



# Stationary distribution and cover time of sparse directed configuration models

Pietro Caputo<sup>1</sup> · Matteo Quattropani<sup>1</sup>

Received: 24 September 2019 / Revised: 28 July 2020  
© The Author(s) 2020

## Abstract

We consider sparse digraphs generated by the configuration model with given in-degree and out-degree sequences. We establish that with high probability the cover time is linear up to a poly-logarithmic correction. For a large class of degree sequences we determine the exponent  $\gamma \geq 1$  of the logarithm and show that the cover time grows as  $n \log^\gamma(n)$ , where  $n$  is the number of vertices. The results are obtained by analysing the extremal values of the stationary distribution of the digraph. In particular, we show that the stationary distribution  $\pi$  is uniform up to a poly-logarithmic factor, and that for a large class of degree sequences the minimal values of  $\pi$  have the form  $\frac{1}{n} \log^{1-\gamma}(n)$ , while the maximal values of  $\pi$  behave as  $\frac{1}{n} \log^{1-\kappa}(n)$  for some other exponent  $\kappa \in [0, 1]$ . In passing, we prove tight bounds on the diameter of the digraphs and show that the latter coincides with the typical distance between two vertices.

**Mathematics Subject Classification** Primary 05C81 · 60J10 · 60C05; Secondary 60G42

## 1 Introduction

The problem of determining the cover time of a graph is a central one in combinatorics and probability [3–5, 18, 19, 21, 26]. In recent years, the cover time of random graphs has been extensively studied [1, 14, 16, 17, 20]. All these works consider undirected graphs, with the notable exception of the paper [17] 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

---

✉ Pietro Caputo  
pietro.caputo@uniroma3.it

Matteo Quattropani  
matteo.quattropani@uniroma3.it

<sup>1</sup> Dipartimento di Matematica e Fisica, Università di Roma Tre, Largo S. Leonardo Murialdo 1, 00146 Rome, Italy

main difficulty in the directed case is that, in contrast with the undirected case, the graph’s stationary distribution is an unknown random variable.

In this paper we address the problem of determining the cover time of sparse random digraphs with bounded degrees. More specifically, we consider random digraphs  $G$  with given in- and out-degree sequences, generated via the configuration model. For the sake of this introductory discussion let us look at the special case where all vertices have either in-degree 2 and out-degree 3 or in-degree 3 and out-degree 2, with the two types evenly represented in the vertex set  $V(G)$ . We refer to this as the  $(2, 3)(3, 2)$  case. With high probability  $G$  is strongly connected and we may ask how long the random walk on  $G$  takes to cover all the nodes. The expectation of this quantity, maximized over the initial point of the walk defines  $T_{\text{cov}}(G)$ , the cover time of  $G$ . We will show that with high probability as the number of vertices  $n$  tends to infinity one has

$$T_{\text{cov}}(G) \asymp n \log^\gamma(n) \tag{1.1}$$

where  $\gamma = \frac{\log 3}{\log 2} \approx 1.58$ , and  $a_n \asymp b_n$  stands for  $C^{-1} \leq a_n/b_n \leq C$  for some constant  $C > 0$ . The constant  $\gamma$  can be understood in connection with the statistics of the extremal values of the stationary distribution  $\pi$  of  $G$ . Indeed, following the theory developed by Cooper and Frieze, if the graphs satisfy suitable requirements, then the problem of determining the cover time can be reformulated in terms of the control of the minimal values of  $\pi$ . In particular, we will see that the hitting time of a vertex  $x \in V(G)$  effectively behaves as an exponential random variable with parameter  $\pi(x)$ , and that to some extent these random variables are weakly dependent. This supports the heuristic picture that represents the cover time as the expected value of  $n$  independent exponential random variables, each with parameter  $\pi(x)$ ,  $x \in V(G)$ . Controlling the stationary distribution is however a rather challenging task, especially if the digraphs have bounded degrees.

Recently, Bordenave, Caputo and Salez [9] analyzed the mixing time of sparse random digraphs with given degree sequences and their work provides some important information on the distribution of the values of  $\pi$ . In particular, in the  $(2, 3)(3, 2)$  case, the empirical distribution of the values  $\{n\pi(x), x \in V(G)\}$  converges as  $n \rightarrow \infty$  to the probability law  $\mu$  on  $[0, \infty)$  of the random variable  $X$  given by

$$X = \frac{2}{5} \sum_{k=1}^N Z_k, \tag{1.2}$$

where  $N$  is the random variable with  $N = 2$  with probability  $\frac{1}{2}$  and  $N = 3$  with probability  $\frac{1}{2}$ , and the  $Z_k$  are independent and identically distributed mean-one random variables uniquely determined by the recursive distributional equation

$$Z_1 \stackrel{d}{=} \frac{1}{M} \sum_{k=1}^{5-M} Z_k, \tag{1.3}$$

where  $M$  is the random variable with  $M = 2$  with probability  $2/5$  and  $M = 3$  with probability  $3/5$ , independent of the  $Z_k$ 's, and  $\stackrel{d}{=}$  denotes equality in distribution.

This gives convergence of the distribution of the bulk values of  $\pi$ , that is of the values of  $\pi$  on the scale  $1/n$ . What enters in the cover time analysis are however the extremal values, notably the minimal ones, and thus what is needed is a local convergence result towards the left tail of  $\mu$ , which cannot be extracted from the analysis in [9]. To obtain a heuristic guess of the size of the minimal values of  $\pi$  at large but finite  $n$  one may pretend that the values of  $n\pi$  are  $n$  i.i.d. samples from  $\mu$ . This would imply that  $\pi_{\min}$ , the minimal value of  $\pi$  is such that  $n\pi_{\min} \sim \varepsilon(n)$  where  $\varepsilon(n)$  is a sequence for which  $n\mu([0, \varepsilon(n)]) \sim 1$ , if  $\mu([0, x])$  denotes the mass given by  $\mu$  to the interval  $[0, x]$ .

Recursive distributional equations of the form (1.3) have been extensively studied, and much is known about the law of random variables satisfying this type of stability and self-similarity; see, e.g., [6,23–25,27] and references therein. In particular, the left and right tail asymptotic of the law  $\mu$  defined by (1.2) and (1.3) can be obtained as a special case of these results. For instance, it is shown in Liu [25, Theorem 2.5] that the left tail of  $\mu$  must have the form

$$\log \mu([0, x]) \asymp -x^{-\alpha}, \quad x \rightarrow 0^+,$$

where  $\alpha = 1/(\gamma - 1)$ , with the coefficient  $\gamma$  taking the value  $\gamma = \frac{\log 3}{\log 2}$  in the  $(2, 3)(3, 2)$  case. Thus, returning to our heuristic reasoning, one has that the minimal value of  $\pi$  should satisfy

$$n\pi_{\min} \asymp \log^{1-\gamma}(n). \tag{1.4}$$

Moreover, this argument also suggests that with high probability there should be at least  $n^\beta$  vertices  $x \in V(G)$ , for some constant  $\beta > 0$ , such that  $n\pi(x)$  is as small as  $O(\log^{1-\gamma}(n))$ .

A similar heuristic argument, this time based on the analysis of the right tail of  $\mu$ , see [23,24] (in particular, see the first display in [24, p. 271]), predicts that  $\pi_{\max}$ , the maximal value of  $\pi$ , should satisfy

$$n\pi_{\max} \asymp \log^{1-\kappa}(n), \tag{1.5}$$

where  $\kappa$  takes the value  $\kappa = \frac{\log 2}{\log 3} \approx 0.63$  in the  $(2, 3)(3, 2)$  case.

Our main results below will confirm these heuristic predictions. The proof involves the analysis of the statistics of the in-neighbourhoods of a node. Roughly speaking, it will be seen that the smallest values of  $\pi$  are achieved at vertices  $x \in V(G)$  whose in-neighbourhood at distance  $\log_2 \log n$  is a directed tree composed entirely of vertices with in-degree 2 and out-degree 3, while the maximal values of  $\pi$  are achieved at  $x \in V(G)$  whose in-neighbourhood at distance  $\log_3 \log n$  is a directed tree composed entirely of vertices with in-degree 3 and out-degree 2. Once the results (1.4) and (1.5) are established, the cover time asymptotic (1.1) will follow from an appropriate implementation of the Cooper-Frieze approach.

We conclude this preliminary discussion by comparing our estimates (1.4) and (1.5) with related results for different random graph models. The asymptotic of extremal values of  $\pi$  has been determined in [17] for the directed Erdős-Renyi random graphs

with logarithmically diverging average degree. There, the authors show that  $n\pi_{\min}$  and  $n\pi_{\max}$  are essentially of order 1, which can be interpreted as a concentration property enforced by the divergence of the degrees. On the other hand, for uniformly random out-regular digraphs, that is with constant out-degrees but random in-degrees, the recent paper [2] shows that the stationary distribution restricted to the strongly connected component satisfies  $n\pi_{\min} = n^{-\eta+o(1)}$ , where  $\eta$  is a computable constant, and  $n\pi_{\max} = n^{o(1)}$ . Indeed, in this model in contrast with our setting one can have in-neighborhoods made by long and thin filaments which determine a power law deviation from uniformity.

We now turn to a more systematic exposition of our results.

### 1.1 Model and statement of results

Set  $[n] = \{1, \dots, n\}$ , and for each integer  $n$ , fix two sequences  $\mathbf{d}^+ = (d_x^+)_{x \in [n]}$  and  $\mathbf{d}^- = (d_x^-)_{x \in [n]}$  of positive integers such that

$$m = \sum_{x=1}^n d_x^+ = \sum_{x=1}^n d_x^-. \tag{1.6}$$

The *directed configuration model*  $\text{DCM}(\mathbf{d}^\pm)$  is the distribution of the random digraph  $G$  with vertex set  $V(G) = [n]$  obtained by the following procedure: 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  $\omega$ ; 3) for all  $x, y \in [n]$ , add a directed edge  $(x, y)$  every time a tail from  $x$  is mapped into a head from  $y$  through  $\omega$ . The resulting digraph  $G$  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  $G$  has the uniform distribution among simple digraphs with in degree sequence  $\mathbf{d}^-$  and out degree sequence  $\mathbf{d}^+$ .

Structural properties of digraphs obtained in this way have been studied in [13]. Here we consider the sparse case corresponding to bounded degree sequences and, in order to avoid non irreducibility issues, we shall assume that all degrees are at least 2. Thus, from now on it will always be assumed that

$$\delta_\pm = \min_{x \in [n]} d_x^\pm \geq 2 \quad \Delta_\pm = \max_{x \in [n]} d_x^\pm = O(1). \tag{1.7}$$

Under the first assumption it is known that  $\text{DCM}(\mathbf{d}^\pm)$  is strongly connected with high probability. Under the second assumption, it is known that  $\text{DCM}(\mathbf{d}^\pm)$  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  $\text{DCM}(\mathbf{d}^\pm)$  will also hold with high probability for a uniformly random simple digraph with degrees given by  $\mathbf{d}^-$  and  $\mathbf{d}^+$  respectively. Here and throughout the rest of the paper we say that a property holds with high probability (w.h.p. for short) if the probability of the corresponding event converges to 1 as  $n \rightarrow \infty$ .

The (directed) distance  $d(x, y)$  from  $x$  to  $y$  is the minimal number of edges that need to be traversed to reach  $y$  from  $x$ . The diameter is the maximal distance between

two distinct vertices, i.e.

$$\text{diam}(G) = \max_{x \neq y} d(x, y). \tag{1.8}$$

We begin by showing that the diameter  $\text{diam}(G)$  concentrates around the value  $c \log n$  within a  $O(\log \log n)$  window, where  $c$  is given by  $c = 1/\log v$  and  $v$  is the parameter defined by

$$v = \frac{1}{m} \sum_{y=1}^n d_y^- d_y^+. \tag{1.9}$$

**Theorem 1.1** *Set  $d_\star = \log_v n$ . There exists  $\varepsilon_n = O\left(\frac{\log \log(n)}{\log(n)}\right)$  such that*

$$\mathbb{P}((1 - \varepsilon_n) d_\star \leq \text{diam}(G) \leq (1 + \varepsilon_n) d_\star) = 1 - o(1). \tag{1.10}$$

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

$$\mathbb{P}((1 - \varepsilon_n) d_\star \leq d(x, y) \leq (1 + \varepsilon_n) d_\star) = 1 - o(1). \tag{1.11}$$

The proof of Theorem 1.1 is a directed version of a classical argument for undirected graphs [8]. It requires controlling the size of in- and out-neighborhoods of a node, which in turn follows ideas from [2] and [9]. The value  $d_\star = \log_v n$  can be interpreted as follows: both the in- and the out-neighborhood of a node are tree-like with average branching given by  $v$ , so that their boundary at depth  $h$  has typically size  $v^h$ , see Lemma 2.10; if the in-neighborhood of  $y$  and the out-neighborhood of  $x$  are exposed up to depth  $h$ , one finds that the value  $h = \frac{1}{2} \log_v(n)$  is critical for the formation of an arc connecting the two neighborhoods.

In particular, Theorem 1.1 shows that w.h.p. the digraph is strongly connected, so there exists a unique stationary distribution  $\pi$  characterized by the equation

$$\pi(x) = \sum_{y=1}^n \pi(y) P(y, x), \quad x \in [n], \tag{1.12}$$

with the normalization  $\sum_{x=1}^n \pi(x) = 1$ . Here  $P$  is the transition matrix of the simple random walk on  $G$ , namely

$$P(y, x) = \frac{m(y, x)}{d_y^+}, \tag{1.13}$$

and we write  $m(y, x)$  for the multiplicity of the edge  $(y, x)$  in the digraph  $G$ . If the sequences  $\mathbf{d}^\pm$  are such that  $d_x^+ = d_x^-$  for all  $x \in [n]$ , then the stationary distribution is given by

$$\pi(x) = \frac{d_x^\pm}{m}. \tag{1.14}$$

The digraph is called Eulerian in this case. In all other cases the stationary distribution is a nontrivial random variable. To discuss our results on the extremal values of  $\pi$  it is convenient to introduce the following notation.

**Definition 1.2** 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 (1.7) 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

$$\liminf_{n \rightarrow \infty} \frac{|\mathcal{V}_{i,j}|}{n} > 0. \tag{1.15}$$

We call  $\mathcal{L} \subset \mathcal{C}$  the set of types with linear size, and define the parameters

$$\gamma_0 := \frac{\log \Delta_+}{\log \delta_-}, \quad \gamma_1 := \max_{(k,\ell) \in \mathcal{L}} \frac{\log \ell}{\log k}, \quad \kappa_1 := \min_{(k,\ell) \in \mathcal{L}} \frac{\log \ell}{\log k}, \quad \kappa_0 := \frac{\log \delta_+}{\log \Delta_-}. \tag{1.16}$$

**Theorem 1.3** Set  $\pi_{\min} = \min_{x \in [n]} \pi(x)$ . There exists a constant  $C > 0$  such that

$$\mathbb{P}\left(C^{-1} \log^{1-\gamma_0}(n) \leq n\pi_{\min} \leq C \log^{1-\gamma_1}(n)\right) = 1 - o(1). \tag{1.17}$$

Moreover, there exists  $\beta > 0$  such that

$$\mathbb{P}\left(\exists S \subset [n], |S| \geq n^\beta, n \max_{y \in S} \pi(y) \leq C \log^{1-\gamma_1}(n)\right) = 1 - o(1). \tag{1.18}$$

Notice that  $\gamma_0 \geq \gamma_1 \geq 1$ . If the sequences  $\mathbf{d}^\pm$  are such that  $(\delta_-, \Delta_+) \in \mathcal{L}$ , then  $\gamma_0 = \gamma_1 =: \gamma$ . This immediately implies the following

**Corollary 1.4** If the sequences  $\mathbf{d}^\pm$  are such that  $(\delta_-, \Delta_+) \in \mathcal{L}$ , then

$$\pi_{\min} \asymp \frac{1}{n} \log^{1-\gamma}(n) \quad \text{w.h.p.} \tag{1.19}$$

**Remark 1.5** If  $(\delta_-, \Delta_+) \notin \mathcal{L}$ , then the estimate (1.17) can be strengthened by replacing  $\gamma_0$  with  $\gamma'_0$  where

$$\gamma'_0 := \frac{\log \Delta'_+}{\log \delta'_-}, \quad \Delta'_+ := \max\{\ell : (k, \ell) \in \mathcal{L}_0\}, \quad \delta'_- := \min\{k : (k, \ell) \in \mathcal{L}_0\}, \tag{1.20}$$

and  $\mathcal{L}_0 \subset \mathcal{C}$  is defined as the set of  $(k, \ell) \in \mathcal{C}$  such that

$$\limsup_{n \rightarrow \infty} \frac{|\mathcal{V}_{k,\ell}|}{n^{1-a}} = +\infty, \quad \forall a > 0. \tag{1.21}$$

We refer to Remark 3.9 below for additional details on this improvement.

Concerning the maximal values of  $\pi$  we establish the following estimates.

