Adv. Appl. Probab. **56,** 785–824 (2024)
*doi:*10.1017/apr.2023.54

LOCAL WEAK LIMIT OF PREFERENTIAL ATTACHMENT RANDOM TREES WITH ADDITIVE FITNESS

TIFFANY Y. Y. LO,* *Uppsala University*

Abstract

We consider linear preferential attachment trees with additive fitness, where fitness is the random initial vertex attractiveness. We show that when the fitnesses are independent and identically distributed and have positive bounded support, the local weak limit can be constructed using a sequence of mixed Poisson point processes. We also provide a rate of convergence for the total variation distance between the r-neighbourhoods of a uniformly chosen vertex in the preferential attachment tree and the root vertex of the local weak limit. The proof uses a Pólya urn representation of the model, for which we give new estimates for the beta and product beta variables in its construction. As applications, we obtain limiting results and convergence rates for the degrees of the uniformly chosen vertex and its ancestors, where the latter are the vertices that are on the path between the uniformly chosen vertex and the initial vertex.

Keywords: Graph limit; complex networks; distributional approximation

2020 Mathematics Subject Classification: Primary 90B15
Secondary 05C80; 60F05

1. Introduction

There has been considerable interest in studying preferential attachment (PA) random graphs ever since [2] used them to explain the power-law degree distribution observed in some real-world networks, such as the World Wide Web. The primary feature of the stochastic mechanism consists of adding vertices sequentially over time, with some number of edges attached to them, and then connecting these edges to the existing graph in such a way that vertices with higher degrees are more likely to receive them. A general overview of PA random graphs can be found in the books [19] and [20]. In the basic models, vertices are born with the same constant 'weight' as their initial vertex attractiveness. To relax this assumption, [15] introduced a class of PA graphs with additive fitness (referred to as Model A in [15]), where *fitness* is defined as the *random* initial attractiveness. When the fitnesses are independent and identically distributed (i.i.d.), this family is the subject of recent works such as [22] and [25], whose results we discuss in Section 1.3.2. In this paper, we study the local weak limit of this family and provide a rate of convergence for the total variation distance between the local neighbourhoods of

Received 21 March 2021; accepted 10 October 2023.

* Postal address: Department of Mathematics, Uppsala University, Lägerhyddsvägen 1, 752 37 Uppsala, Sweden. Email address: yin_yuan.lo@math.uu.se

the PA tree and its local weak limit. The result extends that of [4], which considered PA graphs with constant initial attractiveness. As applications, we obtain limiting results for the degree distributions of the uniformly chosen vertex and its ancestors, where rates of convergence are also provided. Another objective of this article is to present the arguments of [4] in more detail; this is the main reason we consider the PA tree instead of the case where multiple edges are possible.

Before we define the model, note that we view the edges as directed, with a newly born vertex always sending a single outgoing edge to an existing vertex in the graph. We define the *weight* of a vertex as its in-degree plus its fitness; each time a vertex receives an edge from another vertex, its weight increases by one. As we mostly work conditionally on the fitness sequence, we also introduce a conditional version of the model. Note that the seed graph is a single vertex that has a non-random fitness, but as we explain in Section 1.3.1 below, the choice of the seed graph and its fitness has no effect on the local weak limit. It is chosen purely to streamline the argument.

Definition 1. (($\mathbf{x}$, n)-*sequential model and PA tree with additive fitness.*) Given a positive integer n and a sequence $\mathbf{x} := (x_i, i \geqslant 1)$, with $x_1 > -1$ and $x_i > 0$ for $i \geqslant 2$, we construct a sequence of random trees $(G_i, 1 \leqslant i \leqslant n)$ as follows. The seed graph G_1 consists of vertex 1 with initial attractiveness x_1 and degree 0. The graph G_2 is constructed by joining vertex 2 and vertex 1 by an edge, and equipping vertex 2 with initial attractiveness x_2. For $3 \leqslant m \leqslant n$, G_m is constructed from G_{m-1} by attaching one edge between vertex m and $k \in [m-1]$, and the edge is directed towards vertex k with probability

$$\frac{D^{(in)}_{m-1,k} + x_k}{m - 2 + \sum_{j=1}^{m-1} x_j} \quad \text{for } 1 \leqslant k \leqslant m-1,$$

where $D^{(in)}_{m,k}$ is the in-degree of vertex k in G_m, and $D^{(in)}_{m,k} = 0$ whenever $k \geqslant m$. The mth attachment step is completed by assigning vertex m the initial attractiveness x_m. We call the resulting graph G_n an ($\mathbf{x}$, n)-*sequential model*, and its law is denoted by $\mathrm{Seq}(\mathbf{x})_n$. Taking $X_1 = x_1$, we obtain the distribution $\mathrm{PA}(\pi, X_1)_n$ of the PA tree with additive fitness from mixing $\mathrm{Seq}(\mathbf{x})_n$ with the distribution of the i.i.d. fitness sequence $(X_i, i \geqslant 2)$.

1.1. Local weak convergence

The concept of local weak convergence was independently introduced by [1] and [3]. Here, we follow [20, Section 2.3 and 2.4], and also [3]. Informally, this involves exploring some random graph G_n from vertex o_n, chosen uniformly at random from G_n, and studying the distributional limit of the neighbourhoods of radius r rooted at o_n for each $r < \infty$. We begin with a few definitions. A rooted graph is a pair (G,o), where $G = (V(G), E(G))$ is a graph with vertex set $V(G)$ and edge set $E(G)$, and $o \in V(G)$ is the designated root in G. Next, let r be a finite positive integer. For any (G,o), denote by $B_r(G, o)$ the rooted neighbourhood of radius r around o. More formally, $B_r(G, o) = (V(B_r(G, o))), E(B_r(G, o)))$, where

$$V(B_r(G, o)) = \{u \in V(G) : \text{the distance between } u \text{ and } o \text{ is at most } r \text{ edges}\}, \text{ and}$$
$$E(B_r(G, o)) = \{\{u, v\} \in E(G) : u, v \in V(B_r(G, o))\}.$$

We refer to $B_r(G, o)$ as the *r-neighbourhood* of vertex o, or simply as the *local neighbourhood* of o when the reference to r is not needed. Finally, two rooted graphs (G,o) and (H,o') are isomorphic, for which we write $(G, o) \cong (H, o')$, if there is a bijection $\psi : V(G) \to V(H)$ such that $\psi(o) = o'$ and $\{u, v\} \in E(G)$ if and only if $\{\psi(u), \psi(v)\} \in E(H)$. Below we define the local

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

weak convergence of a sequence of finite random graphs $(G_n, n \geqslant 1)$ using a criterion given in [20, Theorem 2.14].

Definition 2. (*Local weak convergence.*) Let $(G_n, n \geqslant 1)$ be a sequence of finite random graphs. The local weak limit of G_n is (G, o) if, for all finite rooted graphs (H, v) and all finite r,

$$\frac{1}{n}\sum_{j=1}^{n} \mathbb{P}[(B_r(G_n, j), j) \cong (H, v)] \overset{n\to\infty}{\longrightarrow} \mathbb{P}[(B_r(G, o), o) \cong (H, v)]. \tag{1.1}$$

The left-hand side of (1.1) is the probability that the r-neighbourhood of a randomly chosen vertex is isomorphic to (H, v). Note that the convergence in (1.1) is equivalent to the convergence of the expectations of all bounded and continuous functions with respect to an appropriate metric on rooted graphs; see e.g. [20, Chapter 2].

1.2. Statement of the main results

In the main result, we fix $X_1 > -1$ and assume that $(X_i, i \geqslant 2)$ are i.i.d. positive bounded variables with distribution π. We write $\mu := \mathbb{E}X_2 < \infty$ and

$$\chi := \frac{\mu}{\mu+1}. \tag{1.2}$$

We first define the local weak limit of the PA tree with additive fitness, which is an infinite rooted random tree that generalises the Pólya point tree introduced in [4]. Hence, we refer to it simply as a π-Pólya point tree, with π being the fitness distribution of the PA tree. Denote this random tree by $(\mathcal{T}, 0)$, so that 0 is its root. We begin by explaining the Ulam–Harris labelling of trees that we use in the construction of $(\mathcal{T}, 0)$. Starting from the root 0, the *children* of any vertex $\bar{v}$ (if any) are generated recursively as $(\bar{v}, j)$, $j \in \mathbb{N}$, and we say that $\bar{v}$ is the *parent* of $(\bar{v}, j)$. With the convention $B_0(\mathcal{T}, 0) = \{0\}$, note that if $\bar{v} := (0, v_1, \ldots, v_r)$, then $(\bar{v}, i) \in \partial B_{r+1} := V(B_{r+1}(\mathcal{T}, 0)) \setminus V(B_r(\mathcal{T}, 0))$. Furthermore, each vertex $\bar{v} \in V((\mathcal{T}, 0))$ has a fitness $X_{\bar{v}}$ and a random *age* $a_{\bar{v}}$, where $0 < a_{\bar{v}} \leqslant 1$. We write $a_{\bar{v},i} := a_{(\bar{v},i)}$ for convenience. Apart from the root vertex 0, there are two types of vertices, namely, type L (for 'left') and type R (for 'right'). Vertex $\bar{v}$ belongs to type L if $a_{\bar{v},i} < a_{\bar{v}}$ for some $i \geqslant 1$, and to type R if $a_{\bar{v},i} \geqslant a_{\bar{v}}$ for all $i \geqslant 1$. There is exactly one type-L vertex in ∂B_r for each $r \geqslant 1$, and the labels $(0, 1, 1, \ldots, 1)$ are assigned to these vertices. For any vertex $\bar{v}$, let

$$R_{\bar{v}} = \text{the number of type-R children of } \bar{v} \text{ in the } (\mathcal{T}, 0). \tag{1.3}$$

We also label the type-R vertices in increasing order of their ages, so that if $\bar{v}$ is the root or belongs to type L, then $a_{\bar{v},1} \leqslant a_{\bar{v}} \leqslant a_{\bar{v},2} \leqslant \ldots \leqslant a_{\bar{v},1+R_{\bar{v}}}$, while if $\bar{v}$ belongs to type R, then $a_{\bar{v}} \leqslant a_{\bar{v},1} \leqslant \ldots \leqslant a_{\bar{v},R_{\bar{v}}}$. See Figure 1 below for an illustration of $(\mathcal{T}, 0)$ and the vertex ages. To understand how the vertex types and the ages above arise in the weak limit, consider the r-neighbourhood of a uniformly chosen vertex k_0 in $G_n \sim \mathrm{PA}(\pi, X_1)_n$. Observe that there is a unique path from k_0 to the initial vertex, unless k_0 is itself the initial vertex. Apart from k_0, the vertices in the r-neighbourhood that belong to this path are called type-L vertices, and the remaining vertices are referred to as type-R vertices. The ages in $(\mathcal{T}, 0)$ encode the (rescaled) arrival times of the vertices in the r-neighbourhood of vertex k_0, which determine their degree distributions. A comparison of the 2-neighbourhoods of the PA tree and $(\mathcal{T}, 0)$ is given in Figure 1.

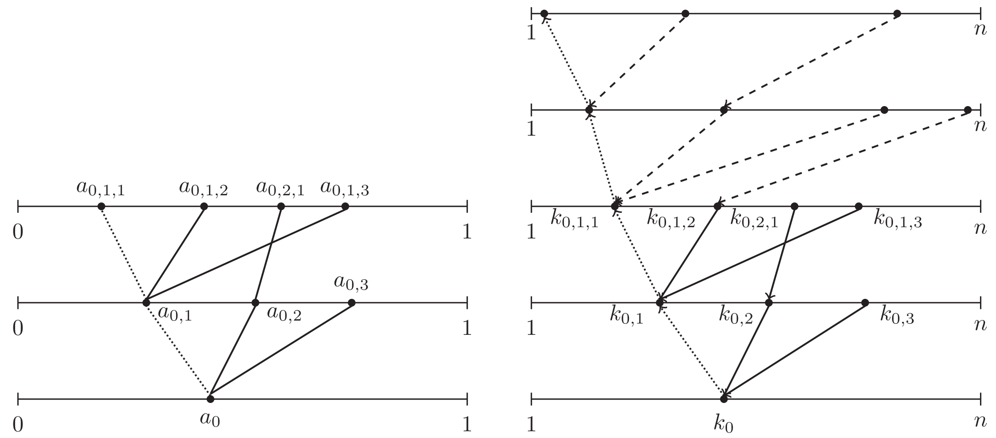

FIGURE 1. A comparison between the 2-neighbourhoods in $(\mathcal{T}, 0)$ (left) and the PA tree G_n rooted at the uniformly chosen vertex k_0 (right). We assign Ulam–Harris labels as subscripts $(k_{\bar{v}})$ to the vertices in (G_n, k_0) to better compare the 2-neighbourhoods. In both figures, the vertex location corresponds to either its arrival time or its age. A vertex is a type-L (type-R) child if it lies on the dotted (solid) path. On the left, $R_0 = 2$, $R_{0,1} = 2$, $R_{0,2} = 1$, and $R_{0,3} = 0$. On the right, the dotted path starting from k_0 leads to the initial vertex. The 2-neighbourhoods are coupled so that they are isomorphic, and the vertex ages and rescaled arrival times are close to each other. The unlabelled vertices and their respective edges (dashed or dotted) are not coupled.

