{x,y\in[n]} |P^T(x,y) - \pi(y)| \leq n^{-3}$.

(ii) $T^2\pi_{\max} = o(1)$ *and* $T\pi_{\min} \geq n^{-2}$.

Suppose that $y \in [n]$ satisfies:

(iii) there exist $K, \psi > 0$ independent of n such that

$$\min_{|z|\leq 1+\frac{1}{KT}} |R_y^T(z)| \geq \psi.$$

Then there exist $\xi_1, \xi_2 = O(T\pi_{\max})$ such that for all $t \geq T$:

$$\max_{x\in[n]} \left| \mathbf{P}_x\left(\mathcal{A}_y^T(t)\right) - \frac{1+\xi_1}{(1+p_y)^{t+1}} \right| \leq e^{-\frac{t}{2KT}}, \tag{6.9}$$

where

$$p_y = (1+\xi_2)\frac{\pi(y)}{R_y^T(1)}. \tag{6.10}$$

We want to apply the above lemma to digraphs from our configuration model. Thus, our first task is to make sure that the assumptions of Lemma 6.1 are satisfied. From now on we fix the sequence $T = T_n$ as

$$T = \log^3(n). \tag{6.11}$$

From Theorem 1.7 and the argument in (5.67) it follows that item (i) of Lemma 6.1 is satisfied with high probability. Moreover, Theorem 5.1 and Theorem 5.2 imply that item (ii) of Lemma 6.1 is also satisfied with high probability. Next, following [27], we define a class of vertices $y \in [n]$ which satisfy item (iii) of Lemma 6.1. We use the convenient notation

$$\vartheta = \log\log\log(n). \tag{6.12}$$

Definition 6.1 *We call* small cycle *a collection of $\ell \leq 3\vartheta$ edges such that their undirected projection forms a simple undirected cycle of length ℓ. We say that $v \in [n]$ is* locally tree-like (LTL) *if its in- and out-neighbourhoods up to depth ϑ are both directed trees and they intersect only at x. We denote by V_1 the set of LTL vertices, and write $V_2 = [n] \setminus V_1$ for the complementary set.*

The next proposition can be proved as in [27, Section 3].

Proposition 6.1 *The following holds with high probability:*

1. *The number of small cycles is at most $\Delta^{9\vartheta}$.*
2. *The number of vertices which are not LTL satisfies $|V_2| \leq \Delta^{15\vartheta}$.*
3. *There are no small cycles which are less than 9ϑ undirected steps away.*

Proposition 6.2 *With high probability, uniformly in $y \in V_1$:*

$$R_y^T(1) = 1 + O(2^{-\vartheta}). \tag{6.13}$$

Moreover, there exist constants $K, \psi > 0$ such that with high probability, every $y \in V_1$ satisfies item (iii) of Lemma 6.1. In particular, (6.9) holds uniformly in $y \in V_1$.

Proof: We first prove (6.13). Fix $y \in V_1$ and consider the neighbourhoods $\mathcal{B}_\vartheta^\pm(y)$ and $\mathcal{B}_\hbar^-(y)$. By Proposition 2.3 we may assume that $\mathcal{B}_\hbar^-(y)$ and $\mathcal{B}_\vartheta^+(y)$ are both directed trees except for at most one extra edge. By the assumption $y \in V_1$ we know that $\mathcal{B}_\vartheta^-(y), \mathcal{B}_\vartheta^+(y)$ are both directed trees with no intersection except y, so that the extra edge in $\mathcal{B}_\hbar^-(y) \cup \mathcal{B}_\vartheta^+(y)$ cannot be in $\mathcal{B}_\vartheta^-(y) \cup \mathcal{B}_\vartheta^+(y)$. Thus, the following cases only need to be considered:

1. there is no extra edge in $\mathcal{B}_\hbar^-(y) \cup \mathcal{B}_\vartheta^+(y)$;
2. the extra edge connects $\mathcal{B}_\hbar^-(y) \setminus \mathcal{B}_\vartheta^-(y)$ to itself
3. the extra edge connects $\mathcal{B}_\vartheta^-(y)$ to $\mathcal{B}_\hbar^-(y) \setminus \mathcal{B}_\vartheta^-(y)$;
4. the extra edge connects $\mathcal{B}_\vartheta^+(y)$ to $\mathcal{B}_\hbar^-(y) \setminus \mathcal{B}_\vartheta^-(y)$.

In all cases but the last, if a walk started at y returns at y at time $t > 0$ then it must exit $\partial\mathcal{B}_\vartheta^+(y)$ and enter $\partial\mathcal{B}_\hbar^-(y)$, and from any vertex of $\partial\mathcal{B}_\hbar^-(y)$ the probability to reach y before exiting $\mathcal{B}_\hbar^-(y)$ is at most $2\delta^{-\hbar}$. Therefore, in these cases the number of visits to y up to T is stochastically dominated by $1 + \mathrm{Bin}(T, 2\delta^{-\hbar})$ and

$$1 \leq R_y^T(1) \leq 1 + 2T\delta^{-\hbar} = 1 + O(n^{-a}),$$

for some $a > 0$. In the last case instead it is possible for the walk to jump from $\mathcal{B}_\vartheta^+(y)$ to $\mathcal{B}_\hbar^-(y) \setminus \mathcal{B}_\vartheta^-(y)$. Let E_k denote the event that the walk visits y exactly k times in the interval