**Theorem 1.6** *Set  $\pi_{\max} = \max_{x \in [n]} \pi(x)$ . There exists a constant  $C > 0$  such that*

$$\mathbb{P}\left(C^{-1} \log^{1-\kappa_1}(n) \leq n\pi_{\max} \leq \log^{1-\kappa_0}(n)\right) = 1 - o(1). \tag{1.22}$$

Moreover, there exists  $\beta > 0$  such that

$$\mathbb{P}\left(\exists S \subset [n], |S| \geq n^\beta, n \min_{y \in S} \pi(y) \geq C^{-1} \log^{1-\kappa_1}(n)\right) = 1 - o(1). \tag{1.23}$$

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 we obtain the following

**Corollary 1.7** *If  $(\Delta_-, \delta_+) \in \mathcal{L}$*

$$\pi_{\max} \asymp \frac{1}{n} \log^{1-\kappa}(n) \quad \text{w.h.p.} \tag{1.24}$$

**Remark 1.8** In analogy with Remark (1.5), if  $(\Delta_-, \delta_+) \notin \mathcal{L}$ , then (1.22) can be improved by replacing  $\kappa_0$  with  $\kappa'_0$  where

$$\kappa'_0 := \frac{\log \delta'_+}{\log \Delta'_-}, \quad \delta'_+ := \min\{\ell : (k, \ell) \in \mathcal{L}_0\}, \quad \Delta'_- := \max\{k : (k, \ell) \in \mathcal{L}_0\}, \tag{1.25}$$

We turn to a description of our results concerning the cover time. A central role here will be played by both statements in Theorem 1.3. On the other hand the information on  $\pi_{\max}$  derived in Theorem 1.6 will not be essential, and the statement (1.23) will never be used in our proof.

Let  $X_t, t = 0, 1, 2, \dots$ , denote the simple random walk on the digraph  $G$ , that is the Markov chain with transition matrix  $P$  defined in (1.13). Consider the hitting times

$$H_y = \inf\{t \geq 0 : X_t = y\}, \quad \tau_{\text{cov}} = \max_{y \in [n]} H_y. \tag{1.26}$$

The cover time  $T_{\text{cov}} = T_{\text{cov}}(G)$  is defined by

$$T_{\text{cov}} = \max_{x \in [n]} \mathbf{E}_x[\tau_{\text{cov}}], \tag{1.27}$$

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  $G$ . Let  $\gamma_0, \gamma_1$  be as in Definition 1.2

**Theorem 1.9** *There exists a constant  $C > 0$  such that*

$$\mathbb{P}\left(C^{-1} n \log^{\gamma_1}(n) \leq T_{\text{cov}} \leq C n \log^{\gamma_0}(n)\right) = 1 - o(1). \tag{1.28}$$

**Corollary 1.10** For sequences  $\mathbf{d}^\pm$  such that  $(\delta_-, \Delta_+) \in \mathcal{L}$  one has  $\gamma_0 = \gamma_1 = \gamma$  and

$$T_{\text{cov}} \asymp n \log^\gamma(n), \quad \text{w.h.p.} \tag{1.29}$$

**Remark 1.11** As in Remark 1.5, if  $(\delta_-, \Delta_+) \notin \mathcal{L}$ , then Theorem 1.9 can be strengthened by replacing  $\gamma_0$  with the constant  $\gamma'_0$  defined in (1.20).

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 1.9 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 [1].

**Theorem 1.12** 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

$$|\mathcal{V}_d| = n^{\alpha_d + o(1)} \tag{1.30}$$

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

$$T_{\text{cov}} = (\beta + o(1)) n \log n, \quad \text{w.h.p.} \tag{1.31}$$

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 (1.31) holds with  $\beta = \bar{d}/\delta$ , where  $\delta = \delta_\pm$  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 [7].

The rest of the paper is divided into three sections. The first is a collection of preliminary structural facts about the directed configuration model. It also includes the proof of Theorem 1.1. The second section is the core of the paper. There we establish Theorem 1.3 and Theorem 1.6. The last section contains the proof of the cover time results Theorem 1.9 and Theorem 1.12.

## 2 Neighborhoods and diameter

We start by recalling some simple facts about the directed configuration model.

### 2.1 Sequential generation

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  $\omega$  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.

### 2.2 In-neighborhoods and out-neighborhoods

We will use the notation

$$\delta = \min\{\delta_-, \delta_+\}, \quad \Delta = \max\{\Delta_-, \Delta_+\}. \tag{2.1}$$

For any  $h \in \mathbb{N}$ , the  $h$ -in-neighborhood 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  $(e_1 f_1, \dots, e_k f_k)$  such that  $v_{f_i} = v_{e_{i+1}}$  for all  $i = 1, \dots, k - 1$ , 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-neighborhood 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  $h \in \mathbb{N}$ .

During the generation process of the in-neighborhood, 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  $k^{\text{th}}$  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  $k^{\text{th}}$  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  $\text{Bin}(k, p_k)$ . In particular,

$$\mathbb{P}(Z_k \geq \ell) \leq \frac{k^\ell p_k^\ell}{\ell!}, \quad \ell \in \mathbb{N}. \tag{2.2}$$

Notice that as long as no collision occurs, the resulting digraph is a directed tree. The same applies to out-neighborhoods simply by inverting the role of heads and tails.

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

$$\text{TX}(G) = 1 + |E| - |V|, \tag{2.3}$$

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

$$\mathcal{G}_x(h) = \{\text{TX}(\mathcal{B}_h^-(x)) \leq 1 \text{ and } \text{TX}(\mathcal{B}_h^+(x)) \leq 1\}, \quad \mathcal{G}(h) = \bigcap_{x \in [n]} \mathcal{G}_x(h). \quad (2.4)$$

Set also

$$\hbar = \frac{1}{5} \log_{\Delta}(n). \quad (2.5)$$

**Proposition 2.1** *There exists  $\chi > 0$  such that  $\mathbb{P}(\mathcal{G}_x(\hbar)) = 1 - O(n^{-1-\chi})$  for any  $x \in [n]$ . In particular,*

$$\mathbb{P}(\mathcal{G}(\hbar)) = 1 - O(n^{-\chi}). \quad (2.6)$$

**Proof** During the generation of  $\mathcal{B}_h^-(x)$  one creates at most  $\Delta^h$  edges. It follows from (2.2) 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  $\chi > 0$ :

$$\mathbb{P}(\mathcal{G}_x(\hbar)) = 1 - O(n^{-1-\chi}). \quad (2.7)$$

The conclusion follows from the union bound. □

We will need to control the size of the boundary of our neighborhoods. 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 2.2** *There exists  $\chi > 0$  such that for all  $y \in [n]$ ,*

$$\mathbb{P}\left(|\partial\mathcal{B}_h^\pm(y)| \geq \frac{1}{2}\delta_\pm^h, \forall h \in [1, \hbar]\right) = 1 - O(n^{-1-\chi}), \quad (2.8)$$

where  $\delta_\pm$  is defined as in (1.7).

**Proof** By symmetry we may restrict to the case of in-neighborhoods. By (2.7) 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-neighborhood  $\mathcal{B}_h^-(y)$  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-neighbors  $w, w' \in \partial\mathcal{B}_{t-1}^-(y)$ ;
- (2) there exists some  $0 \leq t \leq h$  and a  $v \in \partial\mathcal{B}_t^-(y)$  s.t.  $v$  has an in-neighbor  $w$  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. 1.

In case 1a) we note that the  $(h - t)$ -in-neighborhood of  $v$  must be a directed tree with at least  $\delta_-^{h-t}$  elements on its boundary and with no intersection with the  $(h - t)$ -in-neighborhoods of other  $v' \in \partial\mathcal{B}_t^-(y)$ . Moreover,  $\mathcal{B}_{t-1}^-(y)$  must be a directed tree with

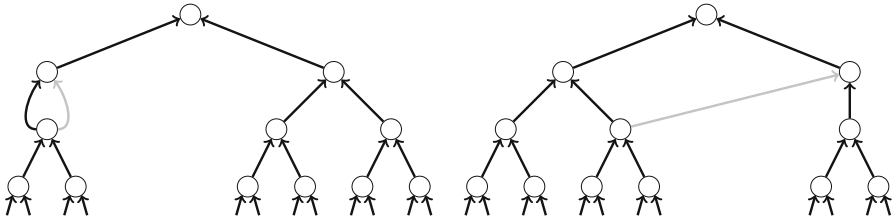


Fig. 1 The light-coloured arrow represents a collision of type (1a) (left) and a collision of type (1b) (right)

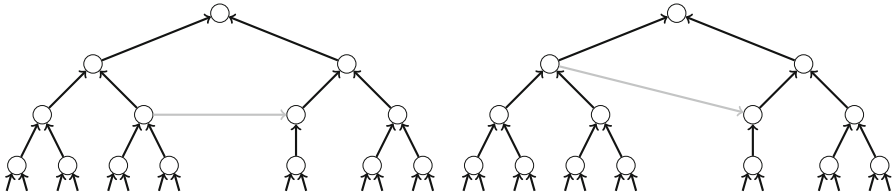


Fig. 2 Two examples of collision of type (2a)

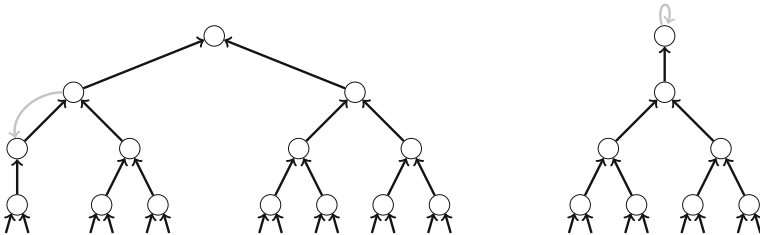


Fig. 3 Two examples of collision of type (2b)

$|\partial\mathcal{B}_{t-1}^-(y)| \geq \delta_-^{t-1}$ , and all elements of  $\partial\mathcal{B}_{t-1}^-(y)$  except one have disjoint  $(h-t+1)$ -in-neighborhoods with  $\delta_-^{h-t+1}$  elements on their boundary. Therefore

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

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

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

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 Figs. 2 and 3.

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

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

□

We shall need a more precise control of the size of  $\partial\mathcal{B}_h^\pm(y)$ , and for values of  $h$  that are larger than  $\tilde{h}$ . Recall the definition (1.9) of the parameter  $\nu$ . We use the following notation in the sequel:

$$\ell_0 = 4 \log_\delta \log(n), \quad h_\eta = (1 - \eta) \log_\nu(n). \tag{2.9}$$

**Lemma 2.3** *For every  $\eta \in (0, 1)$ , there exist constants  $c_1, c_2 > 0, \chi > 0$  such that for all  $y \in [n]$ ,*

$$\mathbb{P}\left(\nu^h \log^{-c_1}(n) \leq |\partial\mathcal{B}_h^\pm(y)| \leq \nu^h \log^{c_2}(n), \forall h \in [\ell_0, h_\eta]\right) = 1 - O(n^{-1-\chi}). \tag{2.10}$$

**Proof** We run the proof for the in-neighborhood only since the case of the out-neighborhood 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  $j$  neighborhood  $\mathcal{B}_j^-(y)$  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

$$\kappa_h := [\nu(1 - u)]^{h-\ell_0} \log^{7/2}(n), \quad \widehat{\kappa}_h := [\nu(1 + u)]^{h-\ell_0} \Delta^{\ell_0}. \tag{2.11}$$

We are going to prove

$$\mathbb{P}(\kappa_h \leq |\mathcal{F}_h| \leq \widehat{\kappa}_h, \forall h \in [\ell_0, h_\eta]) = 1 - O(n^{-1-\chi}). \tag{2.12}$$

Notice that, choosing suitable constants  $c_1, c_2 > 0$ , (2.10) is a consequence of (2.12).

Consider the events

$$A_j = \{|\mathcal{F}_i| \in [\kappa_i, \widehat{\kappa}_i], \forall i \in [\ell_0, j]\}. \tag{2.13}$$

Thus, we need to prove  $\mathbb{P}(A_h) = 1 - O(n^{-1-\chi})$ , for  $h = h_\eta$ . From Lemma 2.2 and the choice of  $\ell_0$ , it follows that

$$\mathbb{P}(A_{\ell_0}) = 1 - O(n^{-1-\chi}). \tag{2.14}$$

For  $h > \ell_0$  we write

$$\mathbb{P}(A_h) = \mathbb{P}(A_{\ell_0}) \prod_{j=\ell_0+1}^h \mathbb{P}(A_j|A_{j-1}). \tag{2.15}$$

To estimate  $\mathbb{P}(A_j|A_{j-1})$ , note that  $A_{j-1}$  depends only on the in-neighborhood  $\mathcal{B}_{j-1}^-(y)$ , so if  $\sigma_{j-1}$  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  $\sigma_{j-1}$ . Then

$$\mathbb{P}(A_j|A_{j-1}) = \frac{\sum_{\sigma_{j-1}} \mathbb{P}(\sigma_{j-1}) \mathbb{P}(A_j|\sigma_{j-1}) \mathbf{1}_{\sigma_{j-1} \in A_{j-1}}}{\mathbb{P}(A_{j-1})}. \tag{2.16}$$

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 neighborhood  $\sigma_{j-1}$  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

$$|\mathcal{F}_j| = \sum_{f \in \mathcal{F}_{j-1}} X_f^{(j)}. \tag{2.17}$$

Therefore,

$$\begin{aligned} \mathbb{P}(A_j|\sigma_{j-1}) &= \mathbb{P}(v(1-u)\kappa_{j-1} \leq |\mathcal{F}_j| \leq v(1+u)\widehat{\kappa}_{j-1} | \sigma_{j-1}) \\ &= 1 - \mathbb{P}\left(\sum_{f \in \mathcal{F}_{j-1}} X_f^{(j)} < v(1-u)\kappa_{j-1} | \sigma_{j-1}\right) \\ &\quad - \mathbb{P}\left(\sum_{f \in \mathcal{F}_{j-1}} X_f^{(j)} > v(1+u)\widehat{\kappa}_{j-1} | \sigma_{j-1}\right). \end{aligned} \tag{2.18}$$

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  $d \in [\delta, \Delta]$ ,  $\sigma_{j-1} \in A_{j-1}$ :

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

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

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

Notice that

$$\mathbb{E}(Y^{(j)}) = \sum_{d=\delta}^{\Delta} dp(d) \geq v - \frac{\Delta^2 K}{m} = v - O(n^{-\eta/2}). \tag{2.19}$$

Similarly,

$$\mathbb{E}(\widehat{Y}^{(j)}) \leq v + \frac{\Delta^2 K}{m} = v + O(n^{-\eta/2}). \tag{2.20}$$

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 (2.17) 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,  $\sum_{f \in \mathcal{F}_{j-1}} X_f^{(j)} < v(1-u)\kappa_{j-1}$  implies that

$$\sum_{i=1}^{\kappa_{j-1}} [Y_i^{(j)} - \mathbb{E}(Y^{(j)})] \leq -\frac{1}{2} u\kappa_{j-1}, \tag{2.21}$$

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

$$\sum_{i=1}^{\widehat{\kappa}_{j-1}} [\widehat{Y}_i^{(j)} - \mathbb{E}(\widehat{Y}^{(j)})] \geq \frac{1}{2} u\widehat{\kappa}_{j-1}. \tag{2.22}$$

An application of Hoeffding’s inequality shows that the probability of the events (2.21) and (2.22) 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 (2.18) we conclude that for some constant  $c > 0$ :

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

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

$$\mathbb{P}(A_j | \sigma_{j-1}) = 1 - O(n^{-3}), \tag{2.23}$$

uniformly in  $j \in [\ell_0, h_\eta]$  and  $\sigma_{j-1} \in A_{j-1}$ . By (2.16) the same bound applies to  $\mathbb{P}(A_j | A_{j-1})$  and going back to (2.15), 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 2.3. Define the events

$$F_y^\pm = F_y^\pm(\eta, c_1, c_2) = \left\{ \nu^h \log^{-c_1}(n) \leq |\partial \mathcal{B}_h^\pm(y)| \leq \nu^h \log^{c_2}(n), \forall h \in [\ell_0, h_\eta] \right\}. \tag{2.24}$$

Lemma 2.3 states that

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

Let  $\mathcal{G}(\hbar)$  be the event from Proposition 2.1.