Definition 3. (*π-Pólya point tree.*) A π-Pólya point tree $(\mathcal{T}, 0)$ is defined recursively as follows. The root 0 has an age $a_0 = U_0^{\chi}$, where $U_0 \sim \mathrm{U}[0, 1]$. Assuming that $\bar{v} \in \partial B_r$ and $a_{\bar{v}}$ have been generated, we define its children $(\bar{v}, j) \in \partial B_{r+1}$ for $j = 1, 2, \ldots$ as follows. Independently of all random variables generated before, let $X_{\bar{v}} \sim \pi$ and

$$Z_{\bar{v}} \sim \begin{cases} \mathrm{Gamma}(X_{\bar{v}}, 1), & \text{if } \bar{v} \text{ is the root or of type R;} \\ \mathrm{Gamma}(X_{\bar{v}} + 1, 1), & \text{if } \bar{v} \text{ is of type L.} \end{cases}$$

If $\bar{v}$ is the root or of type L, let $a_{\bar{v},1} | a_{\bar{v}} \sim \mathrm{U}[0, a_{\bar{v}}]$ and $(a_{\bar{v},i}, 2 \leqslant i \leqslant 1 + R_{\bar{v}})$ be the points of a mixed Poisson point process on $(a_{\bar{v}}, 1]$ with intensity

$$\lambda_{\bar{v}}(y)dy := \frac{Z_{\bar{v}}}{\mu a_{\bar{v}}^{1/\mu}} y^{1/\mu - 1} dy.$$

If $\bar{v}$ is of type R, then $(a_{\bar{v},i}, 1 \leqslant i \leqslant R_{\bar{v}})$ are sampled as the points of a mixed Poisson process on $(a_{\bar{v}}, 1]$ with intensity $\lambda_{\bar{v}}$. We obtain $(\mathcal{T}, 0)$ by continuing this process ad infinitum.

Remark 1. When exploring the r-neighbourhood of the uniformly chosen vertex in the PA tree, the type-L vertices are uncovered via the incoming edge it received from the probed uniformly chosen vertex or type-L vertices. To account for the size-biasing effect of these edges, the type-L gamma variables in the π-Pólya point tree thus have a unit increment in the shape parameter.

We define the total variation distance between two probability distributions ν_1 and ν_2 as

$$d_{\mathrm{TV}}(\nu_1, \nu_2) = \inf\{\mathbb{P}[V \neq W] : (V, W) \text{ is a coupling of } \nu_1 \text{ and } \nu_2\}. \tag{1.4}$$

When proving the local weak convergence, we couple the random elements $(B_r(G_n, k_0), k_0)$ and $(B_r(\mathcal{T}, 0), 0)$ in the space of rooted graphs (modulo isomorphisms) $\mathcal{G}$ so that they are isomorphic with high probability (w.h.p.), thus bounding their total variation distance. In the main theorem below, we emphasise that the convergence does not take into account the ages and fitness of the π-Pólya point tree, but they are essential for the graph construction and are used for the couplings later.

Theorem 1. *Suppose that the fitness distribution π is supported on $(0, \kappa]$ for some $\kappa < \infty$. Let $G_n \sim \mathrm{PA}(\pi, X_1)_n$, let k_0 be a uniformly chosen vertex in G_n, and let $(\mathcal{T}, 0)$ be the π-Pólya point tree. Then, given $r < \infty$, there is a positive constant $C := C(X_1, \mu, r, \kappa)$ such that*

$$d_{\mathrm{TV}}\left(\mathscr{L}((B_r(G_n, k_0), k_0)), \mathscr{L}((B_r(\mathcal{T}, 0), 0))\right) \leqslant C(\log\log n)^{-\chi} \tag{1.5}$$

for all $n \geqslant 3$. This implies that the local weak limit of G_n is the π-Pólya point tree.

Next we state the limiting results for some degree statistics of the PA tree. The connection of the following results to [4, 7, 20, 25, 30] is discussed in detail later in Section 8, where we also give the probability mass functions of the limiting distributions. Recall that $D^{(in)}_{n,j}$ is the in-degree of vertex j in $G_n \sim \mathrm{PA}(\pi, X_1)_n$. Define the degree of vertex j in G_n as

$$D_{n,j} := D^{(in)}_{n,j} + 1, \quad \text{with } D^0_n := D_{n,k_0}, \tag{1.6}$$

so that D^0_n is the degree of the uniformly chosen vertex. Let $L(i)$ be the Ulam–Harris labels $(0, 1, 1, \ldots, 1)$ such that $|(0, 1, \ldots, 1)| = i + 1$, so that $k_{L[i]}$ is the type-L vertex that is exactly i edges away from k_0. Type-L vertices are commonly known as the ancestors of k_0 in fringe tree analysis (see e.g. [21]), where their degrees are often of particular interest. Fixing $r \in \mathbb{N}$, let D^0_n be as in (1.6) and

$$D^i_n = D_{n,k_{L[i]}} \quad \text{if } k_{L[i]} \neq 1 \text{ for all } 1 \leqslant i \leqslant r;$$

and if $k_{L[i]} = 1$ for some $i \leqslant r$, let

$$D^j_n = D_{n,k_{L[j]}} \quad \text{for } 1 \leqslant j < i \qquad \text{and} \quad D^j_n = -1 \quad \text{for } i \leqslant j \leqslant r.$$

As the probability that we see vertex 1 in the r-neighbourhood $B_r(G_n, k_0)$ tends to zero as $n \to \infty$, $(D^0_1, \ldots, D^r_n)$ can be understood as the joint degree sequence of $(k_0, k_{L[1]}, \ldots, k_{L[r]})$. The next result concerning this joint degree sequence follows directly from Theorem 1.

Corollary 1. *Retaining the notation above, assume that π is supported on $(0, \kappa]$ for some $\kappa < \infty$. Define $U_0 \sim \mathrm{U}[0, 1]$, $a_{L[0]} = U_0^{\chi}$, and given $a_{L[i-1]}$, let $a_{L[i]} \sim \mathrm{U}\left[0, a_{L[i-1]}\right]$ for $1 \leqslant i \leqslant r$. Independently from $\left(a_{L[i]}, 0 \leqslant i \leqslant r\right)$, let $X_{L[i]}$ be i.i.d. random variables with distribution π, and let*

$$Z_{L[i]} \sim \begin{cases} \mathrm{Gamma}(X_{L[i]}, 1) & \text{if } i = 0, \\ \mathrm{Gamma}(X_{L[i]} + 1, 1) & \text{if } 1 \leqslant i \leqslant r. \end{cases}$$

Writing $L[0] = 0$, let $R_{L[i]}$ be conditionally independent variables with $R_{L[i]} \sim \mathrm{Po}\left(Z_{L[i]}\left(a_{L[i]}^{-1/\mu} - 1\right)\right)$. Define $\overline{R}^{(r)} := (R_0 + 1, R_{L[1]} + 2, \ldots, R_{L[r]} + 2)$ and $\overline{D}^{(r)}_n := \left(D^0_n, D^1_n, \ldots, D^r_n\right)$. There is a positive constant $C := C(X_1, \mu, r, \kappa)$ such that

$$d_{\mathrm{TV}}\left(\mathscr{L}\left(\overline{D}^{(r)}_n\right), \mathscr{L}\left(\overline{R}^{(r)}\right)\right) \leqslant C(\log\log n)^{-\chi} \quad \textit{for all } n \geqslant 3.$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

We now state a convergence result for D_n^0 in (1.6). In view of Definition 2, the limiting distribution of D_n^0 and the convergence rate can be read from Theorem 1. However, the theorem below holds without the assumption of bounded fitness and has a much sharper rate. The improvement in the rate can be understood as a consequence of the fact that k_0 only needs to be large enough so that w.h.p. it has a small degree, in contrast to having a small enough r-neighbourhood for all $r < \infty$, as required when proving Theorem 1. In the interest of article length, we give the proof of this result only in the supplementary article [24].

Theorem 2. *Assume that the pth moment of the distribution* π *is finite for some* $p > 4$. *Let* $R_0 \sim \mathrm{Po}\Big(Z_0\Big(a_0^{-1/\mu} - 1\Big)\Big)$, *where given* $X_0 \sim \pi$, $Z_0 \sim \mathrm{Gamma}(X_0, 1)$, *and independently of* Z_0, $U_0 \sim \mathrm{U}[0, 1]$ *and* $a_0 := U_0^{\chi}$. *Writing* $\xi_0 = R_0 + 1$, *there are positive constants* $C := C(X_1, \mu, p)$ *and* $0 < d < \chi(1/4 - 1/(2p))$ *such that*

$$d_{\mathrm{TV}}\left(\mathscr{L}(D_n^0), \mathscr{L}(\xi_0)\right) \leqslant Cn^{-d} \quad \textit{for all } n \geqslant 1. \tag{1.7}$$

1.3. Possible extensions and related works

Below we discuss the possible ways of extending Theorem 1, although, to limit the length of this article, we refrain from pursuing these directions. We also give an overview of recent developments concerning PA graphs with additive fitness, and collect some results on the local weak convergence of related PA models.

1.3.1. *Possible extensions.* The convergence rate in (1.5) roughly follows from the fact that $k_0 \geqslant n(\log\log n)^{-1}$ with probability at least $(\log\log n)^{-1}$, and on this event, we can couple the two graphs so that the probability that $(B_r(G_n, k_0), k_0) \cong (B_r(\mathcal{T}, 0), 0)$ tends to one as $n \to \infty$. It is likely possible to improve the rate by optimising this and similar choices of thresholds, as well as by making the dependence on the radius r explicit by carefully keeping track of the coupling errors, but with much added technicality.

[10, Theorem 1] established that when $X_i = 1$ almost surely (a.s.) for all $i \geqslant 2$, the choice of the seed graph has no effect on the local weak limit. This is because w.h.p., the local neighbourhood does not contain any of the seed vertices. By simply replacing the seed graph in the proof, we can show that Theorem 1 holds for more general seed graphs. With some straightforward modifications to the proof, we can also show that when the fitness is bounded, the π-Pólya point tree is the local weak limit of the PA trees with self-loops, and when each vertex in the PA model sends $m \geqslant 2$ outgoing edges, the limit is a variation of the π-Pólya point tree; see e.g. [4] for the non-random unit fitness case.

By adapting the argument of the recent paper [17], it is possible to show that the weak convergence in Theorem 1 holds for fitness distributions with finite pth moment for some $p > 1$. The convergence rate should be valid for fitness distributions with at least exponentially decaying tails, but the assumption of bounded fitness greatly simplifies the proof. The i.i.d. assumption is only needed so that the fitness variables that we see in the local neighbourhood are i.i.d., and for applying the standard moment inequality in Lemma 15. Hence, we believe the theorem to hold at least for a fitness sequence $\mathbf{X}$ such that (1) X_i has the same marginal distribution π for all $i \geqslant 2$, and (2) for some $m \geqslant 1$, the variables in the collection $(X_i, i \in A)$ are independent for any A such that $\{i, j \in A : |i - j| > 2m\}$. If, in the limit, the vertex labels in the local neighbourhood are at least $2m$ apart from each other w.h.p., then (1) and (2) ensure that these vertices have i.i.d. fitness, while (2) alone may

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

be sufficient for proving a suitable analogue of the standard moment inequality, Lemma 15. The weak convergence in Theorem 1 should hold *in probability*, meaning that as $n \to \infty$, $\frac{1}{n}\sum_{j=1}^{n} \mathbb{1}[(B_r(G_n, j), j) \cong (H, v)] \to \mathbb{P}[(B_r(\mathcal{T}, 0), 0) \cong (H, v)]$. This convergence is valid in the case where the initial attractivenesses are equal a.s. [20, Chapter 5], and for certain PA models where each newly added vertex sends a random number of outgoing edges [17]. For the PA tree with additive fitness, this could be proved by adapting the 'second moment' method used in [17, 20], which involves establishing that the r-neighbourhoods of two independently, uniformly chosen vertices in the PA tree are disjoint w.h.p.