$[1, T]$. Let B denote the event that the walk visits y exactly ϑ units of time after its first visit to $\partial\mathcal{B}^-_\vartheta(y)$. Then $\mathbf{P}_y(B) \le \delta^{-\vartheta}$. On the complementary event B^c the walk must enter $\partial\mathcal{B}^-_\hbar(y)$ before visiting y, and each time it visits $\partial\mathcal{B}^-_\hbar(y)$ it has probability at most $2\delta^{-\hbar}$ to visit y before the next visit to $\partial\mathcal{B}^-_\hbar(y)$. Since the number of attempts is at most T one finds

$$\mathbf{P}_y(E_1) \le \mathbf{P}_y(B) + \mathbf{P}_y(E_1, B^c) \le \delta^{-\vartheta} + 2T\delta^{-\hbar} \le 2\delta^{-\vartheta}.$$

By the strong Markov property,

$$\mathbf{P}_y(E_k) \le \mathbf{P}_y(E_1)^k.$$

Therefore

$$R^T_y(1) = 1 + \sum_{k=1}^{\infty} k\mathbf{P}_y(E_k) = 1 + O(\delta^{-\vartheta}).$$

To see that $y \in V_1$ satisfies item (iii) of Lemma 6.1, take $z \in \mathbb{C}$ with $|z| \le 1 + 1/KT$ and write

$$|R^T_y(z)| \ge 1 - \sum_{t=1}^{T} \mathbf{P}_y(X_t = y)|z|^t \ge 1 - e^{1/K}(R^T_y(1) - 1) = 1 - O(\delta^{-\vartheta}).$$

$\square$

6.1.2 Upper bound on the cover time

We prove the following estimate relating the cover time to $\pi_{\min}$. From Theorem 5.1 this implies the upper bound on the cover time in Theorem 6.1.

Lemma 6.2 *For any constant $\varepsilon > 0$, with high probability*

$$\max_{x\in[n]} \mathbf{E}_x(\tau_{\mathrm{cov}}) \le (1+\varepsilon)\frac{\log n}{\pi_{\min}}. \tag{6.14}$$

Proof: Let U_s denote the set of vertices that are not visited in the time interval $[0, s]$. By Markov's inequality, for all $t_* \geq T$:

$$\begin{aligned}\mathbf{E}_x[\tau_{\text{cov}}] &= \sum_{s\geq 0} \mathbf{P}_x(\tau_{\text{cov}} > s) = \sum_{s\geq 0} \mathbf{P}_x(U_s \neq \emptyset) \\ &\leq t_* + \sum_{s\geq t_*} \mathbf{E}_x\left[|U_s|\right] = t_* + \sum_{s\geq t_*}\sum_{y\in[n]} \mathbf{P}_x(y \in U_s) \\ &\leq t_* + \sum_{s\geq t_*}\sum_{y\in[n]} \mathbf{P}_x(\mathcal{A}_y^T(s)). \end{aligned} \tag{6.15}$$

Choose

$$t_* := \frac{(1+\varepsilon)\log n}{\pi_{\min}},$$

for $\varepsilon > 0$ fixed. It is sufficient to prove that the last term in (6.15) is $o(t_*)$ uniformly in $x \in [n]$.

From Proposition 6.2 we can estimate

$$\mathbf{P}_x(\mathcal{A}_y^T(s)) = \frac{(1+\xi')}{(1+\bar{p}_y)^{s+1}}\,, \tag{6.16}$$

where $\bar{p}_y := (1+\xi)\pi(y)$ with $\xi, \xi' = O(T\pi_{\max}) + O(\delta^{-\vartheta}) = o(1)$ uniformly in $x \in [n], y \in V_1$. Therefore,

$$\sum_{s\geq t_*}\sum_{y\in V_1} \mathbf{P}_x(\mathcal{A}_y^T(s)) = (1+o(1))\sum_{y\in V_1} \frac{1}{\bar{p}_y(1+\bar{p}_y)^{t_*}}. \tag{6.17}$$

Using $\pi(y) \geq \pi_{\min}$, (6.17) is bounded by

$$\frac{(1+o(1))n}{\bar{p}_y(1+\bar{p}_y)^{t_*}} \leq \frac{2n}{\pi_{\min}} \exp\left(-\pi_{\min} t_*(1+o(1))\right) \leq \frac{1}{\pi_{\min}} = o(t_*),$$

for all fixed $\varepsilon > 0$ in the definition of t_*.

It remains to control the contribution of $y \in V_2$ to the sum in (6.15). From Proposition 6.1 we may assume that $|V_2| = O(\Delta^{15\vartheta})$. In particular, it is sufficient to show that with high probability uniformly in $x \in [n]$ and $y \in V_2$:

$$\sum_{s\geq t_*} \mathbf{P}_x(\mathcal{A}_y^T(s)) = o(t_*\Delta^{-15\vartheta}). \tag{6.18}$$