**Lemma 2.4** *For every  $\eta \in (0, 1)$ , there exist constants  $c_1, c_2 > 0, \chi > 0$  such that for all  $y \in [n]$ ,*

$$\mathbb{P}\left(\left(F_y^\pm\right)^c; \mathcal{G}(\hbar)\right) = O\left(n^{-2-\chi}\right). \tag{2.25}$$

**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-neighborhoods of  $y$  compatible with the event  $\mathcal{G}(\hbar)$ , that is the set of labeled digraphs  $D$  such that

$$\mathbb{P}(\mathcal{B}_2^-(y) = D; \mathcal{G}(\hbar)) > 0. \tag{2.26}$$

Then

$$\mathbb{P}\left(\left(F_y^-\right)^c; \mathcal{G}(\hbar)\right) \leq \sup_{D \in \mathcal{D}_y^-} \mathbb{P}\left(\left(F_y^\pm\right)^c \mid \mathcal{B}_2^-(y) = D\right). \tag{2.27}$$

Thus it is sufficient to prove that

$$\mathbb{P}\left(\left(F_y^\pm\right)^c \mid \mathcal{B}_2^-(y) = D\right) = O\left(n^{-2-\chi}\right), \tag{2.28}$$

uniformly in  $D \in \mathcal{D}_y^-$ . To this end, we may repeat exactly the same argument as in the proof of Lemma 2.3 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 (2.14) can be strengthened to  $O(n^{-2-\chi})$  if we condition on  $\mathcal{B}_2^-(y) = D$ . That is, for some  $\chi > 0$ , uniformly in  $D \in \mathcal{D}_y^-$ ,

$$\mathbb{P}\left(A_{\ell_0} \mid \mathcal{B}_2^-(y) = D\right) = 1 - O\left(n^{-2-\chi}\right), \tag{2.29}$$

To prove (2.29) notice that if the 2-in-neighborhood 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  $|\mathcal{F}_{\ell_0}^-(y)| \geq \delta^{\ell_0}/C$  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 (2.2) the probability of this event is bounded by  $p_k^3 k^3$  with  $k = \Delta^{\ell_0}$ , which implies (2.29) if  $\chi \in (0, 1)$ . Once (2.29) is established, the rest of the proof is a repetition of the argument in (2.15)-(2.23). □

### 2.3 Upper bound on the diameter

The upper bound in Theorem 1.1 is reformulated as follows.

**Lemma 2.5** *There exist constants  $C, \chi > 0$  such that if  $\varepsilon_n = \frac{C \log \log(n)}{\log(n)}$ ,*

$$\mathbb{P}(\text{diam}(G) > (1 + \varepsilon_n) d_\star) = O(n^{-\chi}). \tag{2.30}$$

**Proof** From Proposition 2.1 we may restrict to the event  $\mathcal{G}(\hbar)$ . From the union bound

$$\mathbb{P}(\text{diam}(G) > (1 + \varepsilon_n) d_\star; \mathcal{G}(\hbar)) \leq \sum_{x,y \in [n]} \mathbb{P}(d(x, y) > (1 + \varepsilon_n) d_\star; \mathcal{G}(\hbar)). \tag{2.31}$$

From Lemma 2.4, for all  $x, y \in [n]$

$$\mathbb{P}(d(x, y) > (1 + \varepsilon_n) d_\star; \mathcal{G}(\hbar)) = \mathbb{P}\left(d(x, y) > (1 + \varepsilon_n) d_\star; F_x^+ \cap F_y^-\right) + O(n^{-2-\chi}). \tag{2.32}$$

Fix

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

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

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

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

$$E_{x,y} = \{\text{There is no path of length } \leq 2k - 1 \text{ from } x \text{ to } y\}. \tag{2.33}$$

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

$$\begin{aligned} \mathbb{P}\left(d(x, y) > (1 + \varepsilon_n) d_\star; F_x^+ \cap F_y^-\right) &\leq \mathbb{P}\left(d(x, y) > (1 + \varepsilon_n) d_\star; U_{x,y} \cap E_{x,y}\right) \\ &\leq \sup_{\sigma \in U_{x,y} \cap E_{x,y}} \mathbb{P}(d(x, y) > (1 + \varepsilon_n) d_\star | \sigma). \end{aligned} \tag{2.34}$$

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  $d(x, y) > (1 + \varepsilon_n)d_\star$  implies that there are no matchings between  $\mathcal{F}_{k-1}$  and  $\mathcal{E}_k$ . The probability of this event is dominated by

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

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

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

Inserting this in (2.31)–(2.32) completes the proof. □

### 2.4 Lower bound on the diameter

We prove the following lower bound on the diameter. Note that Lemma 2.5 and Lemma 2.6 imply Theorem 1.1.

**Lemma 2.6** *There exists  $C > 0$  such that taking  $\varepsilon_n = \frac{C \log \log(n)}{\log(n)}$ , for any  $x, y \in [n]$ ,*

$$\mathbb{P}(d(x, y) \leq (1 - \varepsilon_n)d_\star) = o(1). \tag{2.35}$$

**Proof** Define

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

We start by sampling the out-neighborhood of  $x$  up to distance  $\ell$ . Consider the event

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

From Lemma 2.3,  $\mathbb{P}(J_x) = 1 - O(n^{-1-\chi})$  for suitable constants  $c_2, \chi > 0$ , and therefore

$$\mathbb{P}(y \in \mathcal{B}_\ell^+(x)) = \mathbb{P}(y \in \mathcal{B}_\ell^+(x); J_x) + O(n^{-1-\chi}). \tag{2.36}$$

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

$$\mathbb{P}(y \in \mathcal{B}_\ell^+(x); J_x) \leq \frac{2\Delta}{m} K = O(n^{-\frac{1}{2}}). \tag{2.37}$$

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

$$\mathbb{P}(d(x, y) \leq (1 - \varepsilon_n)d_\star; y \notin \mathcal{B}_\ell^+(x)) = \mathbb{P}(y \notin \mathcal{B}_\ell^+(x); \mathcal{B}_\ell^-(y) \cap \partial\mathcal{B}_\ell^+(x) \neq \emptyset). \tag{2.38}$$

Consider the event

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

From Lemma 2.3 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

$$\mathbb{P}(y \notin \mathcal{B}_\ell^+(x); \mathcal{B}_\ell^-(y) \cap \partial\mathcal{B}_\ell^+(x) \neq \emptyset) \leq \frac{\Delta K^2}{m} = O(n^{-\varepsilon_n/2}) = o(1), \quad (2.39)$$

where we take the constant  $C$  in the definition of  $\varepsilon_n$  sufficiently large. In conclusion,

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

and the inequalities (2.36)–(2.39) end the proof. □

### 3 Stationary distribution

We start by recalling some key facts established in [9].

#### 3.1 Convergence to stationarity

Let  $P^t(x, \cdot)$  denote the distribution after  $t$  steps of the random walk started at  $x$ . The total variation distance between two probabilities  $\mu, \nu$  on  $[n]$  is defined as

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

Let the entropy  $H$  and the associated *entropic time*  $T_{\text{ENT}}$  be defined by

$$H = \sum_{x \in V} \frac{d_x^-}{m} \log d_x^+, \quad T_{\text{ENT}} = \frac{\log n}{H}. \quad (3.1)$$

Note that under our assumptions on  $\mathbf{d}^\pm$ , the deterministic quantities  $H, T_{\text{ENT}}$  satisfy  $H = \Theta(1)$  and  $T_{\text{ENT}} = \Theta(\log n)$ . Theorem 1 of [9] states that for all  $s > 0$  with  $s \neq 1$ ,

$$\max_{x \in [n]} \left\| P^{sT_{\text{ENT}}}(x, \cdot) - \pi \right\|_{\text{TV}} - \vartheta(s) \xrightarrow{\mathbb{P}} 0, \quad (3.2)$$

where  $\vartheta$  denotes the step function  $\vartheta(s) = 1$  if  $s < 1$  and  $\vartheta(s) = 0$  if  $s > 1$ , and we use the notation  $\xrightarrow{\mathbb{P}}$  for convergence in probability as  $n \rightarrow \infty$ . In words, convergence to stationarity for the random walk on the directed configuration model displays with high probability a cutoff phenomenon, uniformly in the starting point, with mixing

time given by the entropic time  $T_{\text{ENT}}$ . By concavity of  $x \mapsto \log(x)$  the mixing time  $T_{\text{ENT}} = \frac{\log n}{H}$  is always larger than the diameter  $d_{\star} = \frac{\log n}{\log v}$  in Theorem 1.1,

$$H = \sum_{x=1}^n \frac{d_x^-}{m} \log d_x^+ \leq \log \left( \sum_{x=1}^n \frac{d_x^-}{m} d_x^+ \right) = \log v, \tag{3.3}$$

with equality if and only if the sequence is out-regular, that is  $d_x^+ \equiv d$ . Thus, the analysis of convergence to stationarity requires investigating the graph on a length scale that may well exceed the diameter. Considering all possible paths on this length scale is not practical, and we shall rely on a powerful construction of [9] that allows one to restrict to a subset of paths with a tree structure, see Sect. 3.3.1 below for the details.

### 3.2 The local approximation

A consequence of the arguments of [9] is that the unknown stationary distribution at a node  $y$  admits an approximation in terms of the in-neighborhood of  $y$  at a distance that is much smaller than the mixing time. More precisely, it follows from [9, Theorem 3] that for any sequence  $t_n \rightarrow \infty$

$$\|\pi - \mu_{\text{in}} P^{t_n}\|_{\text{TV}} \xrightarrow{\mathbb{P}} 0, \tag{3.4}$$

where we use the notation  $\mu_{\text{in}}$  for the in-degree distribution

$$\mu_{\text{in}}(x) = \frac{d_x^-}{m}, \tag{3.5}$$

and for any probability  $\mu$  on  $[n]$ ,  $\mu P^t$  is the distribution

$$\mu P^t(y) = \sum_{x \in [n]} \mu(x) P^t(x, y), \quad y \in [n].$$

We refer to [11, Lemma 1] for a stronger statement than (3.4) where  $\mu_{\text{in}}$  is replaced by any sufficiently widespread probability on  $[n]$ . While these facts are very useful to study the typical values of  $\pi$ , they give very poor information on its extremal values  $\pi_{\text{min}}$  and  $\pi_{\text{max}}$ , and to prove Theorem 1.3 and Theorem 1.6 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

$$\Gamma_h(y) := \sum_{z \in \partial \mathcal{B}_h^-(y)} d_z^- P^h(z, y). \tag{3.6}$$

The definitions (3.6) and (1.13) are such that for any  $y \in [n]$  and  $h \in \mathbb{N}$

$$\Gamma_h(y) \leq m \mu_{\text{in}} P^h(y), \tag{3.7}$$

where  $\mu_{\text{in}}$  is defined in (3.5). If  $\mathcal{B}_h^-(y)$  is a tree, then (3.7) is an equality. In any case,  $\Gamma_h(y)$  satisfies the following rough inequalities.

**Lemma 3.1** *With high probability, for all  $y \in [n]$ , for all  $h \in [1, \bar{h}]$ :*

$$\left(\frac{\delta_-}{\Delta_+}\right)^h \leq \Gamma_h(y) \leq 2\Delta_- \left(\frac{\Delta_-}{\delta_+}\right)^h. \tag{3.8}$$

**Proof** From Proposition 2.1 we may assume that the event  $\mathcal{G}(\bar{h})$  holds. From Lemma 2.2 we know that  $\frac{1}{2}\delta_-^h \leq |\partial\mathcal{B}_h^-(y)| \leq \Delta_-^h$ . Thus it suffices to show that for any  $z \in \partial\mathcal{B}_h^-(y)$ ,  $h \in [1, \bar{h}]$ :

$$\Delta_+^{-h} \leq P^h(z, y) \leq 2\delta_+^{-h}. \tag{3.9}$$

The bounds in (3.9) follow from the observation that any path of length  $h$  from  $z$  to  $y$  has weight at least  $\Delta_+^{-h}$  and at most  $\delta_+^{-h}$ , and that there is at least one and at most two such paths if  $z \in \partial\mathcal{B}_h^-(y)$  and  $\mathcal{G}(\bar{h})$  holds. The latter fact can be seen with the same argument used in the proof of Lemma 2.2. With reference to that proof: in case 1) there are at most two paths from  $z$  to  $y$ , see Fig. 1; in case 2) there is only one path from  $z$  to  $y$ ; see Figs. 2 and 3.  $\square$

Roughly speaking, in what follows the extremal values of  $\pi$  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  $h \in [h_0, \bar{h}]$  where  $h_0$  is of order  $\log \log n$ .

**Lemma 3.2** *There exist constants  $c > 0$  and  $C > 0$  such that:*

$$\mathbb{P}\left(\forall y \in [n], \forall h \in [h_0, \bar{h}], \Gamma_h(y) \geq c \log^{1-\gamma_0}(n)\right) = 1 - o(1), \tag{3.10}$$

where  $\gamma_0$  is the constant from Theorem 1.3 and  $h_0 := \log_{\delta_-} \log(n) + C$ .

**Proof** From Lemma 2.2 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-neighborhood  $\mathcal{B}_{h_0}^-(y)$  we proceed with the generation of the  $(h - h_0)$ -in-neighborhoods 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  $(z_1, \dots, z_R)$  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  $Z_i := \mathcal{B}_{h-h_0}^-(z_i)$ ,  $i = 1, \dots, R$  with a sequence of independent rooted directed random trees  $W_i$ ,  $i = 1, \dots, R$ , 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 Sect. 2. 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  $K := \Delta^h = O(n^{1/5})$ , 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

$$\Gamma_h(y) \geq \Delta_+^{-h_0} \sum_{i \notin \mathcal{I}} \Gamma_{h-h_0}(z_i). \tag{3.11}$$

The above construction shows that the cardinality of the set  $\mathcal{I}$  is stochastically dominated by twice the binomial  $\text{Bin}(R, q)$ . Therefore,

$$\mathbb{P}(|\mathcal{I}| \geq 10) \leq \mathbb{P}(\text{Bin}(R, q) \geq 5) \leq (Rq)^5 = o(n^{-2}). \tag{3.12}$$

On the other hand, notice that for all  $i \notin \mathcal{I}$ :

$$\Gamma_{h-h_0}(z_i) = M_{h-h_0}^i, \tag{3.13}$$

where  $M_t^i, t \in \mathbb{N}$ , 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

$$\mathbf{w}(x) := \mathbf{w}(x \mapsto z_i; \mathcal{T}_i) = \prod_{u=1}^t \frac{1}{d_{x_u}^+}, \tag{3.14}$$

for the weight of the path  $(x_t = x, x_{t-1}, \dots, x_1, x_0 = z_i)$  from  $x$  to  $z_i$  along  $\mathcal{T}_i$ . Then  $M_t^i$  is defined by

$$M_t^i = \sum_{x \in \mathcal{T}_{t,i}} d_x^- \mathbf{w}(x), \quad M_0^i = d_{z_i}^-. \tag{3.15}$$

It is not hard to check (see e.g. [11, Proposition 4]) that for fixed  $n$ ,  $(M_t^i)_{t \geq 0}$  is a martingale with

$$\mathbb{E}[M_t^i] = M_0^i = d_{z_i}^-.$$

In particular, by truncating at a sufficiently large constant  $C_1 > 0$  one has  $M_{h-h_0}^i \geq X_i$ , where

$$X_i := \min\{M_{h-h_0}^i, C_1\}$$

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  $k \in \mathbb{N}$ :

$$\mathbb{P}\left(\sum_{i=1}^k M_{h-h_0}^i \leq k/2\right) \leq e^{-c_1 k}, \tag{3.16}$$

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

$$\sum_{i \notin \mathcal{I}} \Gamma_{h-h_0}(z_i) \geq \sum_{i \in I_{j^*}} M_{h-h_0}^i \geq \min_{\ell=1, \dots, 10} \sum_{i \in I_\ell} M_{h-h_0}^i. \tag{3.17}$$

Using (3.12), and (3.16), (3.17) we conclude that, for a suitable constant  $c_2 > 0$ :

$$\begin{aligned} \mathbb{P}\left(\sum_{i \notin \mathcal{I}} \Gamma_{h-h_0}(z_i) \leq c_2 R\right) &\leq \mathbb{P}\left(\min_{\ell=1, \dots, 10} \sum_{i \in I_\ell} M_{h-h_0}^i \leq c_2 R\right) + \mathbb{P}(|\mathcal{I}| \geq 10) \\ &\leq 10 \exp(-c_1 R/10) + o(n^{-2}). \end{aligned} \tag{3.18}$$

Since  $R = \frac{1}{2} \delta_-^{h_0} = \frac{1}{2} \delta_-^C \log n$ , the probability in (3.18) is  $o(n^{-2})$  if  $C$  is large enough. From (3.11), on the event  $\sum_{i \notin \mathcal{I}} \Gamma_{h-h_0}(z_i) > c_2 R$  one has

$$\Gamma_h(y) \geq \frac{1}{2} c_2 \delta_-^{h_0} \Delta_+^{-h_0} = c \log^{1-\gamma_0}(n), \tag{3.19}$$

where  $c = \frac{1}{2} c_2 (\delta_- / \Delta_+)^C$ . Thus the event (3.19) has probability  $1 - o(n^{-2})$ , and the desired conclusion follows by taking a union bound over  $y \in [n]$  and  $h \in [h_0, \bar{h}]$ .  $\square$

**Lemma 3.3** *There exists a constant  $K > 0$  such that for all  $\varepsilon > 0$ , with high probability:*

$$\max_{y \in [n]} \max_{h \in [h_1, \bar{h}]} \left| \frac{\Gamma_h(y)}{\Gamma_{h_1}(y)} - 1 \right| \leq \varepsilon, \tag{3.20}$$

where  $h_1 := K \log \log(n)$ .

