



Co-evolving dynamic networks

Sayan Banerjee¹ · Shankar Bhamidi¹ · Xiangying Huang¹

Received: 23 March 2022 / Revised: 14 November 2023 / Accepted: 1 March 2024 /

Published online: 2 April 2024

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

Co-evolving network models, wherein dynamics such as random walks on the network influence the evolution of the network structure, which in turn influences the dynamics, are of interest in a range of domains. While much of the literature in this area is currently supported by numerics, providing evidence for fascinating conjectures and phase transitions, proving rigorous results has been quite challenging. We propose a general class of co-evolving tree network models driven by local exploration, started from a single vertex called the root. New vertices attach to the current network via randomly sampling a vertex and then exploring the graph for a random number of steps in the direction of the root, connecting to the terminal vertex. Specific choices of the exploration step distribution lead to the well-studied affine preferential attachment and uniform attachment models, as well as less well understood dynamic network models with global attachment functionals such as PageRank scores (Chebolu and Melsted, in: SODA, 2008). We obtain local weak limits for such networks and use them to derive asymptotics for the limiting empirical degree and PageRank distribution. We also quantify asymptotics for the degree and PageRank of fixed vertices, including the root, and the height of the network. Two distinct regimes are seen to emerge, based on the expected exploration distance of incoming vertices, which we call the ‘fringe’ and ‘non-fringe’ regimes. These regimes are shown to exhibit different qualitative and quantitative properties. In particular, networks in the non-fringe regime undergo ‘condensation’ where the root degree grows at the same rate as the network size. Networks in the fringe regime do not exhibit condensation. Further, non-trivial phase transition phenomena are shown to arise for: (a) height asymptotics in the non-fringe regime, driven by the subtle competition between the condensation at the root and network growth; (b) PageRank distribution in the fringe regime, connecting to the well

✉ Shankar Bhamidi
bhamidi@email.unc.edu

Sayan Banerjee
sayan@email.unc.edu

Xiangying Huang
zoehuang@unc.edu

¹ Department of Statistics and Operations Research, University of North Carolina, 304 Hanes Hall, Chapel Hill, NC 27599, USA

known power-law hypothesis. In the process, we develop a general set of techniques involving local limits, infinite-dimensional urn models, related multitype branching processes and corresponding Perron–Frobenius theory, branching random walks, and in particular relating tail exponents of various functionals to the scaling exponents of quasi-stationary distributions of associated random walks. These techniques are expected to shed light on a variety of other co-evolving network models.

Keywords Continuous time branching processes · Temporal networks · PageRank · Random trees · Stable age distribution theory · Local weak convergence · Multitype branching processes · Perron–Frobenius theory · Quasi-stationary distribution · Phase transition

Mathematics Subject Classification Primary 60K35 · 05C80

1 Introduction

1.1 Motivation

Driven by the explosion in the amount of data on various real world networks, the last few years have seen the emergence of many new mathematical network models. Goals underlying these models include, **(a)** extracting unexpected connectivity patterns within the network (e.g. community detection); **(b)** understanding properties of dynamics on these real world systems such as the spread of epidemics and opinion dynamics; **(c)** understanding mechanistic reasons for the emergence of empirically observed properties of these systems such as heavy tailed degree distribution or the small world property; see e.g. [1, 17, 23, 48, 49, 64] and the references therein. Within this vast research area, dynamic or temporal networks, namely systems that change over time, play an important role both in applications such as understanding social networks or the evolution of gene regulatory networks [33, 34, 45]. One major frontier, especially for developing rigorous understanding of proposed models, are the so-called co-evolutionary (or adaptive) networks, where specific dynamics (e.g. random walk explorations) on the network influence the structure of the network, which in turn influences the dynamics; thus both modalities (dynamics on the network and the network itself) co-evolve [5, 28, 60, 61]. Quoting [57], “*... adaptive networks provide a promising mechanistic approach to simultaneously explain both structural features and temporal features ... and can produce extremely interesting and rich dynamics, such as the spontaneous development of extreme states in opinion models ...*” Despite significant interest in such models, deriving rigorous results has been challenging. Let us describe two concrete motivations behind this paper:

- Evolving networks driven through local exploration:** Motivated by the growth of social networks, there has been significant interest in trying to understand the influence of processes such as search engines or influence ranking mechanisms in the growth of networks. A number of papers [15, 20, 56] have explored the dynamic evolution of networks through new nodes first exploring neighborhoods

of randomly selected vertices before deciding on whom to connect. Even simple microscopic rules seemed to result in non-trivial phase transitions.

(b) **Preferential attachment models driven by global attachment schemes:** Perhaps the most well known class of co-evolving network models in practice are the so-called preferential attachment class. Typically one fixes an attachment function $f_{att} : \mathbb{N} \rightarrow \mathbb{R}_+$, with $f_{att}(k)$ denoting the attractiveness of a degree k vertex for new vertices joining the system. New vertices enter the system and attach to existing vertices with probability proportional to their attractiveness. Such models have been used in diverse settings including positing causal mechanisms for heavy tailed degree distributions [8], understanding robustness of networks to attacks [16, 18], and modeling retweet networks [12]. Such models require global knowledge of the network at each stage, yet use only the degree of each vertex for attachment, eschewing potentially more relevant global attractiveness functions such as centrality measures of vertices like the PageRank score (described in more detail below).

Organization of the paper: We describe the general class of models in Sect. 1.2. Section 2 describes initial constructions required to state the main results. Section 3 contains all the main results, where we also connect the results to existing literature and conjectures. Proofs are commenced in Sect. 4 with individual subsections connecting functionals of the model to continuous time branching processes, branching random walks, quasi-stationary distributions and so on, and deriving relevant results for this paper. The rest of the sections then use the technical foundations in Sect. 4 to complete the proofs of the main theorems.

1.2 Model definition

Fix a probability mass function $\mathbf{p} := \{p_k : k \geq 0\}$ on \mathbb{Z}_+ . For the rest of the paper, let $\mathbf{Z} = \{Z_1, Z_2, \dots\}$ be an *i.i.d* sequence with distribution \mathbf{p} . We now describe the recursive construction of a sequence of random trees $\{\mathcal{T}_n : n \geq 1\}$, always rooted at vertex $\{v_0\}$, with edges pointed from descendants to their parents. Start with two vertices $\{v_0, v_1\}$, with \mathcal{T}_1 a rooted tree at $\{v_0\}$, an oriented edge from v_1 to v_0 . Assume for some $n \geq 1$, we have constructed \mathcal{T}_n . Then to construct \mathcal{T}_{n+1} :

- New vertex $\{v_{n+1}\}$ enters the system at time $n + 1$.
- This new vertex selects a vertex V_n , *uniformly at random*, amongst the existing vertices $\mathcal{V}(\mathcal{T}_n) = \{v_0, \dots, v_n\}$.
- Let $\mathcal{P}(v_0, V_n)$ denote the path from the root to this vertex. This new vertex traverses up this path for a random length Z_{n+1} and attaches to the terminal vertex. If the graph distance to the root, $\text{dist}(v_0, V_n) \leq Z_{n+1}$ then this new vertex attaches to the root v_0 .

See Fig. 1 for a pictorial description. Thus new vertices enter the system and perform local explorations, before attaching themselves to an existing vertex. This, at first sight, simple mechanism, results in a host of important special cases depending on the choice of \mathbf{p} which we describe next. We will let $\{\mathcal{T}_n(\mathbf{p}) : n \geq 1\}$ denote the corresponding tree process and suppress dependence on \mathbf{p} when this is clear from the context.

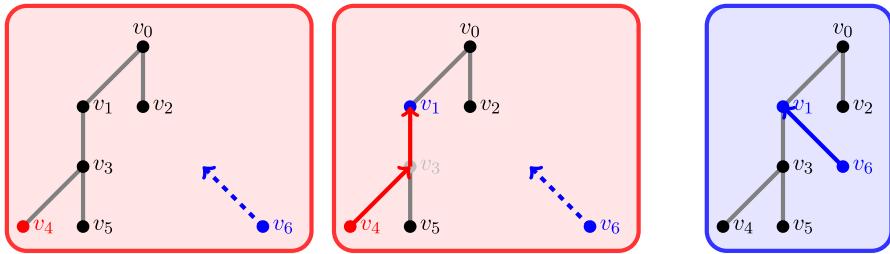


Fig. 1 v_6 is a new incoming vertex, and selects v_4 to start exploring the network, with sampled number of exploration steps $Z_6 = 2$

1.2.1 Random recursive tree

If $p_0 = 1$ then one obtains the random recursive tree where new vertices connect to previously existing vertices uniformly at random. See [44, 62, 63] and the references therein for the extensive use of this model in computer science.

1.2.2 Affine preferential attachment

Suppose \mathbf{p} is Bernoulli(p) for some $0 < p < 1$, namely $p_0 = 1 - p$, $p_1 = p$. Then one can check that the corresponding tree process has the same distribution as an affine preferential attachment model with attachment function $f_{att}(k) = k + (1 - 2p)/p$.

1.2.3 PageRank driven preferential attachment

In trying to understand models where vertices try to game search engines and attempt to increase their popularity by connecting to popular existing vertices in the system, one natural approach is via preferential attachment models where attractiveness of existing vertices is measured by their PageRank score (a global, as opposed to more local degree-only based attachment schemes). The probabilistic models of network evolution that at first sight seem to need global information on the network, but then have equivalent representations as local exploration schemes, has inspired a large body of work especially in statistical physics, under the general area of *network growth with redirection* [9, 24, 25, 40–42, 58, 65].

Definition 1.1 (*PageRank scores*) For a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the PageRank scores of vertices $v \in \mathcal{V}$ with damping factor c , is the stationary distribution $(\mathfrak{R}_{v,c} : v \in \mathcal{G})$ of the following random walk. At each step, with probability c , follow an outgoing edge (uniform amongst available choices) from the current location in the graph while, with probability $1 - c$, restart at a uniformly selected vertex in the entire graph. These scores are given by the linear system of equations:

$$\mathfrak{R}_{v,c} = \frac{1 - c}{n} + c \sum_{u \in \mathcal{N}^-(v)} \frac{\mathfrak{R}_{u,c}}{d^+(u)} \quad (1.1)$$

where $\mathcal{N}^-(v)$ is the set of vertices with edges pointed at v and $d^+(u)$ is the out-degree of vertex u .

At vertices with zero out-degree (e.g. the root of directed tree), the random walk stays in place with probability c and jumps to a uniformly chosen vertex with probability $1 - c$. The stationary probabilities at such vertices are then multiplied by $1 - c$ to keep the formula (1.1) same for all vertices.

Definition 1.2 (*PageRank driven preferential attachment* [56]) Fix damping factor $c \in (0, 1)$. Consider the following sequence of directed random trees, started with a root v_0 and another vertex v_1 with directed edge pointed to v_0 . At each discrete time step $n \geq 2$, a new vertex v_n enters the system and connects to a previously existing vertex with probability proportional to the PageRank of the existing vertex.

Now consider the process $\{\mathcal{T}_n(\mathbf{p}) : n \geq 1\}$, with \mathbf{p} as $\text{Geometric}(p)$, namely for $k \geq 0$, $p_k = p(1 - p)^k$. Thus a new vertex selects an existing vertex at random in the current tree and then walks up the path from the selected vertex to the root, wherein at each step, it decides to attach itself to the present location with probability p or move upwards with probability $1 - p$.

Theorem 1.3 ([20, Thm 1.1]) *The model with \mathbf{p} as $\text{Geometric}(p)$ has the same distribution as the PageRank driven preferential attachment model with damping factor $c = 1 - p$.*

2 Notation and initial constructions

2.1 Mathematical notation

For $J \geq 1$ let $[J] := \{1, 2, \dots, J\}$. If Y has an exponential distribution with rate λ , write this as $Y \sim \exp(\lambda)$. Write \mathbb{Z} for the set of integers, \mathbb{R} for the real line, \mathbb{N} for the natural numbers and let $\mathbb{R}_+ := (0, \infty)$. Write $\xrightarrow{\text{a.s.}}$, \xrightarrow{P} , \xrightarrow{d} for convergence almost everywhere, in probability and in distribution respectively. For a non-negative function $g : \mathbb{Z}_+ \rightarrow [0, \infty)$ and another function $f : \mathbb{Z}_+ \rightarrow \mathbb{R}$, write $f(n) = O(g(n))$ when $f(n)/g(n)$ is uniformly bounded, and $f(n) = o(g(n))$ when $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$. Write $f(n) = \Omega(g(n))$ if $\liminf_{n \rightarrow \infty} f(n)/g(n) > 0$. Furthermore, write $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $g(n) = O(f(n))$. For two real valued stochastic processes on the same space $\{X_t : t \geq 0\}$ and $\{Y_t : t \geq 0\}$, denote $X_t = Y_t + o_{a.s.}(1)$ if the random variable $X_t - Y_t \rightarrow 0$ almost surely as $t \rightarrow \infty$. We write that a sequence of events $(A_n)_{n \geq 1}$ occurs *with high probability* (whp) if $\mathbb{P}(A_n) \rightarrow 1$ as $n \rightarrow \infty$. We use \preceq_{st} for stochastic domination between two real valued probability measures. For a rooted finite tree \mathbf{t} , let $|\mathbf{t}|$ denote the number of vertices in \mathbf{t} and $\text{ht}(\mathbf{t})$ denote its height, namely, the maximal graph distance of the root from any other vertex. For graph \mathcal{G} , we write $\text{dist}_{\mathcal{G}}(\cdot, \cdot)$ for graph distance and in most cases suppress dependence on \mathcal{G} when this is clear from context.

2.2 Continuous time models

Recall the description of the model in Sect. 1.2. For much of the proofs, we will work with the following continuous time versions of the above process.

Definition 2.1 (*Continuous time versions*) We let $\{\mathcal{T}(t, \mathbf{p}) : t \geq 0\}$ denote the continuous time version of the above process wherein $\mathcal{T}(0, \mathbf{p})$ is a tree with vertex set $\{v_0\}$ rooted at v_0 . Each existing vertex v in the tree reproduces at rate one. When vertex v reproduces, a random variable Z following the law \mathbf{p} is sampled independently.

- (a) If $Z \leq \text{dist}(v_0, v)$, then a new vertex \tilde{v} is attached to the unique vertex u lying on the path between v and v_0 that satisfies $\text{dist}(v, u) = Z$ via a directed edge from \tilde{v} to u .
- (b) If $Z > \text{dist}(v_0, v)$, attach the new vertex \tilde{v} to the root v_0 via a directed edge towards v_0 .

Let $\{\mathcal{T}^*(t, \mathbf{p}) : t \geq 0\}$ denote the continuous time process which follows the dynamics as above but with (b) above replaced by,

- (b)' If $Z > \text{dist}(v_0, v)$, nothing happens.

It will turn out later that the process \mathcal{T}^* describes the evolution of the fringe tree below *non-root* vertices. Define the stopping times

$$T_n = \inf\{t \geq 0 : |\mathcal{T}(t, \mathbf{p})| = n + 1\},$$

which is the birth time of the incoming vertex v_n . Let \mathcal{F}_t denote the natural sigma-field of the process up to time t and $\{\mathcal{F}_t : t \geq 0\}$ the corresponding filtration. The following is obvious from construction.

Lemma 2.2 *The processes $\{\mathcal{T}_n(\mathbf{p}) : n \geq 1\}$ and $\{\mathcal{T}(T_n, \mathbf{p}) : n \geq 1\}$ have the same distribution.*

2.3 Fringe convergence

The aim of this section is to formalize the notion of convergence of neighborhoods of large random trees to neighborhoods of limiting infinite discrete structures. Local weak convergence of discrete random structures has now become quite standard in probabilistic combinatorics see e.g. [3, 10]. In the case of trees, following [2], the theory of local graph limits has an equivalent formulation in terms of convergence of fringe distribution.

2.4 Extended fringe decomposition

Given two rooted (unlabeled) finite trees \mathbf{s}, \mathbf{t} , say that $\mathbf{s} \sim \mathbf{t}$ or $\mathbf{s} = \mathbf{t}$ if there exists a root preserving isomorphism between the two trees. For $n \geq 1$, let \mathbb{T}_n be the space of all rooted trees on n vertices and let $\mathbb{T} = \cup_{n=0}^{\infty} \mathbb{T}_n$ be the space of all finite rooted trees. Here $\mathbb{T}_0 = \emptyset$ will be used to represent the empty tree (tree on zero vertices).

Fix a tree $\mathbf{t} \in \mathbb{T}$ with root ρ and a vertex v at distance h from the root. Let $(v_0 = v, v_1, v_2, \dots, v_h = \rho)$ be the unique path from v to ρ . The tree \mathbf{t} can be decomposed as $h+1$ rooted trees $f_0(v, \mathbf{t}), \dots, f_h(v, \mathbf{t})$, where $f_0(v, \mathbf{t})$ is the tree rooted at v consisting of all vertices for which there exists a path from the root passing through v , and for $i \geq 1$, $f_i(v, \mathbf{t})$ is the subtree rooted at v_i consisting of all vertices for which the path from the root passes through v_i but not through v_{i-1} . Call the map $F : \mathbb{T} \rightarrow \mathbb{T}^\infty$ defined by,

$$F(v, \mathbf{t}) = (f_0(v, \mathbf{t}), f_1(v, \mathbf{t}), \dots, f_h(v, \mathbf{t}), \emptyset, \emptyset, \dots),$$

as the fringe decomposition of \mathbf{t} about the vertex v . Call $f_0(v, \mathbf{t})$ the **fringe** of the tree \mathbf{t} at v . For $k \geq 0$, call $F_k(v, \mathbf{t}) = (f_0(v, \mathbf{t}), \dots, f_k(v, \mathbf{t}))$ the **extended fringe** of the tree \mathbf{t} at v truncated at distance k from v on the path to the root.

Now consider the space \mathbb{T}^∞ . An element $\omega = (\mathbf{t}_0, \mathbf{t}_1, \dots) \in \mathbb{T}^\infty$ with $|\mathbf{t}_i| \geq 1$ for all $i \geq 0$, can be thought of as a locally finite infinite rooted tree with a single path to **infinity** (thus called a **sin-tree** [2]) as follows: Identify the sequence of roots of $\{\mathbf{t}_i : i \geq 0\}$ with the integer lattice $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$, equipped with the natural nearest neighbor edge set, rooted at $\rho = 0$. Analogous to the definition of extended fringes for finite trees, for any $k \geq 0$ write $F_k(0, \omega) = (\mathbf{t}_0, \mathbf{t}_1, \dots, \mathbf{t}_k)$. Call this the extended fringe of the tree ω at vertex 0, till distance k , on the infinite path from 0. Call $\mathbf{t}_0 = F_0(0, \omega)$ the **fringe** of the **sin-tree** ω . Now suppose \mathbb{P} is a probability measure on \mathbb{T}^∞ such that if $\mathcal{T} := (\mathbf{t}_0(\mathcal{T}), \mathbf{t}_1(\mathcal{T}), \dots) \sim \mathbb{P}$, then $|\mathbf{t}_i(\mathcal{T})| \geq 1$ a.s. $\forall i \geq 0$. Then \mathcal{T} can be thought of as an infinite **random** **sin-tree**.

2.5 Convergence on the space of trees

Let \mathcal{T}_∞ be a random **sin-tree** with distribution \mathbb{P} on \mathbb{T}^∞ . Suppose $\{\mathcal{T}_n\}_{n \geq 1}$ be a sequence of **finite** rooted random trees all constructed on the same probability space (for notational convenience assume $|\mathcal{T}_n| = n$, all one needs is $|\mathcal{T}_n| \xrightarrow{\text{P}} \infty$). Fix $k \geq 0$ and (non-empty) trees $\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_k \in \mathbb{T}$. Let $\hat{\mathbf{s}} = (\mathbf{s}_0, \dots, \mathbf{s}_k) \in \mathbb{T}^{k+1}$ and define the empirical proportions,

$$H_n^k(\hat{\mathbf{s}}) = \frac{1}{n} \sum_{v \in \mathcal{T}_n} \mathbb{1}_{\{F_k(v, \mathcal{T}_n) = (\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_k)\}}.$$

As before “ $f_j(v, \mathcal{T}_n) = \mathbf{s}_j$ ” implies identical up to a root preserving isomorphism. Consider two (potentially distinct) notions of convergence of $\{\mathcal{T}_n : n \geq 1\}$:

- (a) Fix a probability measure π on \mathbb{T} . Say that a sequence of trees $\{\mathcal{T}_n\}_{n \geq 1}$ converges in probability, in the fringe sense, to π , if for every rooted tree $\mathbf{t} \in \mathbb{T}$, $H_n^0(\mathbf{t}) \xrightarrow{\text{P}} \pi(\mathbf{t})$.
- (b) Say that a sequence of trees $\{\mathcal{T}_n\}_{n \geq 1}$ converges in probability, in the **extended fringe sense**, to \mathcal{T}_∞ if for all $k \geq 0$ and $\hat{\mathbf{s}} \in \mathbb{T}^{k+1}$, one has

$$H_n^k(\hat{\mathbf{s}}) \xrightarrow{\text{P}} \mathbb{P}(F_k(0, \mathcal{T}_\infty) = \hat{\mathbf{s}}).$$

We shall denote this convergence by $\mathcal{T}_n \xrightarrow{\text{P-fr}} \mathcal{T}_\infty$ as $n \rightarrow \infty$.

Letting $\pi_{\mathcal{T}_\infty}(\cdot) = \mathbb{P}(F_0(0, \mathcal{T}_\infty) = \cdot)$ denote the distribution of the fringe of \mathcal{T}_∞ on \mathbb{T} , convergence in (b) above implies convergence in notion (a). Further, both notions imply convergence of functionals such as the degree distribution. For example notion (a) implies, for any $k \geq 0$,

$$\frac{\#\{v \in \mathcal{T}_n, \deg(v) = k + 1\}}{n} \xrightarrow{P} \mathbb{P}(\deg(0, \mathcal{T}_\infty) = k). \quad (2.1)$$

Here $\deg(0, \mathcal{T}_\infty)$ denotes the number of edges connected to 0 in its fringe $\mathbf{t}_0(\mathcal{T}_\infty)$. However, this convergence gives much more information about the asymptotic properties of \mathcal{T}_n , including convergence of global functionals [11].

In terms of going from convergence in (a) to (b), Aldous in [2] showed that, amongst the distributions π that arise as potential limits of the fringe convergence in (a), there is a special sub-class of measures called fringe distributions, which automatically imply the existence of and convergence to a limit infinite sin-tree . For each $\mathbf{s} \in \mathbb{T}$, suppose the root has children v_1, v_2, \dots, v_d for some $d \geq 0$. Let $\{f(\mathbf{s}, v_i) : 1 \leq i \leq d\}$ denote the subtrees at these vertices, with $f(\mathbf{s}, v_i)$ rooted at v_i . For each $\mathbf{t} \in \mathbb{T}$, let $Q(\mathbf{s}, \mathbf{t}) = \sum_i \mathbb{1}\{f(\mathbf{s}, v_i) = \mathbf{t}\}$.

Definition 2.3 Call a probability distribution π on \mathbb{T} a **fringe distribution** if,

$$\sum_{\mathbf{s} \in \mathbb{T}} \pi(\mathbf{s}) Q(\mathbf{s}, \mathbf{t}) = \pi(\mathbf{t}), \quad \forall \mathbf{t} \in \mathbb{T}.$$

It is easy to check that the space of fringe measures is a convex subspace of the space of probability measures on \mathbb{T} . The next result summarizes some of the remarkable findings in [2], relevant to this paper.

Theorem 2.4 ([2]) Fix a probability measure π on \mathbb{T} . Suppose $\{\mathcal{T}_n\}_{n \geq 1}$ converges in probability in the fringe sense to π .

- (a) π is a fringe distribution in the sense of Definition 2.3 iff $\sum_t \pi(t) \text{root-degree}(t) = 1$.
- (b) If π is a fringe distribution then, convergence in probability in the fringe sense (notion (a)), implies the existence of a random sin-tree \mathcal{T}_∞ such that $\{\mathcal{T}_n\}_{n \geq 1}$ converges in probability in the extended fringe sense to \mathcal{T}_∞ (notion (b)).
- (c) If π is a fringe distribution then $\text{ht}(\mathcal{T}_n) \rightarrow \infty$ in probability.

3 Results

We start by describing local weak convergence of the network model in Sect. 3.1. This gives convergence of the degree distribution to an explicit limit. Further the tail exponent of the limit degree distribution is shown to be related to a large deviation rate constant of the quasi-stationary distribution of a random walk associated to the exploration step distribution. Section 3.2 describes asymptotics of the root and other

fixed vertex degrees, in particular deriving necessary and sufficient conditions on \mathbf{p} for *condensation*, namely the root obtaining a fixed density of the total number of edges. Section 3.3 derives asymptotics for the height, appropriately normalized, in terms of extremal statistics of related branching random walks, showing the emergence of a phase transition. We conclude in Sect. 3.4 with asymptotics and phase transitions for the asymptotic PageRank of fixed vertices as well as the limiting PageRank distribution.

3.1 Local weak convergence

Recall the process $\{\mathcal{T}^*(t, \mathbf{p}) : t \geq 0\}$ in Definition 2.1.

Definition 3.1 (*Fringe limit*) Let $\tau \sim \text{Exp}(1)$ independent of the process \mathcal{T}^* . Let $\pi_{\mathbf{p}}(\cdot)$ denote the distribution of $\mathcal{T}^*(\tau, \mathbf{p})$, viewed as a random a.s. finite rooted tree in \mathbb{T} .

Assumption 3.2 Let $Z \sim \mathbf{p}$. Assume $\mathbb{E}[Z] < \infty$.

Theorem 3.3 (*Fringe convergence*)

(a) *Under Assumption 3.2, the sequence of random trees $\{\mathcal{T}_n(\mathbf{p}) : n \geq 1\}$ converges in probability, in the fringe sense, to $\pi_{\mathbf{p}}(\cdot)$. Writing D for the root degree of the random tree sampled using $\pi_{\mathbf{p}}$, for every $k \geq 0$,*

$$\frac{1}{n+1} \sum_{v \in \mathcal{T}_n(\mathbf{p})} \mathbb{1}_{\{\deg(v) = k+1\}} \xrightarrow{P} \mathbb{P}(D = k), \quad \text{as } n \rightarrow \infty.$$

(b) *$\pi_{\mathbf{p}}(\cdot)$ is a fringe distribution, as in Definition 2.3, if and only if $\mathbb{E}[Z] \leq 1$ where $Z \sim \mathbf{p}$.*

(c) *The root degree D and the size of the limit fringe tree $|\mathcal{T}^*(\tau, \mathbf{p})|$ satisfy:*

- (i) *If $\mathbb{E}[Z] \leq 1$ then $\mathbb{E}[D] = 1$ and $\mathbb{E}[|\mathcal{T}^*(\tau, \mathbf{p})|] = \infty$.*
- (ii) *If $\mathbb{E}[Z] > 1$ then $\mathbb{E}[D] < 1$ and $\mathbb{E}[|\mathcal{T}^*(\tau, \mathbf{p})|] < \infty$.*

Remark 1 Owing to the above Theorem, we will refer to the setting $\mathbb{E}[Z] \leq 1$ as the *fringe regime* whilst $\mathbb{E}[Z] > 1$ will be referred to as the *non-fringe regime*.

The next result follows from Theorem 3.3 and results in [2] summarized in Theorem 2.4.

Corollary 3.4 (*Convergence to limiting sin-tree*) *Assume for $Z \sim \mathbf{p}$, $\mathbb{E}[Z] \leq 1$.*

(a) *There exists a limiting infinite sin-tree $\mathcal{T}_\infty(\mathbf{p})$ such that $\{\mathcal{T}_n(\mathbf{p}) : n \geq 1\}$ converges in probability in the extended fringe sense to $\mathcal{T}_\infty(\mathbf{p})$.*

(b) $\text{ht}(\mathcal{T}_n) \xrightarrow{P} \infty$.

Remark 2 (i) By [11], when $\mathbb{E}[Z] \leq 1$, local weak convergence above implies that even global functionals such as the spectral distribution of the adjacency matrix converge (in this case the limit spectral distribution can be shown to be non-random with an infinite set of atoms). In Aldous' terminology [2], it would be interesting to derive a *reduced Markov description* of this limit object.

(ii) In the non-fringe case when $\mathbb{E}(Z) > 1$, whilst Theorem 3.3 shows convergence in probability in the fringe sense, as described below in Theorem 3.9(a), in this regime the degree of the root scales like $\sim (1 - q_*)n$, and thus if one selects a vertex at random, with probability $\sim (1 - q_*)$ the parent of this vertex is the root whose degree $\rightarrow \infty$. In particular, there is **no convergence** in the extended fringe sense to a **sin-tree**, which by definition has to be locally finite.

In order to derive quantitative bounds on explicit functionals such as the degree distribution and PageRank, we will need to make a few more assumptions on \mathbf{p} . Recall that in Sect. 1.2, the case $p_0 + p_1 = 1$ corresponds to either the random recursive tree or affine preferential attachment, which have already been thoroughly analyzed in the literature. While the techniques described below recover many results for these models, we are mainly interested in settings not covered under these two models. Let $\{Z_i\}_{i \geq 1}$ be a collection of i.i.d. random variables following the law \mathbf{p} . Consider the random walk,

$$S_n = S_0 + \sum_{i=1}^n (Z_i - 1), \quad S_0 \in \mathbb{Z}, \quad n \geq 1. \quad (3.1)$$

Define the probability generating function (pgf):

$$f(s) := \sum_{k=0}^{\infty} p_k s^k, \quad s \geq 0. \quad (3.2)$$

Assumption 3.5 Assume that \mathbf{p} satisfies the following: **(a)** $p_0 \in (0, 1)$, $p_0 + p_1 < 1$. **(b)** Assume that the random walk $\{S_n : n \geq 0\}$ in (3.1) is aperiodic, i.e., $\gcd\{j : p_j > 0\} = 1$.

Remark 3 Most of our results can be extended to the case $p_0 + p_1 = 1$ by straightforward modifications of our proof techniques. See Remark 10.

Definition 3.6 (a) Let s_0 be the unique positive root of $sf'(s) = f(s)$ if it exists, otherwise let s_0 be the radius of convergence of the power series $f(\cdot)$. Define $R := \lim_{s \uparrow s_0} s/f(s)$.

(b) Let q_* be the unique positive solution to the fixed point equation $f(q) = q$.

Let us briefly describe probabilistic interpretations of these objects before stating our main results. Recall that q_* denotes the extinction probability of a Galton-Watson branching process with offspring distribution \mathbf{p} . Next define

$$\bar{T}_i := \inf\{n \geq 0 : S_n = 0 | S_0 = i\} \quad (3.3)$$

to be the hitting time of zero of the random walk starting from state $i > 0$. Then q_* has the alternate interpretation $q_* = \mathbb{P}(\bar{T}_1 < \infty)$. Moreover, by [22, 55], under Assumptions 3.2 and 3.5, R arises as the following limit and lies in the asserted interval:

$$\lim_{n \rightarrow \infty} (\mathbb{P}(n < \bar{T}_1 < \infty))^{1/n} = \frac{1}{R} \in (0, 1]. \quad (3.4)$$

Standard results, connecting random walks to branching processes, imply $q_* = \mathbb{P}(\bar{T}_1 < \infty) < 1$ if and only if $\mathbb{E}[Z] > 1$. The next lemma collects some classical facts about R, s_0, q_* arising in the analysis of quasi-stationary distributions for random walks. We provide brief pointers to the literature for completeness. In the following lemma, let r_f denote the radius of convergence of the pgf f .

Lemma 3.7 *Suppose Assumptions 3.2 and 3.5 hold.*

- (a) *$g(s) := f(s)/s$ is strictly decreasing on $(0, s_0)$ and, if $s_0 < r_f$, g is strictly increasing on (s_0, r_f) . In particular, $\inf_{s \in (0, 1)} f(s)/s \geq 1/R$.*
- (b) *For $\mathbb{E}[Z] < 1$, if $f(s)$ is analytic at $s = 1$, then $R \in (1, 1/\mathbb{E}[Z])$, $s_0 \in (1, \infty)$, otherwise $R = s_0 = 1$.*
- (c) *For $\mathbb{E}[Z] = 1$, $R = s_0 = 1$.*
- (d) *For $\mathbb{E}[Z] > 1$, $R > 1$ and $q_* < s_0 < 1$.*

Proof Part (a) follows from the observation that $g'(s) = (sf'(s) - f(s))/s^2$, whose numerator is strictly increasing on $(0, r_f)$ as $p_0 + p_1 < 1$, and negative at $s = 0$ as $p_0 > 0$. For part (b), when $f(s)$ is analytic at $s = 1$, note that $sf'(s) - f(s) < 0$ at $s = 1$ as $\mathbb{E}(Z) < 1$, which implies $s_0 > 1$. Moreover, if $s_0 = \infty$, $r_f = \infty$ and $g(s)$ is strictly decreasing on $(0, \infty)$, which is a contradiction as $\lim_{s \rightarrow \infty} g(s) = \infty$ (as $p_0 + p_1 < 1$). Thus, $s_0 \in (1, \infty)$. Further, this also implies $f(s_0) < \infty$ and hence, using part (a) and the definition of s_0 , $1 < R = s_0/f(s_0) \leq 1/f'(s_0) < 1/\mathbb{E}[Z]$. For the remaining assertions, see [55, Lemma 1] and the Remark following it. \square

Note that Lemma 3.7 implies that, under Assumptions 3.2 and 3.5, $s_0, f(s_0)$ are both finite and $R = s_0/f(s_0)$. Now let the random variable D be as in Theorem 3.3, following the law of the limiting degree distribution of the discrete tree network $\{\mathcal{T}_n(\mathbf{p})\}_{n \geq 1}$.