1.3.2. *Related works.* The fitness variables were assumed to be i.i.d. in [5, 22, 25]. In [25], martingale techniques were used to investigate the maximum degree for fitness distributions with different tail behaviours; the results are applicable to PA graphs with additive fitness that allow for multiple edges. The authors also studied the empirical degree distribution (e.d.d.), whose details we defer to Remark 4. The model is a special case of the PA tree considered in [22], where vertices are chosen with probability proportional to a suitable function of their fitness and degree at each attachment step. Using Crump–Mode–Jagers (CMJ) branching processes, [22] studied the e.d.d. and the condensation phenomenon. The same method was applied in [5] to investigate the e.d.d., the height, and the degree of the initial vertex, assuming that the fitnesses are bounded. Note that these articles focused on 'global' results, which cannot be deduced from the local weak limit.

As observed in [36], the model is closely related to the weighted random recursive trees introduced in [9]. By the Pólya urn representation (Theorem 3 below), the PA tree with additive fitness can be viewed as a special case of this model class, where the weights of the vertices are distributed as $\left(S_{n,j}^{(X)} - S_{n,j-1}^{(X)}, 2 \leqslant j \leqslant n\right)$ with $S_{n,0}^{(X)} = 0$ and $S_{n,n}^{(X)} = 1$, and with $B_j^{(X)}$ and $S_{n,j}^{(X)}$ being the variables B_j and $S_{n,j}$ in (1.8) and (1.9) mixed over the fitness sequence $\mathbf{X}$. For weighted random recursive trees, [9] studied the average degree of a fixed vertex and the distance between a newly added vertex and the initial vertex. The joint degree sequence of fixed vertices, the height, and the profile were investigated in [36], and a more refined result on the height was given in [29]. For $\mathbf{X}$ a deterministic sequence satisfying a certain growth condition, [36] and [29] showed that their results are applicable to the corresponding PA tree with additive fitness.

We now survey the local weak limit results developed for PA graphs. When $x_1 = 0$ and $x_i = 1$ for all $i \geqslant 2$, the $(\mathbf{x}, n)$-sequential model is the pure 'sequential' model in [4] with no multiple edges, and it is a special case of the model considered in [35], where the 'weight' function is the identity function plus one. Using the CMJ embedding method, [35] studied the asymptotic distribution of the subtree rooted at a uniformly chosen vertex, which implies the local weak convergence of the PA family considered in that work. The urn representation of PA models was used in [4], [20, Chapter 5], [16, Chapter 4], and [17] to study local weak limits. In particular, these works showed that the weak limit of several PA models with non-random fitness and possibly random out-degrees is a variant of the Pólya point tree [4]. Again using the CMJ method, it can be shown that the local weak convergence result of [18] for a certain 'continuous-time branching process tree' implies that the PA tree with additive fitness converges in the *directed* local weak sense. Finally, a different PA model was considered in [5, 8, 11, 12], where the probability that a new vertex attaches to an existing vertex is proportional to its fitness times its degree. Currently, there are no results concerning the local weak convergence of this model class.

FIGURE 2. An example of the $(\mathbf{x}, n)$-Pólya urn tree for $n = 5$, where $U_i \sim \mathrm{U}\left[0, S_{n,i-1}\right]$ for $i = 2, \ldots, 5$ and an outgoing edge is drawn from vertices i to j if $U_i \in \left[S_{n,j-1}, S_{n,j}\right)$.

1.4. Proof overview

1.4.1. *Pólya urn representation of the* $(\mathbf{x}, n)$*-sequential model.* In the proof of Theorem 1, the key ingredient is an alternative definition of the $(\mathbf{x}, n)$-sequential model in Definition 1, which relies on the fact that the dynamics of PA graphs can be represented as embedded classical Pólya urns. In a classical Pólya urn initially with a red balls and b black balls, a ball is chosen randomly from the urn, and is returned to the urn along with a new ball of the same colour. It is well-known that the a.s. limit of the proportion of red balls has the Beta(a,b) distribution; see e.g. [27, Theorem 3.2, p.53]. Furthermore, by de Finetti's theorem (see e.g. [27, Theorem 1.2, p.29]), conditional on $\beta \sim \mathrm{Beta}(a, b)$, the indicators that a red ball is chosen at each step are distributed as independent Bernoulli variables with parameter β. In the PA mechanism, an existing vertex i can be represented by some colour i in an urn, and its weight is given by the total weight of the balls in the urn. At each urn step, we choose a ball with probability proportional to its weight, and if vertex i is chosen at some step $j > i$, we return the chosen ball, an extra ball of colour i with weight 1, and a ball of new colour j with weight x_j to the urn. Classical Pólya urns are naturally embedded in this multi-colour urn; see [31]. The attachment steps of the graph can therefore be generated independently when conditioned on the associated beta variables.

Definition 4. (*$(\mathbf{x}, n)$-Pólya urn tree.*) Given $\mathbf{x}$ and n, let $T_j := \sum_{i=1}^{j} x_i$, and $\left(B_j, 1 \leqslant j \leqslant n\right)$ be independent random variables such that $B_1 := 1$ and

$$B_j \sim \mathrm{Beta}\big(x_j, j-1+T_{j-1}\big) \quad \text{for } 2 \leqslant j \leqslant n. \tag{1.8}$$

Moreover, let $S_{n,0} := 0$, $S_{n,n} := 1$, and

$$S_{n,j} := \prod_{i=j+1}^{n} (1 - B_i) \quad \text{for } 1 \leqslant j \leqslant n-1. \tag{1.9}$$

We connect n vertices with labels $[n] := \{1, \ldots, n\}$ as follows. Let $I_j = \left[S_{n,j-1}, S_{n,j}\right)$ for $1 \leqslant j \leqslant n$. Conditionally on $\left(S_{n,j}, 1 \leqslant j \leqslant n-1\right)$, let $(U_j, 2 \leqslant j \leqslant n)$ be independent variables such that $U_j \sim \mathrm{U}[0, S_{n,j-1}]$. If $j < k$ and $U_k \in I_j$, we attach an outgoing edge from vertex k to vertex j. We say that the resulting graph is an $(\mathbf{x}, n)$-Pólya urn tree and denote its law by $\mathrm{PU}(\mathbf{x})_n$.

An example of the $(\mathbf{x}, n)$-Pólya urn tree is given in Figure 2. Note that $1 - B_j$ in Definition 4 is β_{j-1} in [36]. As we only work with B_j and $S_{n,j}$ by fixing the sequence $\mathbf{x}$, we omit $\mathbf{x}$ from their notation throughout this article. The $(\mathbf{x}, n)$-Pólya urn tree is related to the $(\mathbf{x}, n)$-sequential model via the following result of [36].

Theorem 3. ([36, Theorem 1.1].) *Let G_n be an $(\mathbf{x}, n)$-Pólya urn tree. Then G_n has the same law as the $(\mathbf{x}, n)$-sequential model; that is,* $\mathrm{PU}(\mathbf{x})_n \overset{d}{=} \mathrm{Seq}(\mathbf{x})_n$.

From now on, we work with the $(\mathbf{x}, n)$-Pólya urn tree in place of the $(\mathbf{x}, n)$-sequential model.

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

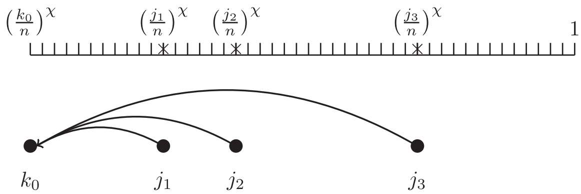

FIGURE 3. An illustration of the relationship between the Bernoulli sequence on $((k_0/n)^\chi, 1]$ constructed using $(\mathbb{1}[U_j \in I_{k_0}], k_0 + 1 \leqslant j \leqslant n)$ and the $(\mathbf{x}, n)$-Pólya urn tree, where k_0 is the uniformly chosen vertex, and U_j and I_{k_0} are as in Definition 4. We put a point on $(j/n)^\chi$ if vertex k_0 receives the outgoing edge from vertex j. Here the type-R children of k_0 are j_1, j_2, and j_3. The rescaled arrival times $((j/n)^\chi, k_0 + 1 \leqslant j \leqslant n)$ are later used to discretise the mixed Poisson process in the coupling step.

1.4.2. *Coupling for the non-random fitness case.* When proving Theorem 1, we couple the (randomised) urn tree and the π-Pólya point tree so that for all positive integers r, the probability that their r-neighbourhoods are not isomorphic is of order at most $(\log\log n)^{-\chi}$. To give a brief overview of the coupling, below we suppose that $X_i = 1$ a.s. for $i \geqslant 2$ and only consider the $r = 1$ case. We couple the children of the uniformly chosen vertex in the urn tree G_n and the root in the π-Pólya point tree $(\mathcal{T}, 0)$ so that, w.h.p., they have the same number of children, and the ages and the rescaled arrival times of these children are close enough. Note that for non-random fitness, the urn tree is simply an alternative definition of the PA tree with additive fitness. For G_n, we use the terms 'age' and 'rescaled arrival times' interchangeably.

I. The ages of the roots. As mentioned before, the ages in $(\mathcal{T}, 0)$ encode the rescaled arrival times in G_n. Since, for any vertex in either G_n or $(\mathcal{T}, 0)$, the number of its children and the ages of its children depend heavily on its own age, we need to couple the uniformly chosen vertex k_0 in G_n and the root of $(\mathcal{T}, 0)$ so that their ages are close enough.

II. The type-R children and a Bernoulli–Poisson coupling. Using the urn representation in Definition 4, the ages and the number of the type-R children of vertex k_0 can be encoded in a sequence of conditionally independent Bernoulli variables, where the success probabilities are given in terms of the variables B_i and $S_{n,i}$ in (1.8) and (1.9). See Figure 3 for an illustration. To couple this Bernoulli sequence to a suitable discretisation of the mixed Poisson point process in Definition 3, we use the beta–gamma algebra and the law of large numbers to approximate B_i and $S_{n,i}$ for large enough i. Once we use these estimates to swap the success probabilities with simpler quantities, we can apply the standard Bernoulli–Poisson coupling.

III. The ages of the type-L children. It is clear that the number of type-L children of vertex k_0 and the number of type-L children of the root of $(\mathcal{T}, 0)$ differ only when $k_0 = 1$, which occurs with probability n^{-1}. To couple their ages so that they are close enough, we use the estimates for B_i and $S_{n,i}$ to approximate the distribution of the type-L child of vertex k_0. This completes the coupling of the 1-neighbourhoods. We reiterate that although the closeness of the ages is not part of the local weak convergence, we need to closely couple the ages of the children of vertex k_0 and of those of the root in $(\mathcal{T}, 0)$, as otherwise we cannot couple the 2-neighbourhoods when we prove the theorem for all finite radii.

1.4.3. *Random fitness and the general local neighbourhood.* Here we summarise the additional ingredients needed for handling the random fitness and for coupling the neighbourhoods of any finite radius. When $\mathbf{x}$ is a realisation of the fitness sequence $\mathbf{X}$, B_i and $S_{n,i}$ can be approximated in the same way as in the case of non-random fitness if, w.h.p., the sums $\sum_{h=2}^{i} X_h$ are close to $(i-1)\mu$ for all i greater than some suitably chosen function $\phi(n)$. As $\mathbf{X}$ is i.i.d., we can apply standard moment inequalities to show that this event occurs w.h.p. The remaining parts of the coupling are similar to the above. When coupling general r-neighbourhoods, we use induction over the neighbourhood radius. For any vertex in the neighbourhood other than the uniformly chosen vertex, the distributions of its degree and the ages of its children cannot be read from Definition 4, owing to the conditioning effects of the discovered edges in the breadth-first search that we define later. However, these effects can be quantified when we work conditionally on the fitness sequence. In particular, their distributions can be deduced from an urn representation of the $(\mathbf{x}, n)$-sequential model, conditional on the discovered edges. Consequently, the coupling argument for the root vertex can be applied to these non-root vertices as well.