**Proof** For any  $h \geq h_1$ , let  $\sigma_h$  denote a realization of the in-neighborhood  $\mathcal{B}_h^-(y)$ , obtained with the usual breadth-first sequential generation. From Proposition 2.1 we may assume that the tree excess of  $\mathcal{B}_h^-(y)$  is at most 1, as long as  $h \leq \tilde{h}$ . Call  $\mathcal{E}_{tot,h}, \mathcal{F}_{tot,h}$  the set of unmatched tails and unmatched heads, respectively, after the generation of  $\sigma_h$ . 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,

$$\Gamma_h(y) = \sum_{f \in \mathcal{F}_h} P^h(v_f, y), \tag{3.21}$$

where  $v_f$  denotes the vertex to which the head  $f$  belongs. To compute  $\Gamma_{h+1}$  given  $\sigma_h$  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,

$$\begin{aligned} \Gamma_{h+1}(y) &= \sum_{e \in \mathcal{E}_h} \frac{d_e^-}{d_e^+} \sum_{f \in \mathcal{F}_h} P^h(v_f, y) \mathbf{1}_{\omega(e)=f} \\ &= \sum_{e \in \mathcal{E}_{tot,h}} c(e, \omega(e)), \end{aligned} \tag{3.22}$$

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

$$c(e, f) = \frac{d_e^-}{d_e^+} P^h(v_f, y) \mathbf{1}_{e \in \mathcal{E}_h, f \in \mathcal{F}_h}. \tag{3.23}$$

Since  $\sigma_h$  is such that  $\text{TX}(\mathcal{B}_h^-(y)) \leq 1$ , we may estimate  $P^h(v_f, y)$  as in (3.9), so that

$$\|c\|_\infty = \max_{e,f} c(e, f) \leq 2\Delta \delta^{-h-1}. \tag{3.24}$$

We now use a version of Bernstein’s inequality proved by Chatterjee ([12, Proposition 1.1]) which applies to any function of a uniform random matching of the form (3.22). It follows that for any fixed  $\sigma_h$ , for any  $s > 0$ :

$$\begin{aligned} &\mathbb{P} \left( |\Gamma_{h+1}(y) - \mathbb{E} [\Gamma_{h+1}(y) | \sigma_h] | \geq s | \sigma_h \right) \\ &\leq 2 \exp \left( - \frac{s^2}{2 \|c\|_\infty (2\mathbb{E} [\Gamma_{h+1}(y) | \sigma_h] + s)} \right). \end{aligned} \tag{3.25}$$

Taking  $s = a\mathbb{E}[\Gamma_{h+1}(y) | \sigma_h]$ ,  $a \in (0, 1)$ , one has

$$\mathbb{P}(|\Gamma_{h+1}(y) - \mathbb{E}[\Gamma_{h+1}(y) | \sigma_h]| \geq s | \sigma_h) \leq 2 \exp\left(-\frac{a^2 \mathbb{E}[\Gamma_{h+1}(y) | \sigma_h]}{6 \|c\|_\infty}\right). \tag{3.26}$$

Since the probability of the event  $\omega(e) = f$  conditioned on  $\sigma_h$  is  $\frac{1}{|\mathcal{E}_{tot,h}|} = \frac{1}{m}(1 + O(\Delta^h/m))$ , we have

$$\begin{aligned} \mathbb{E}[\Gamma_{h+1}(y) | \sigma_h] &= \frac{1}{|\mathcal{E}_{tot,h}|} \sum_{e \in \mathcal{E}_h} \frac{d_e^-}{d_e^+} \Gamma_h(y) \\ &= \frac{1}{m} \left(1 + O(\Delta^h/m)\right) \left(m - \sum_{e \notin \mathcal{E}_h} \frac{d_e^-}{d_e^+}\right) \Gamma_h(y) \\ &= \left(1 + O(\Delta^h/m)\right) \Gamma_h(y) = \left(1 + O(n^{-1/2})\right) \Gamma_h(y), \end{aligned} \tag{3.27}$$

for all  $h \in [h_1, \bar{h}]$ , 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 3.2 it follows that for some constant  $c > 0$ :

$$\mathbb{E}[\Gamma_{h+1}(y) | \sigma_h] \geq c \log^{-\gamma_0+1}(n), \tag{3.28}$$

and therefore, using (3.24), one finds

$$\|c\|_\infty^{-1} \mathbb{E}[\Gamma_{h+1}(y) | \sigma_h] \geq \log^6(n), \tag{3.29}$$

for all  $h \geq h_1$ , if the constant  $K$  in the definition of  $h_1$  is large enough. From (3.26), (3.27) and (3.29) it follows that, letting

$$\mathcal{A} := \{|\Gamma_{h+1}(y) - \Gamma_h(y)| \leq a\Gamma_h(y), \forall h \in [h_1, \bar{h}]\},$$

with  $a := \log^{-2}(n)$ , then

$$\mathbb{P}(\mathcal{A}) = 1 - o(1). \tag{3.30}$$

Moreover, on the event  $\mathcal{A}$ , for all  $h \in [h_1, \bar{h}]$ :

$$|\Gamma_h(y) - \Gamma_{h_1}(y)| \leq \sum_{j=h_1}^{h-1} |\Gamma_{j+1}(y) - \Gamma_j(y)| \leq \varepsilon \Gamma_{h_1}(y).$$

□



### 3.3 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

$$\pi(z) = \sum_{x=1}^n \pi(x) P^t(x, z) \geq a, \tag{3.31}$$

and therefore  $\pi_{\min} \geq a$ . We will prove the lower bound on  $P^t(x, y)$  by choosing  $t$  of the form  $t = (1 + \varepsilon)T_{\text{ENT}}$ , for some small enough  $\varepsilon > 0$ ; see (3.1) for the definition of  $T_{\text{ENT}}$ . More precisely, fix a constant  $\eta > 0$ , set  $\eta' = 3\eta \frac{H}{\log \delta}$ , and define

$$t_\star = h_x + h_y + 1, \quad h_x = (1 - \eta)T_{\text{ENT}}, \quad h_y = \eta' T_{\text{ENT}}. \tag{3.32}$$

Note that  $\eta' \geq 3\eta$  and thus  $t_\star = t_\star(\eta) \geq (1 + 2\eta)T_{\text{ENT}}$ .

**Lemma 3.4** *There exists  $\eta_0 > 0$  such that for all  $\eta \in (0, \eta_0)$ :*

$$\mathbb{P}\left(\forall x, y \in [n], P^{t_\star+1}(x, y) \geq \frac{c}{n} \Gamma_{h_y}(y)\right) = 1 - o(1), \tag{3.33}$$

for some constant  $c = c(\eta, \Delta) > 0$ .

From (3.31) and Lemma 3.4 it follows that w.h.p. for all  $y$

$$\pi(y) \geq \frac{c}{n} \Gamma_{h_y}(y). \tag{3.34}$$

Lemma 3.2 thus implies, for some new constant  $c > 0$

$$\mathbb{P}\left(\pi_{\min} \geq \frac{c}{n} \log^{1-\gamma_0}(n)\right) = 1 - o(1), \tag{3.35}$$

which settles the lower bound in Theorem 1.3.

To prove Lemma 3.4 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.

#### 3.3.1 A concentration result for nice paths

The definition of the nice paths follows a construction introduced in [9], which we now recall. In contrast with [9] however, here we need a lower bound on  $P^{t_\star}(x, y)$  and thus the argument is somewhat different.

Fix  $h_x$  as in (3.32). Following [9, Section 6.2] and [10, Section 4.1], we introduce the rooted directed tree  $\mathcal{T}(x)$ , namely the subgraph of the  $h_x$ -out-neighborhood of  $x$  obtained 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

$$\mathbf{w}^+(e) = \prod_{i=0}^{\mathbf{h}(e)-1} \frac{1}{d_{x_i}^+}, \tag{3.36}$$

where  $(x = x_0, x_1, \dots, x_{\mathbf{h}(e)-1})$  denotes the path in  $\mathcal{T}(x)$  from  $x$  to the endpoint of  $e$ . Hence, at the beginning  $\partial_+ \mathcal{T}(x)$  is the set of tails at  $x$ , which all have height 1 and weight  $1/d_x^+$ . 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  $\mathbf{h}(e) \leq h_x - 1$  and weight  $\mathbf{w}^+(e) \geq \mathbf{w}_{\min}$ . The third item above guarantees that  $\mathcal{T}(x)$  remains a directed tree at each step. The final value of  $\mathcal{T}(x)$  represents the desired directed tree. Notice that this construction applies only to the out-neighborhood of a vertex; a different procedure will be used for the in-neighborhood of a vertex (see the text preceding (3.38) below).

After the generation of the tree  $\mathcal{T}(x)$  a total number  $\kappa$  of edges has been revealed, some of which may not belong to  $\mathcal{T}(x)$ . As in [10, Lemma 7], it is not difficult to see that when exploring the out-neighborhood of  $x$  in this way the random variable  $\kappa$  is deterministically bounded as

$$\kappa \leq n^{1-\frac{\eta^2}{2}}. \tag{3.37}$$

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 3.5** A path  $\mathbf{p} = (x_0 = x, x_1, \dots, x_{t_\star} = y)$  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

$$\prod_{i=0}^{h_x} \frac{1}{d_{x_i}^+} \leq n^{2\eta-1};$$

- (2)  $x_{h_x+1} \in \partial \mathcal{B}_{h_y}^-(y)$ .

We recall that  $h_y$  is defined as in (3.32), and refer to Sect. 2.2 for the definitions of  $\mathcal{B}_{h_y}^-(y)$  and  $\partial \mathcal{B}_{h_y}^-(y)$ . To obtain a useful expression for the probability of going from  $x$  to  $y$  along a nice path, we need to generate  $\mathcal{B}_{h_y}^-(y)$ . To this end, assume that  $\kappa$  edges in the  $h_x$ -out-neighborhood 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 Sect. 2. 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  $\tau$  of edges, satisfying the crude bound  $\tau \leq \Delta^{h_y+1}$ . We call  $\mathcal{F}_{\text{tot}}$  the set of unmatched heads, and  $\mathcal{E}_{\text{tot}}$  the set of unmatched tails after the sampling of these  $\kappa + \tau$  edges. Consider the set  $\mathcal{F}^0 := \mathcal{F}_{h_y} \cap \mathcal{F}_{\text{tot}}$ , 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  $\mathcal{E}^0 := \mathcal{E}^*(x) \cap \mathcal{E}_{\text{tot}}$  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  $\mathcal{E}_{\text{tot}}$  to the  $m - \kappa - \tau$  unmatched heads  $\mathcal{F}_{\text{tot}}$  using a uniformly random bijection  $\omega : \mathcal{E}_{\text{tot}} \mapsto \mathcal{F}_{\text{tot}}$ . For any  $f \in \mathcal{F}_{h_y}$  we introduce the notation

$$\mathbf{w}^-(f) := P^{h_y}(v_f, y), \tag{3.38}$$

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  $t_*$  steps following a nice path can now be written as

$$P_{0,t_*}(x, y) := \sum_{e \in \mathcal{E}^0} \sum_{f \in \mathcal{F}^0} \mathbf{w}^+(e) \mathbf{w}^-(f) \mathbf{1}_{\omega(e)=f} \mathbf{1}_{\mathbf{w}^+(e) \leq n^{2\eta-1}}. \tag{3.39}$$

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  $\sigma$  denotes a fixed realization of the  $\kappa + \tau$  edges, then

$$\mathbb{E} [P_{0,t_*}(x, y) \mid \sigma] \geq \frac{1}{m} A_{x,y}(\sigma) B_{x,y}(\sigma), \tag{3.40}$$

where

$$A_{x,y}(\sigma) := \sum_{e \in \mathcal{E}^0} \mathbf{1}_{\mathbf{w}^+(e) \leq n^{2\eta-1}} \mathbf{w}^+(e), \quad B_{x,y}(\sigma) := \sum_{f \in \mathcal{F}^0} \mathbf{w}^-(f). \tag{3.41}$$

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

$$\mathbb{E} [P_{0,t_*}(x, y) \mid \sigma] \leq \frac{(1 + o(1))}{m} A_{x,y}(\sigma) B_{x,y}(\sigma) \leq \frac{(1 + o(1))}{m} \Gamma_{h_y}(y), \tag{3.42}$$

where we use  $A_{x,y} \leq 1$  and  $B_{x,y} \leq \Gamma_{h_y}(y)$ . Recall the definition of the tree excess, TX, in (2.3) and consider the event

$$\mathcal{Y}_{x,y} = \left\{ \sigma : A_{x,y}(\sigma) \geq \frac{1}{2}, B_{x,y}(\sigma) \geq \log^{-\gamma_0}(n), \text{TX}(\mathcal{B}_{h_y}^-(y)) \leq 1 \right\}, \tag{3.43}$$

where the exponent  $-\gamma_0$  is chosen for convenience only and any exponent  $-c$  with  $c > \gamma_0 - 1$  would be as good.

**Lemma 3.6** *There exists  $\eta_0 > 0$  such that for all  $\eta \in (0, \eta_0)$ , for any  $\sigma \in \mathcal{Y}_{x,y}$ , any  $a \in (0, 1)$ :*

$$\mathbb{P}(|P_{0,t_*}(x, y) - \mathbb{E}[P_{0,t_*}(x, y) | \sigma]| \geq a \mathbb{E}[P_{0,t_*}(x, y) | \sigma] | \sigma) \leq 2 \exp\left(-a^2 n^{\eta/2}\right) \tag{3.44}$$

**Proof** Conditioned on  $\sigma$ ,  $P_{0,t_*}(x, y)$  is a function of the uniform random permutation  $\omega : \mathcal{E}_{\text{tot}} \mapsto \mathcal{F}_{\text{tot}}$ ,

$$P_{0,t_*}(x, y) = \sum_{e \in \mathcal{E}_{\text{tot}}} c(e, \omega(e)), \quad c(e, f) = \mathbf{w}^+(e) \mathbf{w}^-(f) \mathbf{1}_{\mathbf{w}^+(e) \leq n^{2\eta-1}} \mathbf{1}_{e \in \mathcal{E}^0, f \in \mathcal{F}^0}. \tag{3.45}$$

Since we are assuming  $\text{TX}(\mathcal{B}_{h_y}^-(y)) \leq 1$ , we can use (3.9) to estimate  $\mathbf{w}^-(f) \leq 2\delta^{-h_y} = n^{-3\eta}$  for any  $f \in \mathcal{F}^0$ . Therefore

$$\|c\|_\infty = \max_{e,f} c(e, f) \leq 2n^{-1-\eta}. \tag{3.46}$$

As in Lemma 3.3, and as in [9], we use Chatterjee’s concentration inequality for uniform random matchings [12, Proposition 1.1] to obtain for any  $s > 0$ :

$$\begin{aligned} &\mathbb{P}(|P_{0,t_*}(x, y) - \mathbb{E}[P_{0,t_*}(x, y) | \sigma]| \geq s | \sigma) \\ &\leq 2 \exp\left(-\frac{s^2}{2 \|c\|_\infty (2\mathbb{E}[P_{0,t_*}(x, y) | \sigma] + s)}\right). \end{aligned} \tag{3.47}$$

Taking  $s = a \mathbb{E}[P_{0,t_*}(x, y) | \sigma]$ ,  $a \in (0, 1)$ , one has

$$\mathbb{P}(|P_{0,t_*}(x, y) - \mathbb{E}[P_{0,t_*}(x, y) | \sigma]| \geq s | \sigma) \leq 2 \exp\left(-\frac{a^2 \mathbb{E}[P_{0,t_*}(x, y) | \sigma]}{6 \|c\|_\infty}\right). \tag{3.48}$$

Using (3.40), (3.43), and (3.46) one concludes that (3.44) holds for all  $\sigma \in \mathcal{Y}_{x,y}$  and for all  $n$  large enough.  $\square$

### 3.3.2 Proof of Lemma 3.4

Let  $V_*$  denote the set of all  $z \in [n]$  such that  $\mathcal{B}_h^+(z)$  is a directed tree. It is an immediate consequence of Proposition 2.1 that with high probability, for all  $x \in [n]$ :

$$P(x, V_*) = \sum_{z \in V_*} P(x, z) \geq \frac{1}{2}. \tag{3.49}$$

In fact, as observed in [9, Proposition 6], one can show that with high probability  $P^\ell(x, V_*) \geq 1 - 2^{-\ell}$  for any fixed  $\ell \in \mathbb{N}$ . Therefore,

$$P^{t_*+1}(x, y) \geq \frac{1}{2} \min_{x \in V_*} P^{t_*}(x, y). \tag{3.50}$$

Since  $P^{t_\star}(x, y) \geq P_{0,t_\star}(x, y)$  it is sufficient to prove

$$\mathbb{P}(\forall x \in V_\star, \forall y \in [n], P_{0,t_\star}(x, y) \geq \frac{c}{n} \Gamma_{h_y}(y)) = 1 - o(1), \tag{3.51}$$

for some constant  $c = c(\eta, \Delta) > 0$ . The proof of (3.51) is based on Lemma 3.6 and the following two lemmas, which allow us to make sure the events  $\mathcal{Y}_{x,y}$  in Lemma 3.6 have large probability.

