Mixing times of random walks on dynamic configuration models

Luca Avena[*]
Hakan Güldaş[*]
Remco van der Hofstad[†]
Frank den Hollander[*]

June 23, 2016

Abstract

The mixing time of a simple random walk on a random graph generated according to the configuration model is known to be of order $\log n$ when n is the number of vertices. In this paper we investigate what happens when the random graph becomes dynamic, namely, at each unit of time a fraction α_n of the edges is randomly relocated. For degree distributions that converge and have a second moment that is bounded in n, we show that the mixing time is of order $1/\sqrt{\alpha_n}$, provided $\lim_{n\to\infty} \alpha_n (\log n)^2 = \infty$. We identify the sharp asymptotics of the mixing time when we additionally require that $\lim_{n\to\infty} \alpha_n = 0$, and relate the relevant proportionality constant to the average probability of escape from the root by a simple random walk on an augmented Galton-Watson tree which is obtained by takin a Galton-Watson tree whose offspring distribution is the size-biased version of the limiting degree distribution and attaching to its root another Galton-Watson tree with the same offspring distribution. Our proofs are based on a randomised stopping time argument in combination with coupling estimates.

1 Introduction and results

1.1 Motivation and background

The *mixing time* of a Markov chain is the time it needs to approach its stationary distribution. For random walks on *finite graphs*, the characterisation of the mixing time has been the subject of intensive study. One of the main motivations is the fact that the mixing time gives information about the geometry of the graph (see the books by Aldous and Fill [2] and Levin, Peres and Wilson [20] for an account of the state of the art and for applications).

In the last decade, much attention has been then devoted to the analysis of mixing times for random walks on *finite random graphs*. Random graphs are used as models for real-world networks. Three models have been in the focus of attention: the Erdős-Rényi random graph by Benjamini, Kozma and Wormald [5], Ding, Lubetzky and Peres [12], Fountoulakis and

---

[*]Mathematical Institute, Leiden University, P.O. Box 9512, 2300 RA Leiden, The Netherlands
[†]Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Reed [14], and Nachmias and Peres [26], percolation clusters by Benjamini and Mossel [6], and the configuration model by Ben-Hamou and Salez [7], Beresticky, Lubetzky, Peres and Sly [8], Bordenave, Caputo and Salez [11], and Lubetzky and Sly [21].

Many real-world networks are dynamic in nature. It is therefore natural to study random walks on *dynamic finite random graphs*. This line of research started By Peres, Stauffer and Steif in [27], where the authors characterise the mixing time of a simple random walk on dynamical percolation clusters on a discrete torus. The goal of the present paper is to study the mixing time of a simple random walk on a dynamic version of the configuration model.

The *static* configuration model is a random graph with a prescribed degree sequence (possibly random). It is popular because of its mathematical tractability and its flexibility in modelling real-world networks (see van der Hofstad [16] for a recent overview). For random walks on the static configuration model, asymptotics of the mixing time (and related properties such as the presence of the so-called cutoff phenomenon) have been derived recently. In particular, under mild assumptions on the degree sequence, guaranteeing that the graph is an *expander* with high probability, the mixing time is known to be of order $\log n$, with n the number of vertices.

In the present paper we consider a *discrete-time* dynamic version of the configuration model, where at each unit of time a fraction α_n of the edges is sampled and relocated uniformly at random.[1] Our dynamics preserves the degrees of the vertices. Consequently, when considering a simple random walk on this dynamic configuration model, its stationary distribution remains constant over time and the analysis of its mixing time is a well-posed question. In particular, it is natural to expect that, due to the graph dynamics, the random walk *mixes faster* than the $\log n$ order known for the static model. In our main theorems we will make this precise under only mild assumptions. By assuming regularity conditions on the prescribed degree sequence, stated in Condition 1.1 and Remark 1.2 below, and requiring that $\lim_{n \to \infty} \alpha_n (\log n)^2 = \infty$, we show in Theorem 1.5 below that the mixing time is of order $1/\sqrt{\alpha_n}$, with high probability in the sense of Definition 1.4 below. Moreover, under the additional requirement that $\lim_{n \to \infty} \alpha_n = 0$, we obtain in Theorem 1.6 below the sharp asymptotics of the mixing time, with a prefactor that is related to the escape probability from the root of a simple random walk on an associated Galton-Watson tree. This link comes from the fact that the configuration model is locally tree-like.

1.2 Model

For a graph G, we write $V(G)$ to denote the set of vertices of G and $E(G)$ to denote the set of edges of G. We denote by $\mathcal{CM}_n(d_n)$ the set of all graphs on n vertices with degree sequence

$$d_n = (d_n(v))_{v \in [n]}, \qquad [n] = \{1, \ldots, n\}. \tag{1.1}$$

The total degree $\ell_n = \sum_{v \in [n]} d_n(v)$ is assumed to be even, so that the total number of edges $m_n = \ell_n/2$ is integer. The graph need not be simple: it may have self-loops and multiple edges. To each degree sequence we associate a random graph $CM_n(d_n) \in \mathcal{CM}_n(d_n)$, called the *configuration model*. Inspired by Bender and Canfield [4], this model was introduced by Bollobás [9] to study the number of regular graphs of a given size (see also Bollobás [10]). Molloy and Reed [24], [25] introduce the configuration model with general prescribed degrees. We begin by describing how the graph is generated.

Statics. To each vertex $v \in [n]$ we associate a set of half edges $W_n(v) \subset [\ell_n]$ by letting $s \in [\ell_n]$ be in $W_n(i)$, $i \in [n]$, if and only if $\sum_{j=1}^{i-1} d_n(j) < s \leq \sum_{j=1}^{i} d_n(j)$. The edges of the graphs are comprised of pairs of half-edges, and for a half-edge s we say that s is incident to $v \in [n]$ when $s \in W_n(v)$. An edge $e = \{s, t\}$ is incident to $v \in [n]$ when either s or t is.

---

[1] A different dynamic version of the configuration model was considered in the context of *graph sampling*; see [15] and references therein.

Partitions of an even-sized set A into pairs are called *configurations* of A, and the set of all configurations of A is denoted by $\mathcal{C}onf_A$. A configuration of $[\ell_n]$ together with a degree sequence d_n identify a graph by taking the configuration to be the edge set of the graph. We formalize this identification rule via the function $I_{d_n} \colon \mathcal{C}onf_{[\ell_n]} \to \mathcal{CM}_n(d_n)$, where $V(I_{d_n}(\eta)) = [n]$ and $E(I_{d_n}(\eta)) = \eta$ for $\eta \in \mathcal{C}onf_{[\ell_n]}$. Note that different configurations can map to the same graph.

We denote a uniform random configuration of ℓ_n half-edges by $\mathrm{Conf}_{[\ell_n]}$. A configuration can be drawn uniformly at random from $\mathcal{C}onf_{[\ell_n]}$ through the following algorithm:

1. Initialize $S = [\ell_n], \mathrm{Conf}_{[\ell_n]} = \varnothing$.

2. Pick a half-edge, say s, uniformly at random from $S \setminus \{\min S\}$.

3. Update $\mathrm{Conf}_{[\ell_n]} \to \mathrm{Conf}_{[\ell_n]} \cup \{\{s, \min S\}\}$ and $S \to S \setminus \{\min S, s\}$.

4. If $S \neq \varnothing$, then continue from step 2. Else return $\mathrm{Conf}_{[\ell_n]}$.

A random graph in the configuration model is generated by identifying the resulting configuration with a graph using the map I_{d_n}. The above algorithm simulates an exchangeable process that draws a configuration uniformly at random. At each step, it picks one of the free (i.e., not yet paired) half-edges (we pick $\min S$ in our algorithm) and pairs it with another half-edge that is chosen uniformly at random from the remaining free half-edges (denoted by $S \setminus \{\min S\}$). The order in which we pick the half-edges does not affect the distribution of the resulting configuration. Note that two half-edges that belong to same vertex can be paired, which creates a self-loop, or two half-edges that belong to vertices that already have an edge between them can be paired, which creates multiple edges. However, if the degrees are not too large (see, in particular, Condition 1.1 below), then the number of self-loops and the number of multiple edges converge to two independent Poisson random variables (see Janson [18], [19], and the recent approach by Angel, van der Hofstad and Holmgren [3] based on Stein's method, which also gives bounds on the speed of convergence).

Dynamics. Our main object of study is the *Random Walk on the Dynamic Configuration Model* (RWDCM). This is a sequence of Markov chains $(M^n)_{n \in \mathbb{N}}$, where each element of the sequence is a joint Markov chain $(M_t^n)_{t \in \mathbb{N}_0} = (X_t^n, C_t^n)_{t \in \mathbb{N}_0}$ on the Cartesian product of vertex set and configuration set, $[n] \times \mathcal{C}onf_{[\ell_n]}$, parameterized by degree sequences $(d_n)_{n \in \mathbb{N}}$ and a sequence of numbers that control the graph dynamics, $(k_n)_{n \in \mathbb{N}}$ such that $2 \leq k_n \leq m_n = \ell_n/2$, $n \in \mathbb{N}$. At each time, the graph identified by the configuration C_t^n is denoted by G_t^n, i.e. $G_t^n := I_{d_n}(C_t^n)$. For a fixed n, a starting configuration η^n and a starting vertex $u^n \in [n]$, the chain $(M_t^n)_{t \in \mathbb{N}_0}$ starts from $(X_0^n, C_0^n) = (u^n, \eta^n)$ and proceeds as follows:

1. At each time $t \in \mathbb{N}$, pick k_n edges from C_{t-1}^n, say $E_i = \{s_{2i-1}, s_{2i}\}$, $1 \leq i \leq k_n$, uniformly at random from all subsets of C_{t-1}^n of size k_n.

2. Generate a random configuration $\mathrm{Conf}_{[2k_n]} = \{\{i_1, i_2\}, \dots, \{i_{2k_n-1}, i_{2k_n}\}\}$ of $2k_n$ half-edges and set $C_t^n = C_{t-1}^n \setminus \{E_1, \dots, E_{k_n}\} \cup \{\{s_{i_1}, s_{i_2}\}, \dots, \{s_{i_{2k_n-1}}, s_{i_{2k_n}}\}\}$, i.e., we rewire the edges $E_1, \dots, E_{k_n}$ using the configuration model constrained to these edges.

3. Choose a half-edge, say p, from $W_n(X_{t-1}^n)$ uniformly at random. For $q \in [\ell_n]$ such that $\{p, q\} \in C_t^n$, set $X_t^n = x$, where x is the vertex such that $q \in W_n(x)$, i.e., make a random walk move on G_t^n.

The above description can be rephrased as follows: at each time step we pick $m_n - k_n$ edges uniformly at random from the set of current edges and generate a random graph from the configuration model conditioned on those edges that are already present. We first advance the graph process by rewiring the remaining edges and then make a simple random walk move.