1.5. Article outline

The approximation results for B_i and $S_{n,i}$ in (1.8) and (1.9) are stated in Section 2 and proved in Section 9. In Section 3, we define the tree exploration process, and we describe the offspring distributions of the root and of any type-L or type-R parent in the subsequent generations of the local neighbourhood in the $(\mathbf{x}, n)$-Pólya urn tree, using the complementary result derived in [24, Section 10]. We also introduce a conditional analogue of the π-Pólya point tree $(\mathcal{T}, 0)$ in Section 4, which we need for coupling the urn tree and $(\mathcal{T}, 0)$. We couple the 1-neighbourhoods in the urn tree and this analogue in Section 5, and their general r-neighbourhoods in Section 6. In Section 7, to prove Theorem 1, we couple the analogue and $(\mathcal{T}, 0)$. Finally, in Section 8 we discuss in more detail the connection of Corollary 1 and Theorem 2 to the previous works [4, 7, 20, 25].

2. Approximation of the beta and product beta variables

Let B_i and $S_{n,i}$ be as in (1.8) and (1.9), where we treat $\mathbf{x}$ in B_i and $S_{n,i}$ as a realisation of the fitness sequence $\mathbf{X} := (X_i, i \geqslant 1)$. In this section, we state the approximation results for B_i and $S_{n,i}$, assuming that $X_1 := x_1 > -1$ is fixed and $(X_i, i \geqslant 2)$ are i.i.d. positive variables with $\mu := \mathbb{E}X_2 < \infty$ and $\mathbb{E}\big[X_2^p\big] < \infty$ for some $p > 2$. These results are later used to derive the limiting degree distribution of the uniformly chosen vertex of the PA tree and the corresponding convergence rate, where the fitness is not necessarily bounded (Theorem 2). The proofs of the upcoming lemmas are deferred to Section 9. For the approximations, we require that w.h.p., $\mathbf{X}$ is such that $\sum_{h=2}^{i} X_h$ is close enough to its mean $(i-1)\mu$ for all i sufficiently large. So given $0 < \alpha < 1$ and n, we define

$$A_{\alpha,n} = \left\{ \bigcap_{i=\lceil\phi(n)\rceil}^{\infty} \left\{ \left| \sum_{h=2}^{i} X_h - (i-1)\mu \right| \leqslant i^{\alpha} \right\} \right\}, \tag{2.1}$$

where $\phi(n) = \Omega(n^{\chi})$, with χ as in (1.2). The first lemma is due to an application of standard moment inequalities.

Lemma 1. *Assume that* $\mathbb{E}\big[X_2^p\big] < \infty$ *for some* $p > 2$. *Given a positive integer* n *and* $1/2 + 1/p < \alpha < 1$, *there is a constant* $C := C(\mu, \alpha, p)$ *such that* $\mathbb{P}[A_{\alpha,n}] \geqslant 1 - Cn^{\chi[-p(\alpha-1/2)+1]}$.

In the remainder of this article, we mostly work with a realisation $\mathbf{x}$ of $\mathbf{X}$ such that $A_{\alpha,n}$ holds, which we denote (abusively) by $\mathbf{x} \in A_{\alpha,n}$. Below we write $\mathbb{P}_{\mathbf{x}}$ and $\mathbb{E}_{\mathbf{x}}$ to indicate the conditioning on a specific realisation of the fitness sequence $\mathbf{x}$. The next lemma, which extends

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

[4, Lemma 3.1], states that $S_{n,k} \approx (k/n)^{\chi}$ for large enough k and n when $\mathbf{x} \in A_{\alpha,n}$. The proof relies on the fact that we can replace $\mathbb{E}_{\mathbf{x}}[S_{n,k}]$ with $(k/n)^{\chi}$ when $\mathbf{x} \in A_{\alpha,n}$, and approximate $S_{n,k}$ with $\mathbb{E}_{\mathbf{x}}[S_{n,k}]$ using a martingale argument. As $A_{\alpha,n}$ occurs w.h.p., the result allows us to substitute these $S_{n,k}$ with $(k/n)^{\chi}$ when we construct a coupling for the $(\mathbf{x}, n)$-Pólya urn tree (Definition 4).

Lemma 2. *Given a positive integer n and $1/2 < \alpha < 1$, assume that $\mathbf{x} \in A_{\alpha,n}$. Then there are positive constants $C := C(x_1, \mu, \alpha)$ and $c := c(x_1, \mu, \alpha)$ such that*

$$\mathbb{P}_{\mathbf{x}}\left[\max_{\lceil\phi(n)\rceil \leqslant k \leqslant n}\left|S_{n,k} - \left(\frac{k}{n}\right)^{\chi}\right| \leqslant \delta_n\right] \geqslant 1 - \varepsilon_n, \tag{2.2}$$

where $\phi(n) = \Omega(n^{\chi})$, $\delta_n := Cn^{-\chi(1-\alpha)/4}$, and $\varepsilon_n := cn^{-\chi(1-\alpha)/2}$.

Remark 2. For vertices whose arrival time is of order at most n^{χ}, the upper bound δ_n in Lemma 2 is meaningful only when $\chi > (\alpha + 3)/4$, as otherwise δ_n is of order greater than $(k/n)^{\chi}$. However, the bound is still useful for studying the local weak limit of the PA tree, as w.h.p. we do not see vertices with arrival times that are $o(n)$ in the local neighbourhood of the uniformly chosen vertex.

The last lemma is an extension of [4, Lemma 3.2]. It says that on the event $A_{\alpha,n}$, we can construct the urn tree by generating suitable gamma variables in place of B_j in (1.8). These gamma variables are comparable to $Z_{\bar{v}}$ in the construction of the π-Pólya point tree (Definition 3), after assigning Ulam–Harris labels to the vertices of the urn tree. To state the result, we recall a distributional identity. Independently of B_j, let $\big((\mathcal{Z}_j, \widetilde{\mathcal{Z}}_{j-1}), 2 \leqslant j \leqslant n\big)$ be conditionally independent variables such that $\mathcal{Z}_j \sim \mathrm{Gamma}(x_j, 1)$ and $\widetilde{\mathcal{Z}}_j \sim \mathrm{Gamma}(T_j + j, 1)$, where $T_j := \sum_{i=1}^{j} x_i$. Then by the beta–gamma algebra (see e.g. [26]),

$$\left(B_j, \widetilde{\mathcal{Z}}_{j-1} + \mathcal{Z}_j\right) =_d \left(\frac{\mathcal{Z}_j}{\mathcal{Z}_j + \widetilde{\mathcal{Z}}_{j-1}}, \widetilde{\mathcal{Z}}_{j-1} + \mathcal{Z}_j\right) \quad \text{for } 2 \leqslant j \leqslant n,$$

where the two random variables on the right-hand side are independent. Using the law of large numbers, we prove the following.

Lemma 3. *Given a positive integer n and $1/2 < \alpha < 3/4$, let $\mathcal{Z}_j$ and $\widetilde{\mathcal{Z}}_j$ be as above. Define the event*

$$E_{\varepsilon,j} := \left\{\left|\frac{\mathcal{Z}_j}{\mathcal{Z}_j + \widetilde{\mathcal{Z}}_{j-1}} - \frac{\mathcal{Z}_j}{(\mu+1)j}\right| \leqslant \frac{\mathcal{Z}_j}{(\mu+1)j}\varepsilon\right\} \quad \text{for } 2 \leqslant j \leqslant n. \tag{2.3}$$

When $\mathbf{x} \in A_{\alpha,n}$, there is a positive constant $C := C(x_1, \alpha, \mu)$ such that

$$\mathbb{P}_{\mathbf{x}}\left[\bigcap_{j=\lceil\phi(n)\rceil}^{n} E_{\varepsilon,j}\right] \geqslant 1 - C(1+\varepsilon)^4\varepsilon^{-4}n^{\chi(4\alpha-3)}, \tag{2.4}$$

where $\phi(n) = \Omega(n^{\chi})$. In addition,

$$\mathbb{P}_{\mathbf{x}}\left[\bigcap_{j=\lceil\phi(n)\rceil}^{n}\{\mathcal{Z}_j \leqslant j^{1/2}\}\right] \geqslant 1 - \sum_{j=\lceil\phi(n)\rceil}^{n} j^{-2}\prod_{\ell=0}^{3}(x_j + \ell), \tag{2.5}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

and if $x_i \in (0, \kappa]$ *for all* $i \geqslant 2$*, then there is a positive constant* C *such that*

$$\mathbb{P}_{\mathbf{x}}\left[\bigcap_{j=\lceil\phi(n)\rceil}^{n}\left\{\mathcal{Z}_j \leqslant j^{1/2}\right\}\right] \geqslant 1 - C\kappa^4 n^{-\chi}. \tag{2.6}$$

3. Offspring distributions of the type-L and type-R parents in the urn tree

Recall that a vertex in the local neighbourhood of the uniformly chosen vertex k_0 in the PA tree is a type-L parent if it lies on the path from the uniformly chosen vertex to the initial vertex; otherwise it is a type-R parent. In this section, we state two main lemmas for the $(\mathbf{x}, n)$-Pólya urn tree (Definition 4). In Lemma 5, for the root k_0 or any type-L or type-R parent in the subsequent generations of the local neighbourhood, we encode its type-R children in a suitable Bernoulli sequence that we introduce later. In Lemma 6, we construct the distribution of the type-L children of the uniformly chosen vertex and any type-L parent in the subsequent generations. For vertex k_0, these results follow immediately from the urn representation in Definition 4 and Theorem 3. For the non-root vertices, these lemmas cannot be deduced from Definition 4. Instead, we need an urn representation that accounts for the conditioning effects of the edges uncovered in the neighbourhood exploration. For more detail on this variant, we refer to the supplementary article [24, Section 10].

3.1. Breadth-first search

To construct the offspring distributions, we need to keep track of the vertices that we discover in the local neighbourhood. For this purpose, we have to precisely define the exploration process. We start with a definition.

Definition 5. (*Breadth-first order.*) Write $\bar{w} <_{UH} \bar{y}$ if the Ulam–Harris label $\bar{w}$ is smaller than $\bar{y}$ in the breadth-first order. This means that either $|\bar{w}| < |\bar{y}|$, or when $\bar{w} = (0, w_1, \ldots, w_q)$ and $\bar{y} = (0, v_1, \ldots, v_q)$, $w_j < v_j$ for $j = \min\{l : v_l \neq w_l\}$. If $\bar{w}$ is either smaller than or equal to $\bar{y}$ in the breadth-first order, then we write $\bar{w} \leqslant_{UH} \bar{y}$.

As examples, we have $(0, 2, 3) <_{UH} (0, 1, 1, 1)$ and $(0, 3, 1, 5) <_{UH} (0, 3, 4, 2)$. We run a breadth-first search on the $(\mathbf{x}, n)$-Pólya urn tree G_n as follows.

Definition 6. (*Breadth-first search (BFS).*) A BFS of G_n splits the vertex set $V(G_n) = [n]$ into the random subsets $(\mathcal{A}_t, \mathcal{P}_t, \mathcal{N}_t)_{t\geqslant 0}$ as follows, where the letters respectively stand for *active*, *probed*, and *neutral*. We initialise the search with

$$(\mathcal{A}_0, \mathcal{P}_0, \mathcal{N}_0) = (\{k_0\}, \varnothing, V(G_n) \setminus \{k_0\}),$$

where k_0 is uniform in $[n]$. Given $(\mathcal{A}_{t-1}, \mathcal{P}_{t-1}, \mathcal{N}_{t-1})$, where each vertex in $\mathcal{A}_{t-1} \cup \mathcal{P}_{t-1}$ already receives an additional Ulam–Harris label of the form $\bar{w}$, $(\mathcal{A}_t, \mathcal{P}_t, \mathcal{N}_t)$ is generated as follows. Let $v[1] = k_0$, and let $v[t] \in \mathbb{N}$ be the vertex in $\mathcal{A}_{t-1}$ that is the smallest in the breadth-first order:

$$v[t] = k_{\bar{y}} \in \mathcal{A}_{t-1} \quad \text{if } \bar{y} <_{UH} \bar{w} \text{ for all } k_{\bar{w}} \in \mathcal{A}_{t-1} \setminus \{k_{\bar{y}}\}. \tag{3.1}$$

Note that the Ulam–Harris labels appear as subscripts. Denote by $\mathcal{D}_t$ the set of vertices in $\mathcal{N}_{t-1}$ that either receive an incoming edge from $v[t]$ or send an outgoing edge to $v[t]$:

$$\mathcal{D}_t := \{j \in \mathcal{N}_{t-1} : \{j, v[t]\} \text{ or } \{v[t], j\} \in E(G_n)\},$$

where $\{i, j\}$ is the edge directed from vertex j to vertex $i < j$. Then in the tth exploration step, we probe vertex $v[t]$ by marking the neutral vertices attached to $v[t]$ as active. That is,

$$(\mathcal{A}_t, \mathcal{P}_t, \mathcal{N}_t) = (\mathcal{A}_{t-1} \setminus \{v[t]\} \cup \mathcal{D}_t, \mathcal{P}_{t-1} \cup \{v[t]\}, \mathcal{N}_{t-1} \setminus \mathcal{D}_t), \tag{3.2}$$

and if $v[t] = k_{\bar{y}}$, we use the Ulam–Harris scheme to label the newly active vertices as $k_{\bar{y},j} := k_{(\bar{y},j)}$, $j \in \mathbb{N}$, in increasing order of their *vertex arrival times*, so that $k_{\bar{y},i} < k_{\bar{y},j}$ for any $i < j$. If $\mathcal{A}_{t-1} = \varnothing$, we set $(\mathcal{A}_t, \mathcal{P}_t, \mathcal{N}_t) = (\mathcal{A}_{t-1}, \mathcal{P}_{t-1}, \mathcal{N}_{t-1})$.