**Lemma 3.7** *The event  $\mathcal{A}_1 = \{\forall x \in V_\star, \forall y \in [n] : A_{x,y} \geq \frac{1}{2}\}$  has probability*

$$\mathbb{P}(\mathcal{A}_1) = 1 - o(1).$$

**Proof** Let us first note that the event  $\widehat{\mathcal{A}}_1 = \{\forall x \in V_\star : \sum_{e \in \mathcal{E}^\star(x)} \mathbf{w}^+(e) \mathbf{1}_{\mathbf{w}^+(e) \leq n^{2\eta-1}} \geq 0.9\}$  satisfies

$$\mathbb{P}(\widehat{\mathcal{A}}_1) = 1 - o(1).$$

Indeed, this fact is a consequence of [9,10], which established that for any  $\varepsilon > 0$ , with high probability

$$\min_{x \in V_\star} \sum_{e \in \mathcal{E}^\star(x)} \mathbf{w}^+(e) \mathbf{1}_{\mathbf{w}^+(e) \leq n^{2\eta-1}} \geq 1 - \varepsilon, \tag{3.52}$$

see e.g. [10, Theorem 4 and Lemma 11]. Thus, it remains to show that replacing  $\mathcal{E}^\star(x)$  with  $\mathcal{E}^0$  does not alter much the sum. Suppose the  $\kappa$  edges generating  $\mathcal{T}(x)$  have been revealed and then sample the  $\tau$  edges generating the neighborhood  $\mathcal{B}_{h_y}^-(y)$ . 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  $\text{Bin}(N, p)$ , and therefore by Hoeffding’s inequality

$$\mathbb{P}(K > Np + N) \leq \exp(-2N) \leq \exp(-n^{3\eta}).$$

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  $\eta$  is small enough. □

**Lemma 3.8** *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*

$$\mathbb{P}(\mathcal{A}_2) = 1 - o(1).$$

**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 2.1 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  $\mathcal{F}_{\text{tot}}$ . 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 Sect. 3.3.1 but this time we first generate the in-neighborhood  $\mathcal{B}_{h_y}^-(y)$  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  $|V_z| \leq \Delta^T$  and that  $|\mathcal{T}(x)| \leq n^{1-\eta^2/2}$ , so that  $K$  is stochastically dominated by the binomial  $\text{Bin}(N, p)$  where  $N = n^{1-\eta^2/2}$  and  $p = \Delta^{T+1}/n$ . Therefore,

$$\mathbb{P}\left(K > \frac{1}{2} \delta^T\right) \leq (Np)^{\frac{1}{2} \delta^T} \leq \left(\Delta^{T+1} n^{-\eta^2/2}\right)^{\frac{1}{2} \delta^T}.$$

Since  $|V_z| \geq \frac{1}{2} \delta^T$ , 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  $\mathbf{w}^-(f) \geq \Delta^{-T-1} p^{h_y-T}(z, y)$ . Therefore, assuming that there are no bad  $z \in \partial \mathcal{B}_{h_y-T}^-(y)$ :

$$\begin{aligned} B_{x,y}(\sigma) &= \sum_{f \in \mathcal{F}^0} \mathbf{w}^-(f) \\ &\geq \Delta^{-T} \sum_{z \in \partial \mathcal{B}_{h_y-T}^-(y)} p^{h_y-T}(z, y) \geq \Delta^{-T-1} \Gamma_{h_y-T}(y). \end{aligned}$$

From Lemma 3.3 we may finish with the estimate  $\Gamma_{h_y-T}(y) \geq \frac{1}{2} \Gamma_{h_y}(y)$ . □

We can now conclude the proof of (3.51). Consider the event

$$\mathcal{A} = \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{G}(\hbar) \cap \mathcal{R}, \tag{3.53}$$

where  $\mathcal{R}$  denotes the event from Lemma 3.2 and  $\mathcal{G}(\hbar)$  is given by (2.4) and (2.5). For any  $s > 0$ ,

$$\begin{aligned} &\mathbb{P}\left(\forall x, y \in [n], P_{0,t_*}(x, y) \geq \frac{s}{n} \Gamma_{h_y}(y)\right) \\ &\geq \mathbb{P}(\mathcal{A}) - \sum_{x,y \in [n]} \mathbb{P}\left(P_{0,t_*}(x, y) < \frac{s}{n} \Gamma_{h_y}(y); \mathcal{A}\right), \end{aligned} \tag{3.54}$$

where the semicolon represents intersection of events. From Lemma 3.2, Lemma 3.7, Lemma 3.8, and Proposition 2.1 it follows that  $\mathbb{P}(\mathcal{A}) = 1 - o(1)$ . Let  $\mathcal{W}_{x,y}$  denote the event

$$\mathbb{E}[P_{0,t_\star}(x, y) \mid \sigma] \geq \frac{c}{2m} \Gamma_{h_y}(y), \tag{3.55}$$

where  $c$  is the constant from Lemma 3.8. By definition of the events involved

$$\mathcal{A} \subset \mathcal{W}_{x,y} \cap \mathcal{Y}_{x,y},$$

for all  $x, y$ , and for all  $n$  large enough. Therefore,

$$\mathbb{P}(P_{0,t_\star}(x, y) < \frac{s}{n} \Gamma_{h_y}(y); \mathcal{A}) \leq \sup_{\sigma \in \mathcal{W}_{x,y} \cap \mathcal{Y}_{x,y}} \mathbb{P}(P_{0,t_\star}(x, y) < \frac{s}{n} \Gamma_{h_y}(y) \mid \sigma). \tag{3.56}$$

Taking  $s = c/(4\Delta)$  and using (3.55), we see that for every  $\sigma \in \mathcal{W}_{x,y}$ ,  $P_{0,t_\star}(x, y) < \frac{s}{n} \Gamma_{h_y}(y)$  implies:

$$|P_{0,t_\star}(x, y) - \mathbb{E}[P_{0,t_\star}(x, y) \mid \sigma]| \geq \frac{1}{2} \mathbb{E}[P_{0,t_\star}(x, y) \mid \sigma],$$

and therefore from Lemma 3.6

$$\sup_{\sigma \in \mathcal{W}_{x,y} \cap \mathcal{Y}_{x,y}} \mathbb{P}(P_{0,t_\star}(x, y) < \frac{s}{n} \Gamma_{h_y}(y) \mid \sigma) = o(n^{-2}). \tag{3.57}$$

The bounds (3.54) and (3.57) end the proof of (3.51). This ends the proof of Lemma 3.4.

**Remark 3.9** Let us show that if the type  $(\delta_-, \Delta_+)$  is not in the set of linear types  $\mathcal{L}$  one can improve the lower bound on  $\pi_{\min}$  as mentioned in Remark 1.5. The proof given above shows that it is sufficient to replace  $\gamma_0$  by  $\gamma'_0$  in Lemma 3.2, where  $\gamma'_0$  is defined by (1.20). To this end, for any  $\varepsilon > 0$ , let  $\mathcal{L}_\varepsilon$  denote the set of types  $(k, \ell) \in \mathcal{C}$  such that

$$\limsup_{n \rightarrow \infty} \frac{|\mathcal{V}_{k,\ell}|}{n^{1-\varepsilon}} = +\infty, \tag{3.58}$$

where  $\mathcal{V}_{k,\ell}$  denotes the set of vertices of type  $(k, \ell)$ , and define

$$\gamma'_\varepsilon := \frac{\log \Delta'_{\varepsilon,+}}{\log \delta'_{\varepsilon,-}}, \quad \Delta'_{\varepsilon,+} := \max\{\ell : (k, \ell) \in \mathcal{L}_\varepsilon\}, \quad \delta'_{\varepsilon,-} := \min\{k : (k, \ell) \in \mathcal{L}_\varepsilon\}. \tag{3.59}$$

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  $(k, \ell)$  in all in-neighborhoods  $\mathcal{B}_{h_0}^-(y)$ ,  $y \in [n]$ , 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  $\text{Bin}(\Delta^{h_0}, n^{-\varepsilon/2})$ , 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  $(k, \ell)$ . Then we may repeat the argument of Lemma 3.2 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  $\varepsilon$  is small enough then  $\gamma'_0 = \gamma'_\varepsilon$  and the desired conclusion follows.

### 3.4 Upper bound on $\pi_{\min}$

In this section we prove the upper bound for  $\pi_{\min}$  in Theorem 1.3 by establishing the estimate in (1.18). We first show that we can replace  $\pi(y)$  in (1.18) by a more convenient quantity. Define the distances

$$d(s) = \max_{x \in [n]} \|P^s(x, \cdot) - \pi\|_{\text{TV}}, \quad \bar{d}(s) = \max_{x, y \in [n]} \|P^s(x, \cdot) - P^s(y, \cdot)\|_{\text{TV}}. \quad (3.60)$$

It is standard that, for all  $k, s \in \mathbb{N}$ ,

$$d(ks) \leq \bar{d}(ks) \leq \bar{d}(s)^k \leq 2^k d(s)^k, \quad (3.61)$$

see e.g. [22]. In particular, defining

$$\lambda_t(y) = \frac{1}{n} \sum_{x \in [n]} P^t(x, y), \quad (3.62)$$

for any  $k \in \mathbb{N}$ , setting  $t = 2kT_{\text{ENT}}$ , one has

$$\max_{y \in [n]} |\lambda_t(y) - \pi(y)| \leq d(2kT_{\text{ENT}}) \leq 2^k d(2T_{\text{ENT}})^k. \quad (3.63)$$

From (3.2) we know that  $d(2T_{\text{ENT}}) \rightarrow 0$  in probability, or equivalently that for all fixed  $\varepsilon \in (0, 1)$  we have  $d(2T_{\text{ENT}}) \leq \varepsilon$  w.h.p. In particular we can choose  $\varepsilon = \frac{1}{2e}$ , so that w.h.p. 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 (1.18). Thus, it suffices to prove the following statement.

**Lemma 3.10** *For some constants  $\beta > 0$ ,  $C > 0$ , and for any  $t = t_n = \Theta(\log^3(n))$ :*

$$\mathbb{P}\left(\exists S \subset [n], |S| \geq n^\beta, n \max_{y \in S} \lambda_t(y) \leq C \log^{1-\gamma_1}(n)\right) = 1 - o(1). \quad (3.64)$$

**Proof** Let  $(\delta_*, \Delta_*) \in \mathcal{L}$  denote a type realizing the maximum in the definition of  $\gamma_1$ ; see (1.16). 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. Let us fix a constant  $\beta_1 \in (0, \frac{1}{4})$ . This will be related to the constant  $\beta$ , but we shall not look for the optimal exponent  $\beta$  in the statement (3.64). 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-neighborhoods  $\mathcal{B}_{h_0}^-(y_i)$ ,  $i = 1, \dots, N_1$ , where

$$h_0 = \log_{\delta_*} \log n - C_0, \quad (3.65)$$



for some constant  $C_0$  to be fixed later. As in the proof of Lemma 3.2 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,

$$\mathbb{P}(\mathcal{A}) \geq 1 - O(N_1 p) \geq 1 - O(n^{3\beta_1 - 1}) = 1 - o(1). \tag{3.66}$$

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  $(\delta_*, \Delta_*)$ . 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 (3.65). We choose  $C_0$  so large that  $0 < \eta \leq \beta_1 / 4$ . Call  $S_1$  the set of  $y \in \{y_1, \dots, y_{N_1}\}$  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  $\text{Bin}(N_1, q) < n^{\beta_1/2}$ , which by Hoeffding's inequality is at most

$$\exp\left(-n^{\beta_1/3}\right) \tag{3.67}$$

Fix a realization  $\sigma$  of the in-neighborhoods  $\mathcal{B}_{h_0}^-(y_i)$ ,  $i = 1, \dots, N_1$ . Say that  $y_i$  is unlucky if all vertices in  $\mathcal{B}_{h_0}^-(y_i)$  are of type  $(\delta_*, \Delta_*)$ . Thanks to (3.66) 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 (3.67) we may also assume that  $\sigma$  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  $t = \Theta(\log^3(n))$ ,

$$\mathbb{P}\left(\sum_{y \in \bar{S}} \lambda_t(y) > \frac{C\bar{N}}{2n} \log^{1-\gamma_1}(n) \mid \sigma\right) = o(1). \tag{3.68}$$

Notice that (3.68) says that, conditionally on a fixed  $\sigma \in \mathcal{A}'$ , with high probability

$$\sum_{y \in \bar{S}} \lambda_t(y) \leq \frac{C\bar{N}}{2n} \log^{1-\gamma_1}(n),$$

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 (3.68) 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

$$\lambda_t(y) \leq \frac{C}{n} \log^{1-\gamma_1}(n),$$

which implies the desired claim (3.64), taking e.g.  $\beta = \beta_1/3$ .

To prove (3.68), consider the sum

$$\mathcal{X} = \sum_{y \in \bar{S}} \lambda_t(y). \tag{3.69}$$

We first establish that, uniformly in  $\sigma \in \mathcal{A}'$ , for any  $t = \Theta(\log^3(n))$ ,

$$\mathbb{E}(\mathcal{X} \mid \sigma) = (1 + o(1)) \frac{\delta_*}{m} \bar{N} \Delta_*^{-h_0} \delta_*^{h_0}. \tag{3.70}$$

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 (3.70) is reduced to showing that for any  $z \in \partial \mathcal{B}_{h_0}^-(y)$ ,  $t = \Theta(\log^3 n)$ ,

$$\mathbb{E}(\lambda_t(z) \mid \sigma) = (1 + o(1)) \frac{d_z^-}{m}. \tag{3.71}$$

To compute the expected value in (3.71) we use the so called *annealed* process. Namely, observe that

$$\mathbb{E}(\lambda_t(z) \mid \sigma) = \frac{1}{n} \sum_{x \in [n]} \mathbb{E}(P^t(x, z) \mid \sigma) = \frac{1}{n} \sum_{x \in [n]} \mathbb{P}_x^{a, \sigma}(X_t = z), \tag{3.72}$$

where  $X_t$  is the annealed walk with initial environment  $\sigma$ , 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  $\sigma$  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  $\sigma \in \mathcal{A}'$ :

$$\mathbb{P}_x^{a, \sigma}(X_t = z) = (1 + o(1)) \frac{d_z^-}{m}. \tag{3.73}$$

Say that a collision occurs if the walk lands on a vertex that was already visited by using a freshly matched edge. Recall that we fixed  $t = O(\log^3(n))$ . 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 (3.73). 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 again hits  $z$  at some point within time  $t$ . In conclusion, (3.72) and (3.73) imply (3.71) which establishes (3.70).

Let us now show that

$$\mathbb{E}(\mathcal{X}^2 | \sigma) = (1 + o(1))\mathbb{E}(\mathcal{X} | \sigma)^2. \tag{3.74}$$

Once we have (3.74) we can conclude (3.68) by using Chebyshev’s inequality together with (3.70) 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

$$\mathbb{E}(\mathcal{X}^2 | \sigma) = \sum_{y, y' \in \bar{S}} \Delta_*^{-2h_0} \frac{1}{n^2} \sum_{x, x' \in [n]} \sum_{z \in \partial \mathcal{B}_{h_0}^-(y)} \sum_{z' \in \partial \mathcal{B}_{h_0}^-(y')} \mathbb{P}_{x, x'}^{a,\sigma}(X_{t-h_0} = z, X'_{t-h_0} = z'), \tag{3.75}$$

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  $\sigma'$  the environment obtained by adding to  $\sigma$  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  $\mathbb{P}_u^{a,\sigma}$  be defined by

$$\mathbb{P}_u^{a,\sigma} = \frac{1}{n^2} \sum_{x, x' \in [n]} \mathbb{P}_{x, x'}^{a,\sigma}.$$

Thus, under  $\mathbb{P}_u^{a,\sigma}$  the two trajectories have independent uniformly distributed starting points  $x, x'$ . With this notation we write

$$\mathbb{E}(\mathcal{X}^2 | \sigma) = \sum_{y, y' \in \bar{S}} \Delta_*^{-2h_0} \sum_{z \in \partial \mathcal{B}_{h_0}^-(y)} \sum_{z' \in \partial \mathcal{B}_{h_0}^-(y')} \mathbb{P}_u^{a,\sigma}(X_{t-h_0} = z, X'_{t-h_0} = z'). \tag{3.76}$$

Let us show that if  $z \neq z', t = \Theta(\log^3(n))$ :

$$\mathbb{P}_u^{a,\sigma}(X_t = z, X'_t = z') = (1 + o(1)) \frac{d_z^- d_{z'}^-}{m^2}. \tag{3.77}$$

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 (3.73) 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 (3.73) is given by  $(1 + o(1))d_{z'}^-/m$ . This proves (3.77) when  $z \neq z'$ .

If  $z = z', t = \Theta(\log^3(n))$ , let us show that

$$\mathbb{P}_u^{a,\sigma}(X_t = z, X'_t = z) = O(1/m^2). \tag{3.78}$$

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,

$$\mathbb{P}_u^{a,\sigma}(X_t = z, X'_t = z; A^c) = O(t^5/m^3) = O(1/m^2). \tag{3.79}$$

To prove (3.78), notice that to realize  $X'_t = z$  there must be a time  $s = 0, \dots, t$  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

$$\mathbb{P}_u^{a,\sigma}(X_t = z, X'_t = z; A) \leq (1 + o(1)) \frac{d_z^-}{m} \sum_{h=0}^t \frac{2(h+1)}{m} 2\delta^{-h} = O(1/m^2). \tag{3.80}$$

Hence, (3.78) follows from (3.79) and (3.80).