1.3 Main theorems

Let U_n be uniformly distributed on $[n]$. Then

$$D_n = d_n(U_n) \tag{1.2}$$

is the degree of a random vertex on the graph of size n. Write $\mathbb{P}_n$ to denote the law of U_n. Throughout the sequel, we impose the following regularity conditions on $(d_n)_{n \in \mathbb{N}}$:

Condition 1.1 (Regularity of degrees). *There exists a random variable D such that:*

(R1) $D_n \xrightarrow{d} D$ *as* $n \to \infty$ *(convergence in distribution).*

(R2) $\lim_{n \to \infty} \mathbb{E}_n[D_n^2] = \mathbb{E}[D^2] < \infty$.

(R3) $\mathbb{P}_n(D_n \geq 3) = 1$ *for all* $n \in \mathbb{N}$.

Remark 1.2. Conditions (R1) and (R2) ensure that the probability of $\mathrm{CM}_n(d_n)$ being simple has a strictly positive limit [16]. Condition (R3) ensures that the probability of $\mathrm{CM}_n(d_n)$ being connected tends to one as shown by Łuczak [22] and Federico and van der Hofstad [13] (see also [17]).

Let $\mathbb{P}_{u^n,\eta^n}^n$ denote the probability law of the joint Markov chain M_t^n for a starting vertex $u^n \in [n]$ and a starting configuration η^n. Define the random walk ε-mixing time

$$t_{\mathrm{mix}}^n(\varepsilon; u^n, \eta^n) = \inf \left\{ t \in \mathbb{N}_0 \colon \|\mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot) - \pi_n(\cdot)\|_{\mathrm{TV}} < \varepsilon \right\}, \tag{1.3}$$

where $\pi_n(v) = d_n(v)/\ell_n$ is the invariant distribution of the simple random walk on a graph with degree sequence d_n.

Remark 1.3. We use the term "mixing time" even though the random walk component is not Markovian when it is marginalised. However, the term is well-defined because the graph process does not change the degree sequence. Hence the invariant distribution is the same for all realisations of the graph process, and the random walk conditioned on a realisation of the graph process is Markovian.

Note that $t_{\mathrm{mix}}^n(\varepsilon; u^n, \eta^n)$ depends on the initial vertex u^n and the initial configuration η^n. However, we will only prove statements that hold for *typical* choices of (u^n, η^n). To formalise this, we define

$$\mu_n = \mathcal{U}_n \times \mathcal{L}(\mathrm{Conf}_{[\ell_n]}), \tag{1.4}$$

where $\mathcal{U}_n$ is the uniform distribution on $[n]$, and $\mathcal{L}(\mathrm{Conf}_{[\ell_n]})$ denotes the law of the uniform random configuration on ℓ_n half-edges.

Definition 1.4 (With high probability). *A statement that depends on the initial vertex u_n and the initial configuration η_n is said to hold with high probability if the μ_n-measure of the set of pairs (u_n, η_n) for which the statement holds tends to 1 as $n \to \infty$.*

Write

$$\alpha_n = k_n/m_n, \qquad n \in \mathbb{N}, \tag{1.5}$$

to denote the proportion of edges involved in the swapping at each time step. Our first theorem provides an upper and a lower bound on the mixing time.

Theorem 1.5 (Rough asymptotics of mixing time). *Suppose that $\lim_{n \to \infty} \alpha_n (\log n)^2 = \infty$. Then, for every $\varepsilon > 0$, with high probability*

$$[1 + o(1)] \frac{\sqrt{2}}{\sqrt{\alpha_n}} \sqrt{\log(1/\varepsilon)} \leq t_{\mathrm{mix}}^n(\varepsilon; u^n, \eta^n) \leq [1 + o(1)] \frac{2\sqrt{3}}{\sqrt{\alpha_n}} \sqrt{\log(1/\varepsilon)}. \tag{1.6}$$

4

Theorem 1.5 identifies the mixing time as being of order $1/\sqrt{\alpha_n}$, but the constants show a gap. Our second theorem closes this gap when $\alpha_n \downarrow 0$, i.e., the dynamics becomes "slow". Let $\mathbb{P}^{\mathcal{GW}}$ denote the law of the Galton-Watson tree with offspring distribution f given by

$$f(k) = \frac{(k+1)}{\mathbb{E}[D]} \mathbb{P}(D = k+1), \qquad k \in \mathbb{N}_0, \tag{1.7}$$

i.e., the size-biased version of the degree distribution D in Condition 1.1. Given a realisation ω of the tree, consider a simple random walk $X = (X_t)_{t \in \mathbb{N}_0}$ on ω starting from the root, and write $\mathbb{P}_\omega$ to denote its law. Let $R_t = |\{X_0, \ldots, X_t\}|$ denote the number of distinct vertices visited by X up to time $t \in \mathbb{N}_0$. With the help of the techniques developed by Lyons, Pemantle and Peres [23], it is shown by Piau [28] that

$$\mathbb{P}_\omega \left(\lim_{t \to \infty} \frac{1}{t} R_t = a \right) = 1 \quad \text{for } \mathbb{P}^{\mathcal{GW}}\text{-a.e. } \omega, \tag{1.8}$$

where a is given by the formula

$$a = \mathbb{E}^{\mathcal{GW}} \left(\frac{C(\omega)}{1 + C(\omega)} \right), \tag{1.9}$$

where $C(\omega)$ is the effective conductance of ω between the root and infinity. This quantity can also be characterised as the average escape probability of the simple random walk from the root of an augmented Galton-Watson tree, which consists of two Galton-Watson trees joined together at the roots (see Lyons, Pemantle and Peres [23]).

Theorem 1.6 (**Sharp asymptotics of mixing time for slow dynamics**). *Suppose that* $\lim_{n \to \infty} \alpha_n (\log n)^2 = \infty$ *and* $\lim_{n \to \infty} \alpha_n = 0$. *Then, for every* $\varepsilon > 0$, *with high probability*

$$t_{\mathrm{mix}}^n(\varepsilon; u^n, \eta^n) = [1 + o(1)] \frac{\sqrt{2/a}}{\sqrt{\alpha_n}} \sqrt{\log(1/\varepsilon)}. \tag{1.10}$$

1.4 Discussion and outline

Discussion. Theorem 1.5 gives us upper and lower bounds on the mixing time when the dynamics is not too slow: $\alpha_n \gg 1/(\log n)^2$. The mixing time is of order $1/\sqrt{\alpha_n}$, which shows that the dynamics *speeds up* the mixing significantly. Indeed, the critical regime $\alpha_n \approx 1/(\log n)^2$, not captured by Theorem 1.5, corresponds to $1/\sqrt{\alpha_n} \approx \log n$, which is the order of the mixing time found for the static configuration model.

Theorem 1.6 gives the sharp asymptotics of the mixing time in the regime where the dynamics is slow. The proportionality constant involves a non-trivial constant $a \in (0, 1)$ related to simple random walk on a Galton-Watson tree of which the offspring distribution is the size biased version of the empirical degree distribution. This link shows, among others, that the mixing time is controlled by a *combination* of the transient behaviour of simple random walk and the spatial behaviour of the configuration model.

Our proofs can be used to extend our theorems in the following directions:

$\triangleright$ $\mathbb{E}(D^2) = \infty$ and $\lim_{n \to \infty} \alpha_n (\log \log n)^2 = \infty$;

$\triangleright$ $(k_n)_{n \in \mathbb{N}}$ is random with bounded inverse moment;

$\triangleright$ time is continuous and edges are randomly relocated at rate α_n;

$\triangleright$ mixing time of the joint process of random graph and random walk.

We will pursue these directions in future work. Another natural follow up is to derive asymptotics for the *cover time* of the random walk. Robust techniques to study cover times on *static* random graphs have been developed recently (see Abdullah, Cooper and Frieze [1] and the references therein).

The mixing time of the graph process and the random walk *jointly* is clearly larger than the mixing times of each separately. While the graph process helps the random walk to mix, the converse is not true, since the graph process does not depend on the random walk. Observe that once the graph process mixes we have an almost uniform configuration, and hence the random walk mixes immediately. This observation suggests that if the mixing times of the graph process and the random walk are not of the same order, then the mixing time of the joint process will have the same order as the mixing time of the graph process. An upper bound for this can be obtained by a comparison with a coupon collector's problem with group drawings as studied by Stadje [29]. Indeed, the random time at which all the edges are rewired by the process is a strong stationary time, and has the same distribution as the random time it takes to collect all coupons when at each unit of time k_n coupons are drawn from m_n coupons in the coupon collector's problem. This upper bound is of the order $(\ell_n/k_n)\log\ell_n$, which is of the same order as $\log n/\alpha_n$, and is much larger than the mixing time $1/\sqrt{\alpha_n}$ of our random walk in the case where $\alpha_n = o(1)$. This suggests that the random walk mixes much faster than the graph itself.

Outline. The remainder of this paper is organised as follows. In Section 2 we collect various preparatory results. In particular, in Section 2 we introduce a key stopping time τ_n and in Section 2.1 we state and prove a key proposition (Proposition 2.3 below) giving upper and lower bounds on the total variation distance defined in (1.3) in terms of the tail probabilities of τ_n. This proposition will be one of the main ingredients in the proofs of our main theorems along with two more ingredients: (1) control on the tails of τ_n in Section 2.2; (2) a branching process approximation of the random graph in Section 2.3. Section 3 is devoted to the proofs of the main theorems. For the proof of Theorem 1.6 we use Proposition 2.3 in combination with a lemma on the range of a simple random walk on the Galton-Watson with offspring distribution given in (1.7), and a lemma on the convergence of the escape probability as $n \to \infty$. The latter lemmas are proved in Appendix A.

2   Preparations

We employ a *randomized stopping time argument* to get an upper bound on the total variation distance in terms of the distribution function of the randomized stopping time. Before going into details, we define auxiliary random variables that are needed in our proofs.

Definition 2.1 (Auxiliary random variables). *For $n \in \mathbb{N}$ and $t \in \mathbb{N}$, we define:*
(a) $Y_t^n \in [\ell_n]$ is the half-edge chosen by the random walk uniformly at random amongst the half-edges in $W_n(X_{t-1}^n)$.
(b) $\mathcal{E}_t^n \subset C_{t-1}^n$ is the set of k_n edges chosen uniformly at random amongst all subsets of size k_n of C_{t-1}^n by the DCM (Dynamic Configuration Model) Markov chain at time t.
(c) $\mathcal{S}_t^n = \cup_{e \in \mathcal{E}_t^n} e$ is the set of $2k_n$ half-edges involved in the rewiring in the DCM Markov chain at time t.
(d) For $u \in [n]$, $N_t^n(u)$ denotes the number of simple edges (non-loop) in $\mathcal{E}_t^n$ incident to u, and $L_t^n(u)$ denotes the number of loops in $\mathcal{E}_t^n$ incident to u.
(e) For $u, v \in [n]$ and $\eta \in \mathcal{C}onf_{[\ell_n]}$, $N_\eta(u, v)$ denotes the number of edges between u, v in η (in case $u = v$, each loop is counted once).
(f) $Z_{t,i}^n$ denotes the number of distinct edges traversed by the random walker between time $t - i + 1, \ldots, t$, i.e., in the last i steps of a t-step random walk.