The characterisation of the BFS above is standard, and more details can be found in works such as [19, Chapter 4] and [23]. The vertex labelling $v[t]$ is very useful for the construction here, as the offspring distribution of $v[t]$ depends on the vertex partition $(\mathcal{A}_{t-1}, \mathcal{P}_{t-1}, \mathcal{N}_{t-1})$, whereas the vertex arrival times are helpful for identifying a type-L vertex in the BFS (see Lemma 4 below) and constructing the offspring distributions. On the other hand, we shall use the Ulam–Harris labels to match the vertices when we couple G_n and in the π-Pólya point tree (Definition 3). Hereafter we ignore the possibility that $v[t] = 1$, because when t (or equivalently the number of discovered vertices) is not too large, the probability that $v[t] = o(n)$ tends to zero as $n \to \infty$. For $t \geqslant 1$, let $v^{(op)}[t]$ (resp. $v^{(oa)}[t]$) be the vertex in $\mathcal{P}_{t-1}$ (resp. $\mathcal{A}_{t-1}$) that has the earliest arrival time, where *op* and *oa* stand for *oldest probed* and *oldest active*. That is,

$$v^{(op)}[t] := \min\{j : j \in \mathcal{P}_{t-1}\} \quad \text{and} \quad v^{(oa)}[t] := \min\{j : j \in \mathcal{A}_{t-1}\}. \tag{3.3}$$

3.2. Construction of the offspring distributions

To distinguish the urn tree G_n from the other models, we apply Ulam–Harris labels of the form $k_{\bar{y}}^{(\mathrm{u})}$ to its vertices; the superscript $^{(\mathrm{u})}$ stands for 'urn'. An example is given in Figure 4. Define

$$R_{\bar{y}}^{(\mathrm{u})} = \text{the number of type-R children of } k_{\bar{y}}^{(\mathrm{u})} \text{ in the } (\mathbf{x}, n)\text{-Pólya urn tree}, \tag{3.4}$$

noting that $R_0^{(\mathrm{u})}$ is the in-degree of the uniformly chosen vertex $k_0^{(\mathrm{u})}$ in G_n. Before we proceed further, we prove a simple lemma to help us identify when $v[t]$ in (3.1) is a type-L vertex, which will be useful for constructing the offspring distributions later. The result can be understood as a consequence of the facts that in the tree setting, we uncover a new active type-L vertex immediately after we probe an active type-L vertex, and that the oldest probed vertex cannot be rediscovered as a child of another type-R vertex in the subsequent explorations.

Lemma 4. *Assume that $\mathcal{A}_{t-1} \cup \mathcal{P}_{t-1}$ does not contain vertex 1. If $t = 2$, then $v^{(op)}[2] = k_0^{(\mathrm{u})}$ and $v^{(oa)}[2] = k_{0,1}^{(\mathrm{u})}$, while if $2 < i \leqslant t$, then $v^{(op)}[i]$ and $v^{(oa)}[i]$ are type-L children, where $v^{(oa)}[i]$ is the only type-L child in $\mathcal{A}_{i-1}$, and it receives an incoming edge from $v^{(op)}[i]$.*

Proof. We prove the lemma by induction on $2 \leqslant i \leqslant t$. The base case is clear, since $\mathcal{P}_1 = \{k_0^{(\mathrm{u})}\}$ and $\mathcal{A}_1$ consists of its type-L and type-R children. Assume that the lemma holds for some $2 \leqslant i < t$. If we probe a type-L child at time i, then $v[i] = v^{(oa)}[i]$, and there is a vertex $u \in \mathcal{N}_{i-1}$ that receives the incoming edge emanating from $v[i]$. Hence, vertex u belongs to type L and $v^{(oa)}[i+1] = u$. Furthermore, $v^{(op)}[i+1] = v[i]$, as $v^{(op)}[i]$ sends an outgoing edge to $v[i]$ by assumption, implying $v[i] < v^{(op)}[i]$. If we probe a type-R vertex at time i, then $v[i] > v^{(oa)}[i]$ and we uncover vertices in $\mathcal{N}_{i-1}$ that have later arrival times than $v[i]$. Setting $\mathcal{P}_i = \mathcal{P}_{i-1} \cup \{v[i]\}$, we have $v^{(op)}[i+1] = v^{(op)}[i]$ and $v^{(oa)}[i+1] = v^{(oa)}[i]$, which are of type L. □

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

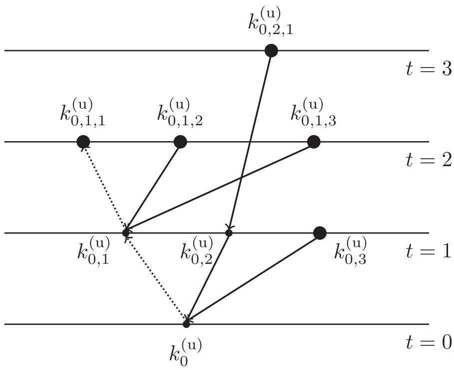

FIGURE 4. Each level corresponds to each time step of the BFS ($\mathbf{x}$, n)-Pólya urn tree G_n. The vertices are arranged from left to right in increasing order of their arrival times. The small and large dots correspond to the probed and active vertices, respectively. The densely dotted path joins the uniformly chosen vertex and the discovered type-L vertices. Here, $\mathcal{P}_3 = \left\{k^{(u)}_0, k^{(u)}_{0,1}, k^{(u)}_{0,2}\right\}$, $\mathcal{A}_3 = \left\{k^{(u)}_{0,3}, k^{(u)}_{0,1,1}, k^{(u)}_{0,1,2}, k^{(u)}_{0,1,3}, k^{(u)}_{0,2,1}\right\}$, $v[4] = k^{(u)}_{0,3}$, $v^{(op)}[4] = k^{(u)}_{0,1}$, and $v^{(oa)}[4] = k^{(u)}_{0,1,1}$.

To construct the offspring distributions for each parent in the local neighbourhood, we also need to define some notation and variables. Let $\mathcal{E}_0 = \varnothing$, and for $t \geqslant 1$, let $\mathcal{E}_t$ be the set of edges connecting the vertices in $\mathcal{A}_t \cup \mathcal{P}_t$. Given a positive integer m, denote the set of vertices and edges in $\mathcal{P}_t$ and $\mathcal{E}_t$ whose arrival time in G_n is earlier than that of vertex m by

$$\mathcal{P}_{t,m} = \mathcal{P}_t \cap [m-1] \quad \text{and} \quad \mathcal{E}_{t,m} = \{\{h, i\} \in \mathcal{E}_t : i < m\}, \tag{3.5}$$

noting that $\{h, i\}$ is the edge directed from i to $h < i$, and $(\mathcal{A}_0, \mathcal{P}_0, \mathcal{N}_0) = (\{k_0\}, \varnothing, V(G_n) \setminus \{k_0\})$ and $\mathcal{E}_0 = \varnothing$. Below we use $[t]$ in the notation to indicate the exploration step, and omit $\mathbf{x}$ for simplicity. Let $\left(\left(\mathcal{Z}_i[t], \widetilde{\mathcal{Z}}_i[t]\right), 2 \leqslant i \leqslant n, i \notin \mathcal{P}_{t-1}\right)$ be independent variables, where

$$\mathcal{Z}_i[1] \sim \text{Gamma}(x_i, 1) \quad \text{and} \quad \widetilde{\mathcal{Z}}_i[1] \sim \text{Gamma}(T_{i-1} + i - 1, 1), \tag{3.6}$$

with $T_i := \sum_{j=1}^{i} x_j$. Now, suppose that $t \geqslant 2$. Because of the size-bias effect of the edge $\{v^{(op)}[t], v^{(oa)}[t]\} \in \mathcal{E}_{t-1}$, we define

$$\mathcal{Z}_i[t] \sim \text{Gamma}\left(x_i + \mathbb{1}\left[i = v^{(oa)}[t]\right], 1\right), \quad i \in \mathcal{A}_{t-1} \cup \mathcal{N}_{t-1}, \tag{3.7}$$

so that the initial attractiveness of the type-L child $v^{(oa)}[t]$ is $x_{v^{(oa)}[t]} + 1$. The shape parameter of $\widetilde{\mathcal{Z}}_i[t]$ defined below is the total weight of the vertices in $\mathcal{A}_{t-1} \cup \mathcal{N}_{t-1}$ whose arrival time is earlier than the ith time step. This is because when a new vertex is added to the ($\mathbf{x}$, n)-sequential model conditional on having the edges $\mathcal{E}_{t-1}$, the recipient of its outgoing edge cannot be a vertex in $\mathcal{P}_{t-1}$, and it is chosen with probability proportional to the current weights of the vertices in $\mathcal{A}_{t-1} \cup \mathcal{N}_{t-1}$ that arrive before the new vertex. Hence, define

$$\widetilde{\mathcal{Z}}_i[t] \sim \begin{cases} \text{Gamma}(T_{i-1} + i - 1, 1), & \text{if } 2 \leqslant i \leqslant v^{(oa)}[t]; \\ \text{Gamma}(T_{i-1} + i, 1), & \text{if } v^{(oa)}[t] < i < v^{(op)}[t]; \\ \text{Gamma}\left(T_{i-1} + i - \sum_{h \in \mathcal{P}_{t-1,i}} x_h - |\mathcal{E}_{t-1,i}|, 1\right), & \text{if } v^{(op)}[t] < i \leqslant n. \end{cases} \tag{3.8}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

Let $B_1[t] := 1$ and $B_i[t] := 0$ for $i \in \mathcal{P}_{t-1}$, as the edges attached to $\mathcal{P}_{t-1}$ are already determined. Furthermore, define

$$B_i[t] := \frac{\mathcal{Z}_i[t]}{\mathcal{Z}_i[t] + \widetilde{\mathcal{Z}}_i[t]} \quad \text{for } i \in \mathcal{A}_{t-1} \cup \mathcal{N}_{t-1}. \tag{3.9}$$

Let $S_{n,0}[t] := 0$, $S_{n,n}[t] := 1$, and

$$S_{n,i}[t] := \prod_{j=i+1}^{n} \big(1 - B_j[t]\big) = \prod_{j=i+1; j \notin \mathcal{P}_{t-1}}^{n} \big(1 - B_j[t]\big) \quad \text{for } 1 \leqslant i \leqslant n-1, \tag{3.10}$$

where the second equality is true because $B_i[t] = 0$ for $i \in \mathcal{P}_{t-1}$. Observe that by the beta–gamma algebra, $(B_i[1], 1 \leqslant i \leqslant n) =_d (B_i, 1 \leqslant i \leqslant n)$ and $(S_{n,i}[1], 1 \leqslant i \leqslant n) =_d (S_{n,i}, 1 \leqslant i \leqslant n)$, where B_i and $S_{n,i}$ are as in (1.8) and (1.9). For $t \geqslant 2$, $(B_i[t], 1 \leqslant i \leqslant n)$ are the beta variables in the urn representation of the $(\mathbf{x}, n)$-sequential model, conditional on having the edges $\mathcal{E}_{t-1}$; see [24, Section 10] for details. We construct a Bernoulli sequence that encodes the type-R children of vertex $v[t]$ as follows. For $j \in \mathcal{N}_{t-1}$ and $v[t] + 1 \leqslant j \leqslant n$, let $\mathbb{1}_{\mathrm{R}}[j \to v[t]]$ be an indicator variable that takes value one if and only if vertex j sends an outgoing edge to $v[t]$; for $j \notin \mathcal{N}_{t-1}$, let $\mathbb{1}_{\mathrm{R}}[j \to v[t]] = 0$ with probability one, since the recipient of the incoming edge from vertex j is already in $\mathcal{P}_{t-1}$. Note that if $v[t] = k_{\bar{y}}^{(\mathrm{u})}$, then $R_{\bar{y}}^{(\mathrm{u})}$ in (3.4) is equal to $\sum_{j=v[t]+1}^{n} \mathbb{1}_{\mathrm{R}}[j \to v[t]]$. We also assume $\mathcal{N}_{t-1} \neq \varnothing$, because for large n, w.h.p. the local neighbourhood of vertex k_0 does not contain all the vertices of G_n. To state the distribution of $(\mathbb{1}_{\mathrm{R}}[j \to v[t]], v[t] + 1 \leqslant j \leqslant n)$, we use (3.9) and (3.10) to define

$$P_{j \to v[t]} := \begin{cases} \frac{S_{n,v[t]}[t]}{S_{n,j-1}[t]} B_{v[t]}[t], & \text{if } j \in \{v[t]+1, \ldots, n\} \cap \mathcal{N}_{t-1}; \\ 0, & \text{if } j \in \{v[t]+1, \ldots, n\} \setminus \mathcal{N}_{t-1}. \end{cases} \tag{3.11}$$