To prove (6.18), fix $y \in V_2$ and notice that by Proposition 6.1 (3), we may assume that there exists $u \in V_1$ s.t. $d(u, y) < 10\vartheta$. If $t_1 = t_0 + 10\vartheta$, $t_0 := 4/\pi_{\min}$, then

$$\begin{aligned}\mathbf{P}_x(\mathcal{A}_y^T(t_1)^c) &= \mathbf{P}_x(y \in \{X_T, X_{T+1}, \ldots, X_{t_1}\}) \\ &\geq \mathbf{P}_x(u \in \{X_T, X_{T+1}, \ldots, X_{t_0}\})\mathbf{P}_u(y \in \{X_1, \ldots, X_{10\vartheta}\}) \\ &\geq \left(1 - \mathbf{P}_x(\mathcal{A}_u^T(t_0))\right) \Delta^{-10\vartheta}.\end{aligned}$$

Since $u \in V_1$, as in (6.16), for n large enough,

$$\mathbf{P}_x(\mathcal{A}_u^T(t_0)) \leq \frac{2}{(1 + \bar{p}_y)^{t_0+1}} \leq \frac{1}{2}. \tag{6.19}$$

Setting $\gamma := \frac{1}{2}\Delta^{-10\vartheta}$, we have shown that $\mathbf{P}_x(\mathcal{A}_y^T(t_1)^c) \geq \gamma$. Since this bound is uniform over x, the Markov property implies, for all $k \in \mathbb{N}$,

$$\mathbf{P}_x(\mathcal{A}_y^T(s)) \leq (1 - \gamma)^k, \;\; s > k(T + t_1). \tag{6.20}$$

Therefore,

$$\begin{aligned}\sum_{s \geq t_*} \mathbf{P}_x(\mathcal{A}_y^T(s)) &\leq \sum_{s \geq t_*} (1-\gamma)^{\lfloor s/(T+t_1) \rfloor} \leq \sum_{s \geq t_*} (1-\gamma)^{s/2t_1} \\ &\leq \frac{\exp\left(-\gamma t_*/2t_1\right)}{1 - \exp\left(-\gamma/2t_1\right)} = O(t_1/\gamma) = o(t_* \Delta^{-15\vartheta}).\end{aligned}$$

□

6.1.3 Lower bound on the cover time

We prove the following stronger statement.

Lemma 6.3 *For some constant $c > 0$, with high probability*

$$\min_{x \in [n]} \mathbf{P}_x(\tau_{\text{cov}} \geq c\, n \log^{\gamma_1} n) = 1 - o(1). \tag{6.21}$$

Clearly, this implies the lower bound on $T_{\text{cov}} = \max_{x \in [n]} \mathbf{E}_x(\tau_{\text{cov}})$ in Theorem 6.1. The proof of Lemma 6.3 is based on the second moment method as in [28]. If $W \subset [n]$ is a set of vertices, let W_t be the set

$$W_t = \{y \in W : \; y \text{ is not visited in } [0, t]\} \tag{6.22}$$

Then

$$\mathbf{P}_x(\tau_{\text{cov}} > t) \geq \mathbf{P}_x(|W_t| > 0) \geq \frac{\mathbf{E}_x[|W_t|]^2}{\mathbf{E}_x[|W_t|^2]}. \tag{6.23}$$

Therefore, Lemma 6.3 is a consequence of the following estimate.

Lemma 6.4 *For some constant $c > 0$, with high probability there exists a nonempty set $W \subset [n]$ such that*

$$\max_{x\in[n]} \frac{\mathbf{E}_x[|W_t|^2]}{\mathbf{E}_x[|W_t|]^2} = 1 + o(1), \qquad t = c\, n \log^{\gamma_1} n. \tag{6.24}$$

We start the proof of Lemma 6.4 by exhibiting a candidate for the set W.

Proposition 6.3 *For any constant $K > 0$, with high probability there exists a set W such that*

1. *$W \subset V_1$, where V_1 is the LTL set from Definition 6.1, and $|W| \geq n^\alpha$ for some constant $\alpha > 0$.*
2. *For some constant $C > 0$, for all $y \in W$,*
$$\pi(y) \leq \tfrac{C}{n} \log^{1-\gamma_1}(n). \tag{6.25}$$
3. *For all $x, y \in W$:*
$$|\pi(x) - \pi(y)| \leq \pi_{\min} \log^{-K}(n). \tag{6.26}$$
4. *For all $x, y \in W$: $\min\{d(x,y), d(y,x)\} > 2\vartheta$.*

Proof: From Theorem 5.1 we know that w.h.p. there exists a set $S \subset [n]$ with $|S| > n^\beta$ such that (6.25) holds. Moreover, a minor modification of the proof of Lemma 5.8 shows that we may also assume that $S \subset V_1$ and that $\min\{d(x,y), d(y,x)\} > 2\vartheta$ for every $x, y \in W$. Indeed, it suffices to generate the out-neighbourhoods $\mathcal{B}^+_\vartheta(y_i)$ for every $i = 1, \ldots, N_1$ and the argument for (5.72) shows that these are disjoint trees with high probability. To conclude, we observe that there is a $W \subset S$ such that $|W| > n^{\beta/2}$ and such that (6.26) holds. Indeed, using $\pi_{\min} \geq n^{-1}\log^{-K_1}(n)$ for some constant K_1, for any constant $K > 0$ we may partition the interval

$$[n^{-1}\log^{-K_1}(n), Cn^{-1}\log^{1-\gamma_1}(n)]$$

in $\log^{2K}(n)$ intervals of equal length and there must be at least one of them containing $n^\beta \log^{-2K}(n) \geq n^{\beta/2}$ elements which, if K is sufficiently large, satisfy (6.26). $\square$