Theorem 3.8 (Tail of limit Degree distribution) *Under Assumptions 3.2 and 3.5:*

- (i) **Fringe regime:** *When $\mathbb{E}[Z] \leq 1$, with R as in Definition 3.6,*

$$\lim_{k \rightarrow \infty} \frac{\log \mathbb{P}(D \geq k)}{\log k} = -R. \quad (3.5)$$

- (ii) **Non-fringe regime:** *When $\mathbb{E}[Z] > 1$,*

$$-R \leq \liminf_{k \rightarrow \infty} \frac{\log \mathbb{P}(D \geq k)}{\log k} \leq \limsup_{k \rightarrow \infty} \frac{\log \mathbb{P}(D \geq k)}{\log k} \leq -\left(R \wedge \frac{\log q_*}{\log s_0}\right). \quad (3.6)$$

Although the upper and lower bounds in part (ii) above are different, they can be checked to be equal in several cases even in the non-fringe regime. See Remark 9 for a discussion.

3.2 Condensation and fixed vertex degree asymptotics

The next result describes asymptotics for the root degree. In particular, part (a) shows that in the non-fringe regime, there is a condensation phenomenon at the root and the root neighbors asymptotically comprise a positive limiting density of all the vertices in the tree.

Theorem 3.9 (Root degree asymptotics) *Let $\deg(v_0, n)$ denote the degree of the root in $T_n(\mathbf{p})$. Under Assumptions 3.2 and 3.5:*

(a) **Non-Fringe regime:** *Assume $\mathbb{E}[Z] > 1$. Then*

$$\frac{\deg(v_0, n)}{n} \xrightarrow{\mathbb{P}} 1 - q_* > 0,$$

where q_* is defined in Definition 3.6(b).

(b) **Fringe regime:** *Assume $\mathbb{E}[Z] \leq 1$. Then for any $\delta > 0$, as $n \rightarrow \infty$,*

$$\frac{\deg(v_0, n)}{n^{\frac{1}{R}-\delta}} \xrightarrow{\text{a.s.}} \infty, \quad \text{and} \quad \frac{\deg(v_0, n)}{n^{\frac{1}{R}}(\log n)^{1+\delta}} \xrightarrow{\text{a.s.}} 0. \quad (3.7)$$

The next result describes the degree evolution of a fixed **non-root** vertex. In particular in the non-fringe regime there is a marked difference between the evolution of the degree of the root v_0 and any non-root vertex.

Theorem 3.10 (Fixed vertex degree asymptotics) *Fix $k \geq 1$ and let $\deg(v_k, n)$ denote the degree of vertex v_k in $T_n(\mathbf{p})$. Then under Assumptions 3.2 and 3.5, for any $\delta > 0$,*

$$\frac{\deg(v_k, n)}{n^{\frac{1}{R}-\delta}} \xrightarrow{\text{a.s.}} \infty, \quad \text{and} \quad \frac{\deg(v_k, n)}{n^{\frac{1}{R}}(\log n)^{1+\delta}} \xrightarrow{\text{a.s.}} 0. \quad (3.8)$$

Remark 4 When \mathbf{p} is Geometric(p) for $p \in (0, 1)$ so that $\mathbb{E}[Z] = (1-p)/p$, the model was first rigorously analyzed in [20] using combinatorial recursions. The authors observed a phase transition for the expected root degree at $p = 1/2$ (precisely as one transitions from $\mathbb{E}[Z] < 1$ to $\mathbb{E}[Z] \geq 1$). In brief, [20] showed that the expected degree of the root in this special case satisfies $\mathbb{E}(\deg(v_0, T_n)) = \Omega(n(\log n)^k)$ if $p \leq 1/2$ whilst $\mathbb{E}(\deg(v_0, T_n)) = O(n^{4pq}(\log n)^{k'})$ if $p > 1/2$, for some $k, k' \in \mathbb{Z}$. One can check that in this case, $R = 1/(4pq)$, thus this paper also clarifies the reason for the mysterious constant $4pq$, in terms of a large deviations exponent of the hitting time of zero by the associated random walk.

3.3 Height asymptotics

Recall that $f(\cdot)$ denotes the pgf of \mathbf{p} . Define

$$\kappa(s) := \frac{f(s)}{s \log(1/s)}, \quad s \in (0, 1).$$

It is shown in Lemma 8.2 that the infimum of $s \mapsto \kappa(s)$ is attained at a unique point in $(0, s_0]$.

Definition 3.11 Define,

$$\kappa_0 := \inf_{s \in (0, 1)} \frac{f(s)}{s \log(1/s)} \quad (3.9)$$

Let $s^* \in (0, 1)$ be the point where the infimum of $\frac{f(s)}{s \log(1/s)}$ is attained, i.e., $\kappa_0 = \frac{f(s^*)}{s^* \log(1/s^*)}$.

The following theorem gives height asymptotics for our model. Interestingly, we observe a phase transition in the limiting behavior of rescaled heights in the non-fringe regime.

Theorem 3.12 (Height asymptotics) *Let \mathcal{H}_n denote the height of $T_n(\mathbf{p})$. Then, under Assumptions 3.2 and 3.5:*

(i) **Fringe regime:** *When $\mathbb{E}[Z] \leq 1$, as $n \rightarrow \infty$,*

$$\frac{\mathcal{H}_n}{\log n} \xrightarrow{a.s.} \kappa_0.$$

(ii) **Non-Fringe regime:** *When $\mathbb{E}[Z] > 1$, as $n \rightarrow \infty$,*

$$\frac{\mathcal{H}_n}{\log n} \xrightarrow{a.s.} \begin{cases} \kappa_0 & \text{if } s^* \in (0, q_*], \\ \frac{1}{\log(1/q_*)} & \text{if } s^* \in (q_*, s_0], \end{cases}$$

where q_* is defined as in Definition 3.6(b).

Three remarks are in order.

Remark 5 (Non-triviality in the non-fringe regime) Theorem 3.9(a) implies that in the non-fringe regime, there are $\Theta(n)$ vertices within distance one of the root whp so it is not obvious that in this case the height should diverge. Thus the result above on the height shows, that even in this case, the height scales like $\log n$ with an appropriate limit constant.

Remark 6 (Probabilistic interpretation of limit and phase transition) The limiting rescaled height and associated phase transition can be probabilistically understood via branching random walks [13, 14].

Definition 3.13 Fix pmf \mathbf{p} . Consider a branching random walk with individuals being born into the population in continuous time, and with spatial locations on \mathbb{Z} , starting with one individual at location zero with dynamics:

(a) Each vertex gives rise to offspring according to a rate one Poisson process.

(b) Writing ρ_v for the location of particle v , if v is born to u then $\rho_v = \rho_u + \zeta_{uv}$, where $\zeta_{uv} \stackrel{d}{=} (1 - Z)$, independent across parent offspring connections and times of birth, where $Z \sim \mathbf{p}$.

Write $\{\text{BRW}(t) : t \geq 0\}$, for the corresponding process keeping track of genealogical structure and locations.

Let $B(t)$ denote the location of the rightmost particle at time t . Then, it turns out, using [13] (see Sect. 4.3), in our setting, $\lim_{t \rightarrow \infty} B(t)/t \xrightarrow{\text{a.s.}} \kappa_0$. Hence, the rescaled height asymptotics agrees with that of $B(\cdot)$ when $\mathbb{E}[Z] \leq 1$, or $\mathbb{E}[Z] > 1$ and $s^* \in (0, q_*]$.

However, when $\mathbb{E}[Z] > 1$ and $s^* \in (q_*, s_0]$, there is a competition between the reproduction rate of the root and the upper tail large deviations behavior of $B(\cdot)$, which is obtained in Lemma 4.9. An inspection of the proof reveals that, for large t , the height of $\mathcal{T}(t, \mathbf{p})$ has the same asymptotics as the maximum of the heights of the subtrees rooted at the children of the root that are born in the time interval $[\frac{1-\varepsilon}{(1-f'(q_*)) \log(1/q_*)}, \frac{1+\varepsilon}{(1-f'(q_*)) \log(1/q_*)}]$ for small $\varepsilon > 0$. This phenomenon manifests itself through the phase transition observed in Theorem 3.12.

Remark 7 (*Explicit height computation for PageRank driven networks*) When \mathbf{p} is $\text{Geometric}(p)$, the pgf is given by $f(s) = \frac{p}{1-qs}$, $s \in [0, 1/q]$. Asymptotics for the height in this model were previously addressed in [46] using quite different techniques. For an explicit constant $\tilde{p} \approx .206$, they show that for $p \in [\tilde{p}, 1)$, $\mathcal{H}_n / \log n$ converges to a limit constant whilst for $p \in (0, \tilde{p})$, there exist constants $c_L(p) < c_U(p)$ such that whp for any $\varepsilon > 0$, $\mathcal{H}_n \in [(c_L(p) - \varepsilon) \log n, (c_U(p) + \varepsilon) \log n]$ as $n \rightarrow \infty$. Our result shows that, contrary to what is conjectured in [46], there is indeed a phase transition in the height asymptotics at $p = \tilde{p}$. More precisely, the minimizer $s^* \in (0, 1)$ in (3.9) can be seen to be the unique solution to the equation

$$(1-p)s = \frac{1 + \log s}{1 + 2 \log s}, \quad s \in (0, 1).$$

Define u^* via the relation $s^* = e^{-1/u^*}$. Then,

$$\kappa_0 = pe^{1/u^*} u^* (2 - u^*),$$

which matches the expression for $c_L(p)$ in [46, Theorem 2]. Moreover, for $p \in (0, 1/2)$, $q_* = p/(1-p)$ and hence

$$\frac{1}{\log(1/q_*)} = \left(\log \left(\frac{1-p}{p} \right) \right)^{-1},$$

which matches $c_U(p)$ in [46, Theorem 2] for $p \in (0, \tilde{p})$. Finally, the value of p where the height phase transition happens in Theorem 3.12 is seen to be the unique such value which gives $s^* = q_*$, or equivalently $\kappa'(q_*) = 0$. This characterizes the value as the unique solution to

$$\log \left(\frac{1-p}{p} \right) = \frac{1-p}{1-2p}, \quad p \in (0, 1/2),$$

which agrees with \tilde{p} obtained in [46, Theorem 2].

Our proofs thus elucidate the connections between limit constants in [46] obtained through subtle combinatorial analysis and extremal statistics in branching random walks [13].

3.4 PageRank asymptotics

Recall the PageRank scores $\{\mathfrak{R}_{v,c}(n) : v \in \mathcal{T}_n(\mathbf{p})\}$ in Definition 1.1. For any $v \in \mathcal{T}_n(\mathbf{p})$, let $P_l(v, n)$ denote the number of *directed* paths of length l that end at v in $\mathcal{T}_n(\mathbf{p})$. Since $\mathcal{T}_n(\mathbf{p})$ is a directed tree, it is easy to check that the PageRank scores have the explicit formulae for any vertex v ,

$$\mathfrak{R}_{v,c}(n) = \frac{(1-c)}{n} \left(1 + \sum_{l=1}^{\infty} c^l P_l(v, n) \right). \quad (3.10)$$

For the sequel, it will be easier to formulate results in terms of the *graph normalized* PageRank scores [27] $\{R_{v,c}(n) : v \in \mathcal{T}_n(\mathbf{p})\} = \{n\mathfrak{R}_{v,c}(n) : v \in \mathcal{T}_n(\mathbf{p})\}$. The first result shows a non-trivial phase transition of the PageRank scores for fixed vertices in the fringe regime. This phase transition carries over to the empirical distribution of PageRank scores which will conclude this section.

Theorem 3.14 (PageRank asymptotics for fixed vertices) *Fix vertex v_k and damping factor $c \in (0, 1)$. Then under Assumptions 3.2 and 3.5,*

(a) **Non-fringe regime:** when $\mathbb{E}[Z] > 1$,

(i) *When $k = 0$ (root PageRank), $R_{v_0,c}(n)/n$ is bounded away from zero in probability: for any $\delta > 0$, there exists $\varepsilon > 0$ such that*

$$\liminf_{n \rightarrow \infty} \mathbb{P} \left(\frac{R_{v_0,c}(n)}{n} \geq c(1-c)(1-q_*) - \varepsilon \right) \geq 1 - \delta.$$

(ii) *For any $k \geq 1$, as $n \rightarrow \infty$,*

$$\frac{R_{v_k,c}(n)}{n^{\frac{1}{R}-\delta}} \xrightarrow{a.s.} \infty, \quad \text{and} \quad \frac{R_{v_k,c}(n)}{n^{\frac{1}{R}}(\log n)^{1+\delta}} \xrightarrow{a.s.} 0. \quad (3.11)$$

(b) **Fringe regime:** When $\mathbb{E}[Z] \leq 1$,

(i) *Fix any $c \in (0, s_0^{-1}]$ with $c < 1$. Then for any $\delta > 0$, and any vertex v_k , $k \geq 0$, as $n \rightarrow \infty$,*

$$\frac{R_{v_k,c}(n)}{n^{\frac{1}{R}-\delta}} \xrightarrow{a.s.} \infty, \quad \text{and} \quad \frac{R_{v_k,c}(n)}{n^{\frac{1}{R}}(\log n)^{1+\delta}} \xrightarrow{a.s.} 0. \quad (3.12)$$

(ii) Suppose $s_0 > 1$. Fix any $c \in (s_0^{-1}, 1)$ and v_k , $k \geq 0$. Then there exists a non-negative random variable $W_{k,c}$ with $\mathbb{P}(W_{k,c} > 0) > 0$ such that,

$$\frac{R_{v_k,c}(n)}{n^{cf(1/c)}} \xrightarrow{a.s.} W_{k,c}, \quad \text{as } n \rightarrow \infty.$$

Next define the empirical distribution of normalized PageRank scores,

$$\hat{\mu}_{n,PR} := n^{-1} \sum_{v \in \mathcal{T}_n(\mathbf{p})} \delta \{ R_{v,c}(n) \}.$$

General results on the implication of local weak convergence of sparse graphs on the convergence of the empirical distribution of PageRank scores was derived in [7, 27]. In particular, the local weak convergence in Theorem 3.3 coupled with [7, 27] leads to the following result. Recall the finite rooted random tree $\mathcal{T}^*(\tau, \mathbf{p})$ from Definition 3.1 and let $P_l(\emptyset)$ denote the number of directed paths of length l that end at the root denoted as \emptyset . Define the *normalized* PageRank score at \emptyset as,

$$\mathcal{R}_{\emptyset,c}(\infty) = (1 - c) \left(1 + \sum_{l=1}^{\infty} c^l P_l(\emptyset) \right). \quad (3.13)$$

Corollary 3.15 (PageRank asymptotics) *Under Assumption 3.2, the random variable $\mathcal{R}_{\emptyset,c}(\infty)$ is finite a.s. Further for every continuity point r of the distribution of $\mathcal{R}_{\emptyset,c}(\infty)$,*

$$n^{-1} \sum_{v \in \mathcal{T}_n(\mathbf{p})} \mathbb{1} \{ R_{v,c}(n) > r \} \xrightarrow{P} \mathbb{P}(\mathcal{R}_{\emptyset,c}(\infty) > r).$$

The next result which, in particular, displays a qualitative phase transition of the tail exponent of the limiting empirical distribution of the PageRank scores in the fringe regime (Theorem 3.16(b)), has not been observed before in the literature and continues the vein of results displayed in Theorem 3.14(b).

Theorem 3.16 (Tail behavior of PageRank distribution) *Under Assumptions 3.2 and 3.5,*

(a) **Non-fringe regime:** When $\mathbb{E}[Z] > 1$, for $c \in (0, 1)$,

$$\begin{aligned} -R &\leq \liminf_{r \rightarrow \infty} \frac{\log(\mathbb{P}(\mathcal{R}_{\emptyset,c}(\infty) \geq r))}{\log r} \leq \limsup_{r \rightarrow \infty} \frac{\log(\mathbb{P}(\mathcal{R}_{\emptyset,c}(\infty) \geq r))}{\log r} \\ &\leq - \left(R \wedge \frac{\log q_*}{\log s_0} \right). \end{aligned}$$

where as before, $q_* < 1$ is the solution to $f(q) = q$.

(b) **Fringe regime:** When $\mathbb{E}[Z] \leq 1$,

$$\lim_{r \rightarrow \infty} \frac{\log(\mathbb{P}(\mathcal{R}_{\emptyset, c}(\infty) \geq r))}{\log r} = \begin{cases} -R & \text{for } c \in (0, s_0^{-1}] \text{ with } c < 1, \\ -\frac{1}{cf(1/c)} & \text{for } c \in (s_0^{-1}, 1), \text{ provided } s_0 > 1. \end{cases}$$

Remark 8 (*PageRank and the power-law hypothesis*) Since its introduction by Brin and Page [54], PageRank has been largely and successfully used to identify influential nodes in a variety of network models [4, 30, 32]. Although PageRank ‘looks beyond’ degrees and captures more intricate local geometry around vertices, it can be computed efficiently in a distributed fashion. The well known power-law hypothesis conjectures that for real world networks with power-law (in)degree distribution, the PageRank also has a power-law distribution with the same exponent as the degree. This has been shown to hold in several *static* network models like the directed configuration model [21, 53] and the inhomogeneous random digraph [43, 53]. Recently, [7] showed that the power-law hypothesis is false for the affine preferential attachment model: the PageRank distribution has a strictly heavier tail than the degrees. This suggests that for *dynamic* networks (evolving over time), the PageRank captures strictly more information than the empirical degree structure even at the level of large deviations.

In Theorem 3.16, we show for the first time a *phase transition* for the limiting PageRank distribution in a random network model: in the fringe regime, the power law hypothesis holds for damping factor $c \in (0, s_0^{-1}]$, but the PageRank tail becomes heavier for larger c . Intuitively, when c crosses a threshold, the PageRank score incorporates information from a large enough local neighborhood of each vertex so as to distinguish the extremal behavior of the limiting PageRank and degree distributions. Identifying the class of network models for which such phase transitions occur should further quantify the efficacy and limits of the power-law hypothesis.

Remark 9 Although we explicitly characterize the tail exponent for the limiting degree and PageRank distribution in the fringe regime (see Theorem 3.8(i) and Theorem 3.16(b)), only upper and lower bounds are established on this exponent in the non-fringe regime (see Theorem 3.8(ii) and Theorem 3.16(a)). However, computation of the quantities R, q_*, s_0 in the specific models below suggests that the upper and lower bounds actually match even in the non-fringe regime unless $\mathbb{E}[Z]$ is much greater than 1.

(i) *PageRank driven preferential attachment:* This model was described in example (c) of Sect. 1.2 (see also Remark 7). In this case, $p_k = p(1-p)^k$, $k \geq 0$, and $\mathbb{E}[Z] = \frac{1}{p} - 1$. Thus, the non-fringe regime corresponds to $p \in [0, 1/2)$. Writing $q = 1-p$, the pgf is $f(s) = p/(1-qs)$, $s \in [0, 1/q)$, which gives $R = 1/(4pq)$, $q_* = (p/q) \wedge 1$ and $s_0 = 1/(2q)$. Hence, the upper bound in the tail exponent in the non-fringe regime is given by

$$-\left(R \wedge \frac{\log q_*}{\log s_0}\right) = -\left(\frac{1}{4pq} \wedge \frac{\log(q/p)}{\log(2q)}\right) = \begin{cases} -\frac{1}{4pq} & \text{for } p \in [p'_0, 1/2), \\ -\frac{\log(q/p)}{\log(2q)} & \text{for } p \in (0, p'_0). \end{cases}$$

where $p'_0 \approx 0.0616$ is obtained using Mathematica. In particular, the tail exponent for the limiting degree distribution is exactly $-1/(4pq)$ for all $p \in [p'_0, 1]$. For the limiting PageRank, as described in Theorem 3.16(b), we see a phase transition in the fringe regime as the damping factor c crosses s_0^{-1} . In the non-fringe regime, for $p \in [p'_0, 1/2]$, the tail exponent equals $-1/(4pq)$ for all values of the damping factor $c \in (0, 1)$.

(ii) *Simple random walk driven attachment*: Here, we take $p_0 = p = 1 - p_2$. Thus, at each attachment, the distance of the new vertex from the root behaves like a (biased) simple random walk with increment ± 1 . $\mathbb{E}[Z] = 2(1 - p)$, and thus the non-fringe regime corresponds to $p \in [0, 1/2]$. In this case, writing $q = 1 - p$, the pgf is $f(s) = p + qs^2$, $s \geq 0$, which gives $R = 1/(2\sqrt{pq})$, $q_* = p/q \wedge 1$ and $s_0 = \sqrt{p/q}$. The upper bound in the tail exponent in the non-fringe regime is given in this case by $-\left(R \wedge \frac{\log q_*}{\log s_0}\right) = -\left(\frac{1}{2\sqrt{pq}} \wedge 2\right)$. In particular, the tail exponent is exactly characterized for $p \in (\frac{1}{2} - \frac{\sqrt{3}}{4}, 1)$.

The disparity between the upper and lower bounds in the non-fringe regime may appear to be an artifact of our proof techniques (see Theorem 6.2(iii) where this discrepancy appears from estimating moments of a functional of the process). However, preliminary calculations using the many-to-few formula for branching random walks [31] (which gives more refined estimates) suggest that this discrepancy might be qualitative in nature due to certain rare events that affect the tail exponent when $\mathbb{E}[Z] \gg 1$. We will investigate this in future work.

Remark 10 (*Affine preferential attachment*) As discussed before, affine preferential attachment with attachment function $f_{att}(k) = k + (1 - 2p)/p$, $p \in (0, 1)$, is a special case of our model, corresponding to $p_0 = 1 - p$, $p_1 = p$. One can easily verify that in this case $s_0 = \infty$ and $R = 1/p$. Although we assume $p_0 + p_1 < 1$, most of our proof techniques can be extended in a straightforward manner to the case $p_0 + p_1 = 1$. Extrapolation of our results to the affine case recovers several known results which we now describe. In this case, since $\mathbb{E}[Z] < 1$, we are always in the fringe regime. Theorem 3.8(i) implies that the limiting degree distribution has a power-law with exponent $R = 1/p$. This is well known (see e.g. [19, 59]). The limiting PageRank distribution, with damping factor $c \in (0, 1)$, turns out to have a heavier tail than the degrees, with exponent $1/((1 - p)c + p)$, which was recently shown in [7].

4 Proofs: Technical foundations

4.1 Roadmap of the proofs

The goal of this section is to build the technical underpinnings for the proofs of the main results, as well as elucidate the connections between the functionals of the model and corresponding core areas in probability.

(i) Section 4.2 deals with properties of the continuous time embeddings in Definition 2.1; these are then used in Sect. 5 to prove Theorem 3.3.

- (ii) Section 4.3 describes stochastic orderings between the height and extremal displacements of associated (upper and lower bounding) branching random walks. Further, large deviations estimates are derived for branching random walks. These results are then used in Sect. 8 to prove Theorem 3.12.
- (iii) While Sect. 4.2 and 4.3 deal with direct embeddings of the process, the next few subsections describe “non-obvious” embeddings. Section 4.4 derives connections between the distance profile (in continuous time) and functionals of the quasi-stationary distribution of the random walk in (3.1) through two infinite dimensional matrices in (4.6), spectral properties of which result in the key role of the constants R, s_0, q_* in the main results.
- (iv) Analysis of truncations of the height profile leads to finite dimensional urn models and their corresponding Athreya–Karlin embeddings in finite dimensional multitype branching processes in Sect. 4.5; asymptotics of these processes, in particular as the level of truncation $K \rightarrow \infty$ needs a careful analysis of the Perron–Frobenius eigenvalue, since the corresponding limit infinite dimensional operator is non-compact; this analysis culminates in Proposition 4.16. These results form the core ingredients in obtaining lower bounds connected to evaluating power-law exponents of the degree (Theorem 3.8) in Sect. 6 and PageRank distribution (Theorem 3.16) in Sect. 9 respectively. They are also used for lower bounds in the analysis of degree and PageRank of fixed (non-root) vertices (Theorem 3.10 in Sect. 7 and Theorem 3.14 in Sect. 9).
- (v) Asymptotics of the degree and PageRank of the root necessitate the construction and analysis of an infinite dimensional multitype branching process (MTBP) in Sect. 4.6. These results play a central role in the proof of Theorem 3.9 in Sect. 7 and Theorem 3.14 in Sect. 9. Technical properties related to α -recurrence and transience of kernels arising in the analysis of the MTBP are proven in Appendix A.

4.2 Size estimates on the continuous time embedding

Recall, from Definition 2.1, the construction of the tree process in continuous time. Let $n(t) = |\mathcal{T}(t, \mathbf{p})|$.

Definition 4.1 (*Rate ν Yule process*) Fix $\nu > 0$. A rate ν Yule process is a pure birth process $\{Y_\nu(t) : t \geq 0\}$ with $Y_\nu(0) = k \in \mathbb{N}$ and where the rate of birth of new individuals is equal to ν times the size of the current population. More precisely, $\mathbb{P}(Y_\nu(t + dt) - Y_\nu(t) = 1 | \mathcal{F}(t)) := \nu Y_\nu(t) dt + o(dt)$ and $\mathbb{P}(Y_\nu(t + dt) - Y_\nu(t) \geq 2 | \mathcal{F}(t)) := o(dt)$, where $\{\mathcal{F}(t) : t \geq 0\}$ is the natural filtration of the process. Write $\{\text{Yule}(t) : t \geq 0\}$ for the corresponding forest valued (tree valued if $k = 1$) rate one process that keeps track of the genealogy of the process.

The following is a standard property of the Yule process.

Lemma 4.2 ([51, Section 2.5]) Fix $t > 0$ and rate $\nu > 0$ and assume $Y_\nu(0) = 1$. Then $Y_\nu(t)$ has a Geometric distribution with parameter $p = e^{-\nu t}$. Precisely, $\mathbb{P}(Y_\nu(t) = k) = e^{-\nu t} (1 - e^{-\nu t})^{k-1}$, $k \geq 1$. The process $\{Y_\nu(t) \exp(-\nu t) : t \geq 0\}$

is an \mathbb{L}^2 bounded martingale and thus $\exists W > 0$ such that $Y_v(t) \exp(-vt) \xrightarrow{a.s.} W$. Further $W \sim \text{Exp}(1)$.

This property leads directly to the next two results.

Lemma 4.3 *Let $n(t) = |\mathcal{T}(t, p)|$. Then $\{n(t) : t \geq 0\}$ has the same distribution as a rate one Yule process started with one individual at time zero. By Lemma 4.2,*

- (a) $\{e^{-t}n(t) : t \geq 0\}$ is an \mathbb{L}^2 bounded martingale, $e^{-t}n(t) \xrightarrow{a.s., \mathbb{L}^2} W$ with $W \sim \text{Exp}(1)$.
- (b) Defining $T_n = \inf\{t \geq 0 : n(t) = n + 1\}$, then $T_n - \log n \xrightarrow{a.s.} -\log W$ as $n \rightarrow \infty$.

Lemma 4.4 *Let $0 \leq s < t$. We have*

$$\begin{aligned}\mathbb{E}[(e^{-t}n(t))^2 | \mathcal{F}_s] &\leq (e^{-s}n(s))^2 + e^{-2s}n(s), \\ \mathbb{E}[(e^{-t}n(t))^3 | \mathcal{F}_s] &\leq 8(e^{-s}n(s))^3.\end{aligned}$$

Proof Applying the generator \mathcal{L} of the Yule process on $(e^{-t}n(t))^2$ we get,

$$\mathcal{L}(e^{-t}n(t))^2 = e^{-2t}n(t) \left((n(t) + 1)^2 - n(t)^2 \right) - 2e^{-2t}n^2(t) = e^{-2t}n(t).$$

It follows that $\{\bar{M}_2(t) : t \geq 0\}$ defined next is a martingale:

$$\bar{M}_2(t) := (e^{-t}n(t))^2 - \int_0^t e^{-2u}n(u)du$$

Thus

$$\begin{aligned}\mathbb{E}[(e^{-t}n(t))^2 | \mathcal{F}_s] &= (e^{-s}n(s))^2 + \int_s^t e^{-2u}\mathbb{E}[n(u) | \mathcal{F}_s]du \\ &\leq (e^{-s}n(s))^2 + e^{-2s}n(s).\end{aligned}$$

The second assertion of the Lemma follows the same reasoning, starting with the application of the generator on $(e^{-t}n(t))^3$. We omit the details. \square

Lemma 4.5 *Consider a rate one Yule process $\{\text{Yule}(t) : t \geq 0\}$ started with a single individual at $t = 0$. Let $\text{ht}(t)$ denote the corresponding height (maximal distance from the root) of the corresponding genealogical tree at time t . For any $\beta \geq 1$, we have $\mathbb{E}(\beta^{\text{ht}(t)}) \leq 2e^{2\beta t} < \infty$.*

Proof Let $\mathcal{Z}_n(t)$ denote the number of n -th generation individuals born before time t and B_n denote the time of the first birth in the n -th generation. Then for $\theta > 0$,

$$\mathbb{E}(\beta^{\text{ht}(t)}) \leq \sum_{n=0}^{\infty} \beta^n \mathbb{P}(\text{ht}(t) \geq n) = \sum_{n=0}^{\infty} \beta^n \mathbb{P}(B_n \leq t) \leq \sum_{n=0}^{\infty} \beta^n e^{\theta t} \mathbb{E}[e^{-\theta B_n}]. \quad (4.1)$$

Theorem 1 in [39] implies that $\mathbb{E}[e^{-\theta B_n}] \leq \psi(\theta)^n$ where

$$\psi(\theta) = \int_0^\infty \theta e^{-\theta t} \mathbb{E}[\mathcal{Z}_1(t)] dt = \frac{1}{\theta}.$$

Thus $\mathbb{E}[e^{-\theta B_n}] \leq \theta^{-n}$. Using this in (4.1) with $\theta = 2\beta$ completes the proof. \square

4.3 Branching random walks

Recall, the definition of BRW, in Definition 3.13. Consider the following variations in step (b) of the dynamics:

- (b)' If the prospective location of a new particle is at zero or below, it is “reflected” to location one. Then in terms of locations describing graph distance to the root, this is precisely the evolution of distances in \mathcal{T} . Thus the height $\mathcal{H}_{\mathcal{T}(t)}$ is precisely the location of the rightmost particle.
- (b)'' If the prospective location of a new particle is zero or below it is killed (removed from the system). This gives the distance process in \mathcal{T}^* . As before the height $\mathcal{H}_{\mathcal{T}^*(t)}$ is precisely the location of the rightmost particle in this process.

From the description of the dynamics, the following is obvious.

Lemma 4.6 *Let $B(t)$ be the rightmost particle in $\text{BRW}(t)$. One can couple \mathcal{T} , \mathcal{T}^* , BRW on a common probability space such that for all $t \geq 0$, $\mathcal{H}_{\mathcal{T}^*(t)} \leq B(t) \leq \mathcal{H}_{\mathcal{T}(t)}$.*

For BRW, the offspring process for each individual is a rate one Poisson process and the corresponding branching process (without location information) is a Yule process. In particular, the offspring process is non-lattice and underlying branching process is supercritical. Following [13], write $\mu(dz, d\tau) = \sum_{k=0}^\infty p_k \delta_{1-k}(dz) \otimes d\tau$ for the mean intensity measure of the walk. Define for $\theta \in \mathbb{R}$, $\phi \geq 0$ the following functionals:

$$\begin{aligned} m(\theta, \phi) &= \int e^{-\theta z - \phi \tau} \mu(dz, d\tau) = \int_0^\infty \sum_{k=0}^\infty e^{-\theta(1-k)} e^{-\phi \tau} p_k d\tau = \frac{f(e^\theta)}{e^\theta \phi}. \\ \alpha(\theta) &= \inf\{\phi : m(\theta, \phi) \leq 1\} = \frac{f(e^\theta)}{e^\theta}. \\ \alpha^*(x) &= \inf_{\theta < 0} \{x\theta + \alpha(\theta)\} = \inf_{s \in (0,1)} \left\{ x \log s + \frac{f(s)}{s} \right\}. \end{aligned} \quad (4.2)$$

We will be using the results in [13] that concern a very general branching random walk model, where particles are allowed to move after birth. In comparison, a particle in our branching random walk BRW performs no further movement beyond the initial displacement from its parent at birth. It would be straightforward for the curious reader to verify that BRW satisfies the mild assumptions in [13], so we will refrain from repeating those detailed assumptions here.

We will rephrase Theorem 4 in [13] into the following proposition, which will be used to prove large deviations results for BRW later.

Proposition 4.7 Let $N_t[xt, \infty)$ denote the number of particles in $\text{BRW}(t)$ that lie in $[xt, \infty)$. For all $x \neq \sup\{y : \alpha^*(y) > -\infty\}$,

$$\frac{\log(\mathbb{E}[N_t[xt, \infty)])}{t} \rightarrow \alpha^*(x) \quad \text{as } t \rightarrow \infty.$$

Recall from Definition 3.11 that

$$\kappa_0 = \inf_{s \in (0, 1)} \frac{f(s)}{s \log(1/s)}.$$

The following result follows from [13, Corollary 2].

Proposition 4.8 For $\text{BRW}(t)$ the rightmost particle satisfies,

$$\frac{B(t)}{t} \rightarrow \inf\{x : \alpha^*(x) < 0\} = \kappa_0 \quad \text{almost surely.} \quad (4.3)$$

Proof The convergence $B(t)/t \rightarrow \inf\{x : \alpha^*(x) < 0\}$ is proven in [13, Corollary 2]. To see that $\kappa_0 = \inf\{x : \alpha^*(x) < 0\}$, note that for $s \in (0, 1)$, $x \log s + \frac{f(s)}{s} < 0$ is equivalent to $x > \frac{f(s)}{s \log(1/s)}$. Thus, $x > \inf_{s \in (0, 1)} \frac{f(s)}{s \log(1/s)}$ if and only if $\alpha^*(x) < 0$. \square

Next we state a large deviations result for $\mathbb{P}(B(t) \geq xt)$ for large t when $x > \kappa_0$, which is the key to our proof of Theorem 3.12, and is interesting in its own right.