Definition 2.2 (Randomized stopping time). *For $n \in \mathbb{N}$, let τ_n to be the first time the walker moves through a half-edge that was rewired before, that is*

$$\tau_n = \inf \left\{ t \in \mathbb{N} \colon Y_t^n \in \cup_{s=1}^t \mathcal{S}_s^n \right\}.$$

2.1 Stopping time decomposition

The main result in this section is the following proposition, in which we give upper and lower bounds on the total variation distance between the law of X_t^n and its stationary distribution π_n:

Proposition 2.3 (Upper and lower bounds on total variation). *For every $t \in \mathbb{N}_0$,*

$$\|\mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot) - \pi_n(\cdot)\|_{\mathrm{TV}} \leq \mathbb{P}_{u^n,\eta^n}^n(\tau_n > t) + O(k_n^{-1}), \tag{2.1}$$

$$\|\mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot) - \pi_n(\cdot)\|_{\mathrm{TV}} \geq \mathbb{P}_{u^n,\eta^n}^n(\tau_n > t) \sum_{v \notin B_t^{\eta^n}(u^n)} \pi_n(v) - O(k_n^{-1}). \tag{2.2}$$

We will study $\mathbb{P}_{u^n,\eta^n}^n(\tau_n > t)$ in Section 2.2, and $\sum_{v \notin B_t^{\eta^n}(u^n)} \pi_n(v)$ in Section 2.3. In the remainder of this section, we prove preliminary bounds leading up to Proposition 2.3. In Lemma 2.7, we investigate the probability mass function of the "stopped" random walk $\mathbb{P}_{u^n,\eta^n}^n(X_t^n = v \mid \tau_n \leq t)$, which will give rise to the $O(k_n^{-1})$ error term.

By the triangle inequality, for any $t \in \mathbb{N}$ we get the following bounds for the total variation distance of X_t^n to its stationary distribution π_n:

$$\|\mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot) - \pi_n(\cdot)\|_{\mathrm{TV}} \leq \|\mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot) - \mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot \mid \tau_n \leq t)\|_{\mathrm{TV}}$$
$$+ \|\mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot \mid \tau_n \leq t) - \pi_n(\cdot)\|_{\mathrm{TV}}, \tag{2.3}$$

$$\|\mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot) - \pi_n(\cdot)\|_{\mathrm{TV}} \geq \|\mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot) - \mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot \mid \tau_n \leq t)\|_{\mathrm{TV}}$$
$$- \|\mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot \mid \tau_n \leq t) - \pi_n(\cdot)\|_{\mathrm{TV}}. \tag{2.4}$$

Equations (2.3)–(2.4) form the starting point of our analysis. In Lemma 2.6 below, we show that the second terms in the right-hand side of both inequalities is $O(k_n^{-1})$. The next lemma gives an upper bound for the first term:

Lemma 2.4 (Stopped versus unstopped walk: upper bound). *Uniformly in $t \in \mathbb{N}$,*

$$\|\mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot) - \mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot \mid \tau_n \leq t)\|_{\mathrm{TV}} \leq \mathbb{P}_{u^n,\eta^n}^n(\tau_n > t). \tag{2.5}$$

Proof. We start by investigating $\mathbb{P}_{u^n,\eta^n}^n(X_t^n = v)$. For any $u^n, v \in [n], \eta^n \in \mathrm{Conf}_{[\ell_n]}$ and $t \in \mathbb{N}$, by conditioning on τ_n we can write $\mathbb{P}_{u^n,\eta^n}^n(X_t = v)$ as follows:

$$\mathbb{P}_{u^n,\eta^n}^n(X_t^n = v) = \mathbb{P}_{u^n,\eta^n}^n(X_t^n = v \mid \tau_n \leq t)\,\mathbb{P}_{u^n,\eta^n}^n(\tau_n \leq t)$$
$$+ \mathbb{P}_{u^n,\eta^n}^n(X_t^n = v \mid \tau_n > t)\,\mathbb{P}_{u^n,\eta^n}^n(\tau_n > t)$$
$$\geq \mathbb{P}_{u^n,\eta^n}^n(X_t^n = v \mid \tau_n \leq t)\,\mathbb{P}_{u^n,\eta^n}^n(\tau_n \leq t). \tag{2.6}$$

The lower bound in (2.6) gives us an upper bound for the first term in the right-hand side of (2.3) and (2.4) as follows:

$$\|\mathbb{P}_{u^n,\eta^n}^n(X_t \in \cdot) - \mathbb{P}_{u^n,\eta^n}^n(X_t^n \in \cdot \mid \tau_n \leq t)\|_{\mathrm{TV}}$$
$$= \sum_{v \in [n]} \left[\mathbb{P}_{u^n,\eta^n}^n(X_t^n = v \mid \tau_n \leq t) - \mathbb{P}_{u^n,\eta^n}^n(X_t^n = v) \right]^+$$
$$\leq \sum_{v \in [n]} \left[\mathbb{P}_{u^n,\eta^n}^n(X_t^n = v \mid \tau_n \leq t) - \mathbb{P}_{u^n,\eta^n}^n(X_t^n = v \mid \tau_n \leq t)\,\mathbb{P}_{u^n,\eta^n}^n(\tau_n \leq t) \right]^+$$
$$= \left(1 - \mathbb{P}_{u^n,\eta^n}^n(\tau_n \leq t)\right) \sum_{v \in [n]} \mathbb{P}_{u^n,\eta^n}^n(X_t^n = v \mid \tau_n \leq t) = \mathbb{P}_{u^n,\eta^n}^n(\tau_n > t). \tag{2.7}$$

This proves the claim. □

Lemma 2.5 (Stopped versus unstopped walk: lower bound). *Uniformly in $t \in \mathbb{N}$,*

$$\|\mathbb{P}^n_{u^n,\eta^n}(X^n_t \in \cdot) - \mathbb{P}^n_{u^n,\eta^n}(X^n_t \in \cdot \mid \tau_n \leq t)\|_{\mathrm{TV}}$$
$$\geq \mathbb{P}^n_{u^n,\eta^n}(\tau_n > t) \sum_{v \notin B^{\eta^n}_t(u^n)} \mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n \leq t). \tag{2.8}$$

Proof. Consider v such that $d_{G^n_0}(u^n, v) > t$, where d_G denotes the graph distance on G. Note that when $\tau_n > t$, the random walk makes all of its first t moves through the edges that are present in η^n and since $d_{G^n_0}(u^n, v) > t$, it cannot reach v by making t moves over the edges in η^n. Hence $\mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n > t) = 0$, and so for such v,

$$\mathbb{P}^n_{u^n,\eta^n}(X^n_t = v) = \mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n \leq t)\, \mathbb{P}^n_{u^n,\eta^n}(\tau_n \leq t). \tag{2.9}$$

Let $B^{\eta^n}_t(u) = \{v \in [n]: d_{G^n_0}(u, v) \leq t\}$. Then

$$\|\mathbb{P}^n_{u^n,\eta^n}(X_t \in \cdot) - \mathbb{P}^n_{u^n,\eta^n}(X^n_t \in \cdot \mid \tau_n \leq t)\|_{\mathrm{TV}} \tag{2.10}$$
$$= \sum_{v \in [n]} \left[\mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n \leq t) - \mathbb{P}^n_{u^n,\eta^n}(X^n_t = v) \right]^+$$
$$= \sum_{v \in B^{\eta^n}_t(u)} \left[\mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n \leq t) - \mathbb{P}^n_{u^n,\eta^n}(X^n_t = v) \right]^+$$
$$\quad + \sum_{v \notin B^{\eta^n}_t(u)} \left[\mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n \leq t) - \mathbb{P}^n_{u^n,\eta^n}(X^n_t = v) \right]^+$$
$$\geq \sum_{v \notin B^{\eta^n}_t(u)} \left[\mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n \leq t) - \mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n \leq t)\, \mathbb{P}^n_{u^n,\eta^n}(\tau_n \leq t) \right]^+$$
$$= \mathbb{P}^n_{u^n,\eta^n}(\tau_n > t) \sum_{v \notin B^{\eta^n}_t(u)} \mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n \leq t), \tag{2.11}$$

which proves the claim. $\qquad\square$

Lemma 2.6 (Total variation distance for stopped random walk). *For any $0 \leq t \leq t_n$ such that $t_n = O(\alpha_n^{-1})$,*

$$\|\mathbb{P}^n_{u^n,\eta^n}(X^n_t \in \cdot \mid \tau_n \leq t) - \pi_n(\cdot)\|_{\mathrm{TV}} = O(k_n^{-1}), \qquad n \to \infty. \tag{2.12}$$

In order to prove Lemma 2.6, we start by proving lower bounds on $\mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n \leq t)$:

Lemma 2.7 (Lower bound for stopped random walk). *Uniformly in $v \in [n]$, for $t \leq t_n$ such that $t_n = O(\alpha_n^{-1})$,*

$$\mathbb{P}^n_{u^n,\eta^n}(X^n_t = v \mid \tau_n \leq t) \geq \pi_n(v) \left[1 - O(k_n^{-1})\right], \qquad n \to \infty. \tag{2.13}$$

Proof. Write

$$\mathbb{P}^n_{u^n,\eta^n}\left(X^n_t = v \mid \tau_n \leq t\right)$$
$$= \sum_{s=1}^{t} \mathbb{P}^n_{u^n,\eta^n}\left(X^n_t = v \mid \tau_n = s\right) \mathbb{P}^n_{u^n,\eta^n}\left(\tau_n = s \mid \tau_n \leq t\right)$$
$$= \sum_{s=1}^{t} \mathbb{P}^n_{u^n,\eta^n}\left(\tau_n = s \mid \tau_n \leq t\right)$$
$$\quad \times \sum_{w \in [n]} \mathbb{P}^n_{u^n,\eta^n}\left(X^n_t = v \mid X^n_s = w, \tau_n = s\right) \mathbb{P}^n_{u^n,\eta^n}\left(X^n_s = w \mid \tau_n = s\right). \tag{2.14}$$

8

On the event $\{\tau_n = s\}$, define $\sigma_{n,s} = \sup\{0 \le t \le s \colon Y_s \in \mathcal{S}_t\}$. By conditioning again we have

$$
\begin{aligned}
&\mathbb{P}^n_{u^n,\eta^n}\big(X^n_s = w \mid \tau_n = s\big) \\
&= \sum_{x \in [n],\, \eta \in \mathcal{C}onf_{l_n}} \mathbb{P}^n_{u^n,\eta^n}\big(X^n_s = w \mid X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big) \\
&\qquad\qquad\qquad \times \mathbb{P}^n_{u^n,\eta^n}\big(X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta \mid \tau_n = s\big).
\end{aligned}
\tag{2.15}
$$