Proof of Lemma 6.4

Consider the first moment $\mathbf{E}_x[|W_t|]$, where W is the set from Proposition 6.3 and t is fixed as $t = c\, n \log^{\gamma_1}(n)$. For $y \in W \subset V_1$ we use Lemma 6.1 and Proposition 6.2. As in (6.16) we have

$$\mathbf{P}_x(\mathcal{A}_y^T(t)) = (1+o(1))(1+\bar{p}_y)^{-(t+1)}, \tag{6.27}$$

where $\bar{p}_y = (1+o(1))\pi(y) \le p_W := 2C\, n^{-1} \log^{1-\gamma_1}(n)$, where C is as in (6.25). Therefore,

$$\begin{aligned}\mathbf{E}_x\left[|W_t|\right] &= \sum_{y \in W} \mathbf{P}_x\left(y \text{ not visited in } [0,t]\right) \\ &\ge -T + \sum_{y\in W} \mathbb{P}(\mathcal{A}_y^T(t)) \ge -T + (1+o(1))|W|(1+p_W)^{-t}.\end{aligned}$$

Taking the constant c in the definition of t sufficiently small, one has $p_W t \le \alpha/2 \log n$ and therefore

$$\mathbf{E}_x\left[|W_t|\right] \ge -T + (1+o(1))|W|n^{-\alpha/2} \ge \tfrac{1}{2}\, n^{\alpha/2}, \tag{6.28}$$

where we use $T = \log^3(n)$ and $|W| \ge n^\alpha$. In particular, since $T = \log^3(n)$, (6.28) shows that

$$\sum_{y\in W} \mathbb{P}(\mathcal{A}_y^T(t)) = (1+o(1))\mathbf{E}_x\left[|W_t|\right]. \tag{6.29}$$

Concerning the second moment $\mathbf{E}_x[|W_t|^2]$, we have

$$\begin{aligned}\mathbf{E}_x\left[|W_t|^2\right] &= \sum_{y,y' \in W} \mathbf{P}_x\left(y \text{ and } y' \text{ not visited in } [0,t]\right) \\ &\le \sum_{y,y'\in W} \mathbf{P}_x\left(\mathcal{A}_y^T(t) \cap \mathcal{A}_{y'}^T(t)\right).\end{aligned}$$

From this and (6.29), the proof of Lemma 6.4 is completed by showing, uniformly in $x \in [n], y, y' \in W$:

$$\mathbf{P}_x\left(\mathcal{A}_y^T(t) \cap \mathcal{A}_{y'}^T(t)\right) = (1+o(1))\mathbf{P}_x\left(\mathcal{A}_y^T(t)\right)\mathbf{P}_x\left(\mathcal{A}_{y'}^T(t)\right). \tag{6.30}$$

We follow the idea of [28]. Let G^* denote the digraph obtained from our digraph G by merging the two vertices y, y' into the single vertex $y_* = \{y, y'\}$. Notice that y^* is LTL in the graph G^* in the sense of Definition 6.1. Moreover, G^* has the law of a directed configuration model with the same degree sequence of G except that at y_* it has $d^{\pm}_{y_*} = d^{\pm}_y + d^{\pm}_{y'}$. It follows that we may apply Lemma 6.1 and Proposition 6.2. Therefore, if $\mathbf{P}^*_x$ denotes the law of the random walk on G^* started at x, as in (6.27) we have

$$\mathbf{P}^*_x(\mathcal{A}^T_{y_*}(t)) = (1+o(1))(1+\bar{p}_{y_*})^{-t}, \tag{6.31}$$

uniformly in $x \in [n], y, y' \in W$, where $\bar{p}_{y_*} = (1+o(1))\pi^*(y_*)$, and π^* is the stationary distribution of G^*. In Lemma 6.5 below we prove that

$$\max_{\substack{v\in[n]:\\ v\neq y,y'}} |\pi(v) - \pi^*(v)| \leq a, \qquad |\pi(y) + \pi(y') - \pi^*(y_*)| \leq a, \tag{6.32}$$

where $a := \pi_{\min}\log^{-1}(n)$. Assuming (6.32), we can conclude the proof of (6.30). Indeed, letting P_* denote the transition matrix of the graph G^*,

$$\begin{aligned}
\mathbf{P}^*_x(\mathcal{A}^T_{y_*}(t)) &= \sum_{v\neq y,y'} P^T_*(x,v)\mathbf{P}^*_v(X_s \neq y_*,\ \forall s \in [1, t-T]) \\
&= \sum_{v\neq y,y'} \left(\pi^*(v) + O(n^{-3})\right)\mathbf{P}^*_v(X_s \neq y_*,\ \forall s \in [1, t-T])
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\mathbf{P}_x(\mathcal{A}^T_y(t) \cap \mathcal{A}^T_{y'}(t)) &= \sum_{v\neq y,y'} P^T(x,v)\mathbf{P}_v(X_s \notin \{y,y'\},\ \forall s \in [1, t-T]) \\
&= \sum_{v\neq y,y'} \left(\pi(v) + O(n^{-3})\right)\mathbf{P}_v(X_s \notin \{y,y'\},\ \forall s \in [1, t-T])
\end{aligned}$$