Lemma 4.9 With κ_0 as in (3.9) and $B(t)$ as the rightmost particle in $\text{BRW}(t)$, for $x > \kappa_0$,

$$\mathbb{P}(B(t) \geq xt) = \exp(\alpha^*(x)t + o(t)) \quad \text{as } t \rightarrow \infty.$$

Proof We start with the upper bound. For fixed $x > \kappa_0$, recall the functional $N_t[xt, \infty)$ from Proposition 4.7. Using Proposition 4.7 gives,

$$\mathbb{P}(B(t) \geq xt) \leq \mathbb{P}(N_t[xt, \infty) \geq 1) \leq \mathbb{E}[N_t[xt, \infty)] \leq \exp(\alpha^*(x)t + o(t)).$$

To prove the lower bound, we will use an argument of induction in time. For $\varepsilon \in (0, 1)$, let $\mathcal{F}_{\varepsilon t}$ denote the filtration generated by the genealogies and locations of all the particles in BRW born up till time εt . For $v \in \text{BRW}(\varepsilon t)$, let S_v denote the location of the particle v in $\text{BRW}(\varepsilon t)$. For each such v consider the branching random walk encoding the genealogy and location of particles born after time εt whose most recent common ancestor in $\text{BRW}(\varepsilon t)$ is v . For $t' \geq 0$, let $B^v(t')$ denote the location of the rightmost particle in this branching random walk originating from particle v observed t' time units after εt . It is easy to see that

$$\begin{aligned} \mathbb{P}(B(t) \geq xt \mid \mathcal{F}_{\varepsilon t}) &= \mathbb{P}(\max_{v \in \text{BRW}(\varepsilon t)} B^v((1 - \varepsilon)t) \geq xt \mid \mathcal{F}_{\varepsilon t}) \\ &= \mathbb{P}\left(\max_{v \in \text{BRW}(\varepsilon t)} \left\{ \frac{B^v((1 - \varepsilon)t) - S_v}{(1 - \varepsilon)t} + \frac{S_v}{(1 - \varepsilon)t} \right\} \geq \frac{x}{1 - \varepsilon} \mid \mathcal{F}_{\varepsilon t}\right). \end{aligned}$$

If there exists a particle $v \in \text{BRW}(\varepsilon t)$ such that both $S_v \geq x\varepsilon t$ and $B^v((1 - \varepsilon)t) - S_v \geq x(1 - \varepsilon)t$, then we will have $B(t) \geq xt$. Based on this observation, we define $D_{\varepsilon t} := \{v \in \text{BRW}(\varepsilon t) : S_v \geq x\varepsilon t\}$, noting that $D_{\varepsilon t}$ is measurable with respect to $\mathcal{F}_{\varepsilon t}$. In addition, the collection $\{B^v((1 - \varepsilon)t) - S_v : v \in \text{BRW}(\varepsilon t)\}$ comprise i.i.d. random variables, independent of $\mathcal{F}_{\varepsilon t}$, each distributed as the location of the rightmost particle in $\text{BRW}((1 - \varepsilon)t)$. Hence,

$$\begin{aligned} \mathbb{P}(B(t) \geq xt \mid \mathcal{F}_{\varepsilon t}) &\geq \mathbb{P}\left(\max_{v \in D_{\varepsilon t}} \frac{B^v((1 - \varepsilon)t) - S_v}{(1 - \varepsilon)t} \geq x \mid \mathcal{F}_{\varepsilon t}\right) \\ &= 1 - \left(1 - \mathbb{P}\left(\frac{B((1 - \varepsilon)t)}{(1 - \varepsilon)t} \geq x\right)\right)^{|D_{\varepsilon t}|} \\ &\geq \mathbb{P}\left(\frac{B((1 - \varepsilon)t)}{(1 - \varepsilon)t} \geq x\right)^{|D_{\varepsilon t}|} \left(1 - \mathbb{P}\left(\frac{B((1 - \varepsilon)t)}{(1 - \varepsilon)t} \geq x\right) \cdot |D_{\varepsilon t}|\right), \end{aligned}$$

where the last line follows from the elementary inequality $1 - (1 - x)^y \geq 1 - e^{-xy} \geq xy(1 - xy)$ for $x \in [0, 1]$, $y \geq 0$. Taking expectations on both sides of the above bound, we have

$$\mathbb{P}(B(t) \geq xt) \geq \mathbb{P}\left(\frac{B((1 - \varepsilon)t)}{(1 - \varepsilon)t} \geq x\right) \left(\mathbb{E}|D_{\varepsilon t}| - \mathbb{P}\left(\frac{B((1 - \varepsilon)t)}{(1 - \varepsilon)t} \geq x\right) \cdot \mathbb{E}(|D_{\varepsilon t}|^2)\right), \quad (4.4)$$

where $|D_{\varepsilon t}|$ denotes the size of $D_{\varepsilon t}$. Noting that $|D_{\varepsilon t}| = N_{\varepsilon t}[x\varepsilon t, \infty)$, it follows directly from Proposition 4.7 that for $x > \kappa_0$,

$$\frac{\log(\mathbb{E}|D_{\varepsilon t}|)}{\varepsilon t} \rightarrow \alpha^*(x) = \inf_{s \in (0, 1)} \{x \log s + f(s)/s\} < 0,$$

i.e.,

$$\mathbb{E}|D_{\varepsilon t}| \geq \exp(\alpha^*(x)\varepsilon t - o(t)). \quad (4.5)$$

We claim that $\mathbb{P}\left(\frac{B((1 - \varepsilon)t)}{(1 - \varepsilon)t} \geq x\right) \cdot \mathbb{E}(|D_{\varepsilon t}|^2) = o(\mathbb{E}|D_{\varepsilon t}|)$ when $\varepsilon > 0$ is chosen to be sufficiently small. To see this, we use the upper bound proved earlier to get

$$\mathbb{P}\left(\frac{B((1 - \varepsilon)t)}{(1 - \varepsilon)t} \geq x\right) \leq \exp(\alpha^*(x)(1 - \varepsilon)t + o(t)).$$

In addition, Lemma 4.4 gives a trivial upper bound on $\mathbb{E}(|D_{\varepsilon t}|^2) \leq \mathbb{E}[n(\varepsilon t)^2] \leq 2e^{2\varepsilon t}$. Knowing (4.5) and noting that $\alpha^*(x) < 0$ for $x > \kappa_0$, we can then choose $\varepsilon < \frac{\alpha^*(x)}{3\alpha^*(x) - 2}$ so that

$$\begin{aligned} \mathbb{P}\left(\frac{B((1 - \varepsilon)t)}{(1 - \varepsilon)t} \geq x\right) \cdot \mathbb{E}(|D_{\varepsilon t}|^2) &\leq 2 \exp((\alpha^*(x)(1 - \varepsilon) + 2\varepsilon)t) \leq \exp(2\alpha^*(x)\varepsilon t) \\ &= o(\mathbb{E}|D_{\varepsilon t}|). \end{aligned}$$

Therefore, when t is sufficiently large, combining the above with (4.4) and (4.5) leads to

$$\begin{aligned}\mathbb{P}(B(t) \geq xt) &\geq \frac{1}{2}\mathbb{P}\left(\frac{B((1-\varepsilon)t)}{(1-\varepsilon)t} \geq x\right) \cdot \mathbb{E}|D_{\varepsilon t}| \\ &\geq \frac{1}{2}\mathbb{P}\left(\frac{B((1-\varepsilon)t)}{(1-\varepsilon)t} \geq x\right) \exp(\alpha^*(x)\varepsilon t - o(t)),\end{aligned}$$

which then implies the following relation

$$\begin{aligned}\liminf_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{P}\left(\frac{B(t)}{t} \geq x\right) &\geq \alpha^*(x)\varepsilon \\ &+ (1-\varepsilon) \liminf_{t \rightarrow \infty} \frac{1}{(1-\varepsilon)t} \log \mathbb{P}\left(\frac{B((1-\varepsilon)t)}{(1-\varepsilon)t} \geq x\right).\end{aligned}$$

This proves our desired conclusion

$$\mathbb{P}\left(\frac{B(t)}{t} \geq x\right) \geq \exp(\alpha^*(x)t - o(t)).$$

□

4.4 Connection to Quasi stationary random walks

We start by clarifying the appearance of the mysterious functionals such as R , s_0 etc in the statement of the main results. Recall the process $\mathcal{T}^*(\cdot)$, $\mathcal{T}(\cdot)$ in Definition 2.1. For $i \geq 0$, let $\mathcal{P}_i(t)$ denote the number of vertices at distance i to the root in $\mathcal{T}^*(t)$ with $\mathcal{P}_0(t) \equiv 1$ for all t . Write $\mathcal{P}(t) = (\mathcal{P}_0(t), \mathcal{P}_1(t), \dots)'$ for the entire column vector. Let $\tilde{\mathcal{P}}_i(t)$ and $\tilde{\mathcal{P}}(t) = (\tilde{\mathcal{P}}_0(t), \tilde{\mathcal{P}}_1(t), \dots)'$ denote the analogous objects for \mathcal{T} . Define the two (infinite dimensional) matrices $\mathbf{A} = (A_{ij})_{i,j \geq 0}$ and $\mathbf{B} = (B_{ij})_{i,j \geq 0}$,

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & \dots \\ p_0 & p_1 & p_2 & \dots \\ 0 & p_0 & p_1 & \dots \\ 0 & 0 & p_0 & \dots \\ 0 & \dots & & \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 0 & 0 & \dots \\ c_0 & c_1 & c_2 & \dots \\ 0 & p_0 & p_1 & \dots \\ 0 & 0 & p_0 & \dots \\ \dots & & & \end{pmatrix}, \quad (4.6)$$

where $c_i = \sum_{k=i}^{\infty} p_k$. In particular, $c_0 = 1$. The following is easy to check from the evolution dynamics. We omit the proof.

Lemma 4.10 *For $t \geq 0$,*

$$\frac{d}{dt} \mathbb{E}[\mathcal{P}(t)] = \mathbf{A} \cdot \mathbb{E}[\mathcal{P}(t)], \quad \frac{d}{dt} \mathbb{E}[\tilde{\mathcal{P}}(t)] = \mathbf{B} \cdot \mathbb{E}[\tilde{\mathcal{P}}(t)].$$

Now we will come to the first connection between the tree valued process and random walk $\{S_n : n \geq 0\}$ as in (3.1) with increments distributed as $Z - 1$. Recall the hitting times of zero started from location k namely \bar{T}_k in (3.3). The next lemma intuitively follows from solving the first differential equation above and noting that \mathbf{A} is the transition matrix for the Markov chain $\{S_n : n \geq 0\}$.

Lemma 4.11 *For all $k \geq 1$,*

$$\mathbb{E}[\mathcal{P}_k(t)] = \sum_{i=0}^{\infty} \frac{t^i}{i!} \mathbb{P}(\bar{T}_k = i).$$

Proof From Lemma 4.10, we have,

$$\frac{d}{dt} \mathbb{E}[\mathcal{P}_k(t)] = \sum_{j=0}^{\infty} A_{kj} \mathbb{E}[\mathcal{P}_j(t)]. \quad (4.7)$$

Let $f_k(t) := \sum_{i=0}^{\infty} \frac{t^i}{i!} \mathbb{P}(\bar{T}_k = i)$ for $k \geq 1$ and write $f_0(t) \equiv 1$. Recalling the matrix $\mathbf{A} = (A_{kj})_{k \geq 0, j \geq 0}$ from (4.6), note that

$$\begin{aligned} f'_k(t) &= \sum_{i=0}^{\infty} \frac{t^i}{i!} \mathbb{P}(\bar{T}_k = i + 1) = \mathbb{P}(\bar{T}_k = 1) + \sum_{i=1}^{\infty} \frac{t^i}{i!} \sum_{j=1}^{\infty} A_{kj} \mathbb{P}(\bar{T}_j = i) \\ &= A_{k0} + \sum_{j=1}^{\infty} A_{kj} f_j(t) = \sum_{j=0}^{\infty} A_{kj} f_j(t). \end{aligned}$$

We then compare this system of ODEs with (4.7). Let $h_k(t) = f_k(t) - \mathbb{E}[\mathcal{P}_k(t)]$ for $t \geq 0$. We have $h_k(0) = 0$ and $h'_k(t) = \sum_{j=0}^{\infty} A_{kj} h_j(t)$ for all $k \geq 1$. Thus,

$$|h_k(t)| \leq \sum_{j=0}^{\infty} A_{kj} \int_0^t |h_j(s)| ds.$$

Let $h^*(t) = \sum_{k=0}^{\infty} |h_k(t)|$. It is not difficult to check $h^*(\cdot)$ is continuous. Hence,

$$h^*(t) \leq \sum_{j=0}^{\infty} \left(\sum_{k=0}^{\infty} A_{kj} \right) \int_0^t |h_j(s)| ds \leq \int_0^t h^*(s) ds.$$

By Grönwall's inequality, $h^*(t) \equiv 0$. This proves the lemma. \square

Given the connection between random walks and the tree evolution in Lemma 4.11, it is clear that spectral properties of \mathbf{A} and \mathbf{B} are key to understanding the evolution of tree functionals. The next result derives some properties. In all the ensuing results in this Subsection, we will **always** make the Assumptions 3.2 and 3.5.

Proposition 4.12 (Spectral properties of \mathbf{A} , \mathbf{B})

- (a) For any positive s such that $f(s) < \infty$, the vector $\mathbf{v}_s = (v_i(s) : i \geq 0)$ with $v_i(s) = s^{-i}$, $i \geq 0$, is a non-negative left sub-invariant eigenvector of \mathbf{A} for eigenvalue $f(s)/s$.
- (b) For any $s \geq 1$ such that $f(s) < \infty$, the vector $\mathbf{v}_s = (v_i(s) : i \geq 0)$ as before is a non-negative left sub-invariant eigenvector of \mathbf{B} for eigenvalue $f(s)/s$.
- (c) Recall q_* from Definition 3.6. When $\mathbb{E}[Z] > 1$, the vector $\mathbf{u} = (u_i : i \geq 0)$ with $u_0 = 0$ and for $i \geq 1$ $u_i = q_*^{i-1}$ is a non-negative right eigenvector of \mathbf{B} .

Proof (a) For $j \geq 0$,

$$\sum_{i=1}^{\infty} s^{-i} A_{ij} = \sum_{i=1}^{j+1} s^{-i} p_{j+1-i} = s^{-(j+1)} \sum_{i=1}^{j+1} s^{j+1-i} p_{j+1-i} \leq s^{-j} \frac{f(s)}{s}.$$

(b) Recall $c_j = \sum_{l=j}^{\infty} p_l$. For $j \geq 0$,

$$\begin{aligned} \sum_{i=1}^{\infty} s^{-i} B_{ij} &= \sum_{i=1}^{j+1} s^{-i} p_{j+1-i} + s^{-1} c_{j+1} = s^{-(j+1)} \left(\sum_{i=1}^{j+1} s^{j+1-i} p_{j+1-i} + s^j c_{j+1} \right) \\ &\leq s^{-(j+1)} \left(\sum_{i=0}^j s^i p_i + \sum_{i=j+1}^{\infty} s^i p_i \right) = s^{-j} \frac{f(s)}{s}, \end{aligned}$$

where, in the first inequality, we have used $s \geq 1$.

(c) Note that since $f(q_*) = q_*$,

$$\begin{aligned} (\mathbf{Bu})_1 &= \sum_{k=0}^{\infty} c_k u_k = \sum_{k=1}^{\infty} \left(1 - \sum_{l=0}^{k-1} p_l \right) q_*^{k-1} = \left(\frac{1}{1 - q_*} - \sum_{l=0}^{\infty} p_l \sum_{k=l+1}^{\infty} q_*^{k-1} \right) \\ &= \frac{1}{1 - q_*} - \frac{f(q_*)}{1 - q_*} = 1 = u_1. \end{aligned}$$

For $j \geq 2$,

$$(\mathbf{Bu})_j = \sum_{k=0}^{\infty} p_k u_{k+j-1} = \sum_{k=0}^{\infty} p_k q_*^{k+j-2} = q_*^{j-2} f(q_*) = q_*^{j-1} = u_j.$$

Clearly, $(\mathbf{Bu})_0 = 0 = u_0$. □

4.5 Urn models and multitype branching processes

This section connects the evolution of \mathcal{P} with urn models and eventually to finite dimensional multitype branching processes. We will consider a truncated version of \mathcal{P} where we keep track of vertices at distance at most $k \geq 1$ from the root for some

fixed k . Recall the matrix \mathbf{A} from (4.6) and let \mathbf{A}_k denote a $k \times k$ matrix such that $(\mathbf{A}_k)_{ij} = A_{ij}$ for $1 \leq i, j \leq k$. Let $\mathbf{e}_j^{(k)}$ denote the unit basis vector in \mathbb{R}^k with one in the j -th co-ordinate and zero elsewhere.

Definition 4.13 (*Urn model encoding distance from the root*) Fix $k \geq 1$. Consider a generalized multitype Polya-urn process with types $\{1, \dots, k\}$ starting with a single ball of type 1. When a ball of type i is drawn, it is returned along with $\xi_i = (\xi_{i1}, \dots, \xi_{ik})$ other balls, where

$$\begin{cases} \mathbb{P}(\xi_i = \mathbf{e}_j^{(k)}) = A_{ji} & (1 \leq j \leq k) \\ \mathbb{P}(\xi_i = \mathbf{0}) = 1 - \sum_{j=1}^k A_{ji} \end{cases}$$

It is easy to see $\mathbb{E}(\xi_{ij}) = A_{ji}$.

A key tool in studying such urn processes is the so-called Athreya-Karlin embedding in finite multitype branching processes [6, 38], which in this case corresponds to the following: we start with one individual (ball) at time zero of type 1. Each individual of type i lives for an exponential mean one unit of time and upon dying gives birth to an offspring of type i and possibly another vertex whose type is determined by an i.i.d. sample of ξ_i (only one offspring of type i is produced if the sampled $\xi_i = 0$). In this continuous time process let $\bar{\mathcal{P}}^k(t) = (\bar{\mathcal{P}}_i(t) : 1 \leq i \leq k)$ denote the individuals (balls) of various types alive at time t .

Lemma 4.14 *Let σ_1 denote the first time that the root reproduces a vertex at distance one in \mathcal{T}^* so that $\mathcal{P}_1(\sigma_1) = 1$. We can couple $\bar{\mathcal{P}}^k(\cdot)$ with the process $\mathcal{P}(\cdot)$ defined in Sect. 4.4 such that $\mathcal{P}_1(t + \sigma_1) \geq \bar{\mathcal{P}}_1(t)$ for all t .*

Proof We describe an explicit coupling:

- (i) If at time $t + \sigma_1$, a vertex in \mathcal{T}^* at distance $1 \leq i \leq k$ from the root reproduces to produce a vertex at distance $1 \leq j \leq k$, a ball of type i is removed from the urn at time t and is returned along with a ball of type j .
- (ii) If the new vertex is at distance greater than k from the root, or if the root or a vertex of distance greater than k reproduces, then no change is made to the urn model.

It is clear from our construction that $\mathcal{P}_1(t + \sigma_1) \geq \bar{\mathcal{P}}_1(t)$ for all $t \geq 0$. \square

The following is obtained by a direct application of [38, Theorem 3.1]. Define $k_0 := \inf\{k \geq 1 : p_k > 0\}$. Note that $k_0 < \infty$ as $p_0 < 1$.

Proposition 4.15 *For $k \geq k_0$, \mathbf{A}_k possesses a positive largest (Perron–Frobenius) eigenvalue α_k and a strictly positive right eigenvector $\bar{\mathbf{v}}_k = (\bar{v}_k(i) : 1 \leq i \leq k)$ such that,*

$$e^{-\alpha_k t} \bar{\mathcal{P}}_1^k(t) \xrightarrow{a.s.} W \bar{\mathbf{v}}_k(1)$$

for a strictly positive random variable W .

Proof Recalling $p_0 + p_1 < 1$ and $p_0 > 0$, it follows that \mathbf{A}_k is irreducible in the sense of [38] for every $k \geq k_0$ and one easily sees that assumptions (A1)-(A6) in [38] are satisfied, required for the application of [38, Theorem 3.1]. The strict positivity of W follows from the fact that extinction is impossible in our case (see [38, Lemma 2.1]). \square

Thus, for $k \geq k_0$, α_k plays a crucial role in the growth of $\tilde{\mathcal{P}}_1^k(\cdot)$. The next result describes its asymptotics as $k \uparrow \infty$.

Proposition 4.16 $\alpha_k \uparrow 1/R$ as $k \rightarrow \infty$.

Remark 11 In principle we are asserting that the maximal eigenvalue of the submatrix \mathbf{A}_k of the infinite dimensional operator \mathbf{A} converges in the limit $k \rightarrow \infty$ to the maximal eigenvalue of \mathbf{A} . If \mathbf{A} was a compact operator then this would follow from standard function analytic methods, however it can be checked that \mathbf{A} is *not a compact operator* and thus we give a proof relying on the specific probabilistic interpretation of \mathbf{A} and the corresponding random walk.

Remark 12 In Sect. 4.6 we will construct an infinite dimensional multitype branching process, driven by the matrix \mathbf{B} to track the **entire** height profile of \mathcal{T} . The analogous argument for \mathcal{T}^* does not work, since the corresponding process with \mathbf{B} replaced by matrix \mathbf{A} is **not** α -recurrent (Lemma A.5) and thus necessitates the truncation scheme here.

Proof Let $\{Z_i\}_{i \geq 1}$ be a collection of i.i.d. random variables distributed as \mathbf{p} and recall that $S_n = S_0 + \sum_{i=1}^n (Z_i - 1)$. Define the stopping times,

$$T = \inf\{n \geq 0 : S_n = 0\}, \quad \tau = \inf\{n \geq 1 : S_n = 1\}, \\ \tau_k = \inf\{n \geq 0 : S_n > k\}, \quad k \geq 1.$$

In this proof, for $j \in \mathbb{Z}$, let \mathbb{P}_j and \mathbb{E}_j denote the probability and expectation operators for the walk started from $S_0 = j$. For $r \geq 0$, define $\Phi(r) = \mathbb{E}_1[e^{r\tau} \cdot 1_{\{\tau < T\}}]$. The proof of the Proposition hinges on the following lemma, whose proof is postponed to the end of this section. \square

Lemma 4.17 $\Phi(\log R) < 1$ and $\Phi(r) = \infty$ when $r > \log R$.

Choose any $k \geq k_0$. As \mathbf{A}_k is substochastic, $\alpha_k \leq 1$ by standard Perron–Frobenius theory. Write $\alpha_k = e^{-\theta_k}$ for $\theta_k \geq 0$. We claim that $\theta_k > 0$ and

$$\mathbb{E}_1[e^{\theta_k \tau} 1_{\{\tau < T \wedge \tau_k\}}] = 1. \quad (4.8)$$

To see this, consider the right eigenvector $\bar{\mathbf{v}}_k$ of \mathbf{A}_k for $k \geq 1$. Then, $\mathbf{P}_{ij} = \frac{A_{ij} \bar{v}_k(j)}{\alpha_k \bar{v}_k(i)}$, $1 \leq i, j \leq k$ defines a probability transition matrix. Denote by \mathbb{Q}_1 the law of the Markov chain associated with \mathbf{P} starting from state 1. By slight abuse of notation,

keep using T, τ, τ_i as above, to denote the associated stopping times for this Markov chain. Then $\mathbb{Q}_1(\tau = 1) = \mathbb{P}_{11} = \frac{A_{11}}{\alpha_k}$. For $i \geq 2$, observe that

$$\begin{aligned}\mathbb{Q}_1(\tau = i) &= \sum_{\{z_j\}_{1 \leq j \leq i-1: 2 \leq z_j \leq k}} \frac{1}{\alpha_k^i} A_{1z_1} A_{z_1 z_2} \cdots A_{z_{i-1} 1} \\ &= e^{i\theta_k} \cdot \mathbb{P}_1(\tau = i, i < T \wedge \tau_k).\end{aligned}$$

Since \mathbf{P} is irreducible (and finite),

$$1 = \mathbb{Q}_1(\tau < \infty) = \mathbb{E}_1 \left[e^{\theta_k \tau} 1_{\{\tau < T \wedge \tau_k\}} \right],$$

i.e., (4.8) holds. Since $p_0 > 0$, $\mathbb{P}_1(\tau < T \wedge \tau_k) < 1$ and this implies that $\theta_k > 0$.

Note that as

$$1 = \mathbb{E}_1 \left[e^{\theta_{k+1} \tau} 1_{\{\tau < T \wedge \tau_{k+1}\}} \right] = \mathbb{E}_1 \left[e^{\theta_k \tau} 1_{\{\tau < T \wedge \tau_k\}} \right] \leq \mathbb{E}_1 \left[e^{\theta_k \tau} 1_{\{\tau < T \wedge \tau_{k+1}\}} \right],$$

we have $\theta_{k+1} \leq \theta_k$ for any $k \geq 1$ and thus $\theta_k \downarrow \theta^*$ for some $\theta^* \geq 0$.

It remains to show that $\theta^* = \log R$, which implies $\alpha_k = e^{-\theta_k} \uparrow e^{-\theta^*} = 1/R$ as $k \rightarrow \infty$. First we prove $\theta^* \geq \log R$. It follows from the monotonicity of $\{\theta_k\}$ and (4.8) that

$$\Phi(\theta_k) = \mathbb{E}_1[e^{\theta_k \tau} 1_{\{\tau < T\}}] = \lim_{i \rightarrow \infty} \mathbb{E}_1[e^{\theta_k \tau} 1_{\{\tau < T \wedge \tau_i\}}] \geq \lim_{i \rightarrow \infty} \mathbb{E}_1[e^{\theta_i \tau} 1_{\{\tau < T \wedge \tau_i\}}] = 1.$$

$\Phi(\cdot)$ is strictly increasing (as $\mathbb{P}_1(\tau < T) > 0$) and $\Phi(\log R) < 1$ by Lemma 4.17, implying that $\theta_k > \log R$. Therefore, $\theta^* = \lim_k \theta_k \geq \log R$.

Next, to prove $\theta^* \leq \log R$, note that

$$\Phi(\theta^*) = \mathbb{E}_1[e^{\theta^* \tau} 1_{\{\tau < T\}}] = \lim_{i \rightarrow \infty} \mathbb{E}_1[e^{\theta^* \tau} 1_{\{\tau < T \wedge \tau_i\}}] \leq \lim_{i \rightarrow \infty} \mathbb{E}_1[e^{\theta_i \tau} 1_{\{\tau < T \wedge \tau_i\}}] = 1.$$

As $\Phi(r) = \infty$ for all $r > \log R$ by Lemma 4.17, the above implies that $\theta^* \leq \log R$. The result follows. \square

Proof of Lemma 4.17 Define $\chi(u) = \sum_{j=0}^{\infty} \mathbb{P}_1(T = j)u^j$ for $u \geq 0$. Note that $\mathbb{E}_1[e^{rT}] = \chi(e^r)$. Defining $\kappa = \inf\{n \geq 0 : S_n = 1\}$, Strong Markov property implies that

$$\begin{aligned}\Phi(r) &= e^r \sum_{j=1}^{\infty} p_j \mathbb{E}_j[e^{r\kappa}] = e^r \sum_{j=1}^{\infty} p_j \left(\mathbb{E}_1[e^{rT}] \right)^{j-1} = e^r \sum_{j=1}^{\infty} p_j \chi(e^r)^{j-1} \\ &= e^r \left(\frac{f(\chi(e^r)) - p_0}{\chi(e^r)} \right) < \infty\end{aligned}\tag{4.9}$$

for all r such that $f(\chi(e^r)) < \infty$ and $\Phi(r) = \infty$ when $\chi(e^r) < \infty$ but $f(\chi(e^r)) = \infty$. We claim that

$$\chi(R) = s_0 \quad \text{and} \quad \chi(u) = \infty \quad \text{when } u > R. \quad (4.10)$$

The desired result now follows from (4.9):

$$\Phi(\log R) = R \cdot \frac{f(s_0) - p_0}{s_0} = \frac{s_0}{f(s_0)} \left(\frac{f(s_0)}{s_0} - \frac{p_0}{s_0} \right) = 1 - \frac{p_0}{f(s_0)} < 1,$$

recalling $f(s_0) < \infty$, and $\Phi(r) = \infty$ when $r > \log R$.

It remains to prove the claim in (4.10). Lemma 1 in [55] implies that

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}_1(T = n + 1)}{\mathbb{P}_1(T = n)} = \frac{1}{R}.$$

By the ratio test for power series, $\chi(u) < \infty$ if $s < R$ and $\chi(u) = \infty$ if $s > R$.

In the cases where $R = 1$, by Lemma 3.7 we have $s_0 = 1$ and $\mathbb{E}[Z] \leq 1$. It follows from [29, Lemma 11.3] that,

$$\chi(1) = \sum_{j=1}^{\infty} \mathbb{P}_1(T = j) = \mathbb{P}_1(T < \infty) = 1.$$

When $R > 1$, to show $\chi(R) = s_0$, observe that for $u < R$,

$$\chi(u) = p_0 u + \sum_{j=1}^{\infty} p_j u (\chi(u))^j = u f(\chi(u)).$$

As $\chi(u)$ is strictly increasing on $[0, R)$, it has an inverse function $\chi^{-1}(\cdot)$ given by

$$\chi^{-1}(z) = \frac{z}{f(z)}$$

for $z \in [0, \chi(R))$, i.e., $\frac{z}{f(z)} \in [0, R)$. Let $\{s_l : l \geq 1\}$ be a sequence of positive numbers such that $\lim_{l \rightarrow \infty} s_l = s_0$ and $s_l < s_0$, $\frac{s_l}{f(s_l)} \in [0, R)$ for all $l \in \mathbb{N}$. Hence,

$$\lim_{l \rightarrow \infty} \chi \left(\frac{s_l}{f(s_l)} \right) = \lim_{l \rightarrow \infty} s_l = s_0.$$

The final step is now to show $\chi(R) = \lim_{l \rightarrow \infty} \chi \left(\frac{s_l}{f(s_l)} \right)$. We define a series of functions $\{\chi_k(u) : k \geq 1\}$ with $\chi_k(u) = \sum_{j=1}^k \mathbb{P}_1(T = j) u^j$. For every $k \geq 1$, $\chi_k(u)$ is continuous and $\{\chi_k(u) : k \geq 1\}$ is a non-decreasing sequence for all $u \geq 0$. As $\chi(\cdot)$

is the limit of a monotone increasing sequence of continuous functions, it is lower semicontinuous, i.e., $\chi(R) \leq \liminf_{s \uparrow R} \chi(u)$. Therefore,

$$\chi(R) \leq \liminf_{s \uparrow R} \chi(u) \leq \lim_{l \rightarrow \infty} \chi\left(\frac{s_l}{f(s_l)}\right) = \lim_{l \rightarrow \infty} s_l = s_0.$$

Further, by the monotonicity of $\chi(\cdot)$, $\chi(R) \geq \chi(\frac{s_l}{f(s_l)})$ for any $l \in \mathbb{N}$ and hence $\chi(R) \geq \lim_{l \rightarrow \infty} \chi(s_l/f(s_l)) = s_0$. This completes the proof. \square

4.6 Infinite dimensional multitype branching processes

In this section we introduce a specific class of continuous time multitype branching processes $\{\text{MBP}(t) : t \geq 0\}$ studied in full generality in [35, 37]. The process $\text{MBP}(\cdot)$ will once again track distances in $\mathcal{T}(\cdot)$ so that a vertex at distance ℓ from the root corresponds to a type ℓ individual in $\text{MBP}(\cdot)$.

The Ulam-Harris set $I = \cup_{n \geq 0} \mathbb{N}^n$, with $N^0 = \emptyset$ representing the root, is used to encode the set of all possible individuals (vertices) in MBP , and let (S, \mathcal{S}) denote the space of types with a countably generated σ -algebra \mathcal{S} , where $S = \{0, 1, \dots\}$. Start with a single individual (the root) of some type $r \in S$. Any existing individual of type $r' \in S$ in the population independently reproduces according to a rate one Poisson process, and each new individual y is independently assigned the type $\rho(y) = s$ with probability $B_{sr'}$, where \mathbf{B} is the matrix defined in (4.6). The reproduction process ξ_x of an individual x is a measure on $S \times \mathbb{R}_+$, with $\xi_x(A \times B)$ denoting the number of children of x of types in A born at times in B . We will write ξ for the reproduction process of the root.