In conclusion, using (3.77) and (3.78) in (3.76), and recalling (3.70), we have obtained (3.74). □

### 3.5 Upper bound on $\pi_{\max}$

As in Sect. 3.4 we start by replacing  $\pi(y)$  with  $\lambda_t(y) = \frac{1}{n} \sum_x P^t(x, y)$ . In (3.63) we have seen that if  $t = 2kT_{\text{ENT}}$ , then w.h.p.

$$\max_{y \in [n]} |\lambda_t(y) - \pi(y)| \leq e^{-k}. \tag{3.81}$$

Thus, using a union bound over  $y \in [n]$ , the upper bound in Theorem 1.6 follows from the next statement.

**Lemma 3.11** *There exists  $C > 0$  such that for any  $t = t_n = \Theta(\log^3(n))$ , uniformly in  $y \in [n]$*

$$\mathbb{P}\left(\lambda_t(y) \geq \frac{C}{n} \log^{1-\kappa_0}(n)\right) = o(n^{-1}). \tag{3.82}$$

**Proof** Fix

$$h_0 = \log_{\Delta_-} \log n,$$

and call  $\sigma$  a realization of the in-neighborhood  $\mathcal{B}_{h_0}^-(y)$ . Clearly,

$$\lambda_{t+h_0}(y) = \sum_{z \in \mathcal{B}_{h_0}^-(y)} \lambda_t(z) P^{h_0}(z, y).$$

From (3.9), under the event  $\mathcal{G}_y(\hbar)$  from Proposition 2.1, we have  $P^{h_0}(z, y) \leq 2\delta_+^{-h_0} = 2 \log^{-\kappa_0}(n)$  for every  $z \in \mathcal{B}_{h_0}^-(y)$ . Define

$$\mathcal{X} := \sum_{z \in \mathcal{B}_{h_0}^-(y)} \lambda_t(z) = \lambda_t(\mathcal{B}_{h_0}^-(y)).$$

Note  $\mathcal{X} = \mathcal{X}^y$  where the vertex  $y$  is fixed. Then it is sufficient to prove that for some constant  $C$ , uniformly in  $\sigma$  and  $y \in [n]$ :

$$\mathbb{P}\left(\mathcal{X} > \frac{C}{n} \log n; \mathcal{G}_y(\hbar) \mid \sigma\right) = o(n^{-1}), \tag{3.83}$$

where  $\mathcal{G}_y(\hbar)$  is defined in (2.4), (2.5). By Markov's inequality, for any  $K \in \mathbb{N}$  and any constant  $C > 0$ :

$$\mathbb{P}\left(\mathcal{X} > \frac{C}{n} \log(n); \mathcal{G}_y(\hbar) \mid \sigma\right) \leq \frac{\mathbb{E}\left[\mathcal{X}^K; \mathcal{G}_y(\hbar) \mid \sigma\right]}{\left(\frac{C}{n} \log n\right)^K}. \tag{3.84}$$

We fix  $K = \log n$ , and claim that there exists an absolute constant  $C_1 > 0$  such that

$$\mathbb{E}\left[\mathcal{X}^K; \mathcal{G}_y(\hbar) \mid \sigma\right] \leq \left(\frac{C_1}{n} \log n\right)^K. \tag{3.85}$$

The desired estimate (3.83) follows from (3.85) and (3.84) by taking  $C$  large enough.

We compute the  $K$ -th moment  $\mathbb{E}\left[\mathcal{X}^K; \mathcal{G}_y(\hbar) \mid \sigma\right]$  by using a version of the annealed process that we used in (3.75) that we now explain. This time we have  $K$  trajectories

instead of 2:

$$\begin{aligned} \mathbb{E} \left[ \mathcal{X}^K; \mathcal{G}_y(\hbar) \mid \sigma \right] &= \frac{1}{n^K} \sum_{x_1, \dots, x_K} \mathbb{E} \left[ P^t(x_1, \mathcal{B}_{h_0}^-(y)) \cdots P^t(x_K, \mathcal{B}_{h_0}^-(y)); \mathcal{G}_y(\hbar) \mid \sigma \right] \\ &= \frac{1}{n^K} \sum_{x_1, \dots, x_K} \mathbb{P}_{x_1, \dots, x_K}^{a, \sigma} \left( X_t^{(1)} \in \mathcal{B}_{h_0}^-(y), \dots, X_t^{(K)} \in \mathcal{B}_{h_0}^-(y); \mathcal{G}_y(\hbar) \right), \end{aligned} \quad (3.86)$$

where  $X^{(j)} := \{X_s^{(j)}, s \in [0, t]\}$ ,  $j = 1, \dots, K$  denote  $K$  annealed walks each with initial point  $x_j$ , and  $\mathbb{P}_{x_1, \dots, x_K}^{a, \sigma}$  denotes the joint law of the trajectories  $X^{(j)}$ ,  $j = 1, \dots, K$ , and the environment, defined as follows. Start with the environment  $\sigma$ , and then run the first random walk  $X^{(1)}$  up to time  $t$  as described after (3.72). After that, run the walk  $X^{(2)}$  up to time  $t$  with initial environment given by the union of edges from  $\sigma$  and the first trajectory, as described in (3.75). 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  $\sigma$  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  $\mathbb{P}_{x_1, \dots, x_K}^{a, \sigma}$  between the environment (the digraph  $G$ ) and  $K$  independent walks in that environment. To ensure the validity of the expression (3.86) it suffices to note that conditionally on the realization of the full digraph  $G$ , under  $\mathbb{P}_{x_1, \dots, x_K}^{a, \sigma}$  the variables  $X_t^{(1)}, \dots, X_t^{(K)}$  are independent random walks on  $G$  with length  $t$ .

It is convenient to introduce the notation

$$\mathbb{P}_u^{a, \sigma} = \frac{1}{n^K} \sum_{x_1, \dots, x_K} \mathbb{P}_{x_1, \dots, x_K}^{a, \sigma},$$

for the annealed law of the  $K$  trajectories such that independently each trajectory starts at a uniformly random point  $X_0^{(j)} = x_j$ . Let  $D_0 = \sigma$  and let  $D_\ell$ , for  $\ell = 1, \dots, K$ , denote the digraph defined by the union of  $\sigma = \mathcal{B}_{h_0}^-(y)$  with the first  $\ell$  paths

$$\{X_s^{(j)}, 0 \leq s \leq t\}, \quad j = 1, \dots, \ell.$$

Call  $D_\ell(\hbar)$  the subgraph of  $D_\ell$  consisting of all directed paths in  $D_\ell$  ending at  $y$  with length at most  $\hbar$ . We define  $\mathcal{G}_y^\ell(\hbar)$  as the event  $\text{TX}(D_\ell(\hbar)) \leq 1$ . Notice that if the final environment has to satisfy  $\mathcal{G}_y(\hbar)$ , then necessarily for every  $\ell$  the digraph  $D_\ell$  must satisfy  $\mathcal{G}_y^\ell(\hbar)$ . Therefore,

$$\mathbb{E} \left[ \mathcal{X}^K; \mathcal{G}_y(\hbar) \mid \sigma \right] \leq \mathbb{P}_u^{a, \sigma} \left( X_t^{(1)} \in \mathcal{B}_{h_0}^-(y), \dots, X_t^{(K)} \in \mathcal{B}_{h_0}^-(y); \mathcal{G}_y^K(\hbar) \right). \quad (3.87)$$

Define

$$\mathcal{W}_\ell = \sum_{x \in V(D_\ell)} [d_x^-(D_\ell) - 1]_+, \quad (3.88)$$

where  $V(D_\ell)$  denotes the vertex set of  $D_\ell$  and  $d_x^-(D_\ell)$  is the in-degree of  $x$  in the digraph  $D_\ell$ . Define also the  $(\ell, s)$  cluster  $\mathcal{C}_\ell^s$  as the digraph given by the union of  $D_{\ell-1}$

and the truncated path  $\{X_u^{(\ell)}, 0 \leq u \leq s\}$  with  $s \leq t$ . We say that the  $\ell$ -th trajectory  $X^{(\ell)}$  has a *collision* at time  $s \geq 1$  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  $\ell$  trajectories.

Let  $\mathcal{Q}_\ell$  denote the total number of collisions after the generation of the first  $\ell$  trajectories. Since  $|\mathcal{B}_{h_0}^-(y)| \leq \Delta \log n$  one must have

$$\mathcal{W}_\ell \leq \Delta \log n + \mathcal{Q}_\ell. \tag{3.89}$$

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  $\text{Bin}(Kt, p)$ . In particular, for any  $k \in \mathbb{N}$ :

$$\mathbb{P}(\mathcal{Q}_K \geq k) \leq (Ktp)^k \leq C_2^k \frac{\log^{8k}(n)}{n^k}, \tag{3.90}$$

for some constant  $C_2 > 0$ . If  $A > 0$  is a large enough constant, then

$$\mathbb{P}(\mathcal{Q}_K \geq A \log n) \leq e^{-\frac{A}{2} \log^2(n)}. \tag{3.91}$$

If  $A \geq 2$  then (3.91) is smaller than the right hand side of (3.85) with e.g.  $C_1 = 1$ , and therefore from now on we may restrict to proving the upper bound

$$\mathbb{P}_u^{a,\sigma} \left( X_t^{(1)} \in \mathcal{B}_{h_0}^-(y), \dots, X_t^{(K)} \in \mathcal{B}_{h_0}^-(y); \mathcal{Q}_K \leq A \log n; \mathcal{G}_y^K(\hbar) \right) \leq \left( \frac{C_1}{n} \log n \right)^K, \tag{3.92}$$

for some constant  $C_1 = C_1(A) > 0$ . To prove (3.92), define the events

$$B_\ell = \{X_t^{(1)} \in \mathcal{B}_{h_0}^-(y), \dots, X_t^{(\ell)} \in \mathcal{B}_{h_0}^-(y); \mathcal{Q}_\ell \leq A \log n; \mathcal{G}_y^\ell(\hbar)\}, \tag{3.93}$$

for  $\ell = 1, \dots, K$ . Since  $B_{\ell+1} \subset B_\ell$ , the left hand side in (3.92) is equal to

$$\mathbb{P}_u^{a,\sigma} (B_1) \prod_{\ell=2}^K \mathbb{P}_u^{a,\sigma} (B_\ell | B_{\ell-1}) \tag{3.94}$$

Thus, it is sufficient to show that for some constant  $C_1$ :

$$\mathbb{P}_u^{a,\sigma} (B_\ell | B_{\ell-1}) \leq \frac{C_1}{n} \log n, \tag{3.95}$$

for all  $\ell = 1, \dots, K$ , where it is understood that  $\mathbb{P}_u^{a,\sigma}(B_1 | B_0) = \mathbb{P}_u^{a,\sigma}(B_1)$ .

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 neighborhood  $\mathcal{B}_{h_0}^-(y)$ . 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,

$$\begin{aligned} \mathbb{P}_u^{a,\sigma}(B_\ell | B_{\ell-1}) &\leq \mathbb{P}_u^{a,\sigma}\left(X_t^{(\ell)} \in \mathcal{B}_{h_0}^-(y) | B_{\ell-1}\right) \\ &\leq 2\delta_+^{-t} \mathbb{P}_u^{a,\sigma}\left(X_0^{(\ell)} \in \mathcal{B}_{h_0}^-(y) | B_{\ell-1}\right) \\ &\quad + \sum_{j=1}^t 2\delta_+^{-(t-j)} \mathbb{P}_u^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y) | B_{\ell-1}\right) \\ &\leq 2t\delta_+^{-t/2} + \sum_{j=t/2+1}^t 2\delta_+^{-(t-j)} \mathbb{P}_u^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y) | B_{\ell-1}\right) \end{aligned}$$

Since  $t = \Theta(\log^3(n))$ , it is enough to show

$$\mathbb{P}_u^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y) | B_{\ell-1}\right) \leq \frac{C_1}{n} \log n, \tag{3.96}$$

uniformly in  $j \in (t/2, t)$  and  $1 \leq \ell \leq K$ .

Let  $\mathcal{H}_0^\ell$  denote the event that the  $\ell$ -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  $\Delta_- |\partial\mathcal{B}_{h_0}^-(y)| \leq \Delta_-^{h_0+1} = \Delta_- \log n$  unmatched heads attached to  $\partial\mathcal{B}_{h_0}^-(y)$ , and therefore

$$\mathbb{P}_u^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y); \mathcal{H}_0^\ell | B_{\ell-1}\right) = O(|\partial\mathcal{B}_{h_0}^-(y)|/m) \leq \frac{C_1}{n} \log n. \tag{3.97}$$

Let  $\mathcal{H}_2^\ell$  denote the event that the  $\ell$ -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  $s_3 \leq j$  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

$$\mathbb{P}_u^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y); \mathcal{H}_2^\ell | B_{\ell-1}\right) = O(K^2 t^4 / m^2) \leq \frac{C_1}{n} \log n. \tag{3.98}$$

It remains to consider the case where the  $\ell$ -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

$$\sum_{s=0}^{j-1} \mathbb{P}_u^{a,\sigma}\left(X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y); \mathcal{H}_{1,s}^\ell | B_{\ell-1}\right) \leq \frac{C_1}{n} \log n. \tag{3.99}$$



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

$$\sum_{s=0}^{j-\hbar+h_0} \mathbb{P}_u^{a,\sigma} \left( X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y); \mathcal{H}_{1,s}^\ell \mid B_{\ell-1} \right) = O(Kt^2n^{-\varepsilon}/m) \leq \frac{C_1}{n} \log n. \tag{3.100}$$

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,

$$\sum_{s=j-\hbar+h_0+1}^{j-1} \mathbb{P}_u^{a,\sigma} \left( X_j^{(\ell)} \in \partial\mathcal{B}_{h_0}^-(y); \mathcal{H}_{1,s}^\ell \mid B_{\ell-1} \right) \leq \sum_{h=1}^{\hbar-h_0-1} \frac{2\Delta|L_h|}{m} 2\delta_+^{-h}, \tag{3.101}$$

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

$$\begin{aligned} |A_h| &\leq \sum_{x \in A_{h-1}} d_x^-(D_{\ell-1}) \\ &\leq |A_{h-1}| + \sum_{x \in A_{h-1}} [d_x^-(D_{\ell-1}) - 1]_+ \\ &\leq |A_{h-2}| + \sum_{x \in A_{h-1} \cup A_{h-2}} [d_x^-(D_{\ell-1}) - 1]_+ \\ &\leq \dots \leq |A_0| + \sum_{x \in A_0 \cup \dots \cup A_{h-1}} [d_x^-(D_{\ell-1}) - 1]_+ \\ &\leq |\partial\mathcal{B}_{h_0}^-(y)| + \mathcal{W}_{\ell-1}. \end{aligned}$$

Since  $h \leq \hbar = O(\log n)$  and  $|\partial\mathcal{B}_{h_0}^-(y)| \leq \log n$ , using (3.89) we have obtained

$$|A_h| \leq C_2 \log n + \mathcal{Q}_{\ell-1}. \tag{3.102}$$

On the event  $B_{\ell-1}$  we know that  $\mathcal{Q}_{\ell-1} \leq A \log n$ , and therefore  $|A_h| \leq C_3 \log n$  for some absolute constant  $C_3 > 0$ . In conclusion, for all  $h \in (0, \hbar - h_0)$

$$|L_h| \leq \sum_{\ell=0}^h |A_\ell| \leq C_3 h \log n. \tag{3.103}$$

Inserting this estimate in (3.101),

$$\sum_{s=j-\tilde{h}+1}^{j-1} \mathbb{P}_u^{a,\sigma} \left( X_j^{(\ell)} \in \partial \mathcal{B}_{h_0}^-(y); \mathcal{H}_{1,s}^\ell \mid B_{\ell-1} \right) \leq \frac{C_4}{n} \log n. \tag{3.104}$$

Combining (3.100) and (3.104) we have proved (3.99) for a suitable constant  $C_1$ .  $\square$

### 3.6 Lower bound on $\pi_{\max}$

**Lemma 3.12** *There exist constants  $\varepsilon, c > 0$  such that*

$$\mathbb{P} \left( \exists S \subset [n], |S| \geq n^\varepsilon, n \min_{y \in S} \pi(y) \geq c \log^{1-\kappa_1}(n) \right) = 1 - o(1). \tag{3.105}$$

**Proof** We argue as in the first part of the proof of Lemma 3.10. Namely, let  $(\Delta_*, \delta_*) \in \mathcal{L}$  denote the type realizing the minimum in the definition of  $\kappa_1$ ; see (1.16). 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-neighborhoods  $\mathcal{B}_{h_0}^-(y_i)$  where

$$h_0 = \log_{\Delta_*} \log n - C, \tag{3.106}$$

and call  $\sigma$  a realization of all these neighborhoods. As in the proof of Lemma 3.10, 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  $(\Delta_*, \delta_*)$ . Then, if  $C$  in (3.106) is large enough we may assume that at least  $n^{\beta_1/2}$  vertices  $y_i$  are lucky; see (3.67). As before, we call  $\mathcal{A}'$  the set of  $\sigma$  realizing these constraints. Given a realization  $\sigma \in \mathcal{A}'$ , and some  $\varepsilon \in (0, \beta_1/2)$  we fix the first  $n^\varepsilon$  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  $c > 0$

$$\max_{\sigma \in \mathcal{A}'} \mathbb{P} \left( \min_{i=1, \dots, n^\varepsilon} n\pi(y_{*,i}) < c \log^{1-\kappa_1}(n) \mid \sigma \right) = o(1). \tag{3.107}$$

To prove (3.107) we first observe that by (3.34) and Lemma 3.3 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

$$\max_{\sigma \in \mathcal{A}'} \mathbb{P} \left( \Gamma_{h_1}(y_*) < c \log^{1-\kappa_1}(n) \mid \sigma \right) \leq n^{-2\varepsilon}, \tag{3.108}$$

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-neighborhood  $\mathcal{B}_{h_1-h_0}^-(z)$ . The same argument of the proof of Lemma 3.2 shows that the probability

that all these neighborhoods are successfully coupled to i.i.d. random directed trees is at least  $1 - O(\Delta^{2h_1}/n)$ . On this event we have

$$\Gamma_{h_1}(y_*) = \delta_*^{-h_0} \sum_{i=1}^{\Delta_*^{h_0}} X_i, \tag{3.109}$$

where  $X_i = M_{h_1-h_0}^i$  is defined by (3.15). Then (3.16) shows that

$$\mathbb{P}\left(\Gamma_{h_1}(y_*) < \frac{1}{2} \Delta_*^{h_0} \delta_*^{-h_0}\right) \leq \exp\left(-c_1 \Delta_*^{h_0}\right), \tag{3.110}$$

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

$$\max_{\sigma \in \mathcal{A}'} \mathbb{P}\left(\Gamma_{h_1}(y_*) < c_2 \log^{1-\kappa_1}(n) \mid \sigma\right) \leq n^{-2\varepsilon}, \tag{3.111}$$

for some new constant  $c_2 > 0$  and for  $\varepsilon = c_1 \Delta_*^{-C}/4$ . This ends the proof of (3.108). □

### 4 Bounds on the cover time

In this section we show how the control of the extremal values of the stationary distribution obtained in previous sections can be turned into the bounds on the cover time presented in Theorem 1.9. To this end we exploit the full strength of the strategy developed by Cooper and Frieze [14–17].