Conditional on $C^n_{\sigma_{s,n}-1} = \eta$, let $A = \{\{Y_s, p\} \in \eta \text{ for some } p \in W_n(w)\}$, i.e., A is the event that the half-edge incident to x chosen by the random walk at time s belonged to an edge between x and w in the configuration η. Then

$$
\begin{aligned}
&\mathbb{P}^n_{u^n,\eta^n}\big(X^n_s = w \mid X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big) \\
&= \mathbb{P}^n_{u^n,\eta^n}\big(X^n_s = w \mid A, X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big) \\
&\quad \times \mathbb{P}^n_{u^n,\eta^n}\big(A \mid X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big) \\
&\quad + \mathbb{P}^n_{u^n,\eta^n}\big(X^n_s = w \mid A^c, X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big) \\
&\quad \times \mathbb{P}^n_{u^n,\eta^n}\big(A^c \mid X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big).
\end{aligned}
\tag{2.16}
$$

Since there are $N_\eta(x, w)$ edges between x and w in η, and one of them is chosen uniformly at random among all the edges incident to x by the random walk, we have

$$
\begin{aligned}
\mathbb{P}^n_{u^n,\eta^n}\big(A \mid X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big) &= \frac{N_\eta(x, w)}{d_n(x)}, \\
\mathbb{P}^n_{u^n,\eta^n}\big(A^c \mid X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big) &= \frac{d_n(x) - N_\eta(x, w)}{d_n(x)}.
\end{aligned}
\tag{2.17}
$$

First consider $w \ne x$. Note that $N^n_{\sigma_{s,n}}(w)$ can take values between 1 and $d_n(w) - N_\eta(w, w)$, and $L^n_{\sigma_{s,n}}(w)$ can take values between 0 and $N_\eta(w, w)$. Letting $k = N_\eta(w, w)$ and $m'_n = m_n - Z_{s,s-\sigma_{s,n}+1}$, we obtain the following expression by conditioning on $N^n_{\sigma_{s,n}}(w)$ and $L^n_{\sigma_{s,n}}(w)$:

$$
\begin{aligned}
&\mathbb{P}^n_{u^n,\eta^n}\big(X^n_s = w \mid A, X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big) \\
&= \sum_{i=1}^{d_n(w)-2k} \sum_{j=0}^{k} \mathbb{P}^n_{u^n,\eta^n}\big(X^n_s = w \mid N^n_{\sigma_{s,n}}(w) = i, L^n_{\sigma_{s,n}}(w) = j, A, X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big) \\
&\qquad\qquad \times \mathbb{P}^n_{u^n,\eta^n}\big(N^n_{\sigma_{s,n}}(w) = i, L^n_{\sigma_{s,n}}(w) = j \mid A, X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big) \\
&= \sum_{i=1}^{d_n(w)-2k} \sum_{j=0}^{k} \frac{i+2j}{2k_n - 1} \frac{\binom{d_n(w)-2k-1}{i-1}\binom{k}{j}\binom{m'_n - d_n(w)+k}{k_n - i - j}}{\binom{m'_n - 1}{k_n - 1}} \\
&= \frac{(d_n(w)-1)\binom{m'_n-2}{k_n-2} + \binom{m'_n-1}{k_n-1}}{(2k_n-1)\binom{m'_n-1}{k_n-1}} = \frac{d_n(w)(k_n - 1)}{(2k_n-1)(m'_n - 1)} + \frac{m'_n - k_n}{(2k_n-1)(m'_n - 1)},
\end{aligned}
\tag{2.18}
$$

since

$$
\mathbb{P}^n_{u^n,\eta^n}\big(X^n_s = w \mid N^n_{\sigma_{s,n}}(w) = i, L^n_{\sigma_{s,n}}(w) = j, A, X^n_{s-1} = x, C^n_{\sigma_{s,n}-1} = \eta, \tau_n = s\big)
\tag{2.19}
$$
$$
= \frac{i+2j}{2k_n - 1}
$$

and

$$
\mathbb{P}^n_{u^n,\eta^n}\big(N^n_{\sigma_{s,n}}(w) = i, L^n_{\sigma_{s,n}}(w) = j \mid A, X^n_{s-1} = x, C^n_{\sigma_{s,n}-1}\eta, \tau_n = s\big)
\tag{2.20}
$$
$$
= \frac{\binom{d_n(w)-2k-1}{i-1}\binom{k}{j}\binom{m'_n - d_n(w)+k}{k_n - i - j}}{\binom{m'_n - 1}{k_n - 1}}.
$$

On the event $A^c \cap \{Y_0 \in \mathcal{S}_0\}$, $N_0^n(w)$ can take values between 0 and $d_n(w) - N_\eta(w, w)$ and $L_0^n(w)$ can take values between 0 and $N_\eta(w, w)$. Hence, as above, we get

$$
\begin{aligned}
\mathbb{P}^n_{u^n, \eta^n} & \left(X_s^n = w \mid A^c, X_{s-1}^n = x, C^n_{\sigma_s, n-1} = \eta, \tau_n = s \right) \\
&= \sum_{i=0}^{d_n(w)-2k} \sum_{j=0}^{k} \frac{i+2j}{2k_n - 1} \frac{\binom{d_n(w)-2k}{i} \binom{k}{j} \binom{m'_n - d_n(w) + k - 1}{k_n - i - j - 1}}{\binom{m'_n - 1}{k_n - 1}} \\
&= \frac{d_n(w) \binom{m'_n - 2}{k_n - 2}}{(2k_n - 1)\binom{m'_n - 1}{k_n - 1}} = \frac{d_n(w)(k_n - 1)}{(2k_n - 1)(m'_n - 1)}.
\end{aligned} \tag{2.21}
$$

Combining (2.16)–(2.21), we get for $w \neq x$,

$$
\begin{aligned}
\mathbb{P}^n_{u^n, \eta^n} & (X_s^n = w \mid X_{s-1}^n = x, C^n_{\sigma_s, n-1} = \eta, \tau_n = s) \\
&= \frac{d_n(w)(k_n - 1)}{(2k_n - 1)(m'_n - 1)} + \frac{N_\eta(x, w)}{d_n(x)} \frac{m'_n - k_n}{(2k_n - 1)(m'_n - 1)}.
\end{aligned} \tag{2.22}
$$

Doing the same calculations for $x = w$, we get

$$
\begin{aligned}
\mathbb{P}^n_{u^n, \eta^n} & \left(X_s^n = w \mid X_{s-1}^n = w, C^n_{\sigma_s, n-1} = \eta, \tau_n = s \right) \\
&= \frac{(d_n(w) - 1)(k_n - 1)}{(2k_n - 1)(m'_n - 1)} + \frac{2N_\eta(w, w)}{d_n(w)} \frac{m'_n - k_n}{(2k_n - 1)(m'_n - 1)}.
\end{aligned} \tag{2.23}
$$

Recalling that $\pi_n(w) = d_n(w)/\ell_n$ and $m_n \geq m'_n$, we get for any $w, x \in [n]$,

$$
\mathbb{P}^n_{u^n, \eta^n} \left(X_s^n = w \mid X_{s-1}^n = x, C^n_{\sigma_s, n-1} = \eta, \tau_n = s \right) \geq \pi_n(w) \left[1 - O(k_n^{-1}) \right], \tag{2.24}
$$

where the error term $O(k_n^{-1})$ is uniform in $w, x \in [n]$ and $\eta \in \mathcal{C}onf_{[\ell_n]}$. Since this result holds for any $x \in [n]$ and $\eta \in \mathcal{C}onf_{l_n}$, it follows from (2.15) that

$$
\begin{aligned}
\mathbb{P}^n_{u^n, \eta^n} \left(X_s^n = w \mid \tau_n = s \right) &\geq \sum_{x \in [n], \eta \in \mathcal{C}onf_{l_n}} \pi_n(w) \left[1 - O(k_n^{-1}) \right] \\
&\qquad \times \mathbb{P}^n_{u^n, \eta^n} (X_{s-1}^n = x, C^n_{\sigma_s, n-1} = \eta \mid \tau_n = s) \\
&= \pi_n(w) \left[1 - O(k_n^{-1}) \right].
\end{aligned} \tag{2.25}
$$

Inserting this into (2.14) and using the stationarity of π_n, we obtain (2.13). $\qquad \square$

Proof of Lemma 2.6. By (2.13),

$$
\begin{aligned}
\| \mathbb{P}^n_{u^n, \eta^n}(X_t^n \in \cdot \mid \tau_n \leq t) - \pi_n(\cdot) \|_{\text{TV}} &= \sum_{v \in [n]} \left[\pi_n(v) - \mathbb{P}^n_{u^n, \eta^n}(X_t^n = v \mid \tau_n \leq t) \right]^+ \\
&\leq \sum_{v \in [n]} \left[\pi_n(v) - (1 - O(k_n^{-1}))\pi_n(v) \right]^+ = O(k_n^{-1}) \sum_{v \in [n]} \pi_n(v) = O(k_n^{-1}).
\end{aligned} \tag{2.26}
$$

$\qquad \square$

Proof of Proposition 2.3. This is immediate from (2.3)–(2.4) combined with Lemmas 2.4, 2.5 and 2.6. $\qquad \square$

2.2 Tail probabilities for τ_n

In this section, we give an explicit formula for the tail probability $\mathbb{P}^n_{u^n,\eta^n}(\tau_n > t_n)$ in terms of the range of the random walk on the configuration model. Let $Z^n_{t,i}$ denote the number of distinct edges in the last i steps of the simple random walk $(\hat{X}^n_s)^t_{s=0}$ on the initial graph G^n_0. In terms of this quantity, we can identify $\mathbb{P}^n_{u^n,\eta^n}(\tau_n > t_n)$ as follows:

Lemma 2.8 (Tail probability of τ_n in terms of range). *For all* $t_n = o(\alpha^{-1}_n)$,

$$\mathbb{P}^n_{u^n,\eta^n}(\tau_n > t_n) = \widehat{\mathbb{E}}^n\left[(1-\alpha_n)^{\sum^{t_n}_{i=1} Z^n_{t_n,i}}\right] + o(1), \tag{2.27}$$

where $\hat{\mathbb{E}}$ *denotes expectation with respect to the law of simple random walk on* G^n_0.

Proof. On the event $\{\tau_n > t_n\}$, the first t_n steps of the random walk are confined to edges in η^n and so, conditioning on the sample path of the random walk, we get

$$\mathbb{P}^n_{u^n,\eta^n}(\tau_n > t_n) \tag{2.28}$$

$$= \sum_{e_1,\ldots,e_{t_n}\in\overleftrightarrow{\eta^n}} \mathbb{P}^n_{u^n,\eta^n}\left(\tau_n > t_n, Y_0 = e^-_1, Y_1 = e^+_1, \ldots, Y_{t_n-1} = e^+_{t_n-1}, Y_{t_n} = e^+_{t_n}\right)$$

$$= \sum_{e_1,\ldots,e_{t_n}\in\overleftrightarrow{\eta^n}} \mathbb{P}^n_{u^n,\eta^n}\left(e'_i \notin \cup^i_{j=1}\mathcal{E}_j, 1 \le i \le t_n, Y_0 = e^-_1, Y_1 = e^+_1, \ldots, Y_{t_n-1} = e^+_{t_n-1}, Y_{t_n} = e^+_{t_n}\right)$$

$$= \sum_{e_1,\ldots,e_{t_n}\in\overleftrightarrow{\eta^n}} \mathbb{P}^n_{u^n,\eta^n}\left(Y_0 = e^-_1, Y_1 = e^+_1, \ldots, Y_{t_n-1} = e^+_{t_n-1}, Y_{t_n} = e^+_{t_n} | e'_i \notin \cup^i_{j=1}\mathcal{E}_j, 1 \le i \le t_n\right)$$

$$\times \mathbb{P}^n_{u^n,\eta^n}\left(e'_i \notin \cup^i_{j=1}\mathcal{E}_j, 1 \le i \le t_n\right),$$

where, for a given configuration η, $\overleftrightarrow{\eta}$ denotes the set of oriented edges in which every edge in η occurs in both directions and, for $e \in \overleftrightarrow{\eta}$, e^-, e^+ denotes the head and the tail of the oriented edge, respectively, and e' denotes the unoriented version of e.