Observe that, if we assign each vertex in our tree process $\mathcal{T}(\cdot)$ the type equalling distance from the root, then $\mathcal{T}(\cdot)$ has the same distribution as $\text{MBP}(\cdot)$ with root of type 0. In particular, for each $k \geq 0$ and all $t \geq 0$,

$$\tilde{\mathcal{P}}_k(\cdot) \stackrel{d}{=} |\{x \in \text{MBP}(\cdot) : \rho(x) = k\}|. \quad (4.11)$$

Let $\{\tau_k : k \geq 1\}$ denote the birth times of the root in $\text{MBP}(\cdot)$. For $r \in S$, we will denote by \mathbb{E}_r the expectation operator when the root is of type r . Define the reproduction kernel $\mu^{\mathbf{B}}$,

$$\mu^{\mathbf{B}}(r, ds \times dt) = \mathbb{E}_r[\xi(ds \times dt)].$$

Note that for $r, s \in S$, $\mu^{\mathbf{B}}(r, s \times dt) := \mu^{\mathbf{B}}(r, \{s\} \times dt) = B_{sr}dt$. For any $\lambda > 0$ define the measure associated with the Laplace transform of $\mu^{\mathbf{B}}$,

$$\mu_{\lambda}^{\mathbf{B}}(r, ds \times dt) = e^{-\lambda t} \mu^{\mathbf{B}}(r, ds \times dt).$$

Define

$$\mu_{\lambda}^{\mathbf{B}}(r, s) := \mu_{\lambda}^{\mathbf{B}}(r, s \times \mathbb{R}_+) = \int_0^{\infty} e^{-\lambda t} \mu^{\mathbf{B}}(r, s \times dt) = \frac{B_{sr}}{\lambda}.$$

The n -th convolution of $\mu_\lambda^{\mathbf{B}}$ is denoted by $\mu_\lambda^{(n)}$ (suppressing dependence on \mathbf{B}) where $\mu_\lambda^{(1)} := \mu_\lambda^{\mathbf{B}}$ and for $n \geq 2$,

$$\mu_\lambda^{(n)}(r, A \times B) = \int_{S \times \mathbb{R}_+} \mu_\lambda^{(n-1)}(s, A \times (B - t)) \mu_\lambda^{\mathbf{B}}(r, ds \times dt).$$

The Malthusian rate associated with $\text{MBP}(\cdot)$ is defined as

$$\alpha := \inf \left\{ \lambda > 0 : \sum_{n=0}^{\infty} \mu_\lambda^{(n)}(s, S \times \mathbb{R}_+) < \infty \text{ for some } s \in S \right\}.$$

Since \mathbf{B}^T is a stochastic matrix, it follows that $\alpha = 1$, and $\text{MBP}(\cdot)$ is Malthusian and supercritical as required for the results of [35, 37]. We show in Appendix A that $\mu_\alpha^{\mathbf{B}}$ is irreducible and α -recurrent in the terminology of [52]. Choosing $\pi(r) = (1 - q_*)u_r$, $r \in S$ (note that this is **non-trivial** if and only if $\mathbb{E}[Z] > 1$ so that $q_* < 1$), where $\mathbf{u} = (u_i : i \geq 0)$ is the right eigenvector of \mathbf{B} in Proposition 4.12, and $h(\cdot) \equiv 1$, note that $\sum_{r \in S} h(r)\pi(r) = 1$ and

$$\sum_{r \in S} \pi(r) \mu_\alpha^{\mathbf{B}}(r, s) = \pi(s) \quad \text{and} \quad \sum_{s \in S} \mu_\alpha^{\mathbf{B}}(r, s) h(s) = h(r).$$

To apply the results in [35, 37], we need to check that, in addition to the above, the following conditions are satisfied for $\text{MBP}(\cdot)$. In the following, for a non-negative random variable U on the probability space on which MBP is defined, write $\mathbb{E}_\pi(U) := \int_S \mathbb{E}_s(U)\pi(ds)$. Although $\alpha = 1$ in our case, we retain this notation to highlight the dependence on α in potentially more general applications.

- (i) The reproduction kernel $\mu^{\mathbf{B}}$ is non-lattice, Malthusian, supercritical and satisfies $\sup_s \mu(s, S \times [0, \varepsilon]) < 1$ for some $\varepsilon > 0$.
- (ii) The homogeneity assumption on h , i.e., $\inf_{s \in S} h(s) > 0$, is satisfied.
- (iii) The kernel $\mu^{\mathbf{B}}$ has strong α -recurrence in the sense that

$$0 < \beta := \int_{S \times S \times \mathbb{R}_+} t e^{-\alpha t} h(s) \mu^{\mathbf{B}}(r, ds \times dt) \pi(dr) < \infty,$$

- (iv) The $x \log x$ -condition is satisfied, i.e., $\mathbb{E}_\pi[\tilde{\xi} \log^+ \tilde{\xi}] < \infty$, where $\tilde{\xi} = \int_{S \times \mathbb{R}_+} e^{-\alpha t} h(s) \xi(ds \times dt)$.

For Condition (i), the Malthusian and supercritical property was discussed before. The remaining conditions are immediate from the fact that births happen according to a rate one Poisson process. Condition (ii) is trivially satisfied due to our choice of $h \equiv 1$. Condition (iii) can be checked by computing,

$$\beta = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \int_0^{\infty} t e^{-\alpha t} h(s) \mu^{\mathbf{B}}(r, s \times dt) \pi(r) dt = \sum_{r=0}^{\infty} \sum_{s=1}^{\infty} \int_0^{\infty} t e^{-\alpha t} B_{sr} \pi(r) dt$$

$$\begin{aligned}
&= \left(\int_0^\infty t e^{-\alpha t} dt \right) \cdot \sum_{s=1}^\infty \sum_{k=-1}^\infty B_{s,s+k} \pi(s+k) \\
&= \frac{1-q_*}{\alpha^2} \cdot \left(\sum_{k=1}^\infty c_k q_*^{k-1} + \sum_{s=2}^\infty \sum_{k=-1}^\infty p_{k+1} q_*^{s+k-1} \right) \\
&= \frac{1}{\alpha^2} \cdot (1-q_*) \left(1 + \frac{q_*}{1-q_*} \right) = \frac{1}{\alpha^2}.
\end{aligned}$$

To check Condition (iv), note that since $h \equiv 1$, $\bar{\xi} = \sum_{k \geq 1} e^{-\alpha \tau_k}$. Let $\mathcal{F}_{(k)}$ denote the natural filtration up to the birth of the k -th child in ξ . Then

$$\begin{aligned}
\mathbb{E}_\pi[\bar{\xi} \log^+ \bar{\xi}] &\leq \mathbb{E}_\pi[\bar{\xi}^2] = \sum_{k,l \geq 1} \mathbb{E}[e^{-\alpha(\tau_k + \tau_l)}] \\
&= \sum_{k \geq 1} \mathbb{E}[e^{-2\alpha\tau_k}] + 2 \sum_{k < l} \mathbb{E}[\mathbb{E}[e^{-\alpha(\tau_k + \tau_l)} | \mathcal{F}_{(k)}]] \\
&= \sum_{k \geq 1} \left(\frac{1}{1+2\alpha} \right)^k + 2 \sum_{k < l} \left(\frac{1}{1+\alpha} \right)^{l-k} \left(\frac{1}{1+2\alpha} \right)^k < \infty. \quad (4.12)
\end{aligned}$$

\mathbb{L}^1 -convergence in the non-fringe regime. Let σ_x denotes the birth time of vertex x in the multitype branching process MBP(\cdot). Consider a *characteristic* $\psi : S \times [0, \infty) \rightarrow [0, \infty)$, which is a random càdlàg measurable function, defined on the probability space on which MBP evolves, thought of as giving a ‘score’ to the root at time t based on its type and its genealogical tree. For $x \in I$ and $t \geq \sigma_x$, write $\psi_x(\rho(x), t - \sigma_x)$ for the corresponding score computed by evaluating the characteristic ψ on the subtree rooted at x , when this vertex is of age $t - \sigma_x$. See [35, Section 7] for a more formal treatment. Define the cumulative ψ -score

$$Z^\psi(t) = \sum_{x: \sigma_x \leq t} \psi_x(\rho(x), t - \sigma_x).$$

We will write $\mathbb{E}_s[\psi(t)] = \mathbb{E}[\psi(s, t)]$. The following theorem gives convergence of expectations for normalized cumulative ψ -scores.

Theorem 4.18 (Theorem 1 in [37]) *Consider a non-lattice strictly Malthusian, supercritical branching population, counted with a bounded characteristic ψ such that the function $t \mapsto e^{-\alpha t} \mathbb{E}_s[\psi(t)]$ is directly Riemann integrable with respect to π . Then, for π -almost every $s \in S$,*

$$e^{-\alpha t} \mathbb{E}[Z^\psi(t)] \rightarrow h(s) \mathbb{E}_\pi[\hat{\psi}(\alpha)] / \alpha \beta.$$

Corollary 4.19 *In the non-fringe regime where $\mathbb{E}[Z] > 1$,*

$$\lim_{t \rightarrow \infty} e^{-t} \mathbb{E}[\tilde{\mathcal{P}}_1(t)] = 1 - q_*.$$

Proof Let $\psi(s, t) = \mathbf{1}\{s = 1, t \geq 0\}$. We can easily verify that $t \mapsto e^{-\alpha t} \mathbb{E}_s[\psi(t)]$ is directly Riemann integrable with respect to π . Moreover,

$$\mathbb{E}_\pi[\hat{\psi}(\alpha)] = \sum_{r=0}^{\infty} \pi(r) \int_0^{\infty} e^{-\alpha t} \mathbb{E}_r[\psi(t)] dt = \frac{\pi(1)}{\alpha} = 1 - q_*$$

as $\pi(1) = 1 - q_*$ and $\alpha = 1$. Recall from previous section that $\beta = 1/\alpha^2 = 1$. Applying Theorem 4.18 with $\psi(s, t) = \mathbf{1}\{s = 1, t \geq 0\}$ and $h \equiv 1$ gives the result. \square

Recall that $n(t)$ denotes the population size at time t . Let \mathcal{F}_t be the filtration generated by the entire life histories of the first $n(t)$ vertices. Let mx denote the mother of vertex x and define for any $t, c \geq 0$,

$$\mathcal{I}(t) = \{x; \sigma_{mx} \leq t < \sigma_x < \infty\}, \quad \mathcal{I}(t, c) = \{x; \sigma_{mx} \leq t, t + c < \sigma_x < \infty\}.$$

Define $w_t = \sum_{x \in \mathcal{I}(t)} e^{-\sigma_x}$ and $w_{t,c} = \sum_{x \in \mathcal{I}(t,c)} e^{-\sigma_x}$. We collect some useful properties of w_t and $w_{t,c}$ in the following lemma. Recall from Lemma 4.3 that $W \sim \text{Exp}(1)$ is the almost sure limit of $e^{-t} n(t)$ as $t \rightarrow \infty$.

Lemma 4.20 (i) $\{w_t\}$ is a non-negative martingale with respect to $\{\mathcal{F}_t\}$ and $\mathbb{E}[w_t] = 1$.

(ii) $w_t \rightarrow W$ almost surely and in \mathbb{L}^2 as $t \rightarrow \infty$.

(iii) $\mathbb{E}[w_{t,c}] \rightarrow k(c)$ as $t \rightarrow \infty$, where $k(c) \downarrow 0$ as $c \rightarrow \infty$.

Proof (i) Follows from Proposition 2.4 and (2.17) in [47]. The \mathbb{L}^2 convergence in (ii) follows from Theorem 4.1 in [36] using the fact that $\mathbb{E}_v[\bar{\xi}^2] < \infty$ derived in (4.12). Further, as w_t is the reproductive martingale (see equation (2.15) in [47]), the almost sure convergence to W follows from [47, Corollary 2.5 and Theorem 5.4]. (iii) follows from [47, Lemma 3.5]. \square

A version of the following theorem is proved for a class of multitype branching processes in [37]. However, it requires the finiteness of $\xi(S \times \mathbb{R}_+)$, which is not satisfied in our case. We thus give a direct proof which bypasses some technicalities in [37] introduced by their generality of hypotheses.

Theorem 4.21 *In the non-fringe regime where $\mathbb{E}[Z] > 1$, as $t \rightarrow \infty$,*

$$e^{-t} \tilde{\mathcal{P}}_1(t) \xrightarrow{\mathbb{L}^1} (1 - q_*)W,$$

where W is as defined in Lemma 4.20.

Proof Recall that $\alpha = \beta = 1$. Let $\psi(s, t) = \mathbf{1}\{s = 1, t \geq 0\}$ so that $Z^\psi(t) = \tilde{\mathcal{P}}_1(t)$. For any $M > 0$, write $\psi = \psi_M + \psi'_M$ where $\psi_M(s, t) = \psi(s, t)\mathbf{1}\{t \leq M\}$ and $\psi'_M(s, t) = \psi(s, t)\mathbf{1}\{t > M\}$. To simplify notation, let

$$\gamma := \mathbb{E}_v[\hat{\psi}(\alpha)]/\alpha\beta = 1 - q_* \quad \text{and} \quad \gamma_M := \mathbb{E}_v[\hat{\psi}_M(\alpha)]/\alpha\beta = (1 - e^{-M})(1 - q_*).$$

Note that

$$\begin{aligned} \mathbb{E}|e^{-t}Z^\psi(t) - \gamma W| &\leq e^{-t}\mathbb{E}|Z^\psi(t) - Z^{\psi_M}(t)| \\ &\quad + \mathbb{E}|e^{-t}Z^{\psi_M}(t) - \gamma_M W| + |\gamma_M - \gamma|\mathbb{E}[W]. \end{aligned} \quad (4.13)$$

The first term is equal to $e^{-t}\mathbb{E}[Z^{\psi_M}(t)]$. Theorem 4.18 implies that

$$\lim_{t \rightarrow \infty} e^{-t}\mathbb{E}[Z^{\psi_M}(t)] = (1 - q_*)e^{-M}.$$

Hence, $e^{-t}\mathbb{E}[Z^{\psi_M}(t)]$ can be made arbitrarily small by taking both t and M to be large. The third term vanishes as $M \rightarrow \infty$ since $\mathbb{E}[W] = 1$ and $\lim_{M \rightarrow \infty} \gamma_M = \gamma$.

It remains to deal with the second term in (4.13). Let $\zeta_t = e^{-t}Z^{\psi_M}(t)$ and $m_s(t) = \mathbb{E}_s[\zeta_t]$. By Theorem 4.18 we have

$$\lim_{t \rightarrow \infty} m_s(t) = \mathbb{E}_v[\hat{\psi}_M(\alpha)]/\alpha\beta =: \gamma_M \quad (4.14)$$

for v -almost sure $s \in S$.

Observe that $\psi_M(u) = 0$ for all $u \geq M$. Hence, for $t' \geq M$ we have

$$\zeta_{t+t'} = \sum_{x \in \mathcal{I}(t)} e^{-\sigma_x} \zeta_{t+t'-\sigma_x} \circ \rho(x),$$

where $\zeta_{t+t_0-\sigma_x} \circ \rho(x)$ denotes the normalized ψ_M score for the vertex x , whose type is $\rho(x)$, at time $t + t' - \sigma_x$. For $t, c \geq 0$ and $t' \geq c$, write $\eta(t + t' - \sigma_x) \circ \rho(x) = \zeta_{t+t'-\sigma_x} \circ \rho(x) - m_{\rho(x)}(t + t' - \sigma_x)$ and define

$$X(t, t', c) = \sum_{x \in \mathcal{I}(t) \setminus \mathcal{I}(t, c)} e^{-\sigma_x} \eta(t + t' - \sigma_x) \circ \rho(x).$$

By triangle inequality we have

$$\begin{aligned} |\zeta_{t+t'} - \gamma_M W| &\leq |X(t, t', c)| + \left| \sum_{x \in \mathcal{I}(t, c)} e^{-\sigma_x} (\zeta_{t+t'-\sigma_x} \circ \rho(x) - \gamma_M) \right| \\ &\quad + \gamma_M |w_t - W| + \sum_{x \in \mathcal{I}(t) \setminus \mathcal{I}(t, c)} e^{-\sigma_x} |m_{\rho(x)}(t + t' - \sigma_x) - \gamma_M|. \end{aligned} \quad (4.15)$$

Fix any $\varepsilon > 0$. Our goal is to show that there exist t, c (depending on ε) sufficiently large and $t'_0 = t'_0(t, c) \in \mathbb{R}_+$ such that for all $t' \geq t'_0$, $\mathbb{E}|\zeta_{t+t'} - \gamma_M W| \leq \varepsilon$. To do this we will deal with each term in (4.15) separately.

Note that $X(t, t', c)$ is a sum of independent random variables each having expectation zero under the conditional law $\mathbb{P}(\cdot | \mathcal{F}_t)$. Hence, $\mathbb{E}[X(t, t', c) | \mathcal{F}_t] = 0$ and

$$\begin{aligned}\mathbb{E}[X(t, t', c)^2 | \mathcal{F}_t] &= \text{Var}[X(t, t', c) | \mathcal{F}_t] \\ &= \sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-2\sigma_x} \text{Var}[\eta(t + t' - \sigma_x) \circ \rho(x) | \mathcal{F}_t] \\ &\leq \sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-2\sigma_x} \mathbb{E}[(\eta(t + t' - \sigma_x) \circ \rho(x))^2 | \mathcal{F}_t]\end{aligned}$$

Since $\zeta_t \leq e^{-t} n(t)$, it follows from Lemma 4.4 that for all $t \geq 0$ and $s \in S$,

$$\mathbb{E}_s[\zeta_t^2] \leq \mathbb{E}[\mathbb{E}[e^{-2t} n(t)^2 | \mathcal{F}_0]] \leq 2.$$

Note that for any $s \in S$, $m_s = \mathbb{E}_s[\zeta_t] \leq e^{-t} \mathbb{E}[n(t)] = 1$. Then, for any $x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)$,

$$\begin{aligned}\mathbb{E}[(\eta(t + t' - \sigma_x) \circ \rho(x))^2 | \mathcal{F}_t] &\leq 2\mathbb{E}[(\zeta_{t+t'-\sigma_x} \circ \rho(x))^2 + m_{\rho(x)}(t + t' - \sigma_x)^2 | \mathcal{F}_t] \\ &\leq 2(2 + 1) = 6.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{E}[X(t, t', c)^2] &= \mathbb{E}\left[\mathbb{E}[X(t, t', c)^2 | \mathcal{F}_t]\right] \\ &\leq 6\mathbb{E}\left[\sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-2\sigma_x}\right] \leq C'e^{-t} \mathbb{E}[w_t] = 6e^{-t}.\end{aligned}$$

Hence,

$$\mathbb{E}|X(t, t', c)| \leq \mathbb{E}[X(t, t', c)^2]^{1/2} \leq \sqrt{6}e^{-t/2}. \quad (4.16)$$

There exists some $N_1 \in \mathbb{R}_+$ such that for $t \geq N_1$, and any choice of t', c , $\mathbb{E}|X(t, t', c)| \leq \varepsilon/4$.

To address the second term in (4.15), simply observe that

$$\begin{aligned}\mathbb{E}\left[\left|\sum_{x \in \mathcal{J}(t, c)} e^{-\sigma_x} (\zeta_{t+t'-\sigma_x} \circ \rho(x) - \gamma_M)\right| | \mathcal{F}_t\right] \\ \leq \sum_{x \in \mathcal{J}(t, c)} e^{-\sigma_x} (m_{\rho(x)}(t + t' - \sigma_x) + \gamma_M) \leq (1 + \gamma_M)w_{t, c}.\end{aligned}$$

Lemma 4.20(iii) implies that there exists some $N_2 \in \mathbb{R}_+$ so that if $c, t \geq N_2$ then, for any choice of t' ,

$$\mathbb{E}\left|\sum_{x \in \mathcal{J}(t, c)} e^{-\sigma_x} (\zeta_{t+t'-\sigma_x} \circ \rho(x) - \gamma_M)\right| \leq (1 + \gamma_M)\mathbb{E}[w_{t, c}] \leq \varepsilon/4. \quad (4.17)$$

For the third term, since w_t converges to W in \mathbb{L}^1 , there exists some $N_3 \in \mathbb{R}_+$ so that

$$\mathbb{E}[\gamma_M | w_t - W |] \leq \varepsilon/4 \quad (4.18)$$

for all $t \geq N_3$.

Finally we upper bound the expectation of the fourth term in (4.15). Let $\mathcal{E}_t = \{\mathcal{H}_{\mathcal{T}(t)} < 2\kappa_0 t\}$. It follows from Proposition 8.1 (whose proof is independent of the tools using MBP) that $\lim_{t \rightarrow \infty} P(\mathcal{E}_t^c) = 0$. Note that on \mathcal{E}_t , the possible types for $x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)$ in MBP(\cdot) are in the set $[2\kappa_0 t] := \{0, 1, \dots, \lceil 2\kappa_0 t \rceil\}$. Observe that

$$\begin{aligned} & \mathbb{E}_0 \left[\sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-\sigma_x} |m_{\rho(x)}(t + t' - \sigma_x) - \gamma_M| \right] \\ & \leq \mathbb{E}_0 \left[\sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-\sigma_x} \sup_{s \in [2\kappa_0 t]} |m_s(t + t' - \sigma_x) - \gamma_M| ; \mathcal{E}_t \right] \\ & \quad + \mathbb{E}_0 \left[\sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-\sigma_x} (1 + \gamma_M) ; \mathcal{E}_t^c \right]. \end{aligned}$$

There exists some $N_4 \in \mathbb{R}_+$ so that for all $t \geq N_4$, and any choice of t', c ,

$$\begin{aligned} \mathbb{E}_0 \left[\sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-\sigma_x} (1 + \gamma_M) ; \mathcal{E}_t^c \right] & \leq (1 + \gamma_M) \mathbb{E}[w_t ; \mathcal{E}_t^c] \\ & \leq (1 + \gamma_M) (\mathbb{E}[|w_t - W|] + \mathbb{E}[W ; \mathcal{E}_t^c]) \leq \varepsilon/8, \end{aligned} \quad (4.19)$$

where the last inequality follows from Lemma 4.20.

Choose and fix any $t \geq \max_{1 \leq i \leq 4} \{N_i\}$ and $c \geq N_2$. Since $x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)$, we have $t + t' - \sigma_x \geq t' - c$. Thus, there exists $t'_0 = t'_0(t, c) \in \mathbb{R}_+$ such that for all $t' \geq t'_0$,

$$\sup_{s \in [2\kappa_0 t]} |m_s(t + t' - \sigma_x) - \gamma_M| \leq \varepsilon/8.$$

Hence, for all $t' \geq t'_0$,

$$\mathbb{E}_0 \left[\sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-\sigma_x} \sup_{s \in [2\kappa_0 t]} |m_s(t + t' - \sigma_x) - \gamma_M| \right] \leq (\varepsilon/8) \mathbb{E}[w_t] = \varepsilon/8. \quad (4.20)$$

Combining (4.19) and (4.20) gives for all $t' \geq t'_0$,

$$\mathbb{E}_0 \left[\sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-\sigma_x} |m_{\rho(x)}(t + t' - \sigma_x) - \gamma_M| \right] \leq \varepsilon/4. \quad (4.21)$$

Using (4.16), (4.17), (4.18), (4.21) in (4.15), we conclude that, for any given $\varepsilon > 0$, there exists $t_0(\varepsilon) \in \mathbb{R}_+$ such that for all $t \geq t_0(\varepsilon)$,

$$\mathbb{E}|\zeta_t - \gamma_M W| \leq \varepsilon.$$

Thus we have established the convergence of $e^{-t} Z^\psi(t)$ to γW in \mathbb{L}^1 . The proof is complete. \square

5 Proofs: Local weak convergence

The goal of this section is to prove Theorem 3.3. Recall the continuous time version of the process $\mathcal{T}(\cdot, \mathbf{p})$ in Definition 2.1; here we label vertices as $\{v_i : i \geq 0\}$ in the order they enter the system starting with v_0 and recall that σ_i denotes the birth time of vertex v_i . For the rest of the section we will suppress dependence on \mathbf{p} . We will start by studying asymptotics of empirical functionals of the fringe distribution of this process as $t \rightarrow \infty$ and then leverage these results to prove Theorem 3.3.

Let $\mathcal{F} := \{\mathcal{F}_t : t \geq 0\}$ denote the natural filtration of the process $\{\mathcal{T}(t) : t \geq 0\}$. Recall the space \mathbb{T} of finite rooted trees from Sect. 2.4. For any $\mathcal{T} \in \mathbb{T}$ and vertex $v \in \mathbb{V}(\mathcal{T})$, we will denote by \mathcal{T}_v the fringe at vertex v namely the subtree rooted at v , consisting of all vertices in \mathcal{T} whose path to the root of \mathcal{T} passes through v . Let $\phi : \mathbb{T} \rightarrow \mathbb{R}$ denote a non-negative bounded measurable function. For any $i \geq 0$, define $\phi_i : \mathbb{R}_+ \rightarrow \mathbb{R}$ by,

$$\phi_i(u) = \begin{cases} \phi(\mathcal{T}_{v_i}(u + \sigma_i)), & u \geq 0, \\ 0, & s < 0. \end{cases}$$

Thus the stochastic process $\{\phi_i(u) : u \geq 0\}$ tracks the evolution of the “score” of the fringe tree below v_i as the age of v_i increases. Now note that for any time $s > 0$ and for v_i such that $\sigma_i \leq s$, the age of vertex v_i at time s is $s - \sigma_i$. Write $\mathcal{Z}^\phi(s) = \sum_{i: \sigma_i \leq s} \phi_i(s - \sigma_i)$; in words we are aggregating the scores of the fringe trees of vertices born before time s . For the rest of this section, fix any \mathcal{F} adapted process $\{a(t) : t \geq 0\}$ with $a(t) \xrightarrow{\text{a.s.}} \infty$, $a(t)/t \xrightarrow{\text{a.s.}} 0$. We will write $y(t) = t + a(t)$. Recall the process $\mathcal{T}^*(\cdot, \mathbf{p})$ in Definition 3.1. When \mathbf{p} is clear from context, we will write $\mathcal{T}^*(\cdot)$ for $\mathcal{T}^*(\cdot, \mathbf{p})$.

Theorem 5.1 *Let $\phi : \mathbb{T} \rightarrow \mathbb{R}$ be a bounded measurable function. Assume that $s \rightarrow \mathbb{E}[\phi(\mathcal{T}^*(s))]$ is Lipschitz on $[0, \infty)$. Then*

$$e^{-y(t)} \mathcal{Z}^\phi(y(t)) \xrightarrow{P} W \int_0^\infty e^{-s} \mathbb{E}[\phi(\mathcal{T}^*(s))] ds \quad \text{as } t \rightarrow \infty,$$

where $W \sim \text{Exp}(1)$ is the almost sure limit of $e^{-t} n(t)$ as $t \rightarrow \infty$.

The main tool to prove this Theorem is the following Proposition.

Proposition 5.2 (Conditional moments of $\mathcal{Z}^\phi(t)$) *In the setting of Theorem 5.1, as $t \rightarrow \infty$,*

$$e^{-y(t)} \mathbb{E} [\mathcal{Z}^\phi(y(t)) | \mathcal{F}_t] \xrightarrow{a.s.} W \int_0^\infty e^{-u} \mathbb{E} [\phi(\mathcal{T}^*(u))] du, \quad (5.1)$$

$$\text{Var} \left[e^{-y(t)} \mathcal{Z}^\phi(y(t)) | \mathcal{F}_t \right] \xrightarrow{a.s.} 0. \quad (5.2)$$

Proof of Theorem 5.1 assuming Proposition 5.2 For any $\delta > 0$, write

$$\begin{aligned} & \mathbb{P} \left[\left| e^{-y(t)} \mathcal{Z}^\phi(y(t)) - W \int_0^\infty e^{-s} \mathbb{E} [\phi(\mathcal{T}^*(s))] ds \right| > \delta \right] \\ & \leq \mathbb{E} \left[\mathbb{P} \left[\left| e^{-y(t)} \mathcal{Z}^\phi(y(t)) - \mathbb{E}[e^{-y(t)} \mathcal{Z}^\phi(y(t)) | \mathcal{F}_t] \right| > \delta/2 \middle| \mathcal{F}_t \right] \right] \\ & \quad + \mathbb{P} \left[\left| \mathbb{E}[e^{-y(t)} \mathcal{Z}^\phi(y(t)) | \mathcal{F}_t] - W \int_0^\infty e^{-s} \mathbb{E} [\phi(\mathcal{T}^*(s))] ds \right| > \delta/2 \right]. \end{aligned}$$

We assert that each of these two terms converge to zero. For the first term, this follows by applying Chebyshev's inequality with (5.2), along with the bounded convergence theorem. For the second term this follows from (5.1). \square

5.1 Proof of Proposition 5.2

We start with the following technical Lemma providing \mathcal{F}_t measurable approximations of birth times of new individuals after large t . Throughout recall the notation $n(t) = |\mathcal{T}(t)|$. As this Lemma lies at the heart of our local limit computations, we provide an intuitive explanation first. Suppose we know the population size $n(t)$ at time t and want to 'guess' the birth time $\alpha_i(t)$ of the $(n(t) + i)$ -th individual based on this information. From Lemma 4.2, we know that the population size (which is a Yule process) grows approximately exponentially with time. Therefore, a good guess is $\alpha_i(t)$ satisfying $n(t)e^{\alpha_i(t)-t} = n(t) + i$. The following lemma shows that, for sufficiently large t , this is a uniformly good guess for all birth times after t .

Lemma 5.3 *Let $\{\tilde{\sigma}_{t,i} : i \geq 1\}$ denote the birth times after time t , i.e., $\tilde{\sigma}_{t,i} = \sigma_{i+n(t)}$. Define $\alpha_i(t) := t + \log \left(1 + \frac{i}{n(t)} \right)$ for $i \geq 1$. Then*

$$\sup_{i \geq 1} |\tilde{\sigma}_{t,i} - \alpha_i(t)| \xrightarrow{a.s.} 0 \quad \text{as } t \rightarrow \infty. \quad (5.3)$$

Moreover,

$$\mathbb{E} \left[\sup_{i \geq 1} |\tilde{\sigma}_{t,i} - \alpha_i(t)| \middle| \mathcal{F}_t \right] \xrightarrow{a.s.} 0 \quad \text{as } t \rightarrow \infty. \quad (5.4)$$

Proof (5.3) follows from [26, Lemma 3.4].

It remains to prove the convergence in expectation. As before, let $\{E_j : j \geq 0\}$ denote an i.i.d. sequence of exponential rate one random variables independent of

\mathcal{F}_t . The dynamics of the Yule process implies that conditional on \mathcal{F}_t , all the joint distributions can be constructed simultaneously as

$$\tilde{\sigma}_{t,i} = t + \sum_{j=0}^{i-1} \frac{E_j}{n(t) + j}, \quad i \geq 1.$$

Thus,

$$|\tilde{\sigma}_{t,i} - \alpha_i(t)| \leq \left| \sum_{j=0}^{i-1} \left(\frac{E_j}{n(t) + j} - \frac{1}{n(t) + j} \right) \right| + \left| \sum_{j=0}^{i-1} \frac{1}{n(t) + j} - \log \left(\frac{n(t) + i}{n(t)} \right) \right|.$$

It is easy to see that there exists some constant $C > 0$ such that the second term is upper bounded by $\frac{C}{n(t)}$ for all $i \geq 1$ and all $t \geq 0$. Since $\{E_j\}_{j \geq 0}$ are independent of \mathcal{F}_t , defining $\mathcal{M}_i := \sum_{j=0}^{i-1} \frac{E_j - 1}{n(t) + j}$, the sequence $\{\mathcal{M}_i : i \geq 1\}$ conditioned on \mathcal{F}_t is a martingale. By Cauchy-Schwarz inequality and Doob's L^2 inequality, for any $k \geq 1$,

$$\begin{aligned} \mathbb{E} \left[\sup_{1 \leq i \leq k} \left| \sum_{j=0}^{i-1} \frac{E_j - 1}{n(t) + j} \right| \middle| \mathcal{F}_t \right] &= \mathbb{E} \left[\sup_{1 \leq i \leq k} |\mathcal{M}_i| \middle| \mathcal{F}_t \right] \leq C' (\mathbb{E}[\mathcal{M}_k^2 | \mathcal{F}_t])^{1/2} \\ &= C' \left(\sum_{j=0}^{k-1} \frac{1}{(n(t) + j)^2} \right)^{1/2} \leq C' \left(\sum_{j=0}^{\infty} \frac{1}{(n(t) + j)^2} \right)^{1/2} \\ &\leq \frac{C''}{\sqrt{n(t)}} \end{aligned}$$

for some positive constants C', C'' not depending on k . Therefore,

$$\mathbb{E} \left[\sup_{i \geq 1} |\tilde{\sigma}_{t,i} - \alpha_i(t)| \middle| \mathcal{F}_t \right] \leq \frac{C}{n(t)} + \frac{C''}{\sqrt{n(t)}} \xrightarrow{a.s.} 0 \text{ as } t \rightarrow \infty.$$

□

We will now commence on the proof of Proposition 5.2.