For all $v \neq y, y'$,

$$\begin{aligned}
\mathbf{P}^*_v(X_s \neq y_*,\ \forall s \in [1, t-T]) &= \mathbf{P}_v(X_s \notin \{y,y'\},\ \forall s \in [1, t-T]) \\
&\leq \frac{(1+o(1))}{\pi_{\min}} P^T(x,v)\mathbf{P}_v(X_s \notin \{y,y'\},\ \forall s \in [1, t-T]),
\end{aligned}$$

uniformly in $x \in [n]$, where we have used condition (i) in Lemma 6.1. Therefore, using (6.32)

$$\begin{aligned}
&\left|\mathbf{P}_x\left(\mathcal{A}_y^T(t) \cap \mathcal{A}_{y'}^T(t)\right) - \mathbf{P}_x^*\left(\mathcal{A}_{y_*}^T(t)\right)\right| \\
&\leq \sum_{v \neq y,y'} \left|\pi(v) - \pi_*(v) + O(n^{-3})\right| \mathbf{P}_v(X_s \notin \{y, y'\},\ \forall s \in [1, t-T]) \\
&\leq (a + O(n^{-3}))\frac{(1+o(1))}{\pi_{\min}} \sum_{v \neq y,y'} P^T(x,v)\mathbf{P}_v(X_s \notin \{y,y'\},\ \forall s \in [1, t-T]) \\
&\leq \frac{2a}{\pi_{\min}} \mathbf{P}_x(\mathcal{A}_y(t) \cap \mathcal{A}_{y'}(t)).
\end{aligned}$$

By definition of a we have $a/\pi_{\min} \to 0$ so that

$$\mathbf{P}_x(\mathcal{A}_y^T(t) \cap \mathcal{A}_{y'}^T(t)) = (1+o(1))\mathbf{P}_x^*(\mathcal{A}_{y_*}^T(t)). \tag{6.33}$$

Using (6.27), (6.31) and (6.32) we conclude that

$$\begin{aligned}
\mathbf{P}_x\left(\mathcal{A}_y^T(t) \cap \mathcal{A}_{y'}^T(t)\right) &= (1+o(1)) \exp\left(-(1+o(1))(\pi(y) + \pi(y'))t\right) \\
&= (1+o(1))\mathbf{P}_x\left(\mathcal{A}_y^T(t)\right)\mathbf{P}_x\left(\mathcal{A}_{y'}^T(t)\right).
\end{aligned}$$

□

Lemma 6.5 *The stationary distributions π, π^* satisfy* (6.32).

Proof: We follow the proof of Eq. (107) in [28]. The stochastic matrix of the simple random walk on G^* is given by

$$P_*(v,w) = \begin{cases} P(v,w) & \text{if } v, w \neq y_* \\ \frac{1}{2}\left(P(y,w) + P(y',w)\right) & \text{if } v = y_* \\ P(v,y) + P(v,y') & \text{if } w = y_*. \end{cases}$$

Let V^* denote the vertices of G^*. Define the vector $\zeta(v)$, $v \in V^*$ via

$$\zeta(v) = \begin{cases} \pi_*(v) - \pi(v) & v \neq y_* \\ \pi_*(y_*) - (\pi(x) + \pi(y)) & v = y_* \end{cases}$$

We are going to show that

$$\max_{v \in V^*} |\zeta(v)| = o(\pi_{\min} \log^{-1}(n)), \tag{6.34}$$

which implies (6.32). A computation shows that

$$\zeta P_*(w) = \sum_{v\in V^*} \zeta(v)P_*(v,w) = \begin{cases} \zeta(w) & \text{if } w \notin \mathcal{B}_1^+(y)\cup\mathcal{B}_1^+(y') \\ \zeta(w) + \frac{\pi(y')-\pi(y)}{2}P(y,w) & \text{if } w \in \mathcal{B}_1^+(y) \\ \zeta(w) + \frac{\pi(y)-\pi(y')}{2}P(y',w) & \text{if } w \in \mathcal{B}_1^+(y'). \end{cases}$$

Therefore, the vector $\phi := \zeta(I-P_*)$ satisfies

$$|\phi(w)| \le \begin{cases} 0 & \text{if } w \notin \mathcal{B}_1^+(y)\cup\mathcal{B}_1^+(y') \\ \frac{|\pi(y)-\pi(y')|}{2\Delta} & \text{otherwise .} \end{cases}$$

Hence $\phi(v) = 0$ for all but at most 2Δ vertices v, and recalling (6.26) we have

$$|\phi(w)| \le (2\Delta)^{-1}\pi_{\min}\log^{-K}(n). \tag{6.35}$$

Next, consider the matrix

$$M = \sum_{s=0}^{T-1} P_*^s,$$

and notice that

$$\zeta(I - P_*^T) = \phi M.$$

Since P_* and π^* satisfy condition (i) in Lemma 6.1,

$$P_*^T = \Pi_* + E, \qquad \text{with} \qquad |E(u,v)| \le n^{-3}, \ \forall u,v \in V^*, \tag{6.36}$$

where Π_* denotes the matrix with all rows equal to π_*. We rewrite the vector ζ as