The event $\{e'_i \notin \cup^i_{j=1}\mathcal{E}_j, 1 \le i \le t_n\}$ can be rewritten as $\{e'_j \notin \mathcal{E}_i, 1 \le i \le j \le t_n\}$. Define $z_{t_n,i} = z_{t_n,i}(e'_1,\ldots,e'_{t_n})$ to be the number of distinct edges amongst $e'_{t_n-i+1},\ldots,e'_{t_n}$, i.e., the last i edges of the path. Then, for $1 \le i < t_n$,

$$\mathbb{P}^n_{u^n,\eta^n}\left(e'_j \notin \mathcal{E}_{i+1}, i < j \le t_n \mid e'_k \notin \mathcal{E}_l, 1 \le l \le i, l \le k \le t_n\right) = \frac{\binom{m_n - z_{t_n,t_n-i}}{k_n}}{\binom{m_n}{k_n}}, \tag{2.29}$$

which gives, after rearrangement of terms,

$$\mathbb{P}^n_{u^n,\eta^n}\left(e_i \notin \cup^i_{j=1}\mathcal{E}_j, 1 \le i \le t_n\right) \tag{2.30}$$

$$= \prod^{t_n}_{i=1} \frac{\binom{m_n - z_{t_n,t_n-i+1}}{k_n}}{\binom{m_n}{k_n}} = \prod^{t_n}_{i=1} \frac{\binom{m_n - z_{t_n,i}}{k_n}}{\binom{m_n}{k_n}}$$

$$= \prod^{t_n}_{i=1} \prod^{z_{t_n,i}-1}_{j=0} \left(1 - \frac{k_n}{m_n - j}\right) = \prod^{t_n}_{i=1} \prod^{z_{t_n,i}-1}_{j=0} \left(1 - \frac{k_n}{m_n} + O\left(\frac{k_n j}{m^2_n}\right)\right).$$

Note that $1 \le z_{t_n,i} \le i$ for all $i \le t_n$. Since, by assumption, $t_n = o(\alpha^{-1}_n)$, we obtain

$$\mathbb{P}^n_{u^n,\eta^n}\left(e'_i \notin \cup^i_{j=1}\mathcal{E}_j, 1 \le i \le t_n\right) = \prod^{t_n}_{i=1}\left[(1-\alpha_n)^{z_{t_n,i}} + o(k^{-1}_n)\right]. \tag{2.31}$$

We note that $\alpha^{-1}_n = o(k_n)$ because $\alpha^{-1}_n = o((\log n)^2)$, and so $t_n = o(k_n)$ and hence

$$\mathbb{P}^n_{u^n,\eta^n}\left(e'_i \notin \cup^i_{j=1}\mathcal{E}_j, 1 \le i \le t_n\right) = \prod^{t_n}_{i=1}(1-\alpha_n)^{z_{t_n,i}} + o(1). \tag{2.32}$$

11

On the other hand, note that

$$\mathbb{P}^n_{u^n,\eta^n}\left(Y_0 = e_1^-, Y_1 = e_1^+, \ldots, Y_{t_n-1} = e_{t_n-1}^+, Y_{t_n} = e_{t_n}^+ \mid e_i' \notin \cup_{j=1}^i \mathcal{E}_j, 1 \le i \le t_n\right) \quad (2.33)$$
$$= \frac{1}{d_n(e_1^-)} \frac{1}{d_n(e_2^-)} \times \cdots \times \frac{1}{d_n(e_{t_n}^-)},$$

so that (2.28) can be rewritten as

$$\mathbb{P}^n_{u^n,\eta^n}(\tau_n > t_n) = \sum_{e_1,\ldots,e_{t_n}\in\eta^n} \prod_{i=1}^{t_n} \frac{(1-\alpha_n)^{z_{t_n,i}}}{d_n(e_i^-)} + o(1). \quad (2.34)$$

Now let $\widehat{X}^n = (\widehat{X}_t^n)_{t\ge0}$ be simple random walk on $G_0^n = I_{d_n}(\eta^n)$ starting from the vertex u^n, $Z_{t,i}^n$ the number of distinct edges traversed by this random walk at times $t-i+1,\ldots,t$, and $\widehat{\mathbb{E}}^n$ expectation under the law of this random walk. With this notation in hand, the tail probability we are interested in can be written as

$$\mathbb{P}^n_{u^n,\eta^n}(\tau_n > t_n) = \widehat{\mathbb{E}}^n\left[(1-\alpha_n)^{\sum_{i=1}^{t_n} Z_{t_n,i}^n}\right] + o(1), \quad (2.35)$$

which completes the proof of Lemma 2.8. $\qquad\square$

2.3 Branching process approximations of configuration model intrinsic balls

In this section, we establish branching process approximations to the intrinsic balls in the configuration model. We start by showing that $\lim_{n\to\infty} \pi_n(B_t^{\eta^n}(u^n)) = 0$ for μ_n with high probability. Define

$$\nu_n = \frac{\mathbb{E}[D_n(D_n-1)]}{\mathbb{E}[D_n]}. \quad (2.36)$$

Lemma 2.9 (Bound on intrinsic ball). *Subject to Condition 1.1, for any $\varepsilon > 0$ and all $t_n = O(\log n)$, there exists a constant $C > 0$ such that*

$$\mu_n\left(\pi_n(B_{t_n}^{\eta^n}(u^n)) > \varepsilon\right) \le C\frac{\nu_n^{t_n}}{n\varepsilon}. \quad (2.37)$$

In particular,

$$\mu_n\left(\sum_{v\in B_{t_n}^{\eta^n}(u^n)} d_n(v) > n\varepsilon\,\mathbb{E}(D_n)\right) \le C\frac{\nu_n^{t_n}}{n\varepsilon}. \quad (2.38)$$

Proof. Letting $\mathbb{P}^n_{\mathrm{CM}}$ to denote the law of configuration model $\mathrm{CM}_n(d_n)$ and $\mathbb{E}^n_{\mathrm{CM}}$ associated expectation, we note that

$$\mathbb{E}_{\mu_n}\left[\pi_n(B_{t_n}^{\eta^n}(u))\right] = \mathbb{E}_{\mu_n}\left[\sum_{v\in B_{t_n}^{\eta^n}(u)} \frac{d_n(v)}{\ell_n}\right] = \frac{1}{n}\sum_{u\in[n]} \mathbb{E}^n_{\mathrm{CM}}\left[\sum_{v\in B_{t_n}^{\eta^n}(u)} \frac{d_n(v)}{\ell_n}\right]$$
$$= \frac{1}{n}\sum_{u,v\in[n]} \frac{d_n(v)}{\ell_n}\,\mathbb{P}^n_{\mathrm{CM}}(d_{\mathrm{CM}}^n(u,v) \le t_n), \quad (2.39)$$

where the first expectation is with respect to the measure μ_n and d_{CM}^n denotes the graph distance in $\mathrm{CM}_n(d_n)$. Note that $v \in B_{t_n}^{\eta^n}(u)$ precisely when there exists a path in the configuration model containing at most t_n edges. Let $m \in [t_n]$, and let $(\gamma_i)_{i=0}^m$ denote a path in $[n]$. Thus, $\gamma_0 = u, \gamma_m = v$ and $\gamma_i \in [n]$ for $i \in [t_n-1]$, and all γ_i are distinct. Then

$$\mathbb{P}^n_{\mathrm{CM}}(d_{\mathrm{CM}}^n(u,v) \le t_n) \le \sum_{m=1}^{t_n} \sum_{\gamma_1,\ldots,\gamma_{m-1}\in[n]} \mathbb{P}^n_{\mathrm{CM}}\left((\gamma_{i-1},\gamma_i) \in E(\mathrm{CM}_n(d_n)) \,\forall\, i \in [t_n-1]\right). \quad (2.40)$$

Now,
$$\mathbb{P}^n_{\mathrm{CM}}\Big((\gamma_{i-1},\gamma_i)\in E(\mathrm{CM}_n(d_n))\ \forall\, i\in[t_n-1]\Big)$$
$$\le d_n(\gamma_0)d_n(\gamma_m)\prod_{i=1}^{m-1}d_n(\gamma_i)(d_n(\gamma_i)-1)\prod_{i=1}^{m}\frac{1}{\ell_n-2i+1}, \tag{2.41}$$

since the product involving d_n overcounts the sequences of half-edges that can form the edges required, and $\prod_{i=1}^{m}\frac{1}{\ell_n-2i+1}$ is the probability that any sequence of such half-edges are paired. We obtain, by Markov's inequality,

$$\mu_n\Big(\pi_n(B_{t_n}^{\eta^n}(u))>\varepsilon\Big) \tag{2.42}$$
$$\le \frac{1}{n\varepsilon}\sum_{u,v\in[n]}\frac{d_n(u)d_n(v)^2}{\ell_n}\sum_{m=1}^{t_n}\Big(\sum_{w\in[n]}d_n(w)(d_n(w)-1)\Big)^{m-1}\prod_{i=1}^{m}\frac{1}{\ell_n-2i+1}$$
$$= \frac{1}{\ell_n n\varepsilon}\sum_{v\in[n]}d_n(v)^2\sum_{m=1}^{t_n}\nu_n^{m-1}\prod_{i=1}^{m}\frac{\ell_n}{\ell_n-2i+1}.$$

By (R1) and (R2) in Condition 1.1,

$$\frac{\ell_n}{n}=\mathbb{E}_n[D_n]\to\mathbb{E}[D]<\infty,\qquad \frac{1}{n}\sum_{v\in[n]}d_n(v)^2\to\mathbb{E}[D^2]<\infty. \tag{2.43}$$

Hence $\sup_{n\in\mathbb{N}}\mathbb{E}_n[D_n^2]/\mathbb{E}_n[D_n]=C<\infty$ and, since $t_n=O(\log n)$,

$$\prod_{i=1}^{m}\frac{\ell_n}{\ell_n-2i+1}=1+O(t_n^2/\ell_n)=1+o(1). \tag{2.44}$$

We conclude that $\mu_n(\pi_n(B_{t_n}^{\eta^n}(u))>\varepsilon)\le C\nu_n^{t_n}/(n\varepsilon)$, as required. $\qquad\square$

We next relate the neighborhood $B_{t_n}^{\eta^n}(u^n)$ to a branching process where the root has offspring distribution D_n, while all other individuals have offspring distribution $D_n^\star-1$, with

$$\mathbb{P}(D_n^\star=k)=\frac{k}{\mathbb{E}[D_n]}\mathbb{P}(D_n=k),\qquad k\in\mathbb{N} \tag{2.45}$$

is the size-biased distribution of D_n. Denote this branching process by $(\mathsf{BP}_n(t))_{t\in\mathbb{N}_0}$. Here, $\mathsf{BP}_n(t)$ denotes the branching process when it contains precisely t vertices, and we explore it in the breadth-first order. We let $(\mathsf{G}_n(t))_{t\in\mathbb{N}_0}$ denote the same quantity for the graph exploration. In particular, from $(\mathsf{G}_n(t))_{t\in\mathbb{N}_0}$ we can retrieve $(B_t^{\eta^n}(u))_{t\in\mathbb{N}_0}$, where $D_n=d_n(u)$. The following lemma proves that we can couple the graph exploration to the branching process in such a way that $(\mathsf{G}_n(t))_{0\le t\le T_n}$ is equal to $(\mathsf{BP}_n(t))_{0\le t\le T_n}$ when $T_n=o(\sqrt{n})$. In the statement, we write $(\widehat{\mathsf{G}}_n(t),\widehat{\mathsf{BP}}_n(t))_{t\in\mathbb{N}_0}$ for the coupling of $(\mathsf{G}_n(t))_{0\le t\le T_n}$ and $(\mathsf{BP}_n(t))_{0\le t\le T_n}$.