Proof of (5.1) For the rest of this proof, let $m(s) := \mathbb{E}[\phi(\mathcal{T}^*(s))]$ for $s \geq 0$ and $m(s) = 0$ for $s < 0$. By boundedness assumption on ϕ and Lipschitz continuity of $m(\cdot)$, there exists constants $M, L < \infty$ such that

$$|\phi(\cdot)| \leq M, \quad |m(s) - m(t)| \leq L|t - s|. \quad (5.5)$$

Note that $\mathcal{Z}^\phi(y(t)) = \sum_{i:\sigma_i \leq t} \phi_i(y(t) - \sigma_i) + \sum_{i:\sigma_i \in (t, y(t)]} \phi_i(y(t) - \sigma_i)$. To simplify notation, we will write the second term as $\tilde{\mathcal{Z}}^\phi(y(t)) := \sum_{i:\sigma_i \in (t, y(t)]} \phi_i(y(t) - \sigma_i)$. By (5.5), $|\phi_i(\cdot)| \leq M$ for all $i \geq 0$. It is straightforward then, observing from Lemma

4.2 that $e^{-t}n(t)$ converges almost surely to a finite random limit,

$$e^{-y(t)}\mathbb{E}\left[\sum_{i:\sigma_i \leq t} \phi_i(y(t) - \sigma_i) \middle| \mathcal{F}_t\right] \leq M \cdot \mathbb{E}[e^{-y(t)}n(t) \mid \mathcal{F}_t] \xrightarrow{a.s.} 0 \quad \text{as } t \rightarrow \infty. \quad (5.6)$$

To address the second term, recalling from Lemma 5.3 that $\tilde{\sigma}_{t,i} = \sigma_{i+n(t)}$, write

$$\mathbb{E}[\tilde{\mathcal{Z}}^\phi(y(t)) \mid \mathcal{F}_t] = \mathbb{E}\left[\sum_{i:\tilde{\sigma}_{t,i} \leq y(t)} m(y(t) - \tilde{\sigma}_{t,i}) \mid \mathcal{F}_t\right]. \quad (5.7)$$

Applying Lemma 5.3 we can approximate (5.7) by using the \mathcal{F}_t -measurable approximations $\{\alpha_i(t)\}_{i \geq 1}$ of the birth times $\{\tilde{\sigma}_{t,i}\}_{i \geq 1}$,

$$\begin{aligned} & e^{-y(t)} \left| \mathbb{E}\left[\sum_{i:\tilde{\sigma}_{t,i} \leq y(t)} m(y(t) - \tilde{\sigma}_{t,i}) - \sum_{i:\alpha_i(t) \leq y(t)} m(y(t) - \tilde{\sigma}_{t,i}) \mid \mathcal{F}_t\right] \right| \\ & \leq e^{-y(t)} \cdot M \mathbb{E}\left[\sum_i 1_{\{\tilde{\sigma}_{t,i} \leq y(t), \alpha_i(t) > y(t)\}} \mid \mathcal{F}_t\right] \\ & \leq M \cdot e^{-y(t)} \mathbb{E}[|n(y(t)) - \lfloor n(t)(e^{a(t)} - 1) \rfloor| \mid \mathcal{F}_t] \\ & \leq M \cdot \left(\mathbb{E}\left[|e^{-y(t)}n(y(t)) - e^{-t}n(t)| \mid \mathcal{F}_t\right] + e^{-y(t)}(n(t) + 1) \right) \\ & \leq M \left(\mathbb{E}\left[(e^{-y(t)}n(y(t)) - e^{-t}n(t))^2 \mid \mathcal{F}_t\right] \right)^{1/2} + M e^{-y(t)}(n(t) + 1). \end{aligned} \quad (5.8)$$

Since $\{e^{-s}n(s) : s \geq 0\}$ is an \mathbb{L}^2 bounded martingale we can compute

$$\mathbb{E}\left[(e^{-y(t)}n(y(t)) - e^{-t}n(t))^2 \mid \mathcal{F}_t\right] = \mathbb{E}\left[(e^{-y(t)}n(y(t)))^2 \mid \mathcal{F}_t\right] - (e^{-t}n(t))^2 \leq e^{-2t}n(t),$$

where the last inequality follows from Lemma 4.4. Therefore,

$$(5.8) \leq 2M e^{-t} \sqrt{n(t)} + 2M e^{-y(t)}(n(t) + 1). \quad (5.9)$$

Now we estimate

$$\begin{aligned} & e^{-y(t)} \left| \mathbb{E}\left[\sum_{i:\alpha_i(t) \leq y(t)} m(y(t) - \tilde{\sigma}_{t,i}) - \sum_{i:\alpha_i(t) \leq y(t)} m(y(t) - \alpha_i(t)) \mid \mathcal{F}_t\right] \right| \\ & \leq e^{-y(t)}n(t)e^{a(t)} \cdot L \mathbb{E}\left[\sup_{i \geq 1} |\tilde{\sigma}_{t,i} - \alpha_i(t)| \mid \mathcal{F}_t\right] \\ & = e^{-t}n(t) \cdot L \mathbb{E}\left[\sup_{i \geq 1} |\tilde{\sigma}_{t,i} - \alpha_i(t)| \mid \mathcal{F}_t\right], \end{aligned} \quad (5.10)$$

where L is as in (5.5). Next, let $\mathcal{M}^\phi(y(t)) := \sum_{i:\alpha_i(t) \leq y(t)} m(y(t) - \alpha_i(t))$. Combining (5.9) and (5.10) yields

$$\begin{aligned} & e^{-y(t)} \left| \mathbb{E}[\tilde{\mathcal{Z}}^\phi(y(t))|\mathcal{F}_t] - \mathbb{E}[\mathcal{M}^\phi(y(t))|\mathcal{F}_t] \right| \\ & \leq 2M(e^{-t}\sqrt{n(t)} + e^{-y(t)}(n(t) + 1)) + e^{-t}n(t) \cdot L \mathbb{E} \left[\sup_{i \geq 1} |\tilde{\sigma}_{t,i} - \alpha_i(t)| \middle| \mathcal{F}_t \right]. \end{aligned} \quad (5.11)$$

The part $e^{-t}\sqrt{n(t)} + e^{-y(t)}(n(t) + 1)$ vanishes almost surely as $t \rightarrow \infty$. Then applying Lemma 5.3 gives that (5.11) goes to 0 almost surely as $t \rightarrow \infty$. It remains to evaluate the almost sure limit of $e^{-y(t)}\mathbb{E}[\mathcal{M}^\phi(y(t))|\mathcal{F}_t]$. Observe that, using the boundedness and Lipschitz continuity of $m(\cdot)$,

$$\begin{aligned} & e^{-y(t)}\mathbb{E}[\mathcal{M}^\phi(y(t))|\mathcal{F}_t] \\ &= e^{-y(t)} \sum_{i:\alpha_i(t) \leq y(t)} m(y(t) - \alpha_i(t)) \\ &= e^{-y(t)} \sum_{i=1}^{\lfloor n(t)e^{a(t)} \rfloor} m \left(a(t) - \log \left(\frac{n(t) + i}{n(t)} \right) \right) + o_{a.s.}(1) \\ &= e^{-y(t)} \sum_{i=1}^{\lfloor e^{a(t)} \rfloor} \sum_{j=1}^{n(t)} m \left(a(t) - \log \left(i + \frac{j}{n(t)} \right) \right) + o_{a.s.}(1) \\ &= e^{-y(t)} \sum_{i=1}^{\lfloor e^{a(t)} \rfloor} n(t)m(a(t) - \log i) + o_{a.s.}(1) \\ &= e^{-y(t)}n(t) \int_1^{e^{a(t)}} m(a(t) - \log x)dx + o_{a.s.}(1) \\ &= e^{-y(t)}n(t) \int_0^{a(t)} m(u)e^{a(t)-u}du + o_{a.s.}(1) \\ &= e^{-t}n(t) \int_0^{a(t)} m(u)e^{-u}du + o_{a.s.}(1) \xrightarrow{a.s.} W \int_0^\infty e^{-u}m(u)du. \end{aligned} \quad (5.12)$$

This proves (5.1). \square

Proof of (5.2) Note that

$$\begin{aligned} \mathbb{E}[(\mathcal{Z}^\phi(y(t)))^2|\mathcal{F}_t] &= \mathbb{E} \left[\sum_{i,j:\sigma_i \leq t, \sigma_j \leq t} \phi_i(y(t) - \sigma_i)\phi_j(y(t) - \sigma_j) \middle| \mathcal{F}_t \right] \\ &\quad + 2\mathbb{E} \left[\sum_{i,j:\sigma_i \leq t < \sigma_j \leq y(t)} \phi_i(y(t) - \sigma_i)\phi_j(y(t) - \sigma_j) \middle| \mathcal{F}_t \right] \end{aligned}$$

$$\begin{aligned}
& + \mathbb{E} \left[\sum_{i,j} \phi_i(y(t) - \tilde{\sigma}_{t,i}) \phi_j(y(t) - \tilde{\sigma}_{t,j}) \middle| \mathcal{F}_t \right] \\
& =: T_1(t) + 2T_2(t) + T_3(t),
\end{aligned}$$

where for the rest of this section, the sum in $T_3(t)$ is implicitly over i, j with $\sigma_i, \sigma_j \in (t, y(t)]$ but we have suppressed this to ease notation. Now, it can be readily checked that

$$e^{-2y(t)} T_1(t) \leq M^2 e^{-2y(t)} (n(t))^2 \xrightarrow{a.s.} 0 \quad \text{as } t \rightarrow \infty,$$

and

$$e^{-2y(t)} T_2(t) \leq M e^{-y(t)} n(t) \cdot e^{-y(t)} \mathbb{E} \left[\tilde{\mathcal{Z}}^\phi(y(t)) \middle| \mathcal{F}_t \right] \xrightarrow{a.s.} 0 \quad \text{as } t \rightarrow \infty,$$

where the second line follows from combining Lemma 4.3 with (5.1). Thus $\mathbb{E}[(e^{-y(t)} \mathcal{Z}^\phi(y(t)))^2 | \mathcal{F}_t] = e^{-2y(t)} T_3(t) + o_{a.s.}(1)$. We write (again recalling that all the ensuing sums are over vertices born in $(t, y(t)]$),

$$\begin{aligned}
T_3(t) & = \mathbb{E} \left[\sum_i \phi_i^2(y(t) - \tilde{\sigma}_{t,i}) \middle| \mathcal{F}_t \right] + 2\mathbb{E} \left[\sum_{i < j} \phi_i(y(t) - \tilde{\sigma}_{t,i}) \phi_j(y(t) - \tilde{\sigma}_{t,j}) \middle| \mathcal{F}_t \right] \\
& = \varepsilon(t) + \tilde{T}_3(t).
\end{aligned}$$

We restrict our attention to the second term since the first term is $o_{a.s.}(e^{-2y(t)})$:

$$\begin{aligned}
e^{-2y(t)} \varepsilon(t) & := e^{-2y(t)} \mathbb{E} \left[\sum_i \phi_i^2(y(t) - \tilde{\sigma}_{t,i}) \middle| \mathcal{F}_t \right] \\
& \leq M^2 e^{-2y(t)} n(t) e^{a(t)} \xrightarrow{a.s.} 0 \quad \text{as } t \rightarrow \infty. \tag{5.13}
\end{aligned}$$

For $i < j$, we write $j \rightarrow i$ if v_j is a descendant of v_i (and write $j \not\rightarrow i$ otherwise). By convention we have $i \rightarrow i$. Then for $i < j$,

$$\begin{aligned}
& \mathbb{E} [\phi_i(y(t) - \tilde{\sigma}_{t,i}) \phi_j(y(t) - \tilde{\sigma}_{t,j}) \middle| \mathcal{F}_t] \\
& \leq M^2 \mathbb{P}[j \rightarrow i, \tilde{\sigma}_{t,j} \leq y(t) | \mathcal{F}_t] + \mathbb{E} [\phi_i(y(t) - \tilde{\sigma}_{t,i}) \phi_j(y(t) - \tilde{\sigma}_{t,j}) \mathbb{1}_{\{j \not\rightarrow i\}} \middle| \mathcal{F}_t]. \tag{5.14}
\end{aligned}$$

Recall that, for the rest of the argument, we are only interested in pairs born in the interval $(t, y(t)]$. For any fixed time T , vertex v_i with $\sigma_i \leq T$, let $\mathcal{D}_v(T)$ denote the number of descendants of v by time T . Summing the bound (5.14) over all pairs of

vertices born in the interval $(t, y(t)]$, one gets,

$$\begin{aligned}\tilde{T}_3(t) &\leq 2M^2 \mathbb{E} \left[\sum_{i: \sigma_i \in (t, y(t)]} \mathcal{D}_{v_i}(y(t)) | \mathcal{F}_t \right] \\ &\quad + 2\mathbb{E} \left[\sum_{i < j: \sigma_i, \sigma_j \in (t, y(t)]} \phi_i(y(t) - \tilde{\sigma}_{t,i}) \phi_j(y(t) - \tilde{\sigma}_{t,j}) \mathbb{1}_{\{j \rightarrow i\}} | \mathcal{F}_t \right] \\ &:= \tilde{T}_{3,1}(t) + \tilde{T}_{3,2}(t).\end{aligned}\tag{5.15}$$

The following lemma completes the proof of (5.2). \square

Lemma 5.4 *As $t \rightarrow \infty$,*

- (a) $e^{-2y(t)} \tilde{T}_{3,1}(t) \xrightarrow{\text{a.s.}} 0$.
- (b) $e^{-2y(t)} \tilde{T}_{3,2}(t) \leq (e^{-y(t)} \mathbb{E}[\mathcal{Z}^\phi(y(t)) | \mathcal{F}_t])^2 + o_{\text{a.s.}}(1)$.

Proof of Lemma 5.4(a) Note that, when a new vertex is attached to some existing vertex v , the number of descendants of v and all its ancestor vertices increases by one. Hence, we can dominate $\sum_{i: \sigma_i \in (t, y(t)]} \mathcal{D}_{v_i}(y(t))$ pathwise by $\text{ht}(\mathcal{T}(y(t), \mathbf{p})) n(y(t))$ where recall that $\text{ht}(\cdot)$ denotes the height of the associated tree. Moreover, observe that, conditionally on \mathcal{F}_t , $\text{ht}(\mathcal{T}(y(t), \mathbf{p})) \leq \text{ht}(\mathcal{T}(t, \mathbf{p})) + \max_{i \leq n(t)} h_i(t, y(t))$, where $h_i(t, y(t))$ denotes the height of the (maximal) tree rooted at the i th vertex formed entirely by its descendants that arrived in the time interval $(t, y(t)]$. Further, note that there exists a collection $\{h_i^*(a(t)) : i \leq n(t)\}$, distributed as the heights of $n(t)$ independent Yule trees run till time $a(t)$, independent of \mathcal{F}_t , so that we can couple to get $h_i(t, y(t)) \leq h_i^*(a(t))$ for every $i \leq n(t)$. Consequently, using Lemma 4.5 with $\beta = e$,

$$\begin{aligned}\mathbb{E} \left((\max_{i \leq n(t)} h_i(t, y(t)))^2 | \mathcal{F}_t \right) &\leq n(t) \mathbb{E}((h_1^*(a(t))^2) \\ &\leq n(t) \left(9e^2 a(t)^2 + \int_{9e^2 a(t)^2}^{\infty} \mathbb{P}(h_1^*(a(t)) \geq \sqrt{x}) dx \right) \\ &\leq n(t) \left(9e^2 a(t)^2 + \int_{9e^2 a(t)^2}^{\infty} 2e^{2ea(t)} e^{-\sqrt{x}} dx \right) \\ &= 9e^2 a(t)^2 n(t) + 2n(t)(1 + 3ea(t)) e^{-ea(t)}.\end{aligned}$$

This implies that $e^{-2y(t)} \mathbb{E} \left((\max_{i \leq n(t)} h_i(t, y(t)))^2 | \mathcal{F}_t \right) \rightarrow 0$ almost surely as $t \rightarrow \infty$. Also, by Lemmas 4.3 and 4.4,

$$e^{-2y(t)} \mathbb{E}(n(y(t))^2 | \mathcal{F}_t) \leq e^{-2t} (n(t)^2 + n(t)) \xrightarrow{\text{a.s.}} W^2$$

as $t \rightarrow \infty$. Moreover, using the monotonicity of $\text{ht}(\mathcal{T}(t, \mathbf{p}))$ in t and Lemma 4.5, it follows that $e^{-y(t)} \text{ht}(\mathcal{T}(t, \mathbf{p})) \rightarrow 0$ almost surely as $t \rightarrow \infty$. Therefore, using

Cauchy-Schwarz inequality,

$$\begin{aligned}
e^{-2y(t)} \tilde{T}_{3,1}(t) &\leq e^{-2y(t)} \mathbb{E}[\text{ht}(\mathcal{T}(t, \mathbf{p})) n(y(t)) | \mathcal{F}_t] \\
&\quad + e^{-2y(t)} \mathbb{E}[\max_{i \leq n(t)} h_i(t, y(t)))^2 n(y(t)) | \mathcal{F}_t] \\
&\leq e^{-2y(t)} \text{ht}(\mathcal{T}(t, \mathbf{p})) n(t) e^{a(t)} \\
&\quad + \left(e^{-2y(t)} \mathbb{E} \left((\max_{i \leq n(t)} h_i(t, y(t)))^2 | \mathcal{F}_t \right) \right)^{1/2} \\
&\quad \times \left(e^{-2y(t)} \mathbb{E} (n(y(t))^2 | \mathcal{F}_t) \right)^{1/2},
\end{aligned}$$

which converges almost surely to 0 as $t \rightarrow \infty$. \square

Proof of Lemma 5.4(b) For $i < j$, let $\mathcal{F}_{i,j}^*$ be the σ -field generated by the tree process up to the $(n(t) + j)$ -th birth time, and the birth times and attachment locations of all vertices that are descendants of v_i . Note that $\phi_i(y(t) - \tilde{\sigma}_{t,i}) \mathbb{1}_{\{j \rightarrow i\}}$ is $\mathcal{F}_{i,j}^*$ -measurable and $\phi_j(y(t) - \tilde{\sigma}_{t,j})$ is independent of $\mathcal{F}_{i,j}^*$ on the event $\{j \not\rightarrow i\}$. Hence first conditioning on $\mathcal{F}_{i,j}^*$ and then using the tower property for conditional expectations we get,

$$\begin{aligned}
&\mathbb{E} [\phi_i(y(t) - \tilde{\sigma}_{t,i}) \phi_j(y(t) - \tilde{\sigma}_{t,j}) \mathbb{1}_{\{j \rightarrow i\}} | \mathcal{F}_t] \\
&= \mathbb{E} [\phi_i(y(t) - \tilde{\sigma}_{t,i}) \mathbb{1}_{\{j \rightarrow i\}} m(y(t) - \tilde{\sigma}_{t,j}) | \mathcal{F}_t] \\
&\leq \mathbb{E} [m(y(t) - \tilde{\sigma}_{t,i}) m(y(t) - \tilde{\sigma}_{t,j}) | \mathcal{F}_t].
\end{aligned} \tag{5.16}$$

The last inequality above follows from a similar conditioning by a sigma field containing information about birth times and attachment locations of all individuals except the descendants of i (excluding i).

By the boundedness and Lipschitz assumption on the mean functional $m(\cdot)$, for $i < j$, where j satisfies $\max\{\tilde{\sigma}_{t,j}, \alpha_j(t)\} \leq y(t)$, there exists some $C > 0$ such that

$$\begin{aligned}
&|m(y(t) - \tilde{\sigma}_{t,i}) m(y(t) - \tilde{\sigma}_{t,j}) - m(y(t) - \alpha_i(t)) m(y(t) - \alpha_j(t))| \\
&\leq C(|\tilde{\sigma}_{i,t} - \alpha_i(t)| + |\tilde{\sigma}_{j,t} - \alpha_j(t)|).
\end{aligned}$$

Hence (and writing $\mathbb{E}_{\mathcal{F}_t}(\cdot) = \mathbb{E}(\cdot | \mathcal{F}_t)$),

$$\begin{aligned}
&e^{-2y(t)} \left| \mathbb{E}_{\mathcal{F}_t} \left[\sum_{\substack{i < j: \\ \tilde{\sigma}_{t,i}, \tilde{\sigma}_{t,j} \in (t, y(t)]}} m(y(t) - \tilde{\sigma}_{t,i}) m(y(t) - \tilde{\sigma}_{t,j}) \right. \right. \\
&\quad \left. \left. - \sum_{\substack{i < j: \\ \alpha_i(t), \alpha_j(t) \in (t, y(t)]}} m(y(t) - \alpha_i(t)) m(y(t) - \alpha_j(t)) \right] \right| \tag{5.17}
\end{aligned}$$

$$\leq e^{-2y(t)} \cdot C \mathbb{E}_{\mathcal{F}_t} \left[\sum_{i < j: \max\{\tilde{\sigma}_{t,j}, \alpha_j(t)\} \leq y(t)} (|\tilde{\sigma}_{i,t} - \alpha_i(t)| + |\tilde{\sigma}_{j,t} - \alpha_j(t)|) \right] \quad (5.18)$$

$$+ e^{-2y(t)} \cdot M^2 \mathbb{E}_{\mathcal{F}_t} \left[\sum_{i < j} 1_{\{\tilde{\sigma}_{t,j} \leq y(t) < \alpha_j(t)\}} \right] \quad (5.19)$$

$$+ e^{-2y(t)} \cdot M^2 \mathbb{E}_{\mathcal{F}_t} \left[\sum_{i < j} 1_{\{\alpha_j(t) \leq y(t) < \tilde{\sigma}_{t,j}\}} \right]. \quad (5.20)$$

By Lemma 5.3,

$$\begin{aligned} (5.18) &\leq e^{-2y(t)} \cdot 2C \mathbb{E} \left[\sum_{i < j: \max\{\tilde{\sigma}_{t,j}, \alpha_j(t)\} \leq y(t)} \sup_{\ell \geq 1} |\tilde{\sigma}_{t,\ell} - \alpha_\ell(t)| \middle| \mathcal{F}_t \right] \\ &\leq 2C e^{-2y(t)} (n(t))^2 e^{2a(t)} \mathbb{E} \left[\sup_{\ell \geq 1} |\tilde{\sigma}_{t,\ell} - \alpha_\ell(t)| \middle| \mathcal{F}_t \right] \\ &= 2C (e^{-t} n(t))^2 \mathbb{E} \left[\sup_{\ell \geq 1} |\tilde{\sigma}_{t,\ell} - \alpha_\ell(t)| \middle| \mathcal{F}_t \right] \xrightarrow{a.s.} 0. \end{aligned}$$

As for (5.19), observe that

$$\begin{aligned} (5.19) &\leq M^2 e^{-2y(t)} \mathbb{E} \left[n(y(t)) |n(y(t)) - \lfloor n(t)(e^{a(t)} - 1) \rfloor| \middle| \mathcal{F}_t \right] \\ &\leq M^2 e^{-2y(t)} n(t) e^{a(t)} \mathbb{E} \left[|n(y(t)) - \lfloor n(t)(e^{a(t)} - 1) \rfloor| \middle| \mathcal{F}_t \right] \\ &\quad + M^2 e^{-2y(t)} \mathbb{E} \left[(n(y(t)) - \lfloor n(t)(e^{a(t)} - 1) \rfloor)^2 \middle| \mathcal{F}_t \right], \end{aligned}$$

which goes to zero almost surely as $t \rightarrow \infty$ following the same reasoning as in (5.9). Similarly we have

$$\begin{aligned} (5.20) &\leq M^2 e^{-2y(t)} \mathbb{E} \left[n(t)(e^{a(t)} - 1) \max\{n(t)(e^{a(t)} - 1) - n(y(t)), 0\} \middle| \mathcal{F}_t \right] \\ &\leq M^2 e^{-2y(t)} n(t) e^{a(t)} \mathbb{E} \left[|n(y(t)) - n(t)(e^{a(t)} - 1)| \middle| \mathcal{F}_t \right] \xrightarrow{a.s.} 0. \end{aligned}$$

Noting that the processes $\alpha_i(t)$ are \mathcal{F}_t adapted so that the conditional expectation of the second sum in (5.17) is itself, the last step is to estimate

$$\mathcal{M}^{\phi, (2)}(y(t)) := \sum_{\substack{i < j: \\ \alpha_i(t), \alpha_j(t) \in (t, y(t)]}} m(y(t) - \alpha_i(t)) m(y(t) - \alpha_j(t))$$

$$\leq \frac{1}{2} \left(\sum_{i:\alpha_i(t) \leq y(t)} m(y(t) - \alpha_i(t)) \right)^2.$$

Recall the definition of $\mathcal{M}^\phi(y(t)) = \sum_{i:\alpha_i(t) \leq y(t)} m(y(t) - \alpha_i(t))$ from the proof of the first moment convergence in (5.1). Then, from (5.11) and (5.6),

$$\begin{aligned} 2e^{-2y(t)} \mathcal{M}^{\phi,(2)}(y(t)) &\leq \left(e^{-y(t)} \mathcal{M}^\phi(y(t)) \right)^2 \\ &= \left(e^{-y(t)} \mathbb{E}[\tilde{Z}^\phi(y(t)) | \mathcal{F}_t] \right)^2 + o_{a.s.}(1) \\ &= \left(e^{-y(t)} \mathbb{E}[\mathcal{Z}^\phi(y(t)) | \mathcal{F}_t] \right)^2 + o_{a.s.}(1). \end{aligned} \quad (5.21)$$

Using (5.16), asymptotics for the terms in (5.18), (5.19), (5.20) and (5.21) completes the proof. \square

This completes the proof of the second moment namely (5.2). \square

5.2 Completing the proof of Theorem 3.3

The goal now is to transfer the continuous time embedding asymptotics in Theorem 5.1 to the discrete time process $\{\mathcal{T}_n(\mathbf{p}) : n \geq 1\}$. In Lemma 2.2, recall the stopping times $T_n = \inf\{t \geq 0 : n(t) = n + 1\}$, connecting the embedding of the discrete process in continuous time.

Theorem 5.5 *Let ϕ be a non-negative functional on \mathbb{T} satisfying the assumptions in Theorem 5.1. Let $v^{(n)}$ be a uniformly chosen vertex in the graph $\mathcal{T}_n(\mathbf{p})$. Let $\mathcal{T}_{v^{(n)}}$ denote the fringe tree of $v^{(n)}$. Then*

$$\begin{aligned} \mathbb{E}[\phi(\mathcal{T}_{v^{(n)}}) | \mathcal{T}_n] &:= \frac{1}{n+1} \sum_{i=0}^n \phi_i(T_n - \sigma_i) \rightarrow \int_0^\infty e^{-s} \mathbb{E}[\phi(\mathcal{T}^*(s))] ds \\ &\text{in probability as } n \rightarrow \infty. \end{aligned}$$

Proof Write

$$\mathbb{E}[\phi(\mathcal{T}_{v^{(n)}}) | \mathcal{T}_n] = \frac{1}{n+1} \sum_{i=0}^n \phi_i(T_n - \sigma_i) := \frac{1}{n+1} \mathcal{Z}^\phi(T_n).$$

Let $\Theta(t) := e^{-t} n(t)$. Note that, by Lemma 4.2, $\Theta(t) \xrightarrow{a.s.} W$ and thus,

$$T_n - \log n + \log W \xrightarrow{a.s.} 0. \quad (5.22)$$

However, working the limit random variable W for finite t approximations of the embedding is difficult so we will work with approximations of the limit W . Define

$t_n = \log n - \sqrt{\log n}$, $\tilde{T}_n = \log n - \log \Theta(t_n)$. Using (5.22) and $\Theta(t_n) \xrightarrow{\text{a.s.}} W$, we get,

$$|\tilde{T}_n - T_n| \leq |\log n - \log W - T_n| + |\log W - \log \Theta(t_n)| \xrightarrow{\text{a.s.}} 0.$$

For $\varepsilon > 0$, define $T_n^\pm(\varepsilon) := \tilde{T}_n \pm \varepsilon$. For any $\varepsilon > 0$ there exists a random n_ε such that for all $n \geq n_\varepsilon$, $T_n^-(\varepsilon) \leq T_n \leq T_n^+(\varepsilon)$ almost surely. Since ϕ is non-negative, for $n > n_\varepsilon$,

$$e^{-T_n^+(\varepsilon)} \mathcal{Z}^\phi(T_n^-(\varepsilon)) \leq e^{-T_n} \mathcal{Z}^\phi(T_n) \leq e^{-T_n^-(\varepsilon)} \mathcal{Z}^\phi(T_n^+(\varepsilon)). \quad (5.23)$$

Recall the limit random variable in Theorem 5.1 and to simplify notation let $X := W \int_0^\infty e^{-s} \mathbb{E}[\phi(\mathcal{T}^*(s))] ds$. Also recall that $W \sim \text{Exp}(1)$. Fix $\delta > 0$. For any $\eta > 0$ we can take $\varepsilon > 0$ such that,

$$\begin{aligned} & \mathbb{P}\left((X + \delta)e^{-2\varepsilon} > X + \delta/2, (X - \delta)e^{2\varepsilon} < X - \delta/2\right) \\ &= \mathbb{P}\left(X < \frac{\delta(e^{2\varepsilon} - 1/2)}{e^{2\varepsilon} - 1} \wedge \frac{\delta(e^{-2\varepsilon} - 1/2)}{1 - e^{-2\varepsilon}}\right) \geq 1 - \eta. \end{aligned} \quad (5.24)$$

Write $A_{\delta, \varepsilon} = \{(X + \delta)e^{-2\varepsilon} > X + \delta/2, (X - \delta)e^{2\varepsilon} < X - \delta/2\}$.

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \mathbb{P}(e^{-T_n} \mathcal{Z}^\phi(T_n) > X + \delta) \\ & \leq \limsup_{n \rightarrow \infty} \mathbb{P}(e^{-T_n^-(\varepsilon)} \mathcal{Z}^\phi(T_n^+(\varepsilon)) > X + \delta) \\ & \quad + \limsup_{n \rightarrow \infty} \mathbb{P}(e^{-T_n} \mathcal{Z}^\phi(T_n) > e^{-T_n^-(\varepsilon)} \mathcal{Z}^\phi(T_n^+(\varepsilon))) \\ &= \limsup_{n \rightarrow \infty} \mathbb{P}(e^{2\varepsilon} e^{-T_n^+(\varepsilon)} \mathcal{Z}^\phi(T_n^+(\varepsilon)) > X + \delta) \\ & \leq \limsup_{n \rightarrow \infty} \mathbb{P}(e^{-T_n^+(\varepsilon)} \mathcal{Z}^\phi(T_n^+(\varepsilon)) > X + \delta/2, A_{\delta, \varepsilon}) + \eta, \end{aligned} \quad (5.25)$$

where the third line comes from (5.23) and the last line follows from (5.24). Note that $T_n^\pm(\varepsilon) = t_n + (\sqrt{\log n} - \log \Theta(t_n) \pm \varepsilon)$ is measurable with respect to \mathcal{F}_{t_n} and

$$\sqrt{\log n} - \log \Theta(t_n) \pm \varepsilon \xrightarrow{\text{a.s.}} \infty \quad \text{and} \quad \frac{\sqrt{\log n} - \log \Theta(t_n) \pm \varepsilon}{t_n} \xrightarrow{\text{a.s.}} 0.$$

Thus applying Theorem 5.1, we get that for any $\delta, \varepsilon > 0$, as $n \rightarrow \infty$,

$$e^{-T_n^\pm(\varepsilon)} \mathcal{Z}^\phi(T_n^\pm(\varepsilon)) \rightarrow X \quad \text{in probability.}$$

Applying this back to (5.25) shows that $\limsup_{n \rightarrow \infty} \mathbb{P}(e^{-T_n} \mathcal{Z}^\phi(T_n) > X + \delta) \leq \eta$ for arbitrary $\eta > 0$, i.e., $\limsup_{n \rightarrow \infty} \mathbb{P}(e^{-T_n} \mathcal{Z}^\phi(T_n) > X + \delta) = 0$. Following a similar argument we have $\limsup_{n \rightarrow \infty} \mathbb{P}(e^{-T_n} \mathcal{Z}^\phi(T_n) < X - \delta) = 0$, thus establishing

$$e^{-T_n} \mathcal{Z}^\phi(T_n) \xrightarrow{\text{P}} X.$$

Since $e^{-T_n} \cdot (n+1) \xrightarrow{\text{a.s.}} W$, combining this with the above equation gives,

$$\begin{aligned}\mathbb{E}[\phi(\mathcal{T}_{v^{(n)}})|\mathcal{T}_n] &= \frac{1}{n+1} \mathcal{Z}^\phi(T_n) = \frac{\mathcal{Z}^\phi(T_n)}{e^{T_n}} \cdot \frac{e^{T_n}}{n+1} \\ &\rightarrow \int_0^\infty e^{-s} \mathbb{E}[\phi(\mathcal{T}^*(s))] ds \quad \text{in probability.}\end{aligned}$$

□

Proof of Theorem 3.3 To complete the proof of part (a), namely, the convergence in probability in the fringe sense, it suffices to show that for any fixed finite rooted tree s_0 , the function $m(u) := \mathbb{P}(\mathcal{T}^*(u) = s_0)$, $u \geq 0$, is Lipschitz continuous in u . But this follows upon noting that there exists a finite positive constant $C(|s_0|)$ depending only on the size of s_0 such that, for any $u > 0$, $\delta > 0$,

$$\begin{aligned}|\mathbb{P}(\mathcal{T}^*(u + \delta) = s_0) - \mathbb{P}(\mathcal{T}^*(u) = s_0)| &\leq \sum_{s \subseteq s_0} \mathbb{P}(\text{There is a birth in the Yule process with initial population size } |s| \text{ before time } \delta) \\ &\leq C(|s_0|)\delta,\end{aligned}$$

where the sum above is over all rooted subtrees of s_0 .

Next we prove part (c). Observe that, from part (a), $\mathbb{E}[D] = \mathbb{E}(\mathcal{P}_1(\tau))$, where $\mathcal{P}_1(\cdot)$ is defined in Sect. 4.4 and τ is an independent $\text{Exp}(1)$ random variable. Using Lemma 4.11,

$$\mathbb{E}[D] = \mathbb{E}(\mathcal{P}_1(\tau)) = \sum_{i=0}^{\infty} \frac{\mathbb{E}(\tau^i)}{i!} \mathbb{P}(\bar{T}_1 = i) = \sum_{i=0}^{\infty} \mathbb{P}(\bar{T}_1 = i) = \mathbb{P}(\bar{T}_1 < \infty),$$

where we have used $\mathbb{E}[\tau^i] = i!$ to obtain the second equality. By standard results on recurrence of random walks, eg. see the Remark after Lemma 1 in [55], the last term above is 1 if and only if $\mathbb{E}[Z] \leq 1$. This proves the claimed assertions on $\mathbb{E}[D]$. Further, again using Lemma 4.11,

$$\begin{aligned}\mathbb{E}(|\mathcal{T}^*(\tau, \mathbf{p})|) &= \sum_{k=0}^{\infty} \mathbb{E}(\mathcal{P}_k(\tau)) = \sum_{k=0}^{\infty} \sum_{i=0}^{\infty} \frac{\mathbb{E}(\tau^i)}{i!} \mathbb{P}(\bar{T}_k = i) = \sum_{k=0}^{\infty} \mathbb{P}(\bar{T}_k < \infty) \\ &= 1 + \sum_{k=1}^{\infty} (\mathbb{P}(\bar{T}_1 < \infty))^k.\end{aligned}$$

The right hand side above is finite if and only if $\mathbb{E}[D] = \mathbb{P}(\bar{T}_1 < \infty) < 1$, proving the assertions on expected tree size.

Part (b) now follows from Theorem 2.4 (a). □

6 Proofs: Degree distribution asymptotics

In this section, we prove Theorem 3.8. The high correlation in the evolution of degrees of different vertices renders conventional tools inapplicable and one has to develop new stochastic analytic techniques to track the degree evolution. The proofs of the degree distribution upper bounds in Theorem 3.8 rely crucially on the asymptotics of a weighted linear combination of vertex counts at different distances from the root of \mathcal{T}^* , summarized in Theorem 6.2. This theorem also plays a key role in subsequent sections involving fixed vertex degree asymptotics and PageRank asymptotics. The lower bounds in Theorem 3.8 rely on a softer analysis involving approximation by multitype branching processes with finitely many types using tools developed in Sect. 4.5.