$$\zeta = \alpha\pi_* + \rho,$$

where $\alpha \in \mathbb{R}$ and ρ is orthogonal to π_*, that is

$$\langle \rho, \pi_*\rangle = \sum_{v\in V^*} \rho(v)\pi_*(v) = 0.$$

Therefore,

$$\langle \phi M, \rho\rangle = \langle \rho, (I-E)\rho\rangle.$$

Moreover,

$$|\langle \phi M, \rho\rangle| \le \sum_{s=0}^{T-1} |\langle \phi, P_*^s\rho\rangle| \le T\frac{\pi^*_{\max}}{\pi^*_{\min}}\|\phi\|_2\|\rho\|_2, \tag{6.37}$$

where we use

$$\langle P_*^s\psi, P_*^s\psi\rangle \le \frac{1}{\pi^*_{\text{min}}}\sum_v \pi^*(v)(P_*^s\psi)^2(v)$$
$$\le \frac{1}{\pi^*_{\text{min}}}\sum_{u,v}\pi^*(v)P_*^s(v,u)\psi^2(u) = \frac{1}{\pi^*_{\text{min}}}\sum_u \pi^*(u)\psi^2(u) \le \frac{\pi^*_{\text{max}}}{\pi^*_{\text{min}}}\|\psi\|_2^2,$$

for any vector $\psi : V^* \mapsto \mathbb{R}$. On the other hand,

$$|\langle \rho, (I-E)\rho\rangle| \ge \|\rho\|_2^2 - n^{-3}\left(\sum_v |\rho(v)|\right)^2 \ge \|\rho\|_2^2(1-n^{-2}). \tag{6.38}$$

Using (6.35), from (6.37) and (6.38) we conclude that

$$\|\rho\|_2 \le 2T\frac{\pi^*_{\text{max}}}{\pi^*_{\text{min}}}\|\phi\|_2 = 2T\frac{\pi^*_{\text{max}}}{\pi^*_{\text{min}}} \times O(\pi_{\text{min}}\log^{-K}(n)).$$

From Theorem 5.1 applied to G^* we can assume that $\frac{\pi^*_{\text{max}}}{\pi^*_{\text{min}}} = O(\log^{K/3}(n))$ if K is a large enough constant. Since $T = \log^3(n)$, with K sufficiently large one has

$$\|\rho\|_2 \le \pi_{\text{min}}\log^{-K/2}(n).$$

Next, notice that

$$0 = \langle \zeta, 1\rangle = \langle \alpha\pi_* + \rho, 1\rangle = \alpha + \langle \rho, 1\rangle.$$

Hence

$$|\alpha| = |\langle \rho, 1\rangle| \le \sqrt{n}\,\|\rho\|_2 \le \sqrt{n}\,\pi_{\text{min}}\log^{-K/2}(n).$$

In conclusion,

$$\zeta(v)^2 \le 2\alpha^2\pi_*(v)^2 + 2\rho(v)^2 \le 2n\pi^2_{\text{min}}\log^{-K}(n)(\pi^*_{\text{max}})^2 + 2\|\rho\|_2^2$$
$$\le 2n\pi^2_{\text{min}}\log^{-K}(n)(\pi^*_{\text{max}})^2 + 2\pi^2_{\text{min}}\log^{-K}(n) \le 4\pi^2_{\text{min}}\log^{-K}(n),$$

which implies (6.34). □

6.2 The Eulerian case

We prove Theorem 6.2. The strategy is the same as for the proof of Theorem 6.1, with some significant simplifications due to the explicit knowledge of the invariant measure $\pi(x) = d_x/m$. For the upper bound, it is then sufficient to prove that, setting $t_* = (1+\varepsilon)\beta n \log n$,

$$\sum_{y\in V_1}\sum_{s\geq t_*}\mathbf{P}_x(\mathcal{A}_y^T(s)) + \sum_{y\in V_2}\sum_{s\geq t_*}\mathbf{P}_x(\mathcal{A}_y^T(s)) = o(n\log n). \tag{6.39}$$

Letting $\mathcal{V}_d$ denote the set of vertices with degree d, reasoning as in (6.17) we have

$$\sum_{y\in V_1}\sum_{s\geq t_*}\mathbf{P}_x(\mathcal{A}_y^T(s)) \leq (1+o(1))\sum_{d=\delta}^{\Delta}|\mathcal{V}_d|\frac{m}{d(1+(1+o(1))d/m)^{t_*}}$$

Since $|\mathcal{V}_d| = n^{\alpha_d+o(1)}$, $m = \bar{d}n$, for any fixed $\varepsilon > 0$ we obtain

$$\sum_{y\in V_1}\sum_{s\geq t_*}\mathbf{P}_x(\mathcal{A}_y^T(s)) \leq \frac{2m}{\delta}\sum_{d=\delta}^{\Delta}\exp\left(-\left(\tfrac{d\beta}{\bar{d}} - \alpha_d\right)\log n\right) = O(n), \tag{6.40}$$

since by definition $\frac{d\beta}{\bar{d}} - \alpha_d \geq 0$. Concerning the vertices $y \in V_2$ one may repeat the argument in (6.20) without modifications, to obtain

$$\sum_{y\in V_2}\sum_{s\geq t_*}\mathbf{P}_x(\mathcal{A}_y^T(s)) = o(n\log n). \tag{6.41}$$

Thus, (6.39) follows from (6.40) and (6.41).