Lemma 2.10 (Coupling graph exploration and branching process). *Subject to Condition 1.1, there exists a coupling $(\widehat{\mathsf{G}}_n(t),\widehat{\mathsf{BP}}_n(t))_{t\in\mathbb{N}_0}$ of $(\mathsf{G}_n(t))_{0\le t\le T_n}$ and $(\mathsf{BP}_n(t))_{0\le t\le T_n}$ such that*

$$\mathbb{P}\Big((\widehat{\mathsf{G}}_n(t))_{0\le t\le T_n}\ne(\widehat{\mathsf{BP}}_n(t))_{0\le t\le T_n}\Big)=o(1), \tag{2.46}$$

whenever $T_n=o(\sqrt{n})$.

Proof. We let the offspring of the root of the branching process $\widehat{D}_n$ be equal to $d_n(\hat{u})$, which is the number of neighbours of the vertex $\hat{u}\in[n]$ that is chosen uniformly at random. By

construction, $\widehat{D}_n = d_n(\hat{u})$, so that also $\widehat{\mathsf{G}}_n(1) = \widehat{\mathsf{BP}}_n(1)$. We next explain how to *jointly* construct $(\widehat{\mathsf{G}}_n(t), \widehat{\mathsf{BP}}_n(t))_{0 \le t \le T}$ *given* that we have already constructed $(\widehat{\mathsf{G}}_n(t), \widehat{\mathsf{BP}}_n(t))_{0 \le t \le T-1}$.

To obtain $(\widehat{\mathsf{G}}_n(t)_{0 \le t \le T}$, we take the first unpaired half-edge p_T. This half-edge needs to be paired to a uniform half-edge that has not been paired so far. We draw a uniform half-edge q_T from the collection of *all* half-edges, independently of the past, and we let the $(T-1)$-st individual in $(\widehat{\mathsf{BP}}_n(t))_{0 \le t \le T-1}$ have precisely $d_n(U_T) - 1$ children. Note that $d_n(U_T) - 1$ has the same distribution as $D_n^\star - 1$ and, by construction, the collection $\big(d_n(U_t) - 1\big)_{t \in \mathbb{N}}$ is i.i.d. When q_T is still free, i.e., has not yet been paired in $(\widehat{\mathsf{G}}_n(t))_{0 \le t \le T-1}$, then we also let p_T be paired to q_T, and we have constructed $(\widehat{\mathsf{G}}_n(t))_{0 \le t \le T}$. However, a problem arises when q_T has already been paired in $(\widehat{\mathsf{G}}_n(t))_{0 \le t \le T-1}$, in which case we draw a uniform *unpaired* half-edge q'_T and pair p_T to q'_T instead. Clearly, this might give rise to a difference between $(\widehat{\mathsf{G}}_n(t))_{t \le T}$ and $(\widehat{\mathsf{BP}}_n(t))_{0 \le t \le T}$. We now provide bounds on the probability that an error occurs before time T_n.

There are two sources of differences between $(\widehat{\mathsf{G}}_n(t))_{t \in \mathbb{N}_0 0}$ and $(\widehat{\mathsf{BP}}_n(t))_{t \in \mathbb{N}_0}$:

Half-edge re-use. In the above coupling q_T had already been paired and is being re-used in the branching process, and we need to redraw q'_T;

Vertex re-use. In the above coupling, this means that q_T is a half-edge that has not yet been paired in $(\widehat{\mathsf{G}}_n(t))_{0 \le t \le T-1}$, but it is incident to a half-edge that has already been paired in $(\widehat{\mathsf{G}}_n(t))_{0 \le t \le T-1}$. In particular, the vertex to which it is incident has already appeared in $(\widehat{\mathsf{G}}_n(t))_{0 \le t \le T-1}$ and it is being re-used in the branching process. In this case, a *copy* of the vertex appears in $(\widehat{\mathsf{BP}}_n(t))_{0 \le t \le T}$, while a *cycle* appears in $(\widehat{\mathsf{G}}_n(t))_{0 \le t \le T}$.

We now provide a bound on both contributions:

Half-edge re-use. Up to time $T - 1$, at most $2T - 1$ half-edges are forbidden to be used by $(\widehat{\mathsf{G}}_n(t)_{t \le T})$. The probability that the half-edge q_T equals one of these two half-edges is at most

$$\frac{2T - 1}{\ell_n}. \tag{2.47}$$

Hence the probability that a half-edge is being re-used before time T_n is at most

$$\sum_{T=1}^{T_n} \frac{2T - 1}{\ell_n} = \frac{T_n^2}{\ell_n} = o(1), \tag{2.48}$$

since $T_n = o(\sqrt{n})$.

Vertex re-use. The probability that vertex i is chosen in the T-th draw is equal to d_i/ℓ_n. The probability that vertex i is drawn twice before time T_n is at most

$$\frac{T_n(T_n - 1)}{2} \frac{d_i^2}{\ell_n^2}. \tag{2.49}$$

By the union bound, the probability that there exists a vertex that is chosen twice up to time T_n is at most

$$\frac{T_n(T_n - 1)}{2\ell_n} \sum_{i \in [n]} \frac{d_i^2}{\ell_n} = o(1), \tag{2.50}$$

by Condition 1.1 because $T_n = o(\sqrt{n})$. $\qquad \square$

3   Proof of main theorems

Proof of lower bound in Theorem 1.5.   We use (2.2) in Proposition 2.3 for $t = t_n$ and write

$$\sum_{v \notin B_{t_n}^{\eta^n}(u^n)} \pi_n(v) = 1 - \sum_{v \in B_{t_n}^{\eta^n}(u^n)} \pi_n(v). \tag{3.1}$$

This leads to

$$\|\mathbb{P}_{u^n,\eta^n}^n(X_{t_n}^n \in \cdot) - \pi_n(\cdot)\|_{\mathrm{TV}} \geq \mathbb{P}_{u^n,\eta^n}^n(\tau_n > t_n) - \sum_{v \in B_{t_n}^{\eta^n}(u^n)} \pi_n(v) - O(k_n^{-1}). \tag{3.2}$$

Lemma 2.9 together with $t_n = O(1/\sqrt{\alpha_n}) = o(\log n)$ shows that, for any $0 < \delta < 1$,

$$\sum_{v \in B_{t_n}^{\eta^n}(u^n)} \pi_n(v) \leq n^{-\delta} \tag{3.3}$$

with μ_n-probability at least $1 - C' n^{\delta-1+o(1)}$ for some constant $C' < \infty$. It remains to prove a lower bound on $\mathbb{P}_{u^n,\eta^n}^n(\tau_n > t_n)$. We use Lemma 2.8 and prove an upper bound on the random walk range functional appearing in the exponent. A trivial upper bound for $Z_{t_n,i}^n$ is i, so that

$$(1-\alpha_n)^{\sum_{i=1}^{t_n} Z_{t_n,i}^n} \geq (1-\alpha_n)^{(t_n+1)^2/2}. \tag{3.4}$$

Since all other terms in the lower bound are $o(1)$ independently of the constant c in $t_n = c/\sqrt{\alpha_n}$, we obtain that $t_{\mathrm{mix}}^n(\varepsilon; u^n, \eta^n) \geq t_n$, where t_n is such that $(1-\alpha_n)^{(t_n+1)^2/2} \leq \varepsilon$. This gives that $\frac{1}{2}(t_n+1)^2 \geq (1/\alpha_n)\log(1/\varepsilon)$, as claimed.

Proof of upper bound in Theorem 1.5.   We now use (2.1) in Proposition 2.3 for $t = t_n$, to see that it suffices to analyse $\mathbb{P}_{u^n,\eta^n}^n(\tau_n > t_n)$. We again use Lemma 2.8, to see that

$$\|\mathbb{P}_{u^n,\eta^n}^n(X_{t_n}^n \in \cdot) - \pi_n(\cdot)\|_{\mathrm{TV}} \leq \widehat{\mathbb{E}}^n\left[(1-\alpha_n)^{\sum_{i=1}^{t_n} Z_{t_n,i}^n}\right] + o(1). \tag{3.5}$$

Note that $Z_{t_n,i}$ is monotone increasing in i, so that

$$\sum_{i=1}^{t_n} Z_{t_n,i}^n \geq \sum_{i=\lfloor t_n/2\rfloor}^{t_n} Z_{t_n,i}^n \geq \tfrac{1}{2} t_n Z_{t_n,\lfloor t_n/2\rfloor}^n. \tag{3.6}$$

We can bound this last term by

$$Z_{t_n,\lfloor t_n/2\rfloor}^n \geq d_{\mathrm{CM}}(\widehat{X}_{t_n}^n, \widehat{X}_{\lfloor t_n/2\rfloor}^n). \tag{3.7}$$

Next, we use the comparison with the branching processes in Lemma 2.10. We note that, with $S_n(t)$ denoting the total number of individuals in generations 1 up to t,

$$\mathbb{E}_{\mathsf{BP}_n}[S_n(t_n)] = \frac{\ell_n}{n}\nu_n^{t_n} = n^{o(1)}, \tag{3.8}$$

since $t_n = o(\log n)$. Therefore, with high probability, $S_n(t_n) = o(\sqrt{n})$, and hence the family tree of the branching process up to generation t_n and $B_{t_n}^{\eta^n}(u)$ agree. With probability 1, $(\widehat{X}_t^n)_{0 \leq t \leq t_n}$ does not leave the ball of radius t_n, so the random walk agrees with a random walk on a tree.