Recall the process $\mathcal{P}(\cdot) = (\mathcal{P}_i(\cdot) : i \geq 0)$ from Sect. 4.4 keeping track of the number of vertices at various levels in the process \mathcal{T}^* . Let $\tau \sim \text{Exp}(1)$ be an independent random variable. The limit in Theorem 3.3 now results in the following description of the limit degree distribution.

Corollary 6.1 *Let D be as in Theorem 3.3. Then $D \stackrel{d}{=} \mathcal{P}_1(\tau)$.*

Thus understanding the evolution of \mathcal{P} will play a key role in the proof of Theorem 3.8. We will begin by stating Theorem 6.2. Assuming this Theorem, we will prove Theorem 3.8. The rest of the section compartmentalized in Sect. 6.1 is then devoted to the proof of Theorem 6.2. Recall the matrix $\mathbf{A} = (A_{ij})$ as in (4.6). With s_0 as in Definition 3.6 let,

$$\mathcal{P}_s^*(t) = \sum_{i=1}^{\infty} s^{-i} \mathcal{P}_i(t), \quad \text{for } t \geq 0, s > 0, \quad \mathcal{P}^*(t) \equiv \mathcal{P}_{s_0}^*(t). \quad (6.1)$$

Recall the p.g.f of \mathbf{p} , $f(\cdot)$.

Theorem 6.2 (i) *For any $s > 0$ such that $f(s) < \infty$ and for all $t \geq 0$,*

$$\mathbb{E}[\mathcal{P}_s^*(t)] \leq \frac{p_0}{f(s)} e^{\frac{f(s)}{s} t}. \quad (6.2)$$

(ii) *When $\mathbb{E}[Z] \leq 1$, for any $s \in [1, s_0]$, $\theta \geq 1$, there exists constant $C_{\theta,s} < \infty$ such that,*

$$\mathbb{E}[(\mathcal{P}_s^*(t))^{\theta}] \leq C_{\theta,s} e^{\frac{f(s)}{s} \theta t}, \quad \forall t > 0. \quad (6.3)$$

(iii) *When $\mathbb{E}[Z] > 1$ so that $s_0 < 1$ by Lemma 3.7, let $\alpha^*(\theta) := \max\{\frac{f(s_0^\theta)}{s_0^\theta}, \frac{\theta}{R}\}$. For any $\theta \geq 1$, there exists some constant $C_\theta > 0$ such that for all $t \geq 0$,*

$$\mathbb{E}[(\mathcal{P}^*(t))^\theta] \leq C_\theta (1 + t^\theta) e^{\alpha^*(\theta)t}. \quad (6.4)$$

Corollary 6.3 *Define $\tilde{\mathcal{P}}_s^*(t) := \sum_{i=1}^{\infty} s^{-i} \tilde{\mathcal{P}}_i(t)$.*

(i) For all $t \geq 0$ and $s \geq 1$ such that $f(s) < \infty$,

$$\mathbb{E}[\tilde{\mathcal{P}}_s^\star(t)] \leq \frac{p_0}{f(s)} e^{\frac{f(s)}{s}t}.$$

(ii) When $\mathbb{E}[Z] \leq 1$, for any $s \in [1, s_0]$, $\theta \geq 1$, there exists constant $C_{\theta,s} < \infty$ such that,

$$\mathbb{E}[(\tilde{\mathcal{P}}_s^\star(t))^\theta] \leq C_{\theta,s} e^{\frac{f(s)}{s}\theta t}, \quad \forall t > 0. \quad (6.5)$$

Proof The result follows from the same proof as of Theorem 6.2 upon noting that $(s^{-i} : i \geq 0)$ is a left subinvariant eigenvector of \mathbf{B} associated with the eigenvalue $f(s)/s$ when $s \geq 1$ (see Proposition 4.12(b)). \square

The proof of Theorem 6.2 is deferred to the end of this section. The above bounds coupled with the preliminary analysis of \mathcal{P}_1 in Sect. 4.5 immediately lead to the following two Corollaries.

Corollary 6.4 For any $\delta > 0$, we have the following limit

$$\lim_{t \rightarrow \infty} e^{-\left(\frac{1}{R} - \delta\right)t} \mathcal{P}_1(t) = \infty \quad a.s. \quad (6.6)$$

$$\lim_{t \rightarrow \infty} t^{-(1+\delta)} e^{-t/R} \mathcal{P}_1(t) = 0 \quad a.s. \quad (6.7)$$

Proof Fix $\delta \in (0, 1/R)$. Using Proposition 4.16 choose $k = k_\delta \in \mathbb{N}$ large enough such that the Perron–Frobenius eigenvalue α_k of the $k \times k$ principal submatrix A_k of A satisfies

$$\frac{1}{R} - \frac{\delta}{2} < \alpha_k \leq \frac{1}{R}.$$

By the stochastic domination in Lemma 4.14 and the limit result for finite urns in Proposition 4.15,

$$\begin{aligned} \lim_{t \rightarrow \infty} e^{-\left(\frac{1}{R} - \delta\right)t} \mathcal{P}_1(t) &= \lim_{t \rightarrow \infty} e^{-\left(\frac{1}{R} - \delta\right)(t + \sigma_1)} \mathcal{P}_1(t + \sigma_1) \\ &\geq e^{-\left(\frac{1}{R} - \delta\right)\sigma_1} \lim_{t \rightarrow \infty} e^{-\left(\frac{1}{R} - \delta\right)t} \bar{\mathcal{P}}_1(t) \\ &\geq e^{-\left(\frac{1}{R} - \delta\right)\sigma_1} \lim_{t \rightarrow \infty} e^{\delta t/2} \cdot (e^{-\alpha_k t} \bar{\mathcal{P}}_1(t)) = \infty \quad a.s. \end{aligned}$$

This proves (6.6). To prove (6.7), for any given $\varepsilon > 0$ and $N \geq 0$ we can define the event

$$E_N = \left\{ \sup_{s \in [N, N+1]} s^{-(1+\delta)} e^{-s/R} \mathcal{P}_1(s) > \varepsilon \right\}.$$

By Theorem 6.2, for any $t \geq 0$

$$\mathbb{E}[\mathcal{P}_1(t)] \leq s_0 \mathbb{E}[\mathcal{P}^*(t)] \leq \frac{p_0 s_0}{f(s_0)} e^{\frac{f(s_0)}{s_0} t} = p_0 R e^{t/R}.$$

Hence,

$$\begin{aligned} \mathbb{P}(E_N) &\leq \mathbb{P}(\mathcal{P}_1(N+1) > \varepsilon e^{N/R} N^{1+\delta}) \leq \frac{\mathbb{E}[\mathcal{P}_1(N+1)]}{\varepsilon e^{N/R} N^{1+\delta}} \\ &\leq \frac{p_0 R e^{(N+1)/R}}{\varepsilon e^{N/R} N^{1+\delta}} = \frac{p_0 R e^{1/R}}{\varepsilon N^{1+\delta}}. \end{aligned}$$

Applying Borel-Cantelli Lemma then gives $\mathbb{P}(\limsup_{N \rightarrow \infty} E_N) = 0$. Hence,

$$t^{-(1+\delta)} e^{-t/R} \mathcal{P}_1(t) \xrightarrow{a.s.} 0,$$

proving (6.7). \square

Corollary 6.5 *Let $\tau \sim \text{Exp}(1)$ independent of \mathcal{T}^* .*

- (i) *When $\mathbb{E}[Z] \leq 1$, for $\theta \in [1, \frac{s}{f(s)})$ and $s \in [1, s_0]$, $\mathbb{E}[(\mathcal{P}_s^*(\tau))^\theta] < \infty$.*
- (ii) *When $\mathbb{E}[Z] > 1$, let $q_* < 1$ be as in Definition 3.6. For $\theta \in [1, R \wedge \frac{\log q_*}{\log s_0})$, $\mathbb{E}[(\mathcal{P}^*(\tau))^\theta] < \infty$.*

Proof (i) Simply follows from (6.3). For $\theta \in (0, \frac{s}{f(s)})$,

$$\mathbb{E}[(\mathcal{P}_s^*(\tau))^\theta] = \int_0^\infty e^{-t} \mathbb{E}[(\mathcal{P}_s^*(t))^\theta] dt \leq C_{\theta,s} \int_0^\infty e^{-t} e^{\frac{f(s)}{s} \theta t} dt < \infty.$$

(ii) We first start by evaluating $\alpha^*(\cdot)$ in Theorem 6.2(iii). For $\mathbb{E}[Z] > 1$ and $\theta \in [1, R \wedge \frac{\log q_*}{\log s_0})$ we obviously have then $\theta/R < 1$. Further, noting that $s_0 \geq s_0^{\theta} > s_0^{(\log q_*)/(\log s_0)}$, Lemma 3.7(a) gives

$$\frac{f(s_0^\theta)}{s_0^\theta} < \frac{f(s_0^{(\log q_*)/(\log s_0)})}{s_0^{(\log q_*)/(\log s_0)}} = \frac{f(q_*)}{q_*} = 1.$$

Hence $\alpha^*(\theta) < 1$. Using (6.5) gives

$$\mathbb{E}[(\mathcal{P}^*(\tau))^\theta] = \int_0^\infty e^{-t} \mathbb{E}[(\mathcal{P}^*(t))^\theta] dt \leq C_\theta \int_0^\infty e^{-t} (1 + t^\theta) e^{\alpha^*(\theta)t} dt < \infty,$$

from which the result follows. \square

Proof of Theorem 3.8 assuming Theorem 6.2 Throughout we use the representation in Corollary 6.1.

Upper bound for the tail exponent: Note that by construction, for any $s > 0$, $\mathcal{P}_1(t) \leq s\mathcal{P}_s^*(t)$. We start with the regime $\mathbb{E}[Z] \leq 1$. Note that $s/f(s) \uparrow R$ as $s \uparrow s_0$. Corollary 6.5(i) gives that for any $\delta \in (0, R)$, there exists some $s_\delta > 0$ such that

$$\mathbb{E}(\mathcal{P}_1(\tau)^{R-\delta}) \leq s_\delta \mathbb{E}(\mathcal{P}_{s_\delta}^*(\tau)^{R-\delta}) \leq s_\delta C_\delta.$$

It follows that

$$\mathbb{P}(D \geq k) = \mathbb{P}(\mathcal{P}_1(\tau) \geq k) \leq \frac{\mathbb{E}[(\mathcal{P}_1(\tau))^{R-\delta}]}{k^{R-\delta}} \leq \frac{s_\delta C_\delta}{k^{R-\delta}}, \quad (6.8)$$

for finite constant C_δ . This implies that,

$$\limsup_{k \rightarrow \infty} \frac{\log \mathbb{P}(D \geq k)}{\log k} \leq \limsup_{k \rightarrow \infty} \frac{\log(s_\delta C_\delta) - (R - \delta) \log k}{\log k} = -R + \delta.$$

As $\delta > 0$ can be chosen arbitrarily small, this completes the upper bound for the tail exponent when $\mathbb{E}[Z] \leq 1$.

The case $\mathbb{E}[Z] > 1$ follows the exact same argument but using Corollary 6.5(ii) with the exponent R in (6.8) replaced by $R \wedge \frac{\log q_*}{\log s_0}$.

Lower bound for the tail exponent: Here we want to show that R is a lower bound on the tail exponent in all regimes. Fix $0 < \delta < R$. Note that for any fixed $k \geq 1$,

$$\begin{aligned} \mathbb{P}(D \geq k) &= \int_0^\infty e^{-s} \mathbb{P}(\mathcal{P}_1(s) \geq k) ds = \int_0^\infty e^{-s} \mathbb{P}\left(e^{-\frac{s}{R+\delta}} \mathcal{P}_1(s) \geq e^{-\frac{s}{R+\delta}} k\right) ds \\ &\geq \int_{\left\{s: e^{-\frac{s}{R+\delta}} k \leq 1\right\}} e^{-s} \mathbb{P}\left(e^{-\frac{s}{R+\delta}} \mathcal{P}_1(s) \geq 1\right) ds \\ &= \int_{(R+\delta) \log k}^\infty e^{-s} \mathbb{P}\left(e^{-\frac{s}{R+\delta}} \mathcal{P}_1(s) \geq 1\right) ds. \end{aligned} \quad (6.9)$$

By (6.6) in Corollary 6.4, $e^{-\frac{s}{R+\delta}} \mathcal{P}_1(s) \rightarrow \infty$ a.s. Thus there exists $k_0 \geq 1$ such that $\mathbb{P}\left(e^{-\frac{s}{R+\delta}} \mathcal{P}_1(s) \geq 1\right) \geq 1/2$ for all $s \geq (R + \delta) \log k_0$. Thus, for $k \geq k_0$, we have

$$\mathbb{P}(D \geq k) \geq \frac{1}{2} \int_{(R+\delta) \log k}^\infty e^{-s} ds = \frac{1}{2k^{R+\delta}},$$

which leads to $\liminf_{k \rightarrow \infty} \log \mathbb{P}(D \geq k)/\log k \geq -(R + \delta)$ for arbitrarily small $\delta > 0$. \square

6.1 Proof of Theorem 6.2

The following lemma gives a tractable formulation for the expectation of powers of \mathcal{P}^* . This will be used to set up differential equations involving $\mathbb{E}[(\mathcal{P}^*(\cdot))^\theta]$ for $\theta \geq 1$ whose analysis will lead to the proof of Theorem 6.2. In the following, a crucially

used object will be the generator $\bar{\mathcal{L}}$ of the continuous time Markov process $\mathcal{P}(\cdot)$ taking values in

$$\bar{\mathcal{S}} := \left\{ \mathbf{x} = (x_i : i \geq 0) \in (\mathbb{N}_0)^{\mathbb{N}_0} : x_0 = 1, \exists l_{\mathbf{x}} \in \mathbb{N} \text{ such that } x_l > 0 \forall l \leq l_{\mathbf{x}} \text{ and } x_l = 0 \forall l > l_{\mathbf{x}} \right\}.$$

For any function $f : \bar{\mathcal{S}} \rightarrow \mathbb{R}$, define the action of the generator $\bar{\mathcal{L}}$ on f as the function $\bar{\mathcal{L}}f : \bar{\mathcal{S}} \rightarrow \mathbb{R}$ given by

$$\bar{\mathcal{L}}f(\mathbf{x}) = \sum_{i=1}^{\infty} [f(\mathbf{x} + \mathbf{e}_i) - f(\mathbf{x})] \sum_{j=0}^{\infty} A_{ij} x_j, \quad \mathbf{x} \in \bar{\mathcal{S}},$$

whenever the right hand side above is well-defined (here \mathbf{e}_i is the i -th coordinate unit vector). For notational convenience, for $g : \bar{\mathcal{S}} \rightarrow \mathbb{R}$, we will write $\mathcal{L}[g(\mathcal{P}(t))] := \bar{\mathcal{L}}g(\mathcal{P}(t))$, $t \geq 0$.

Lemma 6.6 *For any $s > 0, t \geq 0$ and $\theta \geq 1$,*

$$\mathbb{E}[(\mathcal{P}_s^*(t))^{\theta}] = \int_0^t \mathbb{E}[\mathcal{L}[(\mathcal{P}_s^*(r))^{\theta}]] dr < \infty, \quad (6.10)$$

where

$$\mathcal{L}[(\mathcal{P}_s^*(t))^{\theta}] = \sum_{i=1}^{\infty} \left[(\mathcal{P}_s^*(t) + s^{-i})^{\theta} - (\mathcal{P}_s^*(t))^{\theta} \right] \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t). \quad (6.11)$$

Proof Let $\text{Yule}(\cdot)$ denote a rate one Yule process as in Lemma 4.5, let $Y(t) = |\text{Yule}(t)|$ denote the size and $\text{ht}(t)$ the corresponding height of the genealogical tree at time t . Note that the height of $\mathcal{T}^*(t)$ is stochastically dominated by the height $\text{ht}(t)$ of the Yule process. Thus for any $\theta \geq 1$ using Cauchy-Schwartz inequality,

$$\begin{aligned} \mathbb{E}[(\mathcal{P}_s^*(t))^{\theta}] &\leq \mathbb{E}[(\sup_{i \leq \text{ht}(t)} s^{-i} \cdot Y(t))^{\theta}] \leq \mathbb{E}[(1 \vee s^{-\theta \text{ht}(t)})(Y(t))^{\theta}] \\ &\leq \sqrt{\mathbb{E}[(1 \vee s^{-2\theta \text{ht}(t)})]} \sqrt{\mathbb{E}[(Y(t))^{2\theta}]} < \infty, \end{aligned}$$

where the finiteness of the first term follows from Lemma 4.5 and finiteness of the second term follows from distributional identity $Y(t) \sim \text{Geometric}(e^{-t})$. This proves $\mathbb{E}[(\mathcal{P}_s^*(t))^{\theta}] < \infty$. To prove (6.10), first consider the truncation $\mathcal{P}_{s,N}^*(t) = \sum_{i=1}^N s^{-i} \mathcal{P}_i(t)$. Applying the generator of the Markov process $\mathcal{P}(\cdot)$ gives

$$\mathbb{E}[(\mathcal{P}_{s,N}^*(t))^{\theta}] = \int_0^t \mathbb{E}[\mathcal{L}[(\mathcal{P}_{s,N}^*(r))^{\theta}]] dr$$

where,

$$\mathcal{L}[(\mathcal{P}_{s,N}^*(t))^\theta] := \sum_{i=1}^{\infty} \mathbb{1}_{\{i \leq N\}} \left[(\mathcal{P}_{s,N}^*(t) + s^{-i})^\theta - (\mathcal{P}_{s,N}^*(t))^\theta \right] \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t).$$

Note that for each $i \geq 1$,

$$\mathbb{1}_{\{i \leq N\}} \left[(\mathcal{P}_{s,N}^*(t) + s^{-i})^\theta - (\mathcal{P}_{s,N}^*(t))^\theta \right] \nearrow \left[(\mathcal{P}_s^*(t) + s^{-i})^\theta - (\mathcal{P}_s^*(t))^\theta \right]$$

as $N \rightarrow \infty$. By monotone convergence theorem,

$$\mathbb{E}[(\mathcal{P}_s^*(t))^\theta] = \lim_{N \rightarrow \infty} \mathbb{E}[(\mathcal{P}_{s,N}^*(t))^\theta] = \int_0^t \mathbb{E}[\mathcal{L}[(\mathcal{P}_s^*(r))^\theta]] dr.$$

□

Remark 13 By the same argument conditional on \mathcal{F}_u for any fixed $u \geq 0$, it follows that $\mathbb{E}[(\mathcal{P}_s^*(t+u))^\theta - (\mathcal{P}_s^*(u))^\theta | \mathcal{F}_u] = \int_u^{t+u} \mathbb{E}[\mathcal{L}[(\mathcal{P}_s^*(r))^\theta] | \mathcal{F}_u] dr$, $t \geq 0$. This implies that the process $\left(\mathcal{P}_s^*(t)^\theta - \int_0^t \mathcal{L}[(\mathcal{P}_s^*(r))^\theta] dr \right)_{t \geq 0}$ is a martingale.

Proof of Theorem 6.2(i) We begin by proving (6.2) namely $\theta = 1$ case. Using (6.11) and the form of \mathbf{A} gives,

$$\begin{aligned} \mathcal{L}(\mathcal{P}_s^*(t)) &= \sum_{i=1}^{\infty} s^{-i} \sum_{j=i-1}^{\infty} p_{j-i+1} \mathcal{P}_j(t) = \sum_{j=0}^{\infty} \left[\sum_{i=1}^{j+1} s^{-i} p_{j-i+1} \right] \mathcal{P}_j(t), \\ &= \sum_{j=0}^{\infty} \left[\sum_{i=1}^{j+1} s^{j-i+1} p_{j-i+1} \right] s^{-(j+1)} \mathcal{P}_j(t) \leq \frac{p_0}{s} + \frac{f(s)}{s} \sum_{j=1}^{\infty} s^{-j} \mathcal{P}_j(t) \\ &= \frac{p_0}{s} + \frac{f(s)}{s} \mathcal{P}_s^*(t). \end{aligned}$$

Here we have essentially re-derived Proposition 4.12(a) on $(s^{-i} : i \geq 0)$ being a sub-invariant eigenvector with eigenvalue $f(s)/s$. Using Lemma 6.6 gives

$$\frac{d}{dt} \mathbb{E}(\mathcal{P}_s^*(t)) \leq \frac{p_0}{s} + \frac{f(s)}{s} \mathbb{E}(\mathcal{P}_s^*(t)).$$

Integrating completes the proof. □

Proof of Theorem 6.2(ii) To prove (6.3), we will first start by assuming that θ is an integer and argue by induction. After completing the proof for integer θ , we will extend the proof to general θ . By (6.2), the assertion is true when $\theta = 1$. We will use

A_{ij} instead of p_{j-i+1} for ease of notation. Suppose (6.3) holds for $\theta \leq k-1$.

$$\begin{aligned}
\mathcal{L}[(\mathcal{P}_s^*(t))^k] &= \sum_{i=1}^{\infty} \left[(\mathcal{P}_s^*(t) + s^{-i})^k - (\mathcal{P}_s^*(t))^k \right] \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t) \\
&= \sum_{i=1}^{\infty} \sum_{l=1}^k \binom{k}{l} (\mathcal{P}_s^*(t))^{k-l} s^{-il} \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t) \\
&\leq \sum_{i=1}^{\infty} \sum_{l=1}^k \binom{k}{l} (\mathcal{P}_s^*(t))^{k-l} s^{-i} \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t) \\
&= \sum_{l=1}^k \binom{k}{l} (\mathcal{P}_s^*(t))^{k-l} \sum_{j=0}^{\infty} \left(\sum_{i=1}^{\infty} s^{-i} A_{ij} \right) \mathcal{P}_j(t) \\
&\leq \sum_{l=1}^k \binom{k}{l} (\mathcal{P}_s^*(t))^{k-l} \left(s^{-1} p_0 + \frac{f(s)}{s} \mathcal{P}_s^*(t) \right) \\
&= \frac{k f(s)}{s} (\mathcal{P}_s^*(t))^k + \frac{f(s)}{s} \sum_{l=2}^k \binom{k}{l} (\mathcal{P}_s^*(t))^{k-l+1} \\
&\quad + s^{-1} p_0 \sum_{l=1}^k \binom{k}{l} (\mathcal{P}_s^*(t))^{k-l}.
\end{aligned}$$

Here the key inequality in line two follows from assuming that $s \geq 1$ so that $s^{-il} \leq s^{-i}$, whilst the inequalities in the ensuing lines mimic calculations in the $\theta = 1$ case. By the induction hypothesis, for $1 \leq i \leq k-1$, \exists finite constants $C_{i,s}$ such that $\mathbb{E}((\mathcal{P}_s^*(t))^i) \leq C_{i,s} \exp(\frac{f(s)}{s} i \cdot t)$. Let $\tilde{C}_{k,s} := \max_{1 \leq i \leq k-1} \{C_{i,s}\}$. Using Lemma 6.6 gives,

$$\begin{aligned}
\frac{d}{dt} \left[e^{-k \frac{f(s)}{s} t} \mathbb{E}[(\mathcal{P}_s^*(t))^k] \right] &\leq e^{-k \frac{f(s)}{s} t} \left[\frac{f(s)}{s} \sum_{l=2}^k \binom{k}{l} \mathbb{E}[(\mathcal{P}_s^*(t))^{k-l+1}] \right. \\
&\quad \left. + s^{-1} p_0 \sum_{l=1}^k \binom{k}{l} \mathbb{E}[(\mathcal{P}_s^*(t))^{k-l}] \right] \\
&\leq \tilde{C}_{k,s} e^{-\frac{f(s)}{s} t} \cdot \left[\frac{f(s)}{s} \sum_{l=2}^k \binom{k}{l} e^{-(l-2) \frac{f(s)}{s} t} \right. \\
&\quad \left. + s^{-1} p_0 \sum_{l=1}^k \binom{k}{l} e^{-(l-1) \frac{f(s)}{s} t} \right] \\
&\leq \tilde{C}_{k,s} e^{-\frac{f(s)}{s} t} \cdot 2^k \left(\frac{f(s)}{s} + s^{-1} p_0 \right).
\end{aligned}$$

Therefore, there exists $C_{k,s} > 0$ so that

$$\mathbb{E}[(\mathcal{P}_s^\star(t))^k] \leq \tilde{C}_{k,s} e^{k \frac{f(s)}{s} t} \int_0^t e^{-\frac{f(s)}{s} r} dr \cdot 2^k \left(\frac{f(s)}{s} + s^{-1} p_0 \right) \leq C_{k,s} e^{k \frac{f(s)}{s} t},$$

proving (6.3) for $\theta = k$. To extend to $\theta \in (k-1, k]$, we apply Jensen's inequality to obtain,

$$\mathbb{E}[(\mathcal{P}_s^\star(t))^\theta] \leq \mathbb{E}[(\mathcal{P}_s^\star(t))^k]^{\frac{\theta}{k}} \leq C_{k,s}^{\frac{\theta}{k}} e^{\theta \frac{f(s)}{s} t} = C_{\theta,s} e^{\theta \frac{f(s)}{s} t}.$$

□

Proof of Theorem 6.2(iii) It remains to prove (6.5) when $\mathbb{E}[Z] > 1$ so that $s_0 < 1$. Recall that $\mathcal{P}^\star = \mathcal{P}_{s_0}^\star$. Note that for any $\theta > 1$, $i \geq 1$,

$$\begin{aligned} & (\mathcal{P}^\star(t) + s_0^{-i})^\theta - \mathcal{P}^\star(t)^\theta \\ & + \frac{\theta(\theta-1)}{2} s_0^{-2i} [\mathcal{P}^\star(t)^{\theta-2} \vee (\mathcal{P}^\star(t) + s_0^{-i})^{\theta-2}]. \end{aligned}$$

It follows that

$$\begin{aligned} \mathbb{E}[(\mathcal{P}^\star(t))^\theta] &= \sum_{i=1}^{\infty} \left[(\mathcal{P}^\star(t) + s_0^{-i})^\theta - (\mathcal{P}^\star(t))^\theta \right] \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t) \\ &\leq \sum_{i=1}^{\infty} \theta s_0^{-i} \mathcal{P}^\star(t)^{\theta-1} \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t) \\ &+ \sum_{i=1}^{\infty} \frac{\theta(\theta-1)}{2} s_0^{-2i} [\mathcal{P}^\star(t)^{\theta-2} \vee (\mathcal{P}^\star(t) + s_0^{-i})^{\theta-2}] \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t). \end{aligned} \tag{6.12}$$

Since $\sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t) = \sum_{j=i-1}^{\infty} p_{j+1-i} \mathcal{P}_j(t)$, it is easy to see $\sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t) > 0$ only when $\mathcal{P}_{i-1}(t) \geq 1$, which implies $\mathcal{P}^\star(t) \geq s_0^{-(i-1)}$.

Hence for $\theta \geq 2$,

$$\left[\mathcal{P}^\star(t)^{\theta-2} \vee (\mathcal{P}^\star(t) + s_0^{-i})^{\theta-2} \right] \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t) \leq (1 + 1/s_0)^{\theta-2} \mathcal{P}^\star(t)^{\theta-2} \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t).$$

Taking into account the case where $\theta \in (1, 2)$, we have

$$\begin{aligned} \left[\mathcal{P}^\star(t)^{\theta-2} \vee (\mathcal{P}^\star(t) + s_0^{-i})^{\theta-2} \right] \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t) &\leq \left(1 \vee (1 + 1/s_0)^{\theta-2} \right) \mathcal{P}^\star(t)^{\theta-2} \\ &\quad \sum_{j=0}^{\infty} A_{ij} \mathcal{P}_j(t). \end{aligned}$$

Plugging this back into (6.12) gives

$$\begin{aligned}
\mathcal{L}[(\mathcal{P}^*(t))^\theta] &\leq \theta \mathcal{P}^*(t)^{\theta-1} \sum_{j=0}^{\infty} \left(\sum_{i=1}^{\infty} s_0^{-i} A_{ij} \right) \mathcal{P}_j(t) \\
&\quad + \frac{\theta(\theta-1)}{2} (1 \vee (1 + 1/s_0)^{\theta-2}) \mathcal{P}^*(t)^{\theta-2} \sum_{j=0}^{\infty} \left(\sum_{i=1}^{\infty} s_0^{-2i} A_{ij} \right) \mathcal{P}_j(t) \\
&\leq \frac{\theta p_0}{s_0} \mathcal{P}^*(t)^{\theta-1} + \frac{\theta}{R} \mathcal{P}^*(t)^\theta + \frac{\theta(\theta-1)}{2} (1 \vee (1 + 1/s_0)^{\theta-2}) \frac{p_0}{s_0^2} \mathcal{P}^*(t)^{\theta-2} \\
&\quad + \frac{\theta(\theta-1)}{2} (1 \vee (1 + 1/s_0)^{\theta-2}) \frac{f(s_0^2)}{s_0^2} \mathcal{P}^*(t)^{\theta-2} \sum_{j=1}^{\infty} s_0^{-2j} \mathcal{P}_j(t).
\end{aligned}$$

Define $d(t) = \sup\{i \geq 0 : \mathcal{P}_i(t) \geq 1\}$. Observe that, since $s_0 \in (0, 1)$, $\sum_{j=1}^{\infty} s_0^{-2j} \mathcal{P}_j(t) \leq s_0^{-d(t)} \mathcal{P}^*(t)$. Hence,

$$\begin{aligned}
\mathcal{L}[(\mathcal{P}^*(t))^\theta] &\leq \left(\frac{\theta p_0}{s_0} + \frac{\theta(\theta-1)}{2} (1 \vee (1 + 1/s_0)^{\theta-2}) \frac{f(s_0^2)}{s_0^2} s_0^{-d(t)} \right) \mathcal{P}^*(t)^{\theta-1} \\
&\quad + \frac{\theta}{R} \mathcal{P}^*(t)^\theta + \frac{\theta(\theta-1)}{2} (1 \vee (1 + 1/s_0)^{\theta-2}) \frac{p_0}{s_0^2} \mathcal{P}^*(t)^{\theta-2} \\
&\leq \tilde{C}_\theta s_0^{-d(t)} \mathcal{P}^*(t)^{\theta-1} + \frac{\theta}{R} \mathcal{P}^*(t)^\theta,
\end{aligned} \tag{6.13}$$

for some finite constant \tilde{C}_θ that depends on θ , s_0 and p_0 . Notice that the term $\mathcal{P}^*(t)^{\theta-2}$ can be upper bounded by $\mathcal{P}^*(t)^{\theta-1}$ for all $\theta \geq 1$ since $\mathcal{P}^*(t) \geq 1$. For any $\theta \geq 1$,

$$\begin{aligned}
\frac{d}{dt} \mathbb{E}[\mathcal{P}^*(t)^\theta] &= \mathbb{E}[\mathcal{L}[(\mathcal{P}^*(t))^\theta]] \leq \tilde{C}_\theta \mathbb{E}[s_0^{-d(t)} \mathcal{P}^*(t)^{\theta-1}] + \frac{\theta}{R} \mathbb{E}[\mathcal{P}^*(t)^\theta] \\
&\leq \tilde{C}_\theta \left(\mathbb{E}[\mathcal{P}^*(t)^\theta] \right)^{\frac{\theta-1}{\theta}} \left(\mathbb{E}[s_0^{-\theta d(t)}] \right)^{\frac{1}{\theta}} + \frac{\theta}{R} \mathbb{E}[\mathcal{P}^*(t)^\theta].
\end{aligned}$$

By the definition of $\alpha^*(\theta)$ and (6.2),

$$\mathbb{E}[s_0^{-\theta d(t)}] \leq \mathbb{E} \left[\sum_{j=1}^{\infty} s_0^{-j\theta} \mathcal{P}_j(t) \right] \leq \frac{p_0}{f(s_0^\theta)} e^{\frac{f(s_0^\theta)}{s_0^\theta} t} \leq \frac{p_0}{f(s_0^\theta)} e^{\alpha^*(\theta) t}.$$

Let $g(t) = e^{-\alpha^*(\theta)t} \mathbb{E}[\mathcal{P}^*(t)^\theta]$ and $C'_\theta = \tilde{C}_\theta \left(\frac{p_0}{f(s_0^\theta)} \right)^{1/\theta}$. Then

$$g'(t) \leq C'_\theta (g(t))^{\frac{\theta-1}{\theta}}. \tag{6.14}$$

Define $h(t) = 1 + g(0)^{1/\theta} + \frac{2C'_\theta}{\theta}t - g(t)^{1/\theta}$. Then (6.14) implies that,

$$h'(t) = \frac{2C'_\theta}{\theta} - \frac{1}{\theta}g'(t)(g(t))^{\frac{1}{\theta}-1} \geq \frac{C'_\theta}{\theta}.$$

Since $h(0) > 0$ we have $h(t) > 0$. It then follows that there exists some $C_\theta > 0$ so that $g(t) \leq C_\theta(1 + t^\theta)$ for all $t \geq 0$, i.e.,

$$\mathbb{E}[\mathcal{P}^\star(t)^\theta] \leq C_\theta(1 + t^\theta)e^{\alpha^*(\theta)t},$$

which is assertion (iii). \square

7 Proofs: Condensation and fixed vertex degree asymptotics

We abbreviate $d_n(v_k) = \deg(v_k, n)$, $k \geq 0, n \geq k$. We will first prove the non-root fixed vertex asymptotics.