It remains to prove the lower bound. We shall prove that for any fixed d such that $|\mathcal{V}_d| = n^{\alpha_d+o(1)}$, $\alpha_d \in (0,1]$, for any $\varepsilon > 0$,

$$\min_{x\in[n]}\mathbf{P}_x\left(\tau_{\text{cov}} \geq (1-\varepsilon)\frac{\bar{d}\alpha_d}{d}\,n\log^{\gamma_1} n\right) = 1 - o(1). \tag{6.42}$$

We proceed as in the proof of Lemma 6.4. Here we choose W as the subset of $\mathcal{V}_d$ consisting of LTL vertices in the sense of Definition 6.1 and such that for all $x, y \in W$ one has $\min\{d(x,y), d(y,x)\} > 2\vartheta$. Let us check that this set satisfies

$$|W| \geq n^{\alpha_d+o(1)}. \tag{6.43}$$

Indeed, the vertices that are not LTL are at most $\Delta^{9\vartheta}$ by Proposition 6.1. Therefore there are at least $|\mathcal{V}_d| - \Delta^{9\vartheta} = n^{\alpha_d + o(1)}$ LTL vertices in $\mathcal{V}_d$. Moreover, since there are at most $\Delta^{2\vartheta}$ vertices at undirected distance 2ϑ from any vertex, we can take a subset W of LTL vertices of $\mathcal{V}_d$ satisfying the requirement that $\min\{d(x,y), d(y,x)\} > 2\vartheta$ for all $x, y \in W$ and such that $|W| \geq (|\mathcal{V}_d| - \Delta^{9\vartheta})\Delta^{-2\vartheta} = n^{\alpha_d + o(1)}$. From here on all arguments can be repeated without modifications, with the simplification that we no longer need a proof of Lemma 6.5 since a can be taken to be zero in (6.32) in the Eulerian case. The only thing to control is the validity of the bound (6.29) with the choice

$$t = (1-\varepsilon)\frac{\bar{d}\alpha_d}{d}\, n \log n.$$

As in (6.29), it suffices to check that with high probability

$$\sum_{y \in W} \mathbb{P}(\mathcal{A}_y^T(t)) - T \to \infty. \tag{6.44}$$

From (6.27) we obtain

$$\sum_{y \in W} \mathbb{P}\left(\mathcal{A}_y^T(t)\right) = (1+o(1))|W| \exp\left(-\tfrac{(1+o(1))d}{m}\, t\right). \tag{6.45}$$

Using (6.43) and $dt/m = (1-\varepsilon)\alpha_d \log n$, (6.45) is at least $n^{\varepsilon\alpha_d/2}$ for all n large enough. Since $T = \log^3(n)$ this proves (6.44).

BIBLIOGRAPHY

[1] Abdullah, M., Cooper, C., Frieze, A.M.: Cover time of a random graph with given degree sequence. Discrete Mathematics **312**(21), 3146–3163 (2012)

[2] Addario-Berry, L., Balle, B., Perarnau, G.: Diameter and stationary distribution of random r-out digraphs. arXiv preprint arXiv:1504.06840 (2015)

[3] Aldous, D.: On the time taken by random walks on finite groups to visit every state. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **62**(3), 361–374 (1983)

[4] Aldous, D.: An introduction to covering problems for random walks on graphs. Journal of Theoretical Probability **2**(1), 87–89 (1989)

[5] Aldous, D., Diaconis, P.: Shuffling cards and stopping times. The American Mathematical Monthly **93**(5), 333–348 (1986)

[6] Aleliunas, R., Karp, R.M., Lipton, R.J., Lovasz, L., Rackoff, C.: Random walks, universal traversal sequences, and the complexity of maze problems. In: 20th Annual Symposium on Foundations of Computer Science (sfcs 1979). pp. 218–223. IEEE (1979)

[7] Andersen, R., Chung, F., Lang, K.: Local graph partitioning using pagerank vectors. In: 2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06). pp. 475–486. IEEE (2006)

[8] Avena, L., Güldas, H., van der Hofstad, R., den Hollander, F.: Random walks on dynamic configuration models: a trichotomy. Stochastic Processes and their Applications (2018)

[9] Ben-Hamou, A., Salez, J.: Cutoff for nonbacktracking random walks on sparse random graphs. The Annals of Probability **45**(3), 1752–1770 (2017)

[10] Boczkowski, L., Peres, Y., Sousi, P.: Sensitivity of mixing times in eulerian digraphs. SIAM Journal on Discrete Mathematics **32**(1), 624–655 (2018)

[11] Bollobàs, B.: A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. European Journal of Combinatorics **1**(4), 311–316 (1980)

[12] Bollobàs, B., Fernandez de la Vega, W.: The diameter of random regular graphs. Combinatorica **2**(2), 125–134 (1982)

[13] Bordenave, C., Caputo, P., Salez, J.: Random walk on sparse random digraphs. Probability Theory and Related Fields **170**(3), 933–960 (Apr 2018)

[14] Bordenave, C., Caputo, P., Salez, J.: Cutoff at the "entropic time" for sparse markov chains. Probability Theory and Related Fields **173**(1), 261–292 (Feb 2019)

[15] Bordenave, C., Lelarge, M., Massoulié, L.: Nonbacktracking spectrum of random graphs: Community detection and nonregular ramanujan graphs. The Annals of Probability **46**(1), 1–71 (01 2018)

[16] Bressan, M., Peserico, E.: Choose the damping, choose the ranking? Journal of Discrete Algorithms **8**(2), 199 – 213 (2010), selected papers from the 3rd Algorithms and Complexity in Durham Workshop ACiD 2007