Next, we start the random walk from $\widehat{X}_{\lfloor t_n/2\rfloor}^n$ and run it up to time t_n. $B_{t_n}^{\eta^n}(u)$ is a tree and by (R3) in Condition 1.1, wherever the random walk is, at least 2 of the neighbouring vertices

15

are further away from $\widehat{X}^n_{\lfloor t_n/2 \rfloor}$ than the present location, and only one is closer by. Thus, we can stochastically bound $d^n_{\mathrm{CM}}(\widehat{X}^n_{t_n}, \widehat{X}^n_{\lfloor t_n/2 \rfloor})$ from below by $2Y_{\lceil t_n/2 \rceil} - \lceil t_n/2 \rceil$, where Y_t has a binomial distribution with t trials and success probability $\frac{2}{3}$. Thus, with high probability, and for every $\zeta > 0$,

$$d_{\mathrm{CM}}(\widehat{X}^n_{t_n}, \widehat{X}^n_{\lfloor t_n/2 \rfloor}) \geq (\tfrac{1}{3} - \zeta)(t_n/2). \tag{3.9}$$

We conclude that, with high probability,

$$Z^n_{t_n, \lfloor t_n/2 \rfloor} \geq (\tfrac{1}{3} - \zeta)(t_n/2)^2, \tag{3.10}$$

so that

$$\|\mathbb{P}^n_{u^n, \eta^n}(X^n_{t_n} \in \cdot) - \pi_n(\cdot)\|_{\mathrm{TV}} \leq \widehat{\mathbb{E}}^n\left[(1 - \alpha_n)^{\sum_{i=1}^{t_n} Z^n_{t_n, i}}\right] + o(1) \tag{3.11}$$

$$\leq (1 - \alpha_n)^{(\frac{1}{3} - \zeta)(t_n/2)^2} + o(1) \leq \varepsilon,$$

when $(t_n/2)^2(\tfrac{1}{3} - \zeta) \geq (1/\alpha_n)\log(1/\varepsilon)$. This proves the upper bound for any constant exceeding $2\sqrt{3}/\sqrt{1 - 3\zeta}$. Since $\zeta > 0$ is arbitrary, we get the claim. $\qquad\square$

Proof of Theorem 1.6.   We rely on the range property of simple random walk on a Galton-Watson tree derived in the appendix. Let $\mathbb{P}^{\mathcal{GW}}_n$ be the law of the Galton-Watson tree ω whose offspring distribution is the size-biased version of the law of D_n (recall (1.2) and (1.7)). By Lemma A.1, we have that, for any $\delta > 0$ with $\mathbb{P}^{\mathcal{GW}}_n$-probability at least $1 - O(t_n^{-1/2})$,

$$\mathbb{E}_\omega\left[(1 - \alpha_n)^{\sum_{i=1}^{t_n} Z_{t_n, i}}\right] \geq (1 - \alpha_n)^{(1+\delta)a_n t_n^2/2} + O(t_n^{-1/2}), \tag{3.12}$$

and

$$\mathbb{E}_\omega\left[(1 - \alpha_n)^{\sum_{i=1}^{t_n} Z_{t_n, i}}\right] \leq (1 - \alpha_n)^{(1-\delta)a_n t_n^2/2} + O(t_n^{-1/2}), \tag{3.13}$$

where a_n is the average escape probability from the root by the simple random walk.

Recall that in Lemma 2.10 we coupled the neighbourhood $B^{\eta^n}_{t_n}(u^n)$ to the first t_n levels of the family tree of a branching process in which the root had offspring distribution D_n while the other vertices had offspring distribution $D_n^* - 1$, one less than the size biased version of D_n. In order to make use of the bounds given in (3.12),(3.13), we argue that they hold also for the distribution of the family tree of that branching process. Indeed, since D_n is stochastically dominated by D_n^* [16, Chapter 2], we can couple the family tree of the branching process, say ω with root ρ, and the Galton-Watson tree with offspring distribution D_n^* at the root and $D_n^* - 1$ everywhere else, say ω' with root ρ', in such a way that ω is a rooted subtree of ω' and ρ coincides with ρ' almost surely. Now, since the simple random walk on ω' is transient a.s., it will be confined to ω with a positive probability and any statement that holds for the random walk on ω' will hold with high probability for the random walk on ω' conditioned on not leaving ω. Following Remark A.2, we see that the bounds hold also for the random walk on the family tree of the branching process used in Lemma 2.10.

Upon successful coupling of the simple random walk on G^n_0 starting from a uniformly chosen vertex and the simple random walk on the family tree of the branching process starting from the root, which occurs with high probability according to Lemma 2.10, we have

$$\widehat{\mathbb{E}}^n\left[(1 - \alpha_n)^{\sum_{i=1}^{t_n} Z_{t_n, i}}\right] = \mathbb{E}_\omega\left[(1 - \alpha_n)^{\sum_{i=1}^{t_n} Z_{t_n, i}}\right] \tag{3.14}$$

and hence, by Lemma 2.8 with high probability,

$$\mathbb{P}^n_{u^n, \eta^n}(\tau_n > t_n) = \mathbb{E}_\omega\left[(1 - \alpha_n)^{\sum_{i=1}^{t_n} Z_{t_n, i}}\right] + o(1). \tag{3.15}$$

Combining (3.15) with (3.12) and (3.13), we find that for any $\delta > 0$, with high probability,

$$(1 - \alpha_n)^{(1+\delta)a_n t_n^2/2} + o(1) \leq \mathbb{P}^n_{u^n,\eta^n}(\tau_n > t_n) \leq (1 - \alpha_n)^{(1-\delta)a_n t_n^2/2} + o(1). \tag{3.16}$$

Finally, fix $\delta > 0$ and define

$$t_n(\pm)^2 = (1 \pm \delta)\frac{2}{a_n \alpha_n} \log(1/\varepsilon). \tag{3.17}$$

As shown in Lemma A.3, $\lim_{n\to\infty} a_n = a$ with a the escape probability defined in (1.9). Combining Lemma 2.4 and the upper bound in (3.16), we have with high probability,

$$\|\mathbb{P}^n_{u^n,\eta^n}(X^n_{t_n(+)} \in \cdot) - \pi_n(\cdot)\|_{\mathrm{TV}} \leq \varepsilon. \tag{3.18}$$

Combining Lemma 2.5, Lemma 2.9 and the lower bound in (3.16), we have with high probability,

$$\|\mathbb{P}^n_{u^n,\eta^n}(X^n_{t_n(-)} \in \cdot) - \pi_n(\cdot)\|_{\mathrm{TV}} \geq \varepsilon. \tag{3.19}$$

Letting $\delta \downarrow 0$, we obtain the desired result. $\qquad\square$

A   Simple random walk on a Galton–Watson tree

A.1   Range

Let Ω denote the set of rooted locally finite trees. Let $\mathbb{P}^{\mathcal{GW}}$ denote the law on Ω induced by the Galton-Watson tree with a given offspring distribution with support $\mathbb{N} \setminus \{1\}$. For a given tree $\omega \in \Omega$, let $\mathbb{P}_\omega$ denote the law of the simple random walk on ω ("quenched law"), and define

$$\mathbb{P}_{\mathcal{GW}}(\cdot) = \int \mathbb{P}_\omega(\cdot)\,\mathbb{P}^{\mathcal{GW}}(d\omega) \tag{A.1}$$

("annealed law"). Write R_t to denote the number of distinct vertices visited by the random walk up to time t. Piau [28] shows that there exists a constant a (depending on the offspring distribution of the Galton-Watson tree) and a constant c (not depending on the offspring distribution), such that

$$\mathbb{E}_{\mathcal{GW}}(R_t) = at + O(1), \qquad \mathbb{Var}_{\mathcal{GW}}(R_t) \leq ct. \tag{A.2}$$

Here $\mathbb{Var}_{\mathcal{GW}}(X)$ denotes the variance of the random variable X under the law of the Galton-Watson tree. Letting $R_{t,i}$ denote the number of distinct vertices visited by the random walk between time $(t - i + 1)$ and time t, and using the same regeneration time arguments as by Piau [28], we get

$$\mathbb{E}_{\mathcal{GW}}(R_{t,i}) = ai + O(1), \qquad \mathbb{Var}_{\mathcal{GW}}(R_{t,i}) \leq ci, \tag{A.3}$$

from which it follows that, for $\delta > 0$,

$$\mathbb{P}_{\mathcal{GW}}\left(\left|\frac{R_{t,i}}{i} - a\right| > \delta\right) = O(i^{-1}). \tag{A.4}$$

This in turn implies that, with $1 - O(i^{-1/2})$ probability under the $\mathbb{P}^{\mathcal{GW}}$-measure,

$$\mathbb{P}_\omega\left(\left|\frac{R_{t,i}}{i} - a\right| \leq \delta\right) = 1 - O(i^{-1/2}). \tag{A.5}$$

The latter is the crucial ingredient in the proof of the following lemma.

Lemma A.1 (Range asymptotics). *Let $\mathbb{P}_n^{\mathcal{GW}}$ denote the law of the Galton-Watson tree whose offspring distribution is the size-biased version of the degree distribution D_n (recall (1.2) and (1.7)). Let $Z_{t,i}$ denote the number of distinct edges visited during the last i steps of the simple random walk on the tree. For any $\delta > 0$ and any $t_n \to \infty$, with $\mathbb{P}_n^{\mathcal{GW}}$-probability at least $1 - O(t_n^{-1/2})$ as $n \to \infty$,*

$$\mathbb{P}_\omega \left(\left| \frac{1}{t_n^2} \sum_{i=1}^{t_n} Z_{t_n,i} - \frac{a_n}{2} \right| \le \delta \right) = 1 - O(t_n^{-1/2}). \tag{A.6}$$