#### 4.1 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  $t > T$ :

$$\mathcal{A}_y^T(t) = \{X_s \neq y, \forall s \in [T, t]\}. \tag{4.1}$$

Moreover, define the generating function

$$R_y^T(z) = \sum_{t=0}^T z^t \mathbf{P}_y(X_t = y), \quad z \in \mathbb{C}. \tag{4.2}$$

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 [15], see also [17, Lemma 3].

**Lemma 4.1** 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  $t \geq T$ :

$$\max_{x \in [n]} \left| \mathbf{P}_x \left( \mathcal{A}_y^T(t) \right) - \frac{1 + \xi_1}{(1 + p_y)^{t+1}} \right| \leq e^{-\frac{t}{2KT}}, \tag{4.3}$$

where

$$p_y = (1 + \xi_2) \frac{\pi(y)}{R_y^T(1)}. \tag{4.4}$$

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 4.1 are satisfied. From now on we fix the sequence  $T = T_n$  as

$$T = \log^3(n). \tag{4.5}$$

From (3.2) and the argument in (3.61) it follows that item (i) of Lemma 4.1 is satisfied with high probability. Moreover, Theorem 1.3 and Theorem 1.6 imply that item (ii) of Lemma 4.1 is also satisfied with high probability. Next, following [16], we define a class of vertices  $y \in [n]$  which satisfy item (iii) of Lemma 4.1. We use the convenient notation

$$\vartheta = \log \log \log(n). \tag{4.6}$$

**Definition 4.2** We call *small cycle* a collection of  $\ell \leq 3\vartheta$  edges such that their undirected projection forms a simple undirected cycle of length  $\ell$ . We say that  $v \in [n]$  is *locally tree-like (LTL)* if its in- and out-neighborhoods up to depth  $\vartheta$  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 [16, Section 3]. Recall the definition of  $\Delta$  in (2.1).

**Proposition 4.3** 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 two small cycles which are less than  $9\vartheta$  undirected steps away from each other.

**Proposition 4.4** *With high probability, uniformly in  $y \in V_1$ :*

$$R_y^T(1) = 1 + O(2^{-\vartheta}). \tag{4.7}$$

Moreover, there exist constants  $K, \psi > 0$  such that with high probability, every  $y \in V_1$  satisfies item (iii) of Lemma 4.1. In particular, (4.3) holds uniformly in  $y \in V_1$ .

**Proof** We first prove (4.7). Fix  $y \in V_1$  and consider the neighborhoods  $\mathcal{B}_\vartheta^\pm(y)$  and  $\mathcal{B}_h^\pm(y)$ . By Proposition 2.1 we may assume that  $\mathcal{B}_h^-(y)$  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  $\mathcal{B}_h^-(y) \cup \mathcal{B}_\vartheta^+(y)$  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  $\mathcal{B}_h^-(y) \cup \mathcal{B}_\vartheta^+(y)$ ;
- (2) The extra edge connects  $\mathcal{B}_h^-(y) \setminus \mathcal{B}_\vartheta^-(y)$  to itself
- (3) The extra edge connects  $\mathcal{B}_\vartheta^-(y)$  to  $\mathcal{B}_h^-(y) \setminus \mathcal{B}_\vartheta^-(y)$ ;
- (4) The extra edge connects  $\mathcal{B}_\vartheta^+(y)$  to  $\mathcal{B}_h^-(y) \setminus \mathcal{B}_\vartheta^-(y)$ .

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  $\partial\mathcal{B}_h^-(y)$ , and from any vertex of  $\partial\mathcal{B}_h^-(y)$  the probability to reach  $y$  before exiting  $\mathcal{B}_h^-(y)$  is at most  $2\delta^{-h}$ . Therefore, in these cases the number of visits to  $y$  up to  $T$  is stochastically dominated by  $1 + \text{Bin}(T, 2\delta^{-h})$  and

$$1 \leq R_y^T(1) \leq 1 + 2T\delta^{-h} = 1 + O(n^{-a}),$$

for some  $a > 0$ . In the last case instead it is possible for the walk to jump from  $\mathcal{B}_\vartheta^+(y)$  to  $\mathcal{B}_h^-(y) \setminus \mathcal{B}_\vartheta^-(y)$ . 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  $\vartheta$  units of time after its first visit to  $\partial\mathcal{B}_\vartheta^+(y)$ . Then  $\mathbf{P}_y(B) \leq \delta^{-\vartheta}$ . On the complementary event  $B^c$  the walk must enter  $\partial\mathcal{B}_h^-(y)$  before visiting  $y$ , and each time it visits  $\partial\mathcal{B}_h^-(y)$  it has probability at most  $2\delta^{-h}$  to visit  $y$  before the next visit to  $\partial\mathcal{B}_h^-(y)$ . Since the number of attempts is at most  $T$  one finds

$$\mathbf{P}_y(E_1) \leq \mathbf{P}_y(B) + \mathbf{P}_y(E_1, B^c) \leq \delta^{-\vartheta} + 2T\delta^{-h} \leq 2\delta^{-\vartheta}.$$

By the strong Markov property,

$$\mathbf{P}_y(E_k) \leq \mathbf{P}_y(E_1)^k.$$

Therefore

$$R_y^T(1) = 1 + \sum_{k=1}^{\infty} k\mathbf{P}_y(E_k) = 1 + O(\delta^{-\vartheta}).$$

To see that  $y \in V_1$  satisfies item (iii) of Lemma 4.1, take  $z \in \mathbb{C}$  with  $|z| \leq 1 + 1/KT$  and write

$$|R_y^T(z)| \geq 1 - \sum_{t=1}^T \mathbf{P}_y(X_t = y)|z|^t \geq 1 - e^{1/K}(R_y^T(1) - 1) = 1 - O(\delta^{-\vartheta}).$$

□

### 4.2 Upper bound on the cover time

We prove the following estimate relating the cover time to  $\pi_{\min}$ . From Theorem 1.3 this implies the upper bound on the cover time in Theorem 1.9.

**Lemma 4.5** *For any constant  $\varepsilon > 0$ , with high probability*

$$\max_{x \in [n]} \mathbf{E}_x(\tau_{\text{cov}}) \leq (1 + \varepsilon) \frac{\log n}{\pi_{\min}}. \tag{4.8}$$

**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  $t_* \geq T$ :

$$\begin{aligned} \mathbf{E}_x[\tau_{\text{cov}}] &= \sum_{s \geq 0} \mathbf{P}_x(\tau_{\text{cov}} > s) = \sum_{s \geq 0} \mathbf{P}_x(U_s \neq \emptyset) \\ &\leq t_* + \sum_{s \geq t_*} \mathbf{E}_x[|U_s|] = t_* + \sum_{s \geq t_*} \sum_{y \in [n]} \mathbf{P}_x(y \in U_s) \\ &\leq t_* + \sum_{s \geq t_*} \sum_{y \in [n]} \mathbf{P}_x(\mathcal{A}_y^T(s)). \end{aligned} \tag{4.9}$$

Choose

$$t_* := \frac{(1 + \varepsilon) \log n}{\pi_{\min}},$$

for  $\varepsilon > 0$  fixed. It is sufficient to prove that the last term in (4.9) is  $o(t_*)$  uniformly in  $x \in [n]$ .

From Proposition 4.4 we can estimate

$$\mathbf{P}_x(\mathcal{A}_y^T(s)) = \frac{(1 + \xi')}{(1 + \bar{p}_y)^{s+1}}, \tag{4.10}$$

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,

$$\sum_{s \geq t_*} \sum_{y \in V_1} \mathbf{P}_x(\mathcal{A}_y^T(s)) = (1 + o(1)) \sum_{y \in V_1} \frac{1}{\bar{p}_y(1 + \bar{p}_y)^{t_*}}. \tag{4.11}$$

Using  $\pi(y) \geq \pi_{\min}$ , (4.11) is bounded by

$$\frac{(1 + o(1))n}{\bar{p}_y(1 + \bar{p}_y)^{t_*}} \leq \frac{2n}{\pi_{\min}} \exp(-\pi_{\min} t_* (1 + o(1))) \leq \frac{1}{\pi_{\min}} = o(t_*),$$

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 (4.9). From Proposition 4.3 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  $y \in V_2$ :

$$\sum_{s \geq t_*} \mathbf{P}_x(\mathcal{A}_y^T(s)) = o(t_* \Delta^{-15\vartheta}). \tag{4.12}$$

To prove (4.12), fix  $y \in V_2$  and notice that by Proposition 4.3 (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

$$\begin{aligned} \mathbf{P}_x(\mathcal{A}_y^T(t_1)^c) &= \mathbf{P}_x(y \in \{X_T, X_{T+1}, \dots, X_{t_1}\}) \\ &\geq \mathbf{P}_x(u \in \{X_T, X_{T+1}, \dots, X_{t_0}\}) \mathbf{P}_u(y \in \{X_1, \dots, X_{10\vartheta}\}) \\ &\geq \left(1 - \mathbf{P}_x(\mathcal{A}_u^T(t_0))\right) \Delta^{-10\vartheta}. \end{aligned}$$

Since  $u \in V_1$ , as in (4.10), for  $n$  large enough,

$$\mathbf{P}_x(\mathcal{A}_u^T(t_0)) \leq \frac{2}{(1 + \bar{p}_y)^{t_0+1}} \leq \frac{1}{2}. \tag{4.13}$$

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  $k \in \mathbb{N}$ ,

$$\mathbf{P}_x(\mathcal{A}_y^T(s)) \leq (1 - \gamma)^k, \quad s > k(T + t_1). \tag{4.14}$$

Therefore,

$$\begin{aligned} \sum_{s \geq t_*} \mathbf{P}_x(\mathcal{A}_y^T(s)) &\leq \sum_{s \geq t_*} (1 - \gamma)^{\lfloor s/(T+t_1) \rfloor} \leq \sum_{s \geq t_*} (1 - \gamma)^{s/2t_1} \\ &\leq \frac{\exp(-\gamma t_*/2t_1)}{1 - \exp(-\gamma/2t_1)} = O(t_1/\gamma) = o(t_* \Delta^{-15\vartheta}). \end{aligned}$$

□

### 4.3 Lower bound on the cover time

We prove the following stronger statement.

**Lemma 4.6** *For some constant  $c > 0$ , with high probability*

$$\min_{x \in [n]} \mathbf{P}_x(\tau_{\text{cov}} \geq cn \log^{\gamma_1} n) = 1 - o(1). \tag{4.15}$$

Clearly, this implies the lower bound on  $T_{\text{cov}} = \max_{x \in [n]} \mathbf{E}_x(\tau_{\text{cov}})$  in Theorem 1.9. The proof of Lemma 4.6 is based on the second moment method as in [17]. If  $W \subset [n]$  is a set of vertices, let  $W_t$  be the set

$$W_t = \{y \in W : y \text{ is not visited in } [0, t]\} \tag{4.16}$$

Then

$$\mathbf{P}_x(\tau_{\text{cov}} > t) \geq \mathbf{P}_x(|W_t| > 0) \geq \frac{\mathbf{E}_x[|W_t|]^2}{\mathbf{E}_x[|W_t|^2]}. \tag{4.17}$$

Therefore, Lemma 4.6 is a consequence of the following estimate.

**Lemma 4.7** *For some constant  $c > 0$ , with high probability there exists a nonempty set  $W \subset [n]$  such that*

$$\max_{x \in [n]} \frac{\mathbf{E}_x[|W_t|^2]}{\mathbf{E}_x[|W_t|]^2} = 1 + o(1), \quad t = cn \log^{\gamma_1} n. \tag{4.18}$$

We start the proof of Lemma 4.7 by exhibiting a candidate for the set  $W$ .

**Proposition 4.8** *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 4.2, and  $|W| \geq n^\alpha$  for some constant  $\alpha > 0$ .
- (2) For some constant  $C > 0$ , for all  $y \in W$ ,

$$\pi(y) \leq \frac{C}{n} \log^{1-\gamma_1}(n). \tag{4.19}$$

- (3) For all  $x, y \in W$ :

$$|\pi(x) - \pi(y)| \leq \pi_{\min} \log^{-K}(n). \tag{4.20}$$

- (4) For all  $x, y \in W$ :  $\min\{d(x, y), d(y, x)\} > 2\vartheta$ .

**Proof** From Theorem 1.3 we know that w.h.p. there exists a set  $S \subset [n]$  with  $|S| > n^\beta$  such that (4.19) holds. Moreover, a minor modification of the proof of Lemma 3.10 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-neighborhoods  $\mathcal{B}_\vartheta^+(y_i)$  for every  $i = 1, \dots, N_1$  and the argument for (3.66) 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 (4.20) 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 (4.20).  $\square$

**Proof of Lemma 4.7** Consider the first moment  $\mathbf{E}_x[|W_t|]$ , where  $W$  is the set from Proposition 4.8 and  $t$  is fixed as  $t = cn \log^{\gamma_1}(n)$ . For  $y \in W \subset V_1$  we use Lemma 4.1 and Proposition 4.4. As in (4.10) we have

$$\mathbf{P}_x(\mathcal{A}_y^T(t)) = (1 + o(1))(1 + \bar{p}_y)^{-(t+1)}, \tag{4.21}$$

where  $\bar{p}_y = (1 + o(1))\pi(y) \leq p_W := 2Cn^{-1} \log^{1-\gamma_1}(n)$ , where  $C$  is as in (4.19). Therefore,

$$\begin{aligned} \mathbf{E}_x[|W_t|] &= \sum_{y \in W} \mathbf{P}_x(y \text{ not visited in } [0, t]) \\ &\geq -T + \sum_{y \in W} \mathbb{P}(\mathcal{A}_y^T(t)) \geq -T + (1 + o(1))|W|(1 + p_W)^{-t}. \end{aligned}$$

Taking the constant  $c$  in the definition of  $t$  sufficiently small, one has  $p_W t \leq \alpha/2 \log n$  and therefore

$$\mathbf{E}_x[|W_t|] \geq -T + (1 + o(1))|W|n^{-\alpha/2} \geq \frac{1}{2}n^{\alpha/2}, \tag{4.22}$$

where we use  $T = \log^3(n)$  and  $|W| \geq n^\alpha$ . In particular, since  $T = \log^3(n)$ , (4.22) shows that

$$\sum_{y \in W} \mathbb{P}(\mathcal{A}_y^T(t)) = (1 + o(1))\mathbf{E}_x[|W_t|]. \tag{4.23}$$

Concerning the second moment  $\mathbf{E}_x[|W_t|^2]$ , we have

$$\begin{aligned} \mathbf{E}_x[|W_t|^2] &= \sum_{y, y' \in W} \mathbf{P}_x(y \text{ and } y' \text{ not visited in } [0, t]) \\ &\leq \sum_{y, y' \in W} \mathbf{P}_x(\mathcal{A}_y^T(t) \cap \mathcal{A}_{y'}^T(t)). \end{aligned}$$

From this and (4.23), the proof of Lemma 4.7 is completed by showing, uniformly in  $x \in [n]$ ,  $y, y' \in W$ :

$$\mathbf{P}_x(\mathcal{A}_y^T(t) \cap \mathcal{A}_{y'}^T(t)) = (1 + o(1))\mathbf{P}_x(\mathcal{A}_y^T(t))\mathbf{P}_x(\mathcal{A}_{y'}^T(t)). \tag{4.24}$$