Proof of Theorem 3.10 This follows from a direct application of Corollary 6.4. Using the continuous time embedding, note that for any $i \geq 1$,

$$\{d_n(v_i) : n \geq i\} \stackrel{d}{=} \{\mathcal{P}_1^{v_i}(T_n - \sigma_i) : n \geq i\},$$

where as before $T_n = \inf\{t \geq 0 : |\mathcal{T}(t)| = n + 1\}$, σ_i is the birth time of v_i and $\mathcal{P}_1^{v_i}(t)$ denotes the number of children of vertex v_i in $\mathcal{T}(t + \sigma_i)$. Observe that $\mathcal{P}_1^{v_i}(\cdot)$ has the same distribution as the process $\mathcal{P}_1(\cdot)$. By Lemma 4.3, $\frac{e^{T_n}}{n} \xrightarrow{a.s.} \frac{1}{W}$ where $W \sim \text{Exp}(1)$. It follows from (6.6) that

$$\frac{d_n(v_i)}{n^{1/R-\delta}} = \frac{\mathcal{P}_1^{v_i}(T_n - \sigma_i)}{n^{1/R-\delta}} = \frac{\mathcal{P}_1^{v_i}(T_n - \sigma_i)}{e^{(1/R-\delta)(T_n - \sigma_i)}} \cdot \left(\frac{e^{(T_n - \sigma_i)}}{n}\right)^{1/R-\delta} \xrightarrow{a.s.} \infty.$$

Similarly, by (6.7),

$$\begin{aligned} \frac{d_n(v_i)}{n^{1/R}(\log n)^{1+\delta}} &= \frac{\mathcal{P}_1^{v_i}(T_n - \sigma_i)}{n^{1/R}(\log n)^{1+\delta}} \\ &= \frac{\mathcal{P}_1^{v_i}(T_n - \sigma_i)}{e^{(T_n - \sigma_i)/R} \cdot (T_n - \sigma_i)^{1+\delta}} \cdot \left(\frac{e^{(T_n - \sigma_i)}}{n}\right)^{1/R} \cdot \left(\frac{T_n - \sigma_i}{\log n}\right)^{1+\delta} \xrightarrow{a.s.} 0. \end{aligned}$$

\square

Now we proceed to root degree asymptotics.

Proof of Theorem 3.9(a) Recall that T_n is the stopping time when $n(t) = |\mathcal{T}(t)|$ first becomes $n + 1$. Standard properties of the Yule process (see Lemma 4.2) imply that

$$\mathbb{P}(|T_n - \log n| \geq M) \leq 2e^{-M},$$

where $M > 0$ will be chosen later.

For any fixed $\varepsilon > 0$, it follows from Theorem 4.21 that there is some $t_0(\varepsilon) > 0$ such that for $t \geq t_0$,

$$\mathbb{E}|e^{-t}\tilde{\mathcal{P}}_1(t) - W_\infty| \leq \varepsilon^3, \quad (7.1)$$

where $W_\infty := (1 - q_*)W$.

For any $n \geq e^{t_0+M}$ and some $\delta > 0$ to be chosen later, let $a_{n,0} = \log n - M$ and define $a_{n,i} = a_{n,i-1} + \delta$ for $i \geq 1$. Then we can observe that

$$\begin{aligned} \mathbb{E}|e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty| &\leq \sum_{i=0}^{\lceil 2M/\delta \rceil} \mathbb{E} \left[|e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty| \cdot 1_{\{a_i \leq T_n < a_{i+1}\}} \right] \\ &\quad + \mathbb{E} \left[|e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty| \cdot 1_{\{|T_n - \log n| \geq M\}} \right]. \end{aligned} \quad (7.2)$$

To address the second term in (7.2), note that

$$\begin{aligned} &\mathbb{E} \left[|e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty| \cdot 1_{\{|T_n - \log n| \geq M\}} \right] \\ &\leq \left(\mathbb{E}[(e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty)^2] \right)^{1/2} \cdot \mathbb{P}(|T_n - \log n| \geq M)^{1/2} \\ &\leq \sqrt{2} \left(2(n+1)^2 \mathbb{E}[e^{-2T_n}] + 2\mathbb{E}[W_\infty^2] \right)^{1/2} e^{-M/2}. \end{aligned}$$

To upper bound $(n+1)^2 \mathbb{E}[e^{-2T_n}]$, write $X_n = (n+1)^2 e^{-2T_n}$. Observe that

$$\begin{aligned} \mathbb{E}[X_{n+1}] &= (n+2)^2 \mathbb{E}[e^{-2(T_{n+1}-T_n)} e^{-2T_n}] \\ &= \mathbb{E}[e^{-2T_n}] (n+2)^2 \cdot \frac{n+1}{n+3} = \mathbb{E}[X_n] \cdot \frac{(n+2)^2}{(n+1)(n+3)}, \end{aligned}$$

where we used the observation $T_{n+1} - T_n \sim \text{Exp}(n+1)$ and is independent of T_n . Hence, for any $n \geq 1$,

$$\begin{aligned} \mathbb{E}[X_{n+1}] &= \mathbb{E}[X_n] \left(1 + \frac{1}{(n+1)(n+3)} \right) \leq \mathbb{E}[X_n] \left(1 + \frac{1}{n^2} \right) \\ &\leq \mathbb{E}[X_0] \prod_{k=1}^n e^{1/k^2} \leq C_0 \end{aligned}$$

for some constant $C_0 > 0$. This combined, along with the explicit form of W_∞ , shows that there exists some constant $C > 0$ so that $\sqrt{2} (2\mathbb{E}[e^{-2T_n} n(T_n)^2] + 2\mathbb{E}[W_\infty^2])^{1/2} \leq C$. Taking $M = 2 \log(2C/\varepsilon)$ then gives

$$\mathbb{E} \left[|e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty| \cdot 1_{\{|T_n - \log n| \geq M\}} \right] \leq \varepsilon/2. \quad (7.3)$$

For $t \in [a_i, a_{i+1})$, we have $e^{-a_{i+1}}\tilde{\mathcal{P}}_1(a_i) \leq e^{-t}\tilde{\mathcal{P}}_1(t) \leq e^{-a_i}\tilde{\mathcal{P}}_1(a_{i+1})$. Hence,

$$\begin{aligned} & \mathbb{E} \left[|e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty| \cdot 1_{\{a_i \leq T_n < a_{i+1}\}} \right] \\ & \leq \mathbb{E} \left[\max \{|e^{-a_{i+1}}\tilde{\mathcal{P}}_1(a_i) - W_\infty|, |e^{-a_i}\tilde{\mathcal{P}}_1(a_{i+1}) - W_\infty|\} \cdot 1_{\{a_i \leq T_n < a_{i+1}\}} \right] \\ & \leq \mathbb{E} \left[\max \{e^{-\delta}|e^{-a_i}\tilde{\mathcal{P}}_1(a_i) - W_\infty|, e^\delta|e^{-a_{i+1}}\tilde{\mathcal{P}}_1(a_{i+1}) - W_\infty|\} \cdot 1_{\{a_i \leq T_n < a_{i+1}\}} \right] \\ & \quad + \max \{e^\delta - 1, 1 - e^{-\delta}\} \mathbb{E}[W_\infty \cdot 1_{\{a_i \leq T_n < a_{i+1}\}}] \\ & \leq e^{-\delta} \mathbb{E}[|e^{-a_i}\tilde{\mathcal{P}}_1(a_i) - W_\infty|] + e^\delta \mathbb{E}[|e^{-a_{i+1}}\tilde{\mathcal{P}}_1(a_{i+1}) - W_\infty|] \\ & \quad + 2\delta \mathbb{E}[W_\infty \cdot 1_{\{a_i \leq T_n < a_{i+1}\}}] \end{aligned}$$

where the last term follows from choosing $\delta \in (0, 1)$ to be sufficiently small so that $\max\{e^\delta - 1, 1 - e^{-\delta}\} \leq 2\delta$. Since $n \geq e^{t_0+M}$, for all $i \geq 0$ we have $a_i \geq t_0$. Applying (7.1) to $t = a_i$ and $t = a_{i+1}$ yields

$$\begin{aligned} \mathbb{E} \left[|e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty| \cdot 1_{\{a_i \leq T_n < a_{i+1}\}} \right] & \leq (e^\delta + e^{-\delta})\varepsilon^3 + 2\delta \mathbb{E}[W_\infty \cdot 1_{\{a_i \leq T_n < a_{i+1}\}}] \\ & \leq 4\varepsilon^3 + 2\delta \mathbb{E}[W_\infty \cdot 1_{\{a_i \leq T_n < a_{i+1}\}}]. \end{aligned}$$

Hence, the first term in (7.2) satisfies

$$\begin{aligned} \sum_{i=0}^{\lceil 2M/\delta \rceil} \mathbb{E} \left[|e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty| \cdot 1_{\{a_i \leq T_n < a_{i+1}\}} \right] & \leq \sum_{i=0}^{\lceil 2M/\delta \rceil} (4\varepsilon^3 + 2\delta \mathbb{E}[W_\infty \cdot 1_{\{a_i \leq T_n < a_{i+1}\}}]) \\ & \leq (\lceil 2M/\delta \rceil + 1)4\varepsilon^3 + 2\delta \mathbb{E}[W_\infty]. \end{aligned}$$

Take $\delta = \varepsilon/8$. It follows from our choice of $M = 2 \log(2C/\varepsilon)$ that there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ is sufficiently small we have $(\lceil 2M/\delta \rceil + 1)4\varepsilon^3 \leq \varepsilon/4$. Since $\mathbb{E}[W_\infty] \leq 1$,

$$\sum_{i=0}^{\lceil 2M/\delta \rceil} \mathbb{E} \left[|e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty| \cdot 1_{\{a_i \leq T_n < a_{i+1}\}} \right] \leq \varepsilon/4 + 2(\varepsilon/8) = \varepsilon/2. \quad (7.4)$$

Collecting (7.4) and (7.3) in (7.2) gives that for any $\varepsilon \in (0, \varepsilon_0)$, there exist some $t_0, M > 0$ depending on ε so that for all $n \geq e^{t_0+M}$,

$$\mathbb{E}|e^{-T_n}\tilde{\mathcal{P}}_1(T_n) - W_\infty| \leq \varepsilon.$$

Thus we have established the \mathbb{L}^1 -convergence of $e^{-T_n}\tilde{\mathcal{P}}_1(T_n)$ to W_∞ . By Lemma 4.3, $e^{T_n}/(n+1) \xrightarrow{a.s.} 1/W$ where $W \sim \text{Exp}(1)$. Hence,

$$\tilde{\mathcal{P}}_1(T_n)/(n+1) \xrightarrow{\mathbb{P}} W_\infty/W = 1 - q_*. \quad (7.5)$$

The result now follows upon noting that $\{\frac{d_n(v_0)}{n+1} : n \geq 0\} \stackrel{d}{=} \{\frac{\tilde{\mathcal{P}}_1(T_n)}{n+1} : n \geq 0\}$. \square

Proof of Theorem 3.9(b) Let $\delta > 0$ be given. Note that $\{d_n(v_0) : n \geq 1\} \stackrel{d}{=} \{\tilde{\mathcal{P}}_1(T_n) : n \geq 1\}$. It is easy to see $\tilde{\mathcal{P}}_1(\cdot)$ dominates $\mathcal{P}_1(\cdot)$ and hence it follows from the proof of Theorem 3.10 that

$$\frac{d_n(v_0)}{n^{1/R-\delta}} \xrightarrow{a.s.} \infty.$$

In the fringe regime $\mathbb{E}[Z] \leq 1$ we have $s_0 \geq 1$. Using Corollary 6.3(i) with $s = s_0 \geq 1$ to replace the estimate from Theorem 6.2(i) in the proof of (6.7) in Corollary 6.4 yields

$$\lim_{t \rightarrow \infty} t^{-(1+\delta)} e^{-t/R} \tilde{\mathcal{P}}_1(t) = 0 \quad a.s. \quad (7.6)$$

The same argument as in the proof Theorem 3.10 then gives

$$\frac{d_n(v_0)}{n^{1/R} (\log n)^{1+\delta}} \xrightarrow{a.s.} 0.$$

□

8 Proofs: Height

Recall the processes and notation of Lemma 4.6 as well as κ_0, s^* from Definition 3.11. The main goal of this section is to show the following:

Proposition 8.1 (i) **Fringe regime:** When $\mathbb{E}[Z] \leq 1$,

$$\mathcal{H}_{\mathcal{T}(t)}/t \xrightarrow{a.s.} \kappa_0 \quad \text{as } t \rightarrow \infty.$$

(ii) **Non-Fringe regime:** When $\mathbb{E}[Z] > 1$, as $t \rightarrow \infty$,

$$\frac{\mathcal{H}_{\mathcal{T}(t)}}{t} \xrightarrow{a.s.} \begin{cases} \kappa_0 & \text{if } s^* \in (0, q_*], \\ \frac{1}{\log(1/q_*)} & \text{if } s^* \in (q_*, s_0], \end{cases}$$

where q_* is defined as in Definition 3.6(b).

Proof of Theorem 3.12 assuming Proposition 8.1 By the continuous time embedding, $\mathcal{H}_n = \mathcal{H}_{\mathcal{T}(T_n)}$ where T_n as before is the time for \mathcal{T} to get to size $n + 1$. By Lemma 4.3(b) $T_n/\log n \xrightarrow{a.s.} 1$. Combining this with Proposition 8.1 completes the proof. □

The rest of this section is devoted to the proof of Proposition 8.1. Recall from Lemma 3.7 that under Assumptions 3.2 and 3.5, $q_* < s_0 < 1$ when $\mathbb{E}[Z] > 1$. Recall that $\kappa(s) := \frac{f(s)}{s \log(1/s)}$, $s \in (0, 1)$.

Lemma 8.2 (i) Suppose $\mathbb{E}[Z] > 1$. For $x > 0$, the infimum of $s \mapsto x \log s + f(s)/s$ is uniquely attained at some $s_x^* \in (0, s_0]$ that satisfies $x = f(s_x^*)/s_x^* - f'(s_x^*)$. In particular, $s^* = s_{\kappa_0}^* \in (0, s_0]$.

(i) *The infimum of $s \mapsto \kappa(s)$ is uniquely attained at s^* .*

Proof To prove part (i), for $x > 0$, write $h_x(s) = x \log s + f(s)/s$, $s \in (0, 1)$. As both $\log s$ and $f(s)/s$ are increasing on $(s_0, 1)$, we see that

$$\inf_{s \in (0, 1)} h_x(s) = \inf_{s \in (0, s_0]} h_x(s).$$

Differentiating $h_x(s)$ leads to $h'_x(s) = \frac{x - (f(s)/s - f'(s))}{s}$, which implies that the infimum is attained at some $s_x^* \in (0, s_0]$ satisfying the equation $x = f(s_x^*)/s_x^* - f'(s_x^*)$. As $f(s)/s$ and $-f'(s)$ are both strictly decreasing on $(0, s_0]$, the solution s_x^* to this equation is unique.

For part (ii), it is easy to see that $\kappa(s)$ is increasing on $(s_0, 1)$ so we can restrict our attention to $(0, s_0]$. Differentiating $\kappa(s)$ leads to

$$\kappa'(s) = \frac{\log(1/s)(sf'(s) - f(s)) + f(s)}{(s \log(1/s))^2} = \frac{sf'(s) - f(s) + \frac{f(s)}{\log(1/s)}}{s^2 \log(1/s)}.$$

It is straightforward to check that both $sf'(s) - f(s)$ and $\frac{f(s)}{\log(1/s)}$ are strictly increasing on $(0, s_0]$. Hence, the infimum of $\kappa(\cdot)$ is uniquely attained at s such that $\kappa'(s) = 0$, i.e.,

$$f(s)/s - f'(s) = \frac{f(s)}{s \log(1/s)}.$$

By part (i) and (ii), we have $f(s^*)/s^* - f'(s^*) = \kappa_0 = \frac{f(s^*)}{s^* \log(1/s^*)}$, i.e., s^* is the unique point where $\kappa(\cdot)$ is minimized. \square

Proof of Proposition 8.1 Fringe regime. Suppose $\mathbb{E}[Z] \leq 1$. Lemma 4.6 and Proposition 4.8 give $\liminf_{t \rightarrow \infty} \mathcal{H}_{\mathcal{T}(t)}/t \geq \kappa_0$ almost surely. It remains to prove a corresponding upper bound.

Let $s \in (0, 1)$ and recall that we defined $\mathcal{P}_s^*(t) = \sum_{i=1}^{\infty} s^{-i} \mathcal{P}_i(t)$. Using (6.2) in Theorem 6.2 gives for any $s \in (0, 1)$, non-negative x, t and $t' \in (0, t)$,

$$\begin{aligned} \mathbb{P}(\mathcal{H}_{\mathcal{T}^*(t-t')} \geq xt) &= \mathbb{P}(s^{-\mathcal{H}_{\mathcal{T}^*(t-t')}} \geq s^{-xt}) \leq \mathbb{P}(\mathcal{P}_s^*(t-t') \geq s^{-xt}) \\ &\leq s^{xt} \mathbb{E}[\mathcal{P}_s^*(t-t')] \leq \frac{p_0}{f(s)} \exp\left(t(x \log s + f(s)/s) - t' \frac{f(s)}{s}\right). \end{aligned} \quad (8.1)$$

By definition, κ_0 is the infimum of x such that $\inf_{s \in (0, 1)} \{x \log s + f(s)/s\} < 0$. It follows from Lemma 8.2(i) that for any $x > \kappa_0$, there exists $s_x^* \in (0, s_0]$ such that

$$-\delta := \inf_{s \in (0, 1)} \{x \log s + f(s)/s\} = x \log s_x^* + f(s_x^*)/s_x^* < 0. \quad (8.2)$$

Note that in $\mathcal{T}(\cdot)$, all subtrees rooted at level one evolve as $\mathcal{T}^*(\cdot)$. In order for $\mathcal{T}(t)$ to have height larger than $\lceil xt \rceil$, one of the subtrees rooted at level one need to achieve height at least xt .

With R as in Definition 3.6, we partition the time interval $[0, t]$ into segments $\{[t_i, t_{i+1}] : i \geq 0\}$ where $t_0 = 0$ and $t_{i+1} = (t_i + mR) \wedge t$ for some positive constant m to be chosen later. Let $\mathcal{S}_i = \{v_{ij} : j \geq 0\}$ denote the set of vertices at level one that arrived during $(t_i, t_{i+1}]$. Fix any $\varepsilon > 0$. For $i \geq 0$, define the event

$$E_i(t) = \{\text{one of the subtrees } \mathcal{T}_{v_{ij}} \text{ rooted at } v_{ij} \in \mathcal{S}_i \text{ has height at least } (\kappa_0 + \varepsilon)t \text{ at time } t\}. \quad (8.3)$$

In the fringe regime, using Corollary 6.3(i) with $s = s_0 \geq 1$ gives

$$\mathbb{E}(|\mathcal{S}_i|) \leq \mathbb{E}(\tilde{\mathcal{P}}_1(t_{i+1})) \leq p_0 R e^{t_{i+1}/R}.$$

Recall that by Lemma 3.7(a), $f(s)/s \geq 1/R$ for $s \in (0, 1)$. Using (8.1) with $t' = t_i$, $x = \kappa_0 + \varepsilon$ and $s = s_x^*$ as in (8.2) gives, writing $\delta := \delta_x$,

$$\begin{aligned} \mathbb{P}(E_i(t)) &\leq \mathbb{E}(|\mathcal{S}_i|) \cdot \mathbb{P}(\mathcal{H}_{\mathcal{T}^*(t-t_i)} \geq xt) \\ &\leq \frac{p_0^2 R}{f(s_x^*)} \exp\left(\frac{t_{i+1}}{R} + t(x \log s_x^* + f(s_x^*)/s_x^*) - \frac{f(s_x^*)}{s_x^*} t_i\right) \\ &\leq \frac{p_0^2 R}{f(s_x^*)} \exp(-\delta t + \frac{t_{i+1} - t_i}{R}) \leq p_0 R e^{-\delta t + m}. \end{aligned} \quad (8.4)$$

Let $i_0 = \min\{i \geq 0 : t_{i+1} \geq t\}$. Note that $i_0 \leq t/(mR)$. It follows that

$$\mathbb{P}(\mathcal{H}_{\mathcal{T}(t)} > \lceil xt \rceil) \leq \mathbb{P}(\bigcup_{i=0}^{i_0} E_i(t)) \leq \left(\frac{t}{mR} + 1\right) p_0 R e^{-\delta t + m}.$$

Choosing $m = (\delta t)/2$ we have

$$\mathbb{P}(\mathcal{H}_{\mathcal{T}(t)} > \lceil xt \rceil) \leq (2/(\delta R) + 1) p_0 R e^{-\delta t/2}. \quad (8.5)$$

Given any $\varepsilon > 0$ and $N \geq 0$ we can define the event

$$E_N = \{\exists t \in [N, N+1] : \mathcal{H}_{\mathcal{T}(t)} > \lceil (x + \varepsilon)t \rceil\}.$$

Let $d_N = (x + \varepsilon) \frac{N}{N+1}$. When N is sufficiently large, $d_N > x$. Then (8.5) gives

$$\begin{aligned} \mathbb{P}(E_N) &\leq \mathbb{P}(\mathcal{H}_{\mathcal{T}(N+1)} > \lceil (x + \varepsilon)N \rceil) = \mathbb{P}(\mathcal{H}_{\mathcal{T}(N+1)} > \lceil d_N(N+1) \rceil) \\ &\leq (2/(\delta R) + 1) p_0 R e^{-\delta(N+1)/2}. \end{aligned}$$

Applying Borel-Cantelli Lemma then gives $\mathbb{P}(\limsup_{N \rightarrow \infty} E_N) = 0$. As $x = \kappa_0 + \varepsilon$ and $\varepsilon > 0$ is arbitrary,

$$\limsup_{t \rightarrow \infty} \frac{\mathcal{H}_{\mathcal{T}(t)}}{t} \leq \kappa_0 \quad \text{almost surely.}$$

Combining this with the lower bound completes the proof for the case $\mathbb{E}[Z] \leq 1$.

Non-fringe regime with $s^* \in (0, q_*]$. Recall from Lemma 8.2 that $\kappa_0 = f(s^*)/s^* - f'(s^*)$. Using Lemma 3.7(a), $f(s)/s - f'(s)$ is a strictly decreasing function on $(0, s_0]$ and $f(s)/s - f'(s) \uparrow \infty$ as $s \downarrow 0$, for any $x > \kappa_0$, there exists some $s_x^* < s^*$ such that

$$x = f(s_x^*)/s_x^* - f'(s_x^*).$$

As before, fix any $\varepsilon > 0$ and set $x = \kappa_0 + \varepsilon$. It follows from Lemma 8.2(i) that the infimum of $s \mapsto x \log s + f(s)/s$ on $(0, 1)$ is uniquely achieved at s_x^* and by the definition of κ_0 we see that

$$-\delta := \inf_{s \in (0, 1)} \{x \log s + f(s)/s\} = x \log(s_x^*) + f(s_x^*)/s_x^* < 0.$$

Define $\{E_i(t)\}_{i \geq 0}$ as in (8.3). Since $s_x^* < s^* \leq q_*$ we have $f(s_x^*)/s_x^* \geq f(q_*)/q_* = 1$. By a similar calculation as (8.4) with $\mathbb{E}(|\mathcal{S}_i|) \leq \mathbb{E}(|\mathcal{T}(t_{i+1})|) = \exp(t_{i+1})$ and $s = s_x^*$, $t_{i+1} = (t_i + m) \wedge t$, we can obtain

$$\begin{aligned} \mathbb{P}(E_i(t)) &\leq \mathbb{E}(|\mathcal{S}_i|) \cdot \mathbb{P}(\mathcal{H}_{\mathcal{T}^*(t-t_i)} \geq (\kappa_0 + \varepsilon)t) \\ &\leq \frac{p_0}{f(s_x^*)} \exp\left(t_{i+1} + t((\kappa_0 + \varepsilon) \log s_x^* + f(s_x^*)/s_x^*) - \frac{f(s_x^*)}{s_x^*} t_i\right) \\ &\leq \frac{p_0}{f(s_x^*)} \exp(-\delta t + (t_{i+1} - t_i)) \leq e^{-\delta t + m}. \end{aligned}$$

The rest of the proof follows from the same arguments as the case before by choosing $m = \delta t/2$.

Non-fringe regime with $s^* \in (q_*, s_0]$. We begin by proving the upper bound. Fix $\varepsilon > 0$. Let $\{E_i(t)\}_{i \geq 0}$ be events defined as in (8.3) with $\kappa_0 + \varepsilon$ replaced by $(\log(1/q_*))^{-1}(1 + \varepsilon)$. Using the fact that $\mathbb{E}(|\mathcal{S}_i|) \leq \exp(t_{i+1})$, by a similar calculation as (8.4) with $x = (\log(1/q_*))^{-1}(1 + \varepsilon)$, $s = q_*$, $t_{i+1} = (t_i + m) \wedge t$, we have

$$\begin{aligned} \mathbb{P}(E_i(t)) &\leq \mathbb{E}(|\mathcal{S}_i|) \cdot \mathbb{P}(\mathcal{H}_{\mathcal{T}^*(t-t_i)} \geq (\log(1/q_*))^{-1}(1 + \varepsilon)t) \\ &\leq \frac{p_0}{f(q_*)} \exp\left(t_{i+1} + t((\log(1/q_*))^{-1}(1 + \varepsilon) \log q_* + f(q_*)/q_*) - \frac{f(q_*)}{q_*} t_i\right) \\ &= \frac{p_0}{f(q_*)} \exp(-\varepsilon t + (t_{i+1} - t_i)) \leq e^{-\varepsilon t + m}. \end{aligned}$$

Choosing $m = \varepsilon t/2$ and repeating the arguments in previous cases proves the upper bound, i.e.,

$$\limsup_{t \rightarrow \infty} \frac{\mathcal{H}_{\mathcal{T}(t)}}{t} \leq (\log(1/q_*))^{-1} \quad \text{almost surely.}$$

It remains to prove the matching lower bound. Let $\delta \in (0, 1)$ be a constant that we will choose later. Conditional on the tree $\mathcal{T}((1 - \delta)t)$ at time $(1 - \delta)t$, observe that the height of $\mathcal{T}(t)$ is stochastically lower bounded by the maximum of $|\mathcal{T}((1 - \delta)t)|$

many i.i.d. random variables, each distributed as the location of the rightmost particle of a branching random walk $\text{BRW}(\delta t)$, which has the distribution $B(\delta t)$.

Fix any $\varepsilon > 0$. Let $x = (1 - \varepsilon)(\log(1/q_*))^{-1}$. Using the above observation,

$$\mathbb{P}(\mathcal{H}_{\mathcal{T}(t)} < xt) \leq \mathbb{P}(B(\delta t) < xt)^{e^{(1-\delta)t-t^{1/2}}} + \mathbb{P}(|\mathcal{T}((1-\delta)t)| \leq e^{(1-\delta)t-t^{1/2}}). \quad (8.6)$$

Note that $|\mathcal{T}(\cdot)|$ is a rate 1 Yule process and hence by Lemma 4.2 we have for large enough t ,

$$\begin{aligned} \mathbb{P}(|\mathcal{T}((1-\delta)t)| \leq e^{(1-\delta)t-t^{1/2}}) &= \sum_{k=1}^{\lfloor e^{(1-\delta)t-t^{1/2}} \rfloor} e^{-(1-\delta)t} (1 - e^{-(1-\delta)t})^{k-1} \\ &\leq 1 - (1 - e^{-(1-\delta)t})^{e^{(1-\delta)t-t^{1/2}}} \leq 1 - \exp(-2e^{-t^{1/2}}) \\ &\leq 2e^{-t^{1/2}} \end{aligned} \quad (8.7)$$

by the elementary inequality that $1 - 2x \leq e^{-2x} \leq 1 - x$ for $x \in [0, 1/2]$.

It remains to estimate $\mathbb{P}(B(\delta t) < xt)$. For reasons that will become clear soon, we will take $\delta = \frac{1-\varepsilon}{(1-f'(q_*))\log(1/q_*)}$ so that $x/\delta = 1 - f'(q_*)$. To verify that $\delta \in (0, 1)$ it suffices to show $(1 - f'(q_*))\log(1/q_*) \geq 1$. In this case, it follows from Lemma 8.2(ii) that $s^* \in (q_*, s_0]$ is the unique minimizer of $\kappa(s) = \frac{f(s)}{s\log(1/s)}$, which implies that $\kappa'(q_*) \leq 0$. This, in turn, leads to $(1 - f'(q_*))\log(1/q_*) \geq 1$. Moreover,

$$\kappa_0 = \kappa(s^*) < \kappa(q_*) = \frac{1}{\log(1/q_*)}.$$

That is, our choice of x and δ gives $x/\delta = 1 - f'(q_*) \geq (\log(1/q_*))^{-1} > \kappa_0$. Then we can apply Lemma 4.9 to obtain

$$\mathbb{P}(B(\delta t) \geq xt) = \exp(\delta t \inf_{s \in (0, 1)} \{(1 - f'(q_*))\log s + f(s)/s\} + o(t)).$$

Lemma 8.2(i) then shows that the infimum of $s \mapsto (1 - f'(q_*))\log s + f(s)/s$ on $(0, 1)$ is uniquely achieved at $s = q_*$, which leads to

$$\mathbb{P}(B(\delta t) \geq xt) = \exp(\delta t ((1 - f'(q_*))\log q_* + 1) + o(t)) = \exp(-(1 - \varepsilon - \delta)t + o(t)) \quad (8.8)$$

$$\text{as } \delta = \frac{1-\varepsilon}{(1-f'(q_*))\log(1/q_*)}.$$

Combining (8.7) and (8.8) in (8.6) yields for large enough t ,

$$\begin{aligned} \mathbb{P}(\mathcal{H}_{\mathcal{T}(t)} < xt) &\leq (1 - \mathbb{P}(B(\delta t) \geq xt))^{e^{(1-\delta)t-t^{1/2}}} + 2e^{-t^{1/2}} \\ &\leq \exp(-\mathbb{P}(B(\delta t) \geq xt)e^{(1-\delta)t-t^{1/2}}) + 2e^{-t^{1/2}} \\ &= \exp(-\exp(-(1 - \varepsilon - \delta)t + (1 - \delta)t + o(t))) + 2e^{-t^{1/2}} \end{aligned}$$

$$\leq \exp(-\exp(\varepsilon t/2)) + 2e^{-t^{1/2}}.$$

The same argument as in the previous cases with the Borel-Cantelli Lemma would lead to

$$\liminf_{t \rightarrow \infty} \frac{\mathcal{H}_{\mathcal{T}(t)}}{t} \geq (\log(1/q_*))^{-1} \quad \text{almost surely,}$$

which concludes our proof. \square

9 Proofs: PageRank asymptotics

In this section, we prove Theorems 3.14 and 3.16. As before, Assumptions 3.2 and 3.5 continue to hold and are not explicitly stated in the results.

As in the case of degrees, our analysis will rely on continuous time versions of PageRank. Consider the PageRank of the root in $\mathcal{T}^*(t)$ with damping factor c , namely,

$$R_c^*(t) = (1 - c) \left(1 + \sum_{l=1}^{\infty} c^l \mathcal{P}_l(t) \right). \quad (9.1)$$

and the Pagerank of the root in $\mathcal{T}(t)$ given by

$$R_c(t) = (1 - c) \left(1 + \sum_{l=1}^{\infty} c^l \tilde{\mathcal{P}}_l(t) \right). \quad (9.2)$$

We begin with a lower bound on $\mathbb{E}[\mathcal{P}_s^*(t)]$ given in Lemma 9.1, which will play a key role in showing that the limiting random variable $W_{k,c}$ in Theorem 3.14 is non-degenerate. This lower bound involves a ‘change of measure’ argument which we now present.