Definition 7. Given $(\mathcal{A}_{t-1}, \mathcal{P}_{t-1}, \mathcal{N}_{t-1})$ and $(B_j[t], v[t] \leqslant j \leqslant n)$, let $Y_{j \to v[t]}$, $v[t] + 1 \leqslant j \leqslant n$ be conditionally independent Bernoulli variables, each with parameter $P_{j \to v[t]}$. Define this Bernoulli sequence by the random vector

$$\mathbf{Y}_{\mathrm{Be}}^{(v[t],n)} := \big(Y_{(v[t]+1) \to v[t]}, Y_{(v[t]+2) \to v[t]}, \ldots, Y_{n \to v[t]}\big).$$

With the preparations above, we are ready to state the main results of this section. For $t \geqslant 2$, the following lemmas are immediate consequences of the urn representation of the $(\mathbf{x}, n)$-sequential model conditional on the discovered edges; see [24, Section 10]. For $t = 1$, they follow directly from Theorem 3. The first lemma states that we can encode the type-R children of the uniformly chosen vertex (the root) or a non-root parent in the local neighbourhood in a Bernoulli sequence; see Figure 3 in the case of the root.

Lemma 5. *Assume that $\mathcal{N}_{t-1} \neq \varnothing$ and $\mathcal{A}_{t-1} \cup \mathcal{P}_{t-1}$ does not contain vertex 1. Then given $\big(\mathcal{A}_{t-1}, \mathcal{P}_{t-1}, \mathcal{N}_{t-1}\big)$, the random vector $\big(\mathbb{1}_{\mathrm{R}}[j \to v[t]], v[t] + 1 \leqslant j \leqslant n\big)$ is distributed as $\mathbf{Y}_{\mathrm{Be}}^{(v[t],n)}$.*

Remark 3. When $|\mathcal{E}_{t-1}| = o(n)$ for some $t \geqslant 2$, $P_{j \to v[t]}$ in (3.11) is approximately distributed as $\big(S_{n,v[t]}/S_{n,j-1}\big)B_{v[t]}$ for n sufficiently large, with B_j and $S_{n,j}$ as in (1.8) and (1.9). As we shall see later in the proof, $|\mathcal{E}_{t-1}| = o(n)$ indeed occurs w.h.p.

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

The next lemma states that when $v[t]$ is the uniformly chosen vertex or a type-L parent, we can use the beta variables in (3.9) to obtain the distribution of the type-L child of $v[t]$. Observe that $(B_j[t], 2 \leqslant j \leqslant v[t]-1)$ does not appear in Definition 7, but is required for this purpose.

Lemma 6. *Assume that $\mathcal{N}_{t-1} \neq \varnothing$, $\mathcal{A}_{t-1} \cup \mathcal{P}_{t-1}$ does not contain vertex 1, and v[t] is either the uniformly chosen vertex or of type L. Given $(\mathcal{A}_{t-1}, \mathcal{P}_{t-1}, \mathcal{N}_{t-1})$, let $S_{n,j}[t]$ be as in (3.10), and $U \sim \mathrm{U}[0, S_{n,v[t]-1}[t]]$. For $1 \leqslant j \leqslant v[t]-1$, the probability that vertex j receives the only incoming edge from v[t] is given by the probability that $S_{n,j-1}[t] \leqslant U < S_{n,j}[t]$.*

4. An intermediate tree for graph couplings

As the π-Pólya point tree $(\mathcal{T}, 0)$ in Definition 3 does not have vertex labels that are comparable to the vertex arrival times of the $(\mathbf{x}, n)$-Pólya urn tree G_n, we define a suitable conditional analogue of the π-Pólya point tree. We call this analogue the *intermediate Pólya point tree* $(\mathcal{T}_{\mathbf{x},n}, 0)$, where vertex 0 is its root, and the subscripts are the parameters corresponding to $\mathbf{x}$ and n of G_n. All the variables of $(\mathcal{T}_{\mathbf{x},n}, 0)$ have the superscript $^{(\mathrm{i})}$ (for 'intermediate'). Each vertex of $(\mathcal{T}_{\mathbf{x},n}, 0)$ has an Ulam–Harris label $\bar{v}$, an age $a^{(\mathrm{i})}_{\bar{v}}$, and a type (except for the root), which are defined similarly as in Section 1.2 for $(\mathcal{T}, 0)$. Moreover, vertex $\bar{v}$ has an additional *PA label* $k^{(\mathrm{i})}_{\bar{v}}$, which determines its initial attractiveness by taking $x_{k^{(\mathrm{i})}_{\bar{v}}}$. The distributions of the PA labels are constructed using gamma variables similar to (3.6), (3.7), and (3.8), and as we remark after the definition, the $k^{(\mathrm{i})}_{\bar{v}}$ are approximately distributed as the vertices (or equivalently the arrival times) $k^{(\mathrm{u})}_{\bar{v}}$ in the local neighbourhood in G_n. Define

$$R^{(\mathrm{i})}_{\bar{v}} = \text{the number of type-R children of vertex } \bar{v} \text{ in } (\mathcal{T}_{\mathbf{x},n}, 0), \tag{4.1}$$

which is analogous to $R_{\bar{v}}$ in (1.3) and $R^{(\mathrm{u})}_{\bar{v}}$ in (3.4). Finally, recall χ in (1.2) and $\bar{w} \leqslant_{UH} \bar{v}$ whenever the Ulam–Harris label $\bar{w}$ is smaller than $\bar{v}$ in the breadth-first order (Definition 5).

Definition 8. (*Intermediate Pólya point tree.*) Given n and $\mathbf{x}$, $(\mathcal{T}_{\mathbf{x},n}, 0)$ is constructed recursively as follows. The root 0 has an age $a^{(\mathrm{i})}_0 = U_0^{\chi}$ and an initial attractiveness $x_{k^{(\mathrm{i})}_0}$, where $U_0 \sim \mathrm{U}[0, 1]$ and $k^{(\mathrm{i})}_0 = \lceil nU_0 \rceil$. Assume that $\left(\left(a^{(\mathrm{i})}_{\bar{w}}, k^{(\mathrm{i})}_{\bar{w}}\right), \bar{w} \leqslant_{UH} \bar{v}\right)$ and $\left(\left(R^{(\mathrm{i})}_{\bar{w}}, a^{(\mathrm{i})}_{\bar{w},j}, k^{(\mathrm{i})}_{\bar{w},j}\right),\right.$ $\left.\bar{w} <_{UH} \bar{v}\right)$ have been generated, so that $k^{(\mathrm{i})}_{\bar{w}} > 1$ for $\bar{w} \leqslant_{UH} \bar{v}$. If vertex $\bar{v}$ is the root or belongs to type L, we generate $\left(\left(a^{(\mathrm{i})}_{\bar{v},j}, k^{(\mathrm{i})}_{\bar{v},j}\right), 1 \leqslant j \leqslant 1 + R^{(\mathrm{i})}_{\bar{v}}\right)$ as follows:

1. We sample the age of its type-L child $(\bar{v}, 1)$ by letting $U_{\bar{v},1} \sim \mathrm{U}[0, 1]$ and $a^{(\mathrm{i})}_{\bar{v},1} = a^{(\mathrm{i})}_{\bar{v}} U_{\bar{v},1}$.

2. Next we choose the PA label $k^{(\mathrm{i})}_{\bar{v},1}$. Suppose that $t = |\{\bar{w} : \bar{w} \leqslant_{UH} \bar{v}\}|$, so that $t = 1$ if $\bar{v} = 0$. We define the conditionally independent variables $\left(\left(\mathcal{Z}^{(\mathrm{i})}_j[t], \widetilde{\mathcal{Z}}^{(\mathrm{i})}_j[t]\right), 2 \leqslant j \leqslant n, j \notin \left\{k^{(\mathrm{i})}_{\bar{w}} : \bar{w} <_{UH} \bar{v}\right\}\right)$ as follows. Let $T_j := \sum_{\ell=1}^{j} x_\ell$. When $\bar{v} = 0$, let

$$\mathcal{Z}^{(\mathrm{i})}_j[1] \sim \mathrm{Gamma}(x_j, 1) \quad \text{and} \quad \widetilde{\mathcal{Z}}^{(\mathrm{i})}_j[1] \sim \mathrm{Gamma}(T_{j-1} + j - 1, 1);$$

when $\bar{v} = (0, 1, 1, \ldots, 1)$, let $\mathcal{Z}^{(\mathrm{i})}_j[t] \sim \mathrm{Gamma}\big(x_j + \mathbb{1}[j = k^{(\mathrm{i})}_{\bar{v}}], 1\big)$, and

$$\widetilde{\mathcal{Z}}^{(\mathrm{i})}_j[t] \sim \begin{cases} \mathrm{Gamma}(T_{j-1} + j - 1, 1), & 2 \leqslant j \leqslant k^{(\mathrm{i})}_{\bar{v}}, \\ \mathrm{Gamma}(T_{j-1} + j, 1), & k^{(\mathrm{i})}_{\bar{v}} < j < k^{(\mathrm{i})}_{\bar{v}'}, \\ \mathrm{Gamma}\big(T_{j-1} + j + 1 - |\bar{v}| - W_{\bar{v},j}, 1\big), & k^{(\mathrm{i})}_{\bar{v}'} < j \leqslant n, \end{cases}$$

where

$$W_{\bar{v},j} := \sum_{\left\{k^{(\mathrm{i})}_{\bar{w}} < j : \bar{w} <_{UH} \bar{v}\right\}} \left\{x_{k^{(\mathrm{i})}_{\bar{w}}} + R^{(\mathrm{i})}_{\bar{w}}\right\},$$

and $\bar{v}' = (0, 1, 1, \ldots, 1)$ with $|\bar{v}'| = |\bar{v}| - 1$. For either the root or the type-L parent, define $B^{(\mathrm{i})}_1[t] := 1$, $B^{(\mathrm{i})}_j[t] := 0$ for $j \in \left\{k^{(\mathrm{i})}_{\bar{w}} : \bar{w} <_{UH} \bar{v}\right\}$ and

$$B^{(\mathrm{i})}_j[t] := \frac{\mathcal{Z}^{(\mathrm{i})}_j[t]}{\mathcal{Z}^{(\mathrm{i})}_j[t] + \widetilde{\mathcal{Z}}^{(\mathrm{i})}_j[t]} \quad \text{for } j \in [n] \setminus \left\{k^{(\mathrm{i})}_{\bar{w}} : \bar{w} <_{UH} \bar{v}\right\}; \tag{4.2}$$

then let $S^{(\mathrm{i})}_{n,0}[t] := 0$, $S^{(\mathrm{i})}_{n,n}[t] := 1$, and $S^{(\mathrm{i})}_{n,j}[t] := \prod_{\ell=j+1}^{n} \left(1 - B^{(\mathrm{i})}_\ell[t]\right)$ for $1 \leqslant j \leqslant n-1$. Momentarily define $b := k^{(\mathrm{i})}_{\bar{v}}$ and $c := k^{(\mathrm{i})}_{\bar{v},1}$; we choose the PA label c so that $S^{(\mathrm{i})}_{n,c-1}[t] \leqslant U_{\bar{v},1} S^{(\mathrm{i})}_{n,b-1}[t] < S^{(\mathrm{i})}_{n,c}[t]$.