[17] Brin, S., Page, L.: The anatomy of a large-scale hypertextual web search engine. Computer networks and ISDN systems **30**(1-7), 107–117 (1998)

[18] Caputo, P., Quattropani, M.: Mixing time of pagerank surfers on sparse random digraphs. arXiv preprint arXiv:1905.04993 (2019)

[19] Caputo, P., Quattropani, M.: Mixing time trichotomy in regenerating dynamic digraphs. In preparation (2019+)

[20] Caputo, P., Quattropani, M.: Stationary distribution and cover time of sparse directed configuration models. arXiv:1909.05752 (2019)

[21] Chatterjee, S.: Stein's method for concentration inequalities. Probability theory and related fields **138**(1), 305–321 (2007)

[22] Chen, N., Litvak, N., Olvera-Cravioto, M.: Generalized pagerank on directed configuration networks. Random Structures & Algorithms **51**(2), 237–274 (2017)

[23] Combes, R.: An extension of McDiarmid's inequality. arXiv preprint arXiv:1511.05240 (2015)

[24] Cooper, C., Frieze, A.: The size of the largest strongly connected component of a random digraph with a given degree sequence. Combinatorics, Probability and Computing **13** (2002)

[25] Cooper, C., Frieze, A.: The cover time of sparse random graphs. Random Structures & Algorithms **30**(1-2), 1–16 (2007)

[26] Cooper, C., Frieze, A.: The cover time of the giant component of a random graph. Random Structures & Algorithms **32**(4), 401–439 (2008)

[27] Cooper, C., Frieze, A.M.: The cover time of random regular graphs. SIAM J. Discrete Math. **18**(4), 728–740 (2005)

[28] Cooper, C., Frieze, A.M.: Stationary distribution and cover time of random walks on random digraphs. J. Comb. Theory, Ser. B **102**(2), 329–362 (2012)

[29] Diaconis, P.: The cutoff phenomenon in finite markov chains. Proceedings of the National Academy of Sciences **93**(4), 1659–1664 (1996)

[30] Easley, D., Kleinberg, J.: Networks, crowds, and markets, vol. 8. Cambridge university press Cambridge (2010)

[31] Feige, U.: A tight lower bound on the cover time for random walks on graphs. Random Structures & Algorithms **6**(4), 433–438 (1995)

[32] Feige, U.: A tight upper bound on the cover time for random walks on graphs. Random Structures & Algorithms **6**(1), 51–54 (1995)

[33] Freedman, D.A.: On tail probabilities for martingales. The Annals of Probability **3**(1), 100–118 (1975)

[34] Garavaglia, A., van der Hofstad, R., Litvak, N.: Local weak convergence for pagerank. arXiv preprint arXiv:1803.06146 (2018)

[35] Gleich, D.F.: Pagerank beyond the web. SIAM Review **57**(3), 321–363 (2015)

[36] Grusho, A.A.: Limit distributions of certain characteristics of random automaton graphs. Mathematical notes of the Academy of Sciences of the USSR **14**(1), 633–637 (1973)

[37] Jeh, G., Widom, J.: Scaling personalized web search. In: Proceedings of the 12th international conference on World Wide Web. pp. 271–279. Acm (2003)

[38] Jonasson, J.: On the cover time for random walks on random graphs. Combinatorics, Probability and Computing **7**(3), 265–279 (1998)

[39] Kahn, J.D., Linial, N., Nisan, N., Saks, M.E.: On the cover time of random walks on graphs. Journal of Theoretical Probability **2**(1), 121–128 (1989)

[40] Levin, D.A., Peres, Y.: Markov Chains and Mixing Times. American Mathematical Society, Providence, RI (2017), second edition. With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson

[41] Liu, Q.: The growth of an entire characteristic fonction and the tail probabilities of the limit of a tree martingale. In: Chauvin, B., Cohen, S., Rouault, A. (eds.) Trees. pp. 51–80. Birkhäuser Basel, Basel (1996)

[42] Liu, Q.: Products of random weights indexed by galton-watson trees. Publications mathématiques et informatique de Rennes (2) (1996-1997)

[43] Liu, Q.: On generalized multiplicative cascades. Stochastic Processes and their Applications **86**(2), 263–286 (2000)

[44] Liu, Q.: Asymptotic properties and absolute continuity of laws stable by random weighted mean. Stochastic Processes and their Applications **95**(1), 83–107 (2001)

[45] Lovász, L.: Random walks on graphs: A survey. Combinatorics, Paul erdos is eighty **2**(1), 1–46 (1993)

[46] Lubetzky, E., Sly, A.: Cutoff phenomena for random walks on random regular graphs. Duke Mathematical Journal **153**(3), 475–510 (2010)

[47] Matthews, P.: Covering problems for brownian motion on spheres. Ann. Probab. **16**(1), 189–199 (1988)

[48] Matthews, P.: Covering problems for markov chains. Ann. Probab. **16**(3), 1215–1228 (1988)

[49] McDiarmid, C.: Concentration. In: Probabilistic methods for algorithmic discrete mathematics, pp. 195–248. Springer (1998)

[50] Vial, D., Subramanian, V.: Restart perturbations for lazy, reversible markov chains: trichotomy and pre-cutoff equivalence. arXiv preprint arXiv:1907.02926 (2019)