We follow the idea of [17]. 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 4.2. Moreover,  $G^*$  has the law of a directed configuration model with the same degree sequence of  $G$  except that at  $y_*$  it has  $d_{y_*}^\pm = d_y^\pm + d_{y'}^\pm$ . It follows that we may apply Lemma 4.1 and Proposition 4.4. Therefore, if  $\mathbf{P}_x^*$  denotes the law of the random walk on  $G^*$  started at  $x$ , as in (4.21) we have

$$\mathbf{P}_x^*(\mathcal{A}_{y_*}^T(t)) = (1 + o(1))(1 + \bar{p}_{y_*})^{-t}, \tag{4.25}$$

uniformly in  $x \in [n], y, y' \in W$ , where  $\bar{p}_{y_*} = (1 + o(1))\pi^*(y_*)$ , and  $\pi^*$  is the stationary distribution of  $G^*$ . In Lemma 4.9 below we prove that

$$\max_{\substack{v \in [n]: \\ v \neq y, y'}} |\pi(v) - \pi^*(v)| \leq a, \quad |\pi(y) + \pi(y') - \pi^*(y_*)| \leq a, \tag{4.26}$$

where  $a := \pi_{\min} \log^{-1}(n)$ . Assuming (4.26), we can conclude the proof of (4.24). Indeed, letting  $P_*$  denote the transition matrix of the graph  $G^*$ ,

$$\begin{aligned} \mathbf{P}_x^*(\mathcal{A}_{y_*}^T(t)) &= \sum_{v \neq y, y'} P_*^T(x, v) \mathbf{P}_v^*(X_s \neq y_*, \forall s \in [1, t - T]) \\ &= \sum_{v \neq y, y'} \left( \pi^*(v) + O(n^{-3}) \right) \mathbf{P}_v^*(X_s \neq y_*, \forall s \in [1, t - T]) \end{aligned}$$

On the other hand,

$$\begin{aligned} \mathbf{P}_x(\mathcal{A}_y^T(t) \cap \mathcal{A}_{y'}^T(t)) &= \sum_{v \neq y, y'} P^T(x, v) \mathbf{P}_v(X_s \notin \{y, y'\}, \forall s \in [1, t - T]) \\ &= \sum_{v \neq y, y'} \left( \pi(v) + O(n^{-3}) \right) \mathbf{P}_v(X_s \notin \{y, y'\}, \forall s \in [1, t - T]) \end{aligned}$$

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 used condition (i) in Lemma 4.1 to estimate  $1 \leq \frac{(1+o(1))}{\pi_{\min}} P^T(x, v)$ . Therefore, using (4.26)

$$\begin{aligned} & \left| \mathbf{P}_x \left( \mathcal{A}_y^T(t) \cap \mathcal{A}_{y'}^T(t) \right) - \mathbf{P}_x^* \left( \mathcal{A}_{y_*}^T(t) \right) \right| \\ & \leq \sum_{v \neq y, y'} |\pi(v) - \pi_*(v) + O(n^{-3})| \mathbf{P}_v(X_s \notin \{y, y'\}, \forall s \in [1, t - T]) \\ & \leq (a + O(n^{-3})) \frac{(1 + o(1))}{\pi_{\min}} \sum_{v \neq y, y'} P^T(x, v) \mathbf{P}_v(X_s \notin \{y, y'\}, \forall s \in [1, t - T]) \\ & \leq \frac{2a}{\pi_{\min}} \mathbf{P}_x(\mathcal{A}_y(t) \cap \mathcal{A}_{y'}(t)). \end{aligned}$$

By definition of  $a$  we have  $a/\pi_{\min} \rightarrow 0$  so that

$$\mathbf{P}_x(\mathcal{A}_y^T(t) \cap \mathcal{A}_{y'}^T(t)) = (1 + o(1)) \mathbf{P}_x^*(\mathcal{A}_{y_*}^T(t)). \tag{4.27}$$

Using (4.21), (4.25) and (4.26) 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 4.9** *The stationary distributions  $\pi, \pi^*$  satisfy (4.26).*

**Proof** We follow the proof of Eq. (107) in [17]. The stochastic matrix of the simple random walk on  $G^*$  is given by

$$P_*(v, w) = \begin{cases} P(v, w) & \text{if } v, w \neq y_* \\ \frac{1}{2} (P(y, w) + P(y', w)) & \text{if } v = y_* \\ P(v, y) + P(v, y') & \text{if } w = y_*. \end{cases}$$

Let  $V^*$  denote the vertices of  $G^*$ . Define the vector  $\zeta(v), v \in V^*$  via

$$\zeta(v) = \begin{cases} \pi_*(v) - \pi(v) & v \neq y_* \\ \pi_*(y_*) - (\pi(y) + \pi(y')) & v = y_* \end{cases}$$

We are going to show that

$$\max_{v \in V^*} |\zeta(v)| = o(\pi_{\min} \log^{-1}(n)), \tag{4.28}$$

which implies (4.26). A computation shows that

$$\begin{aligned} \zeta P_*(w) &= \sum_{v \in V^*} \zeta(v) P_*(v, w) \\ &= \begin{cases} \zeta(w) & \text{if } w \notin \mathcal{B}_1^+(y) \cup \mathcal{B}_1^+(y') \\ \zeta(w) + \frac{\pi(y') - \pi(y)}{2} P(y, w) & \text{if } w \in \mathcal{B}_1^+(y) \\ \zeta(w) + \frac{\pi(y) - \pi(y')}{2} P(y', w) & \text{if } w \in \mathcal{B}_1^+(y'). \end{cases} \end{aligned}$$

Therefore, the vector  $\phi := \zeta(I - P_*)$  satisfies

$$|\phi(w)| \leq \begin{cases} 0 & \text{if } w \notin \mathcal{B}_1^+(y) \cup \mathcal{B}_1^+(y') \\ |\pi(y) - \pi(y')| & \text{otherwise.} \end{cases}$$

Hence  $\phi(v) = 0$  for all but at most  $2\Delta$  vertices  $v$ , and recalling (4.20) we have

$$|\phi(w)| \leq \pi_{\min} \log^{-K}(n). \tag{4.29}$$

Next, consider the matrix

$$M = \sum_{s=0}^{T-1} P_*^s,$$

and notice that

$$\zeta(I - P_*^T) = \phi M.$$

Since  $P_*$  and  $\pi_*$  satisfy condition (i) in Lemma 4.1,

$$P_*^T = \Pi_* + E, \quad \text{with } |E(u, v)| \leq n^{-3}, \quad \forall u, v \in V^*, \tag{4.30}$$

where  $\Pi_*$  denotes the matrix with all rows equal to  $\pi_*$ . We rewrite the vector  $\zeta$  as

$$\zeta = \alpha \pi_* + \rho,$$

where  $\alpha \in \mathbb{R}$  and  $\rho$  is orthogonal to  $\pi_*$ , 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,

$$|\langle \phi M, \rho \rangle| \leq \sum_{s=0}^{T-1} |\langle \phi, P_*^s \rho \rangle| \leq T \frac{\pi_{\max}^*}{\pi_{\min}^*} \|\phi\|_2 \|\rho\|_2, \tag{4.31}$$

where we use

$$\begin{aligned} \langle P_*^s \psi, P_*^s \psi \rangle &\leq \frac{1}{\pi_{\min}^*} \sum_v \pi^*(v) (P_*^s \psi)^2(v) \\ &\leq \frac{1}{\pi_{\min}^*} \sum_{u,v} \pi^*(v) P_*^s(v, u) \psi^2(u) \\ &= \frac{1}{\pi_{\min}^*} \sum_u \pi^*(u) \psi^2(u) \leq \frac{\pi_{\max}^*}{\pi_{\min}^*} \|\psi\|_2^2, \end{aligned}$$

for any vector  $\psi : V^* \mapsto \mathbb{R}$ . On the other hand,

$$|\langle \rho, (I - E)\rho \rangle| \geq \|\rho\|_2^2 - n^{-3} \left( \sum_v |\rho(v)| \right)^2 \geq \|\rho\|_2^2 (1 - n^{-2}). \tag{4.32}$$

Using (4.29), from (4.31) and (4.32) we conclude that

$$\|\rho\|_2 \leq 2T \frac{\pi_{\max}^*}{\pi_{\min}^*} \|\phi\|_2 = 2T \frac{\pi_{\max}^*}{\pi_{\min}^*} \times O(\pi_{\min} \log^{-K}(n)).$$

From Theorem 1.3 applied to  $G^*$  we can assume that  $\frac{\pi_{\max}^*}{\pi_{\min}^*} = O(\log^{K/3}(n))$  if  $K$  is a large enough constant. Since  $T = \log^3(n)$ , with  $K$  sufficiently large one has

$$\|\rho\|_2 \leq \pi_{\min} \log^{-K/2}(n).$$

Next, notice that

$$0 = \langle \zeta, 1 \rangle = \langle \alpha \pi_* + \rho, 1 \rangle = \alpha + \langle \rho, 1 \rangle.$$

Hence

$$|\alpha| = |\langle \rho, 1 \rangle| \leq \sqrt{n} \|\rho\|_2 \leq \sqrt{n} \pi_{\min} \log^{-K/2}(n).$$

In conclusion,

$$\begin{aligned} \zeta(v)^2 &\leq 2\alpha^2 \pi_*(v)^2 + 2\rho(v)^2 \leq 2n\pi_{\min}^2 \log^{-K}(n) (\pi_{\max}^*)^2 + 2\|\rho\|_2^2 \\ &\leq 2n\pi_{\min}^2 \log^{-K}(n) (\pi_{\max}^*)^2 + 2\pi_{\min}^2 \log^{-K}(n) \leq 4\pi_{\min}^2 \log^{-K}(n), \end{aligned}$$

which implies (4.28). □

#### 4.4 The Eulerian case

We prove Theorem 1.12. The strategy is the same as for the proof of Theorem 1.9, 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  $t_* = (1 + \varepsilon)\beta n \log n$ ,

$$\sum_{y \in V_1} \sum_{s \geq t_*} \mathbf{P}_x(\mathcal{A}_y^T(s)) + \sum_{y \in V_2} \sum_{s \geq t_*} \mathbf{P}_x(\mathcal{A}_y^T(s)) = o(n \log n). \tag{4.33}$$

Letting  $\mathcal{V}_d$  denote the set of vertices with degree  $d$ , reasoning as in (4.11) 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

$$\sum_{y \in V_1} \sum_{s \geq t_*} \mathbf{P}_x(\mathcal{A}_y^T(s)) \leq \frac{2m}{\delta} \sum_{d=\delta}^{\Delta} \exp\left(-\left(\frac{d\beta}{d} - \alpha_d\right) \log n\right) = O(n), \tag{4.34}$$

since by definition  $\frac{d\beta}{d} - \alpha_d \geq 0$ . Concerning the vertices  $y \in V_2$  one may repeat the argument in (4.14) without modifications, to obtain

$$\sum_{y \in V_2} \sum_{s \geq t_*} \mathbf{P}_x(\mathcal{A}_y^T(s)) = o(n \log n). \tag{4.35}$$

Thus, (4.33) follows from (4.34) and (4.35).

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  $\varepsilon > 0$ ,

$$\min_{x \in [n]} \mathbf{P}_x\left(\tau_{\text{cov}} \geq (1 - \varepsilon) \frac{\bar{d}\alpha_d}{d} n \log^{\gamma_1} n\right) = 1 - o(1). \tag{4.36}$$

We proceed as in the proof of Lemma 4.7. Here we choose  $W$  as the subset of  $\mathcal{V}_d$  consisting of LTL vertices in the sense of Definition 4.2 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

$$|W| \geq n^{\alpha_d + o(1)}. \tag{4.37}$$

Indeed, the vertices that are not LTL are at most  $\Delta^{9\vartheta}$  by Proposition 4.3. 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  $2\vartheta$  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 4.9 since  $a$  can be taken to be zero in (4.26) in the

Eulerian case. The only thing to control is the validity of the bound (4.23) with the choice

$$t = (1 - \varepsilon) \frac{\bar{d}\alpha_d}{d} n \log n.$$

As in (4.23), it suffices to check that with high probability

$$\sum_{y \in W} \mathbb{P}(\mathcal{A}_y^T(t)) - T \rightarrow \infty. \quad (4.38)$$

From (4.21) we obtain

$$\sum_{y \in W} \mathbb{P}(\mathcal{A}_y^T(t)) = (1 + o(1))|W| \exp\left(-\frac{(1+o(1))d}{m} t\right). \quad (4.39)$$

Using (4.37) and  $dt/m = (1 - \varepsilon)\alpha_d \log n$ , (4.39) is at least  $n^{\varepsilon\alpha_d/2}$  for all  $n$  large enough. Since  $T = \log^3(n)$  this proves (4.38).

**Acknowledgements** We thank Charles Bordenave and Justin Salez for useful discussions. We acknowledge support of PRIN 2015 5PAWZB “Large Scale Random Structures”, and of INdAM-GNAMPa Project 2019 “Markov chains and games on networks”.

**Funding** Open access funding provided by Università degli Studi Roma Tre within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Abdullah, M., Cooper, C., Frieze, A.M.: Cover time of a random graph with given degree sequence. *Discret. Math.* **312**(21), 3146–3163 (2012)
2. Addario-Berry, L., Balle, B., Perarnau, G.: Diameter and stationary distribution of random  $r$ -out digraphs. (2015). ArXiv preprint [arXiv:1504.06840](https://arxiv.org/abs/1504.06840)
3. Aldous, D.: An introduction to covering problems for random walks on graphs. *J. Theoret. Prob.* **2**(1), 87–89 (1989)
4. Aldous, D.J.: 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)
5. 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)
6. Barral, J.: Mandelbrot cascades and related topics. In: *Geometry and Analysis of Fractals*, pp. 1–45. Springer, New York (2014)

7. Boczkowski, L., Peres, Y., Sousi, P.: Sensitivity of mixing times in Eulerian digraphs. *SIAM J. Discret. Math.* **32**(1), 624–655 (2018)
8. Bollobas, B., Fernandez de la Vega, W.: The diameter of random regular graphs. *Combinatorica* **2**(2), 125–134 (1982)
9. Bordenave, C., Caputo, P., Salez, J.: Random walk on sparse random digraphs. *Prob. Theory Relat. Fields* **170**(3), 933–960 (2018)
10. Bordenave, C., Caputo, P., Salez, J.: Cutoff at the “entropic time” for sparse Markov chains. *Prob. Theory Relat. Fields* **173**(1), 261–292 (2019)
11. Caputo, P., Quattropani, M.: Mixing time of pagerank surfers on sparse random digraphs. (2019). ArXiv preprint [arXiv:1905.04993](https://arxiv.org/abs/1905.04993)
12. Chatterjee, S.: Stein’s method for concentration inequalities. *Prob. Theory Relat. fields* **138**(1), 305–321 (2007)
13. Cooper, C., Frieze, A.: The size of the largest strongly connected component of a random digraph with a given degree sequence. *Combinat. Prob. Comput.* **13**, 05 (2002)
14. Cooper, C., Frieze, A.: The cover time of sparse random graphs. *Random Struct. Algorithms* **30**(1–2), 1–16 (2007)
15. Cooper, C., Frieze, A.: The cover time of the giant component of a random graph. *Random Struct. Algorithms* **32**(4), 401–439 (2008)
16. Cooper, C., Frieze, A.M.: The cover time of random regular graphs. *SIAM J. Discret. Math.* **18**(4), 728–740 (2005)
17. 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)
18. Feige, U.: A tight lower bound on the cover time for random walks on graphs. *Random Struct. Algorithms* **6**(4), 433–438 (1995)
19. Feige, U.: A tight upper bound on the cover time for random walks on graphs. *Random Struct. Algorithms* **6**(1), 51–54 (1995)
20. Jonasson, J.: On the cover time for random walks on random graphs. *Comb. Prob. Comput.* **7**(3), 265–279 (1998)
21. Kahn, J.D., Linial, N., Nisan, N., Saks, M.E.: On the cover time of random walks on graphs. *J. Theoret. Prob.* **2**(1), 121–128 (1989)
22. Levin, D.A., Peres, Y.: Markov Chains and Mixing Times. In: Propp, J.G., & Wilson, D.B. (Eds.) With contributions by Elizabeth L. Wilmer, With a chapter on “Coupling from the past”, Second edition. American Mathematical Society, Providence, RI (2017)
23. Liu, Q.: The growth of an entire characteristic function 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 (1996)
24. Liu, Q.: On generalized multiplicative cascades. *Stochastic Processes and their Applications* **86**(2), 263–286 (2000)
25. Liu, Q.: Asymptotic properties and absolute continuity of laws stable by random weighted mean. *Stoch. Process. Appl.* **95**(1), 83–107 (2001)
26. Lovász, L.: Random walks on graphs: a survey. *Comb. Paul Erdos Eighty* **2**(1), 1–46 (1993)
27. Rösler, U.: A fixed point theorem for distributions. *Stoch. Process. Appl.* **42**(2), 195–214 (1992)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.