3. We generate the ages and the PA labels of the type-R children. Let $\left(a^{(\mathrm{i})}_{\bar{v},j}, 2 \leqslant j \leqslant 1 + R^{(\mathrm{i})}_{\bar{v}}\right)$ be the points of a mixed Poisson process on $\left(a^{(\mathrm{i})}_{\bar{v}}, 1\right]$ with intensity

$$\lambda^{(\mathrm{i})}_{\bar{v}}(y)dy := \frac{\mathcal{Z}^{(\mathrm{i})}_b[t]}{\mu\left(a^{(\mathrm{i})}_{\bar{v}}\right)^{1/\mu}} y^{1/\mu - 1} dy. \tag{4.3}$$

Then choose $k^{(\mathrm{i})}_{\bar{v},j}$, $2 \leqslant j \leqslant 1 + R^{(\mathrm{i})}_{\bar{v}}$, such that

$$\left(\left(k^{(\mathrm{i})}_{\bar{v},j} - 1\right)/n\right)^{\chi} < a^{(\mathrm{i})}_{\bar{v},j} \leqslant \left(k^{(\mathrm{i})}_{\bar{v},j}/n\right)^{\chi}. \tag{4.4}$$

If $\bar{v}$ is of type R, let $\mathcal{Z}_b[t] \sim \mathrm{Gamma}(x_b, 1)$ and apply Step 3 only to obtain $\left(\left(a^{(\mathrm{i})}_{\bar{v},j}, k^{(\mathrm{i})}_{\bar{v},j}\right), 1 \leqslant j \leqslant R^{(\mathrm{i})}_{\bar{v}}\right)$. We build $(\mathcal{T}_{\mathbf{x},n}, 0)$ by iterating this process, and terminate the construction whenever there is some vertex $\bar{v}$ such that $b = 1$.

We now discuss the distributions of the PA labels above. Clearly, $k^{(\mathrm{i})}_0$ is uniform in $[n]$. Observe that t above is essentially the number of breadth-first exploration steps in $(\mathcal{T}_{\mathbf{x},n}, 0)$, starting from the root 0. Suppose that we have completed $t-1$ exploration steps in G_n, starting from its uniformly chosen vertex $k^{(\mathrm{u})}_0$, and the resulting neighbourhood is coupled to that of vertex $0 \in V((\mathcal{T}_{\mathbf{x},n}, 0))$ so that they are isomorphic and $k^{(\mathrm{i})}_{\bar{w}} = k^{(\mathrm{u})}_{\bar{w}}$ for all $\bar{w}$ in the neighbourhoods. If $v[t] = k^{(\mathrm{u})}_{\bar{v}}$, a moment's thought shows that the total weight of the vertices $\left\{k^{(\mathrm{i})}_{\bar{w}} < j : \bar{w} <_{UH} \bar{v}\right\}$ and the vertex set $\mathcal{P}_{t-1,j}$ satisfy

$$W_{\bar{v},j} + |\bar{v}| - 1 = \sum_{h \in \mathcal{P}_{t-1,j}} x_h + |\mathcal{E}_{t-1,j}|,$$

where $\mathcal{P}_{t-1,j}$ and $\mathcal{E}_{t-1,j}$ are as in (3.5). Hence, recalling the gammas in (3.7) and (3.8), it follows that $\mathcal{Z}_j[t] =_d \mathcal{Z}^{(\mathrm{i})}_j[t]$ and $\widetilde{\mathcal{Z}}_j[t] =_d \widetilde{\mathcal{Z}}^{(\mathrm{i})}_j[t]$ for any j. So by Lemma 6, $k^{(\mathrm{i})}_{\bar{v}} =_d k^{(\mathrm{u})}_{\bar{v}}$ if $\bar{v}$ is a type-L child. However, $k^{(\mathrm{i})}_{\bar{v}}$ is only approximately distributed as $k^{(\mathrm{u})}_{\bar{v}}$ if $\bar{v}$ is of type R. The discretised Poisson point process that generates these PA labels will be coupled to the Bernoulli sequence in Lemma 5 that encodes the type-R children of $k_{\bar{v}}^{(\mathrm{u})}$. When $k_{\bar{v}}^{(\mathrm{i})}=1$, it must be the case that $\bar{v}=0$ or $\bar{v}=(0, 1, \ldots, 1)$. We stop the construction in this case to avoid an ill-defined $\mathcal{Z}_1^{(\mathrm{i})}[t]$ when $-1<x_1\leqslant 0$ and $k_0^{(\mathrm{i})}=1$, and also because vertex 1 cannot have a type-L neighbour, so Steps 1 and 2 are unnecessary in this case. Nevertheless, for $r<\infty$ and any vertex $\bar{v}$ in the r-neighbourhood of vertex $0\in V((\mathcal{T}_{\mathbf{x},n}, 0))$, the probability that $k_{\bar{v}}^{(\mathrm{i})}=1$ tends to zero as $n\to\infty$.

5. Proof of the base case

Recall that, as in Section 3.2, $k_{\bar{v}}^{(\mathrm{u})}$ are the vertices in the local neighbourhood of the uniformly chosen vertex $k_0^{(\mathrm{u})}$ of the $(\mathbf{x}, n)$-Pólya urn tree G_n in Definition 4. Let χ be as in (1.2), and let $(\mathcal{T}_{\mathbf{x},n}, 0)$ be the intermediate Pólya point tree in Definition 8, with $k_{\bar{v}}^{(\mathrm{i})}$ and $a_{\bar{v}}^{(\mathrm{i})}$ being the PA label and the age of vertex $\bar{v}$ in the tree. Here, we couple $\big(G_n, k_0^{(\mathrm{u})}\big)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$ so that with probability tending to one, $(B_1\big(G_n, k_0^{(\mathrm{u})}\big), k_0^{(\mathrm{u})})\cong(B_1(\mathcal{T}_{\mathbf{x},n}, 0), 0)$, $\big(k_{0,j}^{(\mathrm{u})}/n\big)^{\chi}\approx a_{0,j}^{(\mathrm{i})}$, and $k_{0,j}^{(\mathrm{u})}=k_{0,j}^{(\mathrm{i})}$. For convenience, we also refer to the rescaled arrival time $(k/n)^{\chi}$ of vertex k in G_n simply as its *age*. Let $U_0\sim \mathrm{U}[0, 1]$, $a_0^{(\mathrm{i})}=U_0^{\chi}$, and $k_0^{(\mathrm{u})}=\lceil nU_0\rceil$, where $a_0^{(\mathrm{i})}$ is the age of vertex 0. By a direct comparison to Definition 8, $a_0^{(\mathrm{i})}\approx\big(k_0^{(\mathrm{u})}/n\big)^{\chi}$ and the initial attractiveness of vertex 0 is $x_{k_0^{(\mathrm{u})}}$. Under this coupling we define the event

$$\mathcal{H}_{1,0}=\big\{a_0^{(\mathrm{i})}>(\log\log n)^{-\chi}\big\},\tag{5.1}$$

which guarantees that we choose a vertex of low degree in G_n. To prepare for the coupling, let $\big(\big(\mathcal{Z}_j[1], \widetilde{\mathcal{Z}}_j[1]\big), 2\leqslant j\leqslant n\big)$ and $\big(S_{n,j}[1], 1\leqslant j\leqslant n\big)$ be as in (3.6) and (3.10). We use $\big(S_{n,j}[1], 1\leqslant j\leqslant n\big)$ to construct the distribution of the type-R children of $k_0^{(\mathrm{u})}$, and for sampling the initial attractiveness of vertex $(0, 1)\in V((\mathcal{T}_{\mathbf{x},n}, 0))$. Furthermore, let the ages $\Big(a_{0,j}^{(\mathrm{i})}, 1\leqslant j\leqslant 1+R_0^{(\mathrm{i})}\Big)$ and $R_0^{(\mathrm{u})}$ be as in Definition 8 and (3.4). Below we define a coupling of $\big(G_n, k_0^{(\mathrm{u})}\big)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$, and for this coupling we define the events that the children of $k_0^{(\mathrm{u})}$ have low degrees, the 1-neighbourhoods are isomorphic and of size at most $(\log n)^{1/r}$, and the ages are close enough as

$$\begin{aligned}\mathcal{H}_{1,1}&=\Big\{a_{0,1}^{(\mathrm{i})}>(\log\log n)^{-2\chi}\Big\},\\ \mathcal{H}_{1,2}&=\left\{R_0^{(\mathrm{u})}=R_0^{(\mathrm{i})},\ \text{for } \bar{v}\in V(B_1(\mathcal{T}_{\mathbf{x},n}, 0)), k_{\bar{v}}^{(\mathrm{u})}=k_{\bar{v}}^{(\mathrm{i})}, \left|a_{\bar{v}}^{(\mathrm{i})}-\left(\frac{k_{\bar{v}}^{(\mathrm{u})}}{n}\right)^{\chi}\right|\leqslant C_1 b(n)\right\},\\ \mathcal{H}_{1,3}&=\Big\{R_0^{(\mathrm{i})}<(\log n)^{1/r}\Big\},\end{aligned}\tag{5.2}$$

where $R_0^{(\mathrm{i})}$ and $R_0^{(\mathrm{u})}$ are as in (3.4) and (4.1), $b(n):=n^{-\frac{\chi}{12}}(\log\log n)^{\chi}$, and $C_1:=C_1(x_1, \mu)$ will be chosen in the proof of Lemma 8 below. Because the initial attractivenesses of the vertices $(0,j)$ and $k_{0,j}^{(\mathrm{u})}$ match and their ages are close enough on the event $\mathcal{H}_{1,2}$, we can couple the Bernoulli and the mixed Poisson sequences that encode their type-R children, and hence the 2-neighbourhoods. The event $\mathcal{H}_{1,1}$ ensures that the local neighbourhood of $k_0^{(\mathrm{u})}$ grows slowly, and on the event $\mathcal{H}_{1,3}$, the number of vertex pairs that we need to couple for the 2-neighbourhoods is not too large. So by a union bound argument, the probability that any of the subsequent couplings fail tends to zero as $n\to\infty$. The aim of this section is to show that when

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

$\sum_{i=2}^{j} x_i \approx (j-1)\mu$ for all j sufficiently large, we can couple the two graphs so that $\bigcap_{i=0}^{3} \mathcal{H}_{1,i}$ occurs w.h.p. Thus, let $A_n := A_{2/3,n}$ be as in (2.1) with $\alpha = 2/3$; this α is chosen to simplify the calculation, and can be chosen under the assumption of bounded fitness. Recall that $\mathbb{P}_{\mathbf{x}}$ indicates the conditioning on a specific realisation of the fitness sequence $\mathbf{x}$. The main result is the lemma below; it serves as the base case when we inductively prove an analogous result for the general radius $r < \infty$. Because $B_1(\mathcal{T}_{\mathbf{x},n}, 0)$ approximately distributes as the 1-neighbourhood of the π-Pólya point tree after randomisation of $\mathbf{x}$, and Lemma 1 says that $\mathbb{P}[A_n^c] = O\big(n^{-b}\big)$ for some $b > 0$, it follows from the lemma below and (1.4) that (1.5) of Theorem 1 holds for $r = 1$.

Lemma 7. *Assume that $x_j \in (0, \kappa]$ for $j \geqslant 2$ and $\mathbf{x} \in A_n$. Let $\mathcal{H}_{1,j}$, $j = 0, \ldots, 3$, be as in (5.1) and (5.2). Then there is a coupling of $\big(G_n, k_0^{(\mathrm{u})}\big)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$, with a positive constant $C := C(x_1, \mu, \kappa)$ such that*

$$\mathbb{P}_{\mathbf{x}}\left[\left(\bigcap_{j=0}^{3} \mathcal{H}_{1,j}\right)^c\right] \leqslant C(\log\log n)^{-\chi} \quad \textit{for all } n \geqslant 3. \tag{5.3}$$

The proof of Lemma 7 consists of several lemmas, which we now develop. From the definitions of $a_0^{(\mathrm{i})}$ and $a_{0,1}^{(\mathrm{i})}$, it is obvious that the probabilities of the events $\mathcal{H}_{1,0}$ and $\mathcal{H}_{1,1}$ tend to one as $n \to \infty$, and by Chebyshev's inequality, we can show that this is also true for the event $\mathcal{H}_{1,3}$. We take care of $\mathcal{H}_{1,2}$ in the lemma below, whose proof is the core of this section.