Proof. First note that, since the simple random walk lives on a tree, we have $Z_{t_n,i} = R_{t_n,i} - 1$. The result follows from monotonicity of $i \mapsto Z_{t_n,i}$, as well as a decomposition into finitely many pieces. Indeed, let $M \in \mathbb{N}$, $i_k = k \lceil \frac{t_n}{M} \rceil$, $j_k = k \lfloor \frac{t_n}{M} \rfloor$, $k = 0, \dots, M-1$. By monotonicity,

$$Z_{t_n,j_{k-1}} \le Z_{t_n,i} \le Z_{t_n,i_k}, \quad j_{k-1} \le i \le i_k, \tag{A.7}$$

and so with $\mathbb{P}_n^{\mathcal{GW}}$-probability at most $O(t_n^{-1/2})$ we have

$$\mathbb{P}_\omega \left(\frac{1}{t_n^2} \sum_{i=1}^{t_n} Z_{t_n,i} < (1-\delta)\frac{a_n}{2} \right) \le \mathbb{P}_\omega \left(\frac{1}{t_n^2} \sum_{k=1}^{M-1} \left\lfloor \frac{t_n}{M} \right\rfloor Z_{t_n,j_k} < (1-\delta)\frac{a_n}{2} \right) \tag{A.8}$$

$$\le \mathbb{P}_\omega \left(\frac{1}{t_n^2} \sum_{k=1}^{M-1} k \left\lfloor \frac{t_n}{M} \right\rfloor^2 \frac{Z_{t_n,j_k}}{j_k} < (1-\delta)\frac{a_n}{2} \right) \le \sum_{k=1}^{M-1} \mathbb{P}_\omega \left(\frac{Z_{t_n,j_k}}{j_k} < (1-\delta')a_n \right) = O(t_n^{-1/2}),$$

where $\delta' > 0$ is a fixed number such that $M(M-1)(1-\delta')/(M+1)^2 > 1 - \delta$, and the last line follows from (A.5) and the fact that M is finite. Similarly, for the upper bound, with $\mathbb{P}_n^{\mathcal{GW}}$-probability at most $O(t_n^{-1/2})$ we have

$$\mathbb{P}_\omega \left(\frac{1}{t_n^2} \sum_{i=1}^{t_n} Z_{t_n,i} > (1+\delta)\frac{a_n}{2} \right) \le \mathbb{P}_\omega \left(\frac{1}{t_n^2} \sum_{k=1}^{M} \left\lceil \frac{t_n}{M} \right\rceil Z_{t_n,i_k} > (1+\delta)\frac{a_n}{2} \right) \tag{A.9}$$

$$\le \mathbb{P}_\omega \left(\frac{1}{t_n^2} \sum_{k=1}^{M} k \left\lceil \frac{t_n}{M} \right\rceil^2 \frac{Z_{t_n,i_k}}{j_k} > (1+\delta)\frac{a_n}{2} \right) \le \sum_{k=1}^{M} \mathbb{P}_\omega \left(\frac{Z_{t_n,j_k}}{j_k} > (1+\delta')a_n \right) = O(t_n^{-1/2}).$$

$\square$

Remark A.2. This result holds also for the augmented Galton-Watson tree as defined by Lyons, Pemantle and Peres [23] that is formed by taking a Galton-Watson tree and attaching to its root another Galton-Watson tree with the same offspring distribution. This construction coincides with modifying the offspring distribution of the root so that it has one more child. For both cases the constant a in (A.2) is the same.

A.2 Escape probability

Lemma A.3 (Convergence of escape probability). $\lim_{n \to \infty} a_n = a$, *where a is as defined in (1.9) and a_n is the analogous quantity for $\mathbb{P}_n^{\mathcal{GW}}$.*

Proof. Write ω_n to denote the random graph generated according to $\mathbb{P}_n^{\mathcal{GW}}$. The effective conductance on ω_n between 0 and infinity is given by the Dirichlet principle:

$$C_n = \inf_{\substack{f: \ V_n \to [0,1] \\ f(0)=1, f(\infty)=0}} \sum_{e \in E_n} |\nabla f(e)|^2, \tag{A.10}$$

where $\omega_n = (V_n, E_n)$ and $\nabla f(e)$ is the gradient of f along e. The infimum is uniquely attained at f_n^* given by

$$f_n^*(v) = \mathbb{P}_{\omega_n}^v(\tau_0 < \infty), \qquad v \in V_n, \tag{A.11}$$

where $\mathbb{P}^v_{\omega_n}$ is the law of the random walk on ω_n starting from v, and τ_0 is the first hitting time of 0.

Let ω_n^m denote the truncation of ω_n obtained by cutting away all the vertices at distance $> m$ from 0. Let C_n^m denote the effective conductance of ω_n^m between 0 and $\partial_m V_n$, the set of leaves of ω_n^m at distance m from 0. Then, again by the Dirichlet principle,

$$C_n^m = \inf_{\substack{f\colon V_n \to [0,1] \\ f(0)=1, f(v)=0\,\forall v:|v|\geq m}} \sum_{e \in E_n} |\nabla f(e)|^2. \tag{A.12}$$

Clearly, $C_n \leq C_n^m$ for all m, n. Moreover, $C_n^m \downarrow C_n$ as $m \to \infty$ for every n. To show that the convergence is uniform in n, we estimate

$$\begin{aligned}
C_n^m &\leq \sum_{e \in E_n \cap (V_n^m \times V_n^m)} |\nabla f_n^*(e)|^2 + \sum_{\substack{v \in V_n^m, v' \notin V_n^m \\ (v,v') \in E_n}} f_n^*(v)^2 \\
&\leq C_n + \sum_{v \in \partial_m V_n} (d_v - 1) f_n^*(v)^2,
\end{aligned} \tag{A.13}$$

where d_v is the degree of v. By Condition 1.1, all degrees in ω_n are at least 3, and so we have

$$f_n^*(v) \leq A\, \mathbb{P}^{v,\text{taboo}}_{\omega_n}(\tau_0 < \infty), \tag{A.14}$$

where taboo stands for the requirement that the walk moves from v to 0 without returning to v, and A is the average number of returns to a vertex in a 3-tree. By reversing time, we see that

$$d_v\, \mathbb{P}^{v,\text{taboo}}_{\omega_n}(\tau_0 < \infty) = d_0\, \mathbb{P}^{0,\text{taboo}}_{\omega_n}(\tau_v < \infty). \tag{A.15}$$

Hence

$$\begin{aligned}
\sum_{v \in \partial_m V_n} (d_v - 1) f_n^*(v)^2 &\leq A^2 d_0 \sum_{v \in \partial_m V_n} \frac{d_v - 1}{d_v} \mathbb{P}^{0,\text{taboo}}_{\omega_n}(\tau_v < \infty)\, \mathbb{P}^{v,\text{taboo}}_{\omega_n}(\tau_0 < \infty) \\
&\leq A^2 d_0\, \mathbb{P}^0_{\omega_n}(\tau_{\partial_m V_n} < \tau_0).
\end{aligned} \tag{A.16}$$

By Condition 1.1, the latter probability in turn is bounded from above, uniformly in n and ω_n, by the probability p_m that simple random walk on the 3-tree moves a distance m away from its starting vertex and afterwards returns to that vertex. Thus, we obtain the sandwich

$$C_n \leq C_n^m \leq C_n + A^2 d_0\, p_m. \tag{A.17}$$

Since $\lim_{m\to\infty} p_m = 0$, it follows that

$$\lim_{m\to\infty} \sup_n |C_n^m - C_n| = 0. \tag{A.18}$$

Finally, for every m,

$$\lim_{n\to\infty} C_n^m = C^m, \tag{A.19}$$

where C^m is the effective conductance of the m-truncation of the Galton-Watson tree ω with law $\mathbb{P}^{\mathcal{GW}}$. Moreover, $C^m \downarrow C$ as $m \to \infty$, with C the effective conductance of ω. Since

$$a_n = \mathbb{E}_n^{\mathcal{GW}}\left(\frac{C_n}{1+C_n}\right), \quad a_n^m = \mathbb{E}_n^{\mathcal{GW}}\left(\frac{C_n^m}{1+C_n^m}\right), \tag{A.20}$$

with similar expressions for a and a^m, and $\lim_{n\to\infty} \mathbb{E}_n^{\mathcal{GW}}(d_0) = \mathbb{E}^{\mathcal{GW}}(d_0) < \infty$ by Condition 1.1, we conclude that

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \lim_{m\to\infty} a_n^m = \lim_{m\to\infty} \lim_{n\to\infty} a_n^m = \lim_{m\to\infty} a^m = a. \tag{A.21}$$

which proves the claim. $\qquad\square$

References

[1] M. Abdullah, C. Cooper and A. Frieze, Cover time of a random graph with given degree sequence, *Discr. Math.* 312, 3146–3163, 2012.

[2] A. Aldous and J. A. Fill, *Reversible Markov Chains and Random Walks on Graphs*, in preparation: http://www.stat.berkeley.edu/∼aldous/RWG/book.html.

[3] O. Angel, R. van der Hofstad and C. Holmgren, Limit laws for self-loops and multiple edges in the configuration model, arXiv: 1603.07172.

[4] E.A. Bender and E.R. Canfield, The asymptotic number of labelled graphs with given degree sequences. *J. Comb. Theory (A)*, 24, 296–307, 1978.

[5] I. Benjamini, G. Kozma and N. Wormald, The mixing time of the giant component of a random graph. *Random Struct. Algor.* 45, 383–407, 2014.

[6] I. Benjamini and E. Mossel, On the mixing time of a simple random walk on the super-critical percolation cluster, *Probab. Theory Relat. Fields* 125, 408–420, 2003.

[7] A. Ben-Hamou and J. Salez, Cuttoff for non-backtracking random walks on sparse random graphs, arXiv:1504.02429.

[8] N. Berestycki, E. Lubetzky, Y. Peres and A. Sly, Random walks on the random graph, arXiv:1504.01999.

[9] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* 1, 311–316, 1980.

[10] B. Bollobás, *Random Graphs*, Cambridge University Press, Cambridge, second edition, 2001.

[11] C. Bordenave, P. Caputo and J. Salez, Random walk on sparse random digraphs, arXiv:1508.06600.

[12] J. Ding, E. Lubetzky and Y. Peres, Mixing time of near-critical random graphs, *Ann. Probab.* 40, 979–1008, 2012.

[13] L. Federico and R. van der Hofstad, Critical window for connectivity in the configuration model, arXiv:1603.03254.

[14] N. Fountoulakis and B.A. Reed, The evolution of the mixing rate of a simple random walk on the giant component of a random graph, *Random Struct. Algo.* 33, 68–86, 2008.

[15] C. Greenhill, The switch Markov chain for sampling irregular graphs, *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, 1564–1572, 2015.

[16] R. van der Hofstad, *Random Graphs and Complex Networks. Vol. I*, Cambridge University Press, to appear.

[17] R. van der Hofstad, *Random Graphs and Complex Networks. Vol. II*, Cambridge University Press, in preparation.

[18] S. Janson, The probability that a random multigraph is simple, *Combinatorics, Probability and Computing* 18, 205–225, 2009.

[19] S. Janson, The probability that a random multigraph is simple. II. *J. Appl. Probab.* 51A, 123–137, 2014.

[20] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Times*, American Mathematical Society, 2009.

[21] E. Lubetzky and A. Sly, Cut-off phenomena for random walks on random regular graphs, *Duke Math. J.* 153, 475–510, 2010.

[22] T. Łuczak. Sparse random graphs with a given degree sequence. In: *Random graphs*, Vol. 2 (Poznán, 1989), pp. 165–182. Wiley, 1992.

[23] R. Lyons, R. Pemantle and Y. Peres, Ergodic theory on Galton-Watson trees: speed of random walk and dimension of harmonic measure, *Erg. Th. Dyn. Syst.* 15, 593–619, 1995.

[24] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, *Random Struct. Algo.* 6, 161–179, 1995.

[25] M. Molloy and B. Reed, The size of the giant component of a random graph with a given degree sequence, *Combin. Probab. Comput.* 7, 295–305, 1998.

[26] A. Nachmias and Y. Peres, Critical random graphs: Diameter and mixing time, *Ann. Probab.* 36, 1267–1286, 2008.

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

[28] D. Piau, Functional limit theorems for the simple random walk on a supercritical Galton-Watson tree, *Ann. Probab.* 26, 1016–1040, 1998.

[29] W. Stadje, The collector's problem with group drawings, *Adv. Appl. Prob.* 22, 866–82, 1990.