For any $s > 0$, define a probability transition kernel

$$p_s^*(x, y) = \frac{s^{-y} p_{x+1-y}}{\alpha(s) s^{-x}}, \quad x, y \in \mathbb{Z} \text{ and } -\infty < y \leq x + 1,$$

where $\alpha(s) = \frac{f(s)}{s}$. We can check that p_s^* is indeed a probability transition kernel by

$$\sum_{y \leq x+1} \frac{s^{-y} p_{x+1-y}}{\alpha(s) s^{-x}} = \frac{s^{-(x+1)} f(s)}{\alpha(s) s^{-x}} = 1.$$

Let S_s^* be a discrete time random walk following the transition kernel p_s^* . Set $S_s^*(0) = 0$ and define

$$\tau_s^* = \inf\{n \geq 1 : S_s^*(n) \leq 0\}.$$

Lemma 9.1 For any $s > 0$ and $t \geq 0$,

$$\mathbb{E}[\mathcal{P}_s^*(t)] \geq \mathbb{P}(\tau_s^* = \infty) e^{\frac{f(s)}{s}t}.$$

In particular, when $\mathbb{E}[Z] < 1$ and $s_0 > 1$, for any $s \in [1, s_0)$ and $t \geq 0$,

$$e^{-\frac{f(s)}{s}t} \mathbb{E}[\mathcal{P}_s^*(t)] \geq \mathbb{P}(\tau_s^* = \infty) > 0.$$

Proof Recall the random walk S defined in (3.1), and recall from Lemma 4.11 that for all $k \geq 1$,

$$\mathbb{E}[\mathcal{P}_k(t)] = \sum_{i=0}^{\infty} \frac{t^i}{i!} \mathbb{P}(\tilde{T}_k = i), \quad (9.3)$$

where $\tilde{T}_k := \inf\{n \geq 0 : S_n = 0 | S_0 = k\}$. Consider the random walk \tilde{S} given by $\tilde{S}_0 = 0$ and

$$\tilde{S}_n := \sum_{j=1}^n (1 - Z_j), \quad n \geq 1.$$

Also, define $\tilde{\tau} := \inf\{j \geq 1 : \tilde{S}_j \leq 0\}$. By a time-reversal argument, it readily follows that for any $i, k \geq 1$,

$$\mathbb{P}(\tilde{T}_k = i) = \mathbb{P}(\tilde{S}_i = k, \tilde{\tau} > i).$$

Using this and (9.3), we obtain

$$\mathbb{E}[\mathcal{P}_s^*(t)] = \sum_{i=0}^{\infty} \frac{t^i}{i!} \sum_{k=1}^{\infty} s^{-k} \mathbb{P}(\tilde{S}_i = k, \tilde{\tau} > i). \quad (9.4)$$

The crucial elementary algebraic identity connecting \tilde{S} to the random walk S_s^* defined before the lemma is the following:

$$s^{-k} \mathbb{P}(\tilde{S}_i = k, \tilde{\tau} > i) = (\alpha(s))^i \mathbb{P}(S_s^*(i) = k, \tau_s^* > i), \quad i \geq 0, k \geq 1, \quad (9.5)$$

where $\alpha(s) = \frac{f(s)}{s}$. Using this observation in (9.4), we obtain

$$\begin{aligned} \mathbb{E}[\mathcal{P}_s^*(t)] &= \sum_{i=0}^{\infty} \frac{t^i}{i!} (\alpha(s))^i \sum_{k=1}^{\infty} \mathbb{P}(S_s^*(i) = k, \tau_s^* > i) \\ &= \sum_{i=0}^{\infty} \frac{(\alpha(s)t)^i}{i!} \mathbb{P}(\tau_s^* > i) \geq e^{\alpha(s)t} \mathbb{P}(\tau_s^* = \infty). \end{aligned}$$

This proves the first assertion in the lemma. To prove the second assertion, namely $\mathbb{P}(\tau_s^* = \infty) > 0$ for $s \in [1, s_0)$ in the case $\mathbb{E}[Z] < 1$, it suffices to show $\mathbb{E}[S_s^*(1)] > 0$ (this implies the result, e.g., by Lemma 11.3 of [29]). We compute the probability generating function $f^*(\cdot)$ (suppressing dependence on s) of $S_s^*(1)$:

$$f^*(\psi) := \sum_{j=-\infty}^1 \psi^j p_s^*(0, j) = \sum_{j=-\infty}^1 \psi^j \frac{s^{-j} p_{1-j}}{\alpha(s)} = \frac{\psi f(\frac{s}{\psi})}{f(s)}, \quad \psi \geq 0.$$

Then,

$$\mathbb{E}[S_s^*(1)] = (f^*)'(1) = \frac{1}{f(s)} \left[f\left(\frac{s}{\psi}\right) - \frac{s}{\psi} f'\left(\frac{s}{\psi}\right) \right] \Big|_{\psi=1} = \frac{f(s) - sf'(s)}{f(s)}.$$

Let $g(s) = f(s) - sf'(s)$. Since $g(1) = 1 - \mathbb{E}[Z] > 0$, by definition of s_0 , we have $g(s) > 0$ for any $s \in [1, s_0)$ and hence $\mathbb{E}[S_s^*(1)] > 0$. This completes the proof. \square

The following theorem proves the analogous assertions of Theorem 3.14 for the continuous time versions of the PageRank defined in (9.1) and (9.2). Recall $T_n = \inf\{t \geq 0 : |\mathcal{T}(t)| = n + 1\}$

Theorem 9.2 *Fix $c \in (0, 1)$.*

(a) **Non-fringe regime:** *When $\mathbb{E}[Z] > 1$,*

(i) *For any $\delta > 0$, there exists $\varepsilon > 0$ such that*

$$\liminf_{n \rightarrow \infty} \mathbb{P}\left(\frac{R_c(T_n)}{n} \geq c(1-c)(1-q_*) - \varepsilon\right) \geq 1 - \delta. \quad (9.6)$$

(ii) *For any $\delta > 0$,*

$$\lim_{t \rightarrow \infty} e^{-\left(\frac{1}{R} - \delta\right)t} R_c^*(t) = \infty \quad \text{and} \quad \lim_{t \rightarrow \infty} t^{-(1+\delta)} e^{-t/R} R_c^*(t) = 0 \quad \text{a.s.} \quad (9.7)$$

(b) **Fringe regime:** *When $\mathbb{E}[Z] \leq 1$,*

(i) *Fix any $c \in (0, s_0^{-1}]$ with $c < 1$. Then for any $\delta > 0$,*

$$\lim_{t \rightarrow \infty} e^{-\left(\frac{1}{R} - \delta\right)t} R_c(t) = \infty \quad \text{and} \quad \lim_{t \rightarrow \infty} t^{-(1+\delta)} e^{-t/R} R_c(t) = 0 \quad \text{a.s.} \quad (9.8)$$

$$\lim_{t \rightarrow \infty} e^{-\left(\frac{1}{R} - \delta\right)t} R_c^*(t) = \infty \quad \text{and} \quad \lim_{t \rightarrow \infty} t^{-(1+\delta)} e^{-t/R} R_c^*(t) = 0 \quad \text{a.s.} \quad (9.9)$$

(ii) *Suppose $s_0 > 1$. For any $c \in (s_0^{-1}, 1)$, there exist non-negative random variables W_c and W_c^* with $\mathbb{P}(W_c > 0) > 0$ and $\mathbb{P}(W_c^* > 0) > 0$ so that as*

$t \rightarrow \infty$,

$$e^{-cf(1/c)t} R_c(t) \xrightarrow{a.s.} W_c \quad \text{and} \quad e^{-cf(1/c)t} R_c^*(t) \xrightarrow{a.s.} W_c^*.$$

Proof We first focus on the non-fringe regime.

Denoting the size of $\mathcal{T}(t)$ by $|\mathcal{T}(t)|$, observe that

$$R_c(t) \geq (1 - c)c\tilde{\mathcal{P}}_1(t).$$

(9.6) then follows from (7.5).

For $R_c^*(t)$ with any $c \in (0, 1)$, recalling from Lemma 3.7(d) that $s_0^{-1} > 1$,

$$(1 - c)c\mathcal{P}_1(t) \leq R_c^*(t) \leq (1 - c) \left(1 + \sum_{l=1}^{\infty} s_0^{-l} \mathcal{P}_l(t) \right) = (1 - c)(1 + \mathcal{P}^*(t)). \quad (9.10)$$

Hence, the first limit in (9.7) follows from (6.6) and the first inequality in (9.10).

To show the second limit in (9.7) take any $\delta > 0$. For $\varepsilon > 0$ and $N \geq 0$, define the event

$$E_N = \left\{ \sup_{u \in [N, N+1]} u^{-(1+\delta)} e^{-u/R} R_c^*(u) > \varepsilon \right\}.$$

It follows from the upper bound in (9.10) and (6.2) that for any $c \in (0, 1)$ and $N \geq 1$,

$$\begin{aligned} \mathbb{P}(E_N) &\leq \mathbb{P}(R_c^*(N+1) > \varepsilon e^{N/R} N^{1+\delta}) \leq \frac{\mathbb{E}[R_c^*(N+1)]}{\varepsilon e^{N/R} N^{1+\delta}} \\ &\leq \frac{(1 - c)(1 + \mathbb{E}[\mathcal{P}^*(N+1)])}{\varepsilon e^{N/R} N^{1+\delta}} \leq \frac{(1 - c) \left(1 + \frac{p_0}{f(s_0)} e^{(N+1)/R} \right)}{\varepsilon e^{N/R} N^{1+\delta}} \leq \frac{C e^{1/R}}{\varepsilon N^{1+\delta}} \end{aligned}$$

for some constant $C > 0$. Applying Borel-Cantelli Lemma then gives $\mathbb{P}(\limsup_{N \rightarrow \infty} E_N) = 0$. Hence,

$$t^{-(1+\delta)} e^{-t/R} R_c^*(t) \xrightarrow{a.s.} 0,$$

proving the second limit in (9.7).

We now turn to the fringe regime. Define $\tilde{\mathcal{P}}^*(t) = \sum_{i=1}^{\infty} s_0^{-i} \tilde{\mathcal{P}}_i(t)$,

$$(1 - c)c\tilde{\mathcal{P}}_1(t) \leq R_c(t) \leq (1 - c) \left(1 + \tilde{\mathcal{P}}^*(t) \right).$$

The first limit in (9.8) comes from the observation that $\tilde{\mathcal{P}}_1(t)$ dominates $\mathcal{P}_1(t)$ and Corollary 6.4. The second limit follows from an argument that is essentially the same as that of the second limit in (9.7) except that we apply Corollary 6.3(i) for the expectation of $\tilde{\mathcal{P}}^*(t)$.

The proof of (9.9) is essentially the same as that of (9.7) upon noting that the bounds in (9.10) remain valid for $c \in (0, s_0^{-1}]$.

It remains to prove (ii) of part (b). For $t \geq 0$, define

$$M_c^*(t) = e^{-cf(1/c)t} R_c^*(t).$$

Using the generator expression in (6.11) (see also Remark 13), observe that for any $t \geq 0$,

$$\begin{aligned} \mathcal{L}(M_c^*(t)) &= -cf(1/c)e^{-cf(1/c)t} R_c^*(t) + e^{-cf(1/c)t}(1-c) \sum_{i=1}^{\infty} c^i \sum_{j=i-1}^{\infty} A_{ij} \mathcal{P}_j(t) \\ &= -cf(1/c)e^{-cf(1/c)t} R_c^*(t) + e^{-cf(1/c)t}(1-c) \sum_{j=0}^{\infty} \left(\sum_{i=1}^{j+1} c^i A_{ij} \right) \mathcal{P}_j(t) \\ &= -cf(1/c)e^{-cf(1/c)t} R_c^*(t) + e^{-cf(1/c)t}(1-c)c \\ &\quad \times \sum_{j=0}^{\infty} \left(\sum_{i=1}^{j+1} c^{i-j-1} p_{j-i+1} \right) c^j \mathcal{P}_j(t) \\ &\leq -cf(1/c)e^{-cf(1/c)t} R_c^*(t) + cf(1/c)e^{-cf(1/c)t} R_c^*(t) = 0. \end{aligned}$$

Hence, $M_c^*(t)$ is a non-negative supermartingale and thus

$$M_c^*(t) \xrightarrow{a.s.} W_c^* \quad \text{as } t \rightarrow \infty, \quad (9.11)$$

for some non-negative random variable W_c^* .

By Theorem 6.2(ii), we conclude that $\sup_{t < \infty} \mathbb{E}(M_c^*(t))^2 < \infty$. Using this observation and Lemma 9.1,

$$\mathbb{E}(W_c^*) = \lim_{t \rightarrow \infty} \mathbb{E}(M_c^*(t)) \geq \mathbb{P}(\tau_{1/c}^* = \infty) > 0.$$

This implies that $\mathbb{P}(W_c^* > 0) > 0$.

The proof for the convergence of $M_c(t) := e^{-cf(1/c)t} R_c(t)$ to some W_c follows similarly upon noting that when $\mathbb{E}[Z] \leq 1$, $(c^i : i \geq 0)$ is a sub-invariant left eigenvector of \mathbf{B} with eigenvalue $cf(1/c)$, see Proposition 4.12(b). We can show $\mathcal{L}(M_c(t)) \leq 0$, which implies that $M_c(t)$ is a non-negative supermartingale that converges to a random variable W_c . By Corollary 6.3(ii), $\sup_{t < \infty} \mathbb{E}(M_c(t)^2) < \infty$. Combined with the fact that $\tilde{\mathcal{P}}_i(t)$ dominates $\mathcal{P}_i(t)$ for all $i \geq 0$, $t \geq 0$, we have

$$\mathbb{E}(W_c) = \lim_{t \rightarrow \infty} \mathbb{E}(M_c(t)) \geq \lim_{t \rightarrow \infty} \mathbb{E}(M_c^*(t)) > 0,$$

proving $\mathbb{P}(W_c > 0) > 0$. This concludes the proof of (ii) of part (b), and the Theorem. \square

Proof of Theorem 3.14 Using the continuous time embedding, observe that for any $k \geq 0$,

$$\{R_{v_k, c}(n), n \geq v\} \stackrel{d}{=} \{R_{v_k, c}^\circ(T_n - \sigma_k), n \geq v\},$$

where σ_k is the birth time of v_k , $T_n = \inf\{t \geq 0 : |\mathcal{T}(t)| = n + 1\}$ and $R_{v_k, c}^\circ(t)$ is the PageRank of v_k in $\mathcal{T}(t + \sigma_k)$. Part (a) and part (b)(i) of the Theorem now follow from Theorem 9.2 upon noting that $\{R_{v_k, c}^\circ(t) : t \geq 0\}$ has the same distribution as $\{R_c(t) : t \geq 0\}$ for $k = 0$ and $\{R_c^*(t) : t \geq 0\}$ for $k \geq 1$.

To prove (b)(ii), note that for $k \geq 1$,

$$\begin{aligned} \frac{R_{v_k, c}(n)}{n^{cf(1/c)}} &= \frac{R_{v_k, c}^\circ(T_n - \sigma_k)}{e^{cf(1/c)(T_n - \sigma_k)}} \cdot \left(\frac{e^{T_n - \sigma_k}}{n} \right)^{cf(1/c)} \\ &\xrightarrow{a.s.} e^{-\sigma_k cf(1/c)} \tilde{W}_{k, c} W^{-cf(1/c)} =: W_{k, c}, \end{aligned}$$

where the limit $W := \lim_{n \rightarrow \infty} n e^{-T_n}$ almost surely exists by Lemma 4.3. The random variable $\tilde{W}_{k, c} \stackrel{d}{=} W_c^*$ for $k \geq 1$ and $\tilde{W}_{k, c} \stackrel{d}{=} W_c$ for $k = 0$, where W_c^* and W_c are obtained in Theorem 9.2 and The result follows. \square

Proof of Theorem 3.16 Recall the form of the limiting PageRank distribution given in (3.13), from which we conclude that $\mathcal{R}_{\emptyset, c}(\infty)$ has the same distribution as $R_c^*(\tau)$ where τ is a unit rate exponential random variable independent of the random tree process $\mathcal{T}(\cdot)$.

Part (a) and the assertion in part (b) for $c \in (0, s_0^{-1}]$ with $c < 1$ now follow along the same argument as in the proof of Theorem 3.8 upon using the bounds in (9.10). The analogue of (6.8) in part (a) uses Corollary 6.5(ii) and that in part (b) uses Corollary 6.5(i). The other direction follows from (6.9) and the lower bound in (9.10).

It only remains to prove the assertion in part (b) for $c \in (s_0^{-1}, 1)$. Observe that for any $\delta \in (0, \frac{1}{cf(1/c)})$, $r > 0$, using Markov's inequality and Corollary 6.5(i),

$$\begin{aligned} r^{\frac{1}{cf(1/c)} - \delta} \mathbb{P}(\mathcal{R}_{\emptyset, c}(\infty) \geq r) &\leq \mathbb{E} \left[(\mathcal{R}_{\emptyset, c}(\infty))^{\frac{1}{cf(1/c)} - \delta} \right] \\ &\leq \mathbb{E} \left[\left((1 - c)(1 + \mathcal{P}_{1/c}^*(\tau)) \right)^{\frac{1}{cf(1/c)} - \delta} \right] < \infty, \end{aligned}$$

which implies

$$\limsup_{r \rightarrow \infty} \frac{\log \mathbb{P}(\mathcal{R}_{\emptyset, c}(\infty) \geq r)}{\log r} \leq -\frac{1}{cf(1/c)}.$$

Moreover, by the almost sure convergence in Theorem 9.2 (b)(ii) to a non-degenerate non-negative random variable, there exist positive η_1, η_2, t_0 such that

$$\mathbb{P} \left(e^{-cf(1/c)t} R_c^*(t) \geq \eta_1 \right) \geq \eta_2, \quad t \geq t_0.$$

Now, proceeding as in Theorem 3.8, for $r \geq \eta_1 e^{cf(1/c)t_0}$,

$$\begin{aligned}\mathbb{P}(\mathcal{R}_{\emptyset,c}(\infty) \geq r) &= \int_0^\infty e^{-s} \mathbb{P}(R_c^*(s) \geq r) ds \\ &= \int_0^\infty e^{-s} \mathbb{P}\left(e^{-cf(1/c)s} R_c^*(s) \geq e^{-cf(1/c)s} r\right) ds \\ &\geq \int_{\frac{1}{cf(1/c)} \log(r/\eta_1)}^\infty e^{-s} \mathbb{P}\left(e^{-cf(1/c)s} R_c^*(s) \geq \eta_1\right) ds \\ &\geq \eta_2 \int_{\frac{1}{cf(1/c)} \log(r/\eta_1)}^\infty e^{-s} ds = \eta_2 \left(\frac{\eta_1}{r}\right)^{\frac{1}{cf(1/c)}}.\end{aligned}$$

This implies

$$\liminf_{r \rightarrow \infty} \frac{\log \mathbb{P}(\mathcal{R}_{\emptyset,c}(\infty) \geq r)}{\log r} \geq -\frac{1}{cf(1/c)},$$

completing the proof of the theorem. \square

Acknowledgements Banerjee was supported in part by the NSF CAREER award DMS-2141621. Bhamidi was supported in part by NSF DMS-2113662. Research was partially supported by NSF RTG grant DMS-2134107. The authors thank an anonymous referee and an editor whose valuable comments led to a major improvement in the paper, in particular allowing us to realize the existence of potential phase transition behavior for the height in the non-fringe regime.

Appendix A: α -recurrence and transience of μ^B and μ^A

A full account of α -recurrence/transience for Markov chains can be found in Chapter 2 of [52]. To be concise we only state what is needed for the rigorous definition of α -recurrence.

Let φ be a σ -finite measure on S and let $A \in \mathcal{S}$ be a φ -positive set (i.e., $\varphi(A) > 0$).

Definition A.1 (*irreducible measure*) The set A is called φ -communicating for a Markov kernel K on $S \times S$ if every φ -positive subset $B \subseteq A$ is attainable from A , i.e., $K^{(n)}(x, B) > 0$ for some $n \geq 0$ for all $x \in A$ and all φ -positive $B \subseteq A$.

If the whole state space S is φ -communicating, then the kernel μ is called φ -irreducible and φ is called an irreducibility measure for K .

It is clear that any measure ψ which is absolutely continuous with respect to an irreducibility measure φ is itself an irreducibility measure.

Proposition A.2 (Proposition 2.4 in [52]) *Suppose kernel K is φ -irreducible. Then there exists a maximal irreducibility measure ψ in the sense that all other irreducibility measures are absolutely continuous with respect to ψ .*

Now we present a definition of α -recurrence in the context of our problem. It follows from the discussion on page 194 of [35] (see also Proposition 2.1 of [50]) that

Definition A.3 (α -recurrence) Let φ be the maximal irreducibility measure for the kernel $\mu(\cdot, \cdot \times \mathbb{R}_+)$. The kernel $\mu(\cdot, \cdot \times \mathbb{R}_+)$ is said to be α -recurrent if

$$\sum_{n=0}^{\infty} \mu_{\alpha}^{(n)}(s, A \times \mathbb{R}_+) = \infty$$

for all $s \in S$ and $A \in \mathcal{S}$ with $\varphi(A) > 0$. Otherwise it is said to be α -transient.

For the rest of this section we will discuss the α -recurrence/transience of the continuous time branching process embedding of our model.

Lemma A.4 When $\mathbb{E}[Z] > 1$, the kernel $\mu^{\mathbf{B}}(s, r \times dt) := B_r dt$ is α -recurrent.

Proof In this case the Malthusian parameter $\alpha = 1$. Hence

$$\mu_{\alpha}^{\mathbf{B}}(s, r \times \mathbb{R}_+) = \sum_{n=0}^{\infty} \frac{(\mathbf{B}^n)_{rs}}{\alpha^n} = \sum_{n=0}^{\infty} (\mathbf{B}^n)_{rs}.$$

Let $\hat{\mathbf{B}}$ denote the matrix obtained by restricting \mathbf{B} to states $\{1, 2, \dots\}$. As $\mathbb{E}[Z] > 1$ implies $p_0 + p_1 < 1$, it can be checked that $\hat{\mathbf{B}}$ is irreducible, which follows from the fact that for any $r, s \in \mathbb{N}$, there exist $m_1, m_2 \in \mathbb{N}$ such that $(\hat{\mathbf{B}}^{m_1})_{r1} > 0$ and $(\hat{\mathbf{B}}^{m_2})_{1s} > 0$.

First, we will show $\sum_{n=0}^{\infty} (\hat{\mathbf{B}}^n)_{11} = \infty$ which, by irreducibility, will imply $\sum_{n=0}^{\infty} (\hat{\mathbf{B}}^n)_{rs} = \infty$ for any $r, s \in \mathbb{N}$. Notice that $\hat{\mathbf{B}}_{11}^n = ((\hat{\mathbf{B}}^T)^n)_{11}$ and $\hat{\mathbf{B}}^T$ is a probability transition matrix. Let $\{X_n\}_{n \geq 1}$ denote the Markov chain following the transition matrix $\hat{\mathbf{B}}^T$ started at $X_0 = 1$ and let $\tau = \inf\{n \geq 1 : X_n = 1\}$ be the first returning time to 1. It is well known that $\sum_{n=0}^{\infty} ((\hat{\mathbf{B}}^T)^n)_{11} = \infty$ if and only if X_n is recurrent. Therefore, to show the α -recurrence of $\mu^{\mathbf{B}}$ we only need to show $P(\tau < \infty | X_0 = 1) = 1$. Note that

$$\begin{aligned} \mathbb{E}[X_n - X_{n-1} | X_{n-1} = i] &= c_i + \sum_{k=0}^{i-1} (i+1-k)p_k - i \\ &= c_i + (i+1) \sum_{k=0}^{i-1} p_k - \sum_{k=0}^{i-1} kp_k - i = 1 - \left(c_i + \sum_{k=0}^{i-1} kp_k \right) \\ &= 1 - \mathbb{E}[Z] + \sum_{k=i}^{\infty} (k-i)p_k \downarrow 1 - \mathbb{E}[Z] \quad \text{as } i \rightarrow \infty. \end{aligned}$$

When $\mathbb{E}[Z] > 1$, there exists $i_0 \in \mathbb{N}$ such that for all $i \geq i_0$, $\mathbb{E}[X_n - X_{n-1} | X_{n-1} = i] < 0$. Let $\tau_{i_0} = \inf\{n \geq 0 : X_n \leq i_0\}$. We claim that for all $j > i_0$ we have $\mathbb{P}_j(\tau_{i_0} < \infty) = 1$. If we view the set $A = \{1, 2, \dots, i_0\}$ as an absorbing state and let Y_n be a Markov chain obtained by projecting X_n onto $\{A, i_0 + 1, \dots\}$ with $Y_0 = j$, then Y_n is a non-negative supermartingale and has to converge almost surely. Note that Y_n can only converge by getting absorbed at A , which implies $\mathbb{P}_j(\tau_{i_0} < \infty) = 1$.

This means that X_n returns to the set A infinitely often almost surely and thus, by irreducibility of $\hat{\mathbf{B}}^T$, will return to 1 in finite time almost surely.

For $r \in \mathbb{N}$ and $s = 0$, note that $(\mathbf{B}^n)_{r0} \geq (\mathbf{B}^{m_1})_{r1}(\mathbf{B}^{n-(m_1+1)})_{11}\mathbf{B}_{10}$ for $n \geq m_1 + 1$ and the above argument implies $\sum_{n=0}^{\infty} (\mathbf{B}^n)_{r0} = \infty$. It follows from Definition A.1 that the maximal irreducibility measure φ for $\mu^{\mathbf{B}}$ must satisfy $\varphi(\{0\}) = 0$. Hence we don't need to consider the case where $r = 0$. \square

Lemma A.5 *The kernel $\mu^A(s, r \times dt) := A_{rs}dt$ is α -transient.*

Proof Recall the function $\chi(u) = \sum_{n=0}^{\infty} u^n \mathbb{P}_1(T = n)$ as defined in the proof of Lemma 4.17.

It is easily checked that for $s, r \in \{0, 1, \dots\}$,

$$\sum_{n=0}^{\infty} (\mu_{\alpha}^A)^{(n)}(s, r \times \mathbb{R}_+) = \sum_{n=0}^{\infty} \frac{(\mathbf{A}^n)_{rs}}{\alpha^n}.$$

Take $(s, r) = (0, 1)$ and note that $\varphi(1) > 0$. Since $(\mathbf{A}^n)_{10} = \mathbb{P}_1(T = n)$ we can observe that

$$\sum_{n=0}^{\infty} (\mu_{\alpha}^A)^{(n)}(0, 1 \times \mathbb{R}_+) = \sum_{n=0}^{\infty} \frac{\mathbb{P}_1(T = n)}{\alpha^n} = \chi(1/\alpha).$$

Note that the Perron root of \mathbf{A} is $1/R$ and hence the corresponding Malthusian rate is $\alpha = 1/R$. It follows from the proof of Lemma 4.17 (see (4.10)) that $\chi(1/\alpha) = \chi(R) < \infty$, which immediately implies the α -transience for μ^A . \square

References

- Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**(1), 47 (2002)
- Aldous, D.: Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.* **1**, 228–266 (1991)
- Aldous, D., Steele, J.M.: The objective method: probabilistic combinatorial optimization and local weak convergence. In: Kesten, H. (ed.) *Probability on Discrete Structures*, pp. 1–72. Springer, Berlin (2004)
- Andersen, R., Chung, F., Lang, K.: Local graph partitioning using PageRank vectors. In: *Proceedings of FOCS 2006*, pp. 475–486 (2006)
- Aoki, T., Rocha, L.E., Gross, T.: Temporal and structural heterogeneities emerging in adaptive temporal networks. *Phys. Rev. E* **93**(4), 040301 (2016)
- Athreya, K.B., Karlin, S.: Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Stat.* **39**(6), 1801–1817 (1968)
- Banerjee, S., Olvera-Cravioto, M.: PageRank asymptotics on directed preferential attachment networks. *Ann. Appl. Probab.* **32**(4), 3060–3084 (2022)
- Barabási, A.L., Albert, R.: Emergence of scaling in random networks. *Science* **286**(5439), 509–512 (1999)
- Ben-Naim, E., Krapivsky, P.: Random ancestor trees. *J. Stat. Mech. Theory Exp.* **06**, P06004 (2010)
- Benjamini, I., Schramm, O.: Recurrence of distributional limits of finite planar graphs. *Selected Works of Oded Schramm*, pp. 533–545 (2011)
- Bhamidi, S., Evans, S.N., Sen, A.: Spectra of large random trees. *J. Theor. Probab.* **25**(3), 613–654 (2012)

12. Bhamidi, S., Steele, J.M., Zaman, T.: Twitter event networks and the superstar model. *Ann. Appl. Probab.* **25**(5), 2462–2502 (2015)
13. Biggins, J.: The growth and spread of the general branching random walk. *Ann. Appl. Probab.* **5**, 1008–1024 (1995)
14. Biggins, J.: How fast does a general branching random walk spread? In: Athreya, K.B., Jagers, P. (eds.) *Classical and Modern Branching Processes*, pp. 19–39. Springer, Berlin (1997)
15. Blum, A., Chan, T.H., Rwebangira, M.R.: A random-surfer web-graph model. In: *Proceedings of the Third Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, vol. 2006, pp. 238–246 (2006)
16. Bollobás, B., Riordan, O.M.: Mathematical results on scale-free random graphs. In: Bornholdt, S., Schuster, H.G. (eds.) *Handbook of Graphs and Networks*, pp. 1–34. Wiley-VCH, Weinheim (2003)
17. Bollobás, B.: *Random Graphs*. Cambridge University Press, Cambridge (2001)
18. Bollobás, B., Riordan, O.: Robustness and vulnerability of scale-free random graphs. *Internet Math.* **1**(1), 1–35 (2004)
19. Bollobás, B., Riordan, O., Spencer, J., Tusnády, G.: The degree sequence of a scale-free random graph process. *Random Struct. Algorithms* **18**(3), 279–290 (2001)
20. Chebolu, P., Melsted, P.: PageRank and the random surfer model. In: *SODA*, pp. 1010–1018 (2008)
21. Chen, N., Litvak, N., Olvera-Cravioto, M.: Generalized PageRank on directed configuration networks. *Random Struct. Algorithms* **51**(2), 237–274 (2017)
22. Daley, D.: Quasi-stationary behaviour of a left-continuous random walk. *Ann. Math. Stat.* **40**(2), 532–539 (1969)
23. Durrett, R.: *Random Graph Dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge (2007)
24. Gabel, A., Krapivsky, P., Redner, S.: Highly dispersed networks by enhanced redirection. *Phys. Rev. E* **88**(5), 050802 (2013)
25. Gabel, A., Redner, S.: Sublinear but never superlinear preferential attachment by local network growth. *J. Stat. Mech. Theory Exp.* **02**, P02043 (2013)
26. Garavaglia, A., van der Hofstad, R.: From trees to graphs: collapsing continuous-time branching processes. *J. Appl. Probab.* **55**(3), 900–919 (2018)
27. Garavaglia, A., van der Hofstad, R., Litvak, N.: Local weak convergence for PageRank. *Ann. Appl. Probab.* **30**(1), 40–79 (2020)
28. Gross, T., Blasius, B.: Adaptive coevolutionary networks: a review. *J. R. Soc. Interface* **5**(20), 259–271 (2008)
29. Gut, A.: *Stopped Random Walks*. Springer, Berlin (2009)
30. Gyöngyi, Z., Garcia-Molina, H., Pedersen, J.: Combating Web spam with TrustRank. In: *Proceeding of VLDB2004*, pp. 576–587 (2004)
31. Harris, S.C., Roberts, M.I.: The many-to-few lemma and multiple spines. *Annales de l'institut henri poincaré, probabilités et statistiques*, pp. 226–242 (2017)
32. Haveliwala, T.H.: Topic-sensitive PageRank. In: *Proceedings of the 11th International Conference on World Wide Web*, pp. 517–526 (2002)
33. Holme, P., Saramäki, J.: Temporal networks. *Phys. Rep.* **519**(3), 97–125 (2012)
34. Holme, P., Saramäki, J.: *Temporal Network Theory*, vol. 2. Springer, Berlin (2019)
35. Jagers, P.: General branching processes as Markov fields. *Stoch. Processes Appl.* **32**(2), 183–212 (1989)
36. Jagers, P., Nerman, O.: The growth and composition of branching populations. *Adv. Appl. Probab.* **16**(2), 221–259 (1984)
37. Jagers, P., Nerman, O.: The asymptotic composition of supercritical, multi-type branching populations. In: Azéma, J., Yor, M., Emery, M. (eds.) *Séminaire de probabilités de Strasbourg*, vol. 30, pp. 40–54. Springer, Berlin (1996)
38. Janson, S.: Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Processes Appl.* **110**(2), 177–245 (2004)
39. Kingman, J.F.C.: The first birth problem for an age-dependent branching process. *Ann. Probab.* **3**, 790–801 (1975)
40. Krapivsky, P.L., Redner, S.: Network growth by copying. *Phys. Rev. E* **71**(3), 036118 (2005)
41. Krapivsky, P., Redner, S.: Emergent network modularity. *J. Stat. Mech. Theory Exp.* **7**, 073405 (2017)
42. Krapivsky, P., Redner, S.: The magic of networks grown by redirection. *arXiv preprint arXiv:2305.10628* (2023)

43. Lee, J., Olvera-Cravioto, M.: PageRank on inhomogeneous random digraphs. *Stoch. Processes Appl.* **130**(4), 2312–2348 (2020)
44. Mahmoud, H.: *Pólya urn Models*. Chapman and Hall/CRC, Boca Raton (2008)
45. Masuda, N., Lambiotte, R.: *A Guidance to Temporal Networks*. World Scientific, Singapore (2016)
46. Mehrabian, A., Wormald, N.: It's a small world for random surfers. *Algorithmica* **76**(2), 344–380 (2016)
47. Nerman, O.: On the convergence of supercritical general (CMJ) branching processes. *Probab. Theory Relat. Fields* **57**(3), 365–395 (1981)
48. Newman, M.: *Networks: An Introduction*. Oxford University Press, Oxford (2010)
49. Newman, M.E.J.: The structure and function of complex networks. *SIAM Rev.* **45**, 167–256 (2003)
50. Niemi, S., Nummelin, E.: On non-singular renewal kernels with an application to a semigroup of transition kernels. *Stoch. Processes Appl.* **22**(2), 177–202 (1986)
51. Norris, J.R.: *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics, vol. 2. Cambridge University Press, Cambridge (1998). Reprint of 1997 original
52. Nummelin, E.: *General irreducible Markov chains and non-negative operators*. Cambridge University Press, Cambridge (2004)
53. Olvera-Cravioto, M.: PageRank's behavior under degree correlations. *Ann. Appl. Probab.* **3**(1), 1403–1442 (2021)
54. Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank Citation Ranking: Bringing Order to the Web. Stanford InfoLab, Stanford (1999)
55. Pakes, A.: Conditional limit theorems for a left-continuous random walk. *J. Appl. Probab.* **10**(1), 39–53 (1973)
56. Pandurangan, G., Raghavan, P., Upfal, E.: Using PageRank to characterize web structure. In: *International Computing and Combinatorics Conference*, pp. 330–339 (2002)
57. Porter, M.A.: Nonlinearity+ networks: a 2020 vision. In: Kevrekidis, P., Cuevas-Maraver, J., Saxena, A. (eds.) *Emerging Frontiers in Nonlinear Science*, pp. 131–159. Springer, Berlin (2020)
58. Rozenfeld, H.D., Avraham, D. ben: Designer nets from local strategies. *Phys. Rev. E* **70**, 056107 (2004)
59. Rudas, A., Tóth, B., Valkó, B.: Random trees and general branching processes. *Random Struct. Algorithms* **31**(2), 186–202 (2007)
60. Sayama, H., Pestov, I., Schmidt, J., Bush, B.J., Wong, C., Yamanoi, J., Gross, T.: Modeling complex systems with adaptive networks. *Comput. Math. Appl.* **65**(10), 1645–1664 (2013)
61. Sayama, H., Sinatra, R.: Social diffusion and global drift on networks. *Phys. Rev. E* **91**(3), 032809 (2015)
62. Smythe, R.T., Mahmoud, H.M.: A survey of recursive trees. *Theory Probab. Math. Stat.* **51**, 1–28 (1995)
63. Szymański, J.: On a nonuniform random recursive tree. In: Barlotti, A., Biliotti, M., Cossu, A., Korchmaros, G., Tallini, G. (eds.) *North-Holland Mathematics Studies*, vol. 144, pp. 297–306. Elsevier, Amsterdam (1987)
64. van der Hofstad, R.: *Random Graphs and Complex Networks*, vol. 43. Cambridge University Press, Cambridge (2016)
65. Vázquez, A.: Growing network with local rules: preferential attachment, clustering hierarchy, and degree correlations. *Phys. Rev. E* **67**(5), 056104 (2003)

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.