Lemma 8. *Retaining the assumptions and the notation of Lemma 7, there is a coupling of $\big(G_n, k_0^{(\mathrm{u})}\big)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$, with a positive constant $C := C(x_1, \mu, \kappa)$ such that*

$$\mathbb{P}_{\mathbf{x}}\left[\mathcal{H}_{1,0} \cap \mathcal{H}_{1,1} \cap \mathcal{H}_{1,2}^c\right] \leqslant Cn^{-\beta}(\log\log n)^{1-\chi} \quad \textit{for all } n \geqslant 3,$$

where $0 < \gamma < \chi/12$ and $\beta = \min\{\chi/3 - 4\gamma, \gamma\}$.

Given that the vertices $k_0^{(\mathrm{u})} \in V\big(\big(G_n, k_0^{(\mathrm{u})}\big)\big)$ and $0 \in V((\mathcal{T}_{\mathbf{x},n}, 0))$ are coupled, we prove Lemma 8 as follows. Note that $v[1] = k_0^{(\mathrm{u})}$, where $v[t]$ is as in (3.1). We first couple $\mathbf{Y}_{\mathrm{Be}}^{(v[1],n)}$ in Definition 7 and a discretisation of the mixed Poisson process that encodes the ages and the PA labels that are of type R. Then we couple the type-L vertices $k_{0,1}^{(\mathrm{u})}$ and $(0, 1) \in V((\mathcal{T}_{\mathbf{x},n}, 0))$ so that on the event $\mathcal{H}_{1,0} \cap \mathcal{H}_{1,1}$, we have $k_{0,1}^{(\mathrm{i})} = k_{0,1}^{(\mathrm{u})}$ and $\big(k_{0,1}^{(\mathrm{u})}/n\big)^{\chi} \approx a_{0,1}^{(\mathrm{i})}$ w.h.p. For the means of the discretised Poisson process, we define

$$\lambda_{v[1]+1}^{[1]} := \int_{a_0^{(\mathrm{i})}}^{\left(\frac{v[1]+1}{n}\right)^{\chi}} \frac{\mathcal{Z}_{v[1]}[1]}{\big(a_0^{(\mathrm{i})}\big)^{1/\mu}\mu} y^{1/\mu-1} dy \quad \text{and} \quad \lambda_j^{[1]} := \int_{\left(\frac{j-1}{n}\right)^{\chi}}^{\left(\frac{j}{n}\right)^{\chi}} \frac{\mathcal{Z}_{v[1]}[1]}{\big(a_0^{(\mathrm{i})}\big)^{1/\mu}\mu} y^{1/\mu-1} dy \tag{5.4}$$

for $v[1] + 2 \leqslant j \leqslant n$, where $\mathcal{Z}_{v[1]}[1]$ is the gamma variable in (3.6).

Definition 9. Given $v[1] = k_0^{(\mathrm{u})}$, $a_0^{(\mathrm{i})}$, and $\mathcal{Z}_{v[1]}[1]$, let $V_{j\to v[1]}$, $v[1] + 1 \leqslant j \leqslant n$, be conditionally independent Poisson random variables, each with parameter $\lambda_j^{[1]}$ as in (5.4). Define this Poisson sequence by the random vector

$$\mathbf{V}_{\mathrm{Po}}^{(v[1],n)} := \big(V_{(v[1]+1)\to v[1]}, V_{(v[1]+2)\to v[1]}, \ldots, V_{n\to v[1]}\big).$$

Next, we define the events that ensure that $P_{j\to v[1]}$ is close enough to $\lambda_j^{[1]}$. Let $\phi(n) = \Omega(n^{\chi})$, and let $C^\star := C^\star(x_1, \mu)$ be a positive constant such that (2.2) of Lemma 2 holds with $\delta_n =$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

$C^\star n^{-\frac{\chi}{12}}$. Let $\mathcal{Z}_j[1]$, $B_j[1]$, and $S_j[1]$ be as in (3.6), (3.9), and (3.10). Given $0<\gamma<\chi/12$, define the events

$$\begin{aligned} F_{1,1} &= \left\{ \max_{\lceil \phi(n)\rceil \leqslant j\leqslant n} \left| S_{n,j}[1] - \left(\frac{j}{n}\right)^{\chi} \right| \leqslant C^\star n^{-\frac{\chi}{12}} \right\}, \\ F_{1,2} &= \bigcap_{j=\lceil \phi(n)\rceil}^{n} \left\{ \left| B_j[1] - \frac{\mathcal{Z}_j[1]}{(\mu+1)j} \right| \leqslant \frac{\mathcal{Z}_j[1] n^{-\gamma}}{(\mu+1)j} \right\}, \\ F_{1,3} &= \bigcap_{j=\lceil \phi(n)\rceil}^{n} \left\{ \mathcal{Z}_j[1] \leqslant j^{1/2} \right\}. \end{aligned} \tag{5.5}$$

The next lemma is the major step towards proving Lemma 8, as it implies that we can couple the ages and the initial attractivenesses of the type-R children of the uniformly chosen vertex $k_0^{(\mathrm{u})} := v[1] \in V\big(G_n, k_0^{(\mathrm{u})}\big)$ and those of the root $0 \in V(\mathcal{T}_{\mathbf{x},n}, 0)$.

Lemma 9. *Retaining the assumptions and the notation in Lemma 8, let* $\mathbf{Y}_{\mathrm{Be}}^{(v[1],n)}$ *and* $\mathbf{V}_{\mathrm{Po}}^{(v[1],n)}$ *be as in Definitions 7 and 9, and let* $F_{1,i}$, $i=1,2,3$, *be as in (5.5). Then we can couple the random vectors so that there is a positive constant* $C := C(x_1, \mu, \kappa)$ *such that*

$$\mathbb{P}_{\mathbf{x}}\left[\left\{\mathbf{Y}_{\mathrm{Be}}^{(v[1],n)} \neq \mathbf{V}_{\mathrm{Po}}^{(v[1],n)}\right\} \cap \bigcap_{i=1}^{3} F_{1,i} \cap \mathcal{H}_{1,0}\right] \leqslant C n^{-\gamma} (\log\log n)^{1-\chi} \quad \textit{for all } n \geqslant 3.$$

Sketch of proof. We summarise the proof here and defer the details to [24]. On the event $F_{1,1} \cap F_{1,2} \cap \mathcal{H}_{1,0}$, a little calculation shows that the Bernoulli success probability $P_{j\to v[1]}$ in (3.11) is close enough to

$$\widehat{P}_{j\to v[1]} := \left(\frac{v[1]}{j}\right)^{\chi} \frac{\mathcal{Z}_{v[1]}[1]}{(\mu+1)v[1]}, \tag{5.6}$$

while the event $F_{1,3}$ ensures that $\widehat{P}_{j\to v[1]} \leqslant 1$. The Poisson mean $\lambda_j^{[1]} \approx \widehat{P}_{j\to v[1]}$ as $a_0^{(\mathrm{i})} \approx (v[1]/n)^{\chi}$. Hence, we use $\widehat{P}_{j\to v[1]}$ in (5.6) as means of constructing two intermediate Bernoulli and Poisson random vectors, and then explicitly couple the four processes. It is enough to consider the coupling under the event $\bigcap_{j=1}^{3} F_{1,j} \cap \mathcal{H}_{1,0}$, because when $\mathbf{x} \in A_n$, Lemmas 2 and 3 imply that $F_{1,1}$, $F_{1,2}$, and $F_{1,3}$ occur with probability tending to one as $n\to\infty$. □

Below we use Lemmas 2, 3, and 9 to prove Lemma 8.

Proof of Lemma 8. We bound the right-hand side of

$$\begin{aligned} &\mathbb{P}_{\mathbf{x}}\big[\mathcal{H}_{1,2}^c \cap \mathcal{H}_{1,1} \cap \mathcal{H}_{1,0}\big] \\ &\leqslant \mathbb{P}_{\mathbf{x}}\left[\bigcap_{j=1}^{3} F_{1,j} \cap \bigcap_{j=0}^{1} \mathcal{H}_{1,j} \cap \mathcal{H}_{1,2}^c\right] + \mathbb{P}_{\mathbf{x}}\left[\left(\bigcap_{j=1}^{3} F_{1,j}\right)^c\right] \\ &\leqslant \mathbb{P}_{\mathbf{x}}\left[\bigcap_{j=1}^{3} F_{1,j} \cap \bigcap_{j=0}^{1} \mathcal{H}_{1,j} \cap \mathcal{H}_{1,2}^c\right] + \mathbb{P}_{\mathbf{x}}\big[F_{1,3}^c\big] + \mathbb{P}_{\mathbf{x}}\big[F_{1,2}^c\big] + \mathbb{P}_{\mathbf{x}}\big[F_{1,1}^c\big], \end{aligned} \tag{5.7}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

under a suitable coupling of $\big(G_n, k_0^{(\mathrm{u})}\big)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$. We start by handling the last three terms. First, apply (2.6) of Lemma 3 to obtain

$$\mathbb{P}_{\mathbf{x}}\big[F_{1,3}^c\big] = \mathbb{P}_{\mathbf{x}}\left[\bigcup_{j=\phi(n)}^{n} \{\mathcal{Z}_j[1] \geqslant j^{1/2}\}\right] \leqslant C\kappa^4 n^{-\chi}, \tag{5.8}$$

where C is a positive constant. Since $\mathbf{x} \in A_n$, we can apply (2.4) of Lemma 3 (with $\varepsilon = n^{-\gamma}$, $0 < \gamma < \chi/12$, and $\alpha = 2/3$) and Lemma 2 to deduce that

$$\mathbb{P}_{\mathbf{x}}\big[F_{1,2}^c\big] = \mathbb{P}_{\mathbf{x}}\left[\bigcup_{j=\lceil\phi(n)\rceil}^{n} \left\{\left|B_j[1] - \frac{\mathcal{Z}_j[1]}{(\mu+1)j}\right| \geqslant \frac{\mathcal{Z}_j[1]n^{-\gamma}}{(\mu+1)j}\right\}\right] \leqslant Cn^{4\gamma-\frac{\chi}{3}},$$

$$\mathbb{P}_{\mathbf{x}}\big[F_{1,1}^c\big] \leqslant \mathbb{P}_{\mathbf{x}}\left[\max_{\lceil\phi(n)\rceil \leqslant k \leqslant n}\left|S_{n,k}[1] - \left(\frac{k}{n}\right)^{\chi}\right| \geqslant C^{\star} n^{-\frac{\chi}{12}}\right] \leqslant cn^{-\frac{\chi}{6}}, \tag{5.9}$$

where $C := C(x_1, \gamma, \mu)$ and $c := c(x_1, \mu)$. Note that $\mathbb{P}_{\mathbf{x}}\big[F_{1,2}^c\big] \to 0$ as $n \to \infty$ thanks to our choice of γ. Next, we give the appropriate coupling to bound the first probability of (5.7). Let the vertices $k_0^{(\mathrm{u})} \in V\big(G_n, k_0^{(\mathrm{u})}\big)$ and $0 \in V(\mathcal{T}_{\mathbf{x},n}, 0)$ be coupled as in the beginning of this section. Assume that they are such that the event $\mathcal{H}_{1,0}$ occurs, and the variables $\big(\big(\mathcal{Z}_j[1], \widetilde{\mathcal{Z}}_j[1]\big), 2 \leqslant j \leqslant n\big)$ are such that $\bigcap_{j=1}^{3} F_{1,j}$ holds. We first argue that under the event $\bigcap_{j=1}^{3} F_{1,j} \cap \mathcal{H}_{1,0}$, their type-R children can be coupled so that w.h.p., $R_0^{(\mathrm{u})} = R_0^{(\mathrm{i})}$ and $k_{0,j}^{(\mathrm{u})} = k_{0,j}^{(\mathrm{i})}$ for $j = 2, \ldots, 1 + R_0^{(\mathrm{i})}$. In view of Definitions 4 and 8, this follows readily from Lemma 9. It remains to prove that under this coupling, $\big|(k_0^{(\mathrm{u})}/n)^{\chi} - a_0^{(\mathrm{i})}\big|$ and $\big|a_{0,j}^{(\mathrm{i})} - (k_{0,j}^{(\mathrm{u})}/n)^{\chi}\big|$ are sufficiently small. Since $k_0^{(\mathrm{u})} = \left\lceil n\big(a_0^{(\mathrm{i})}\big)^{1/\chi}\right\rceil$, and for $j = 2, \ldots, 1 + R_0^{(\mathrm{i})}$, $k_{0,j}^{(\mathrm{u})}$ satisfies $k_{0,j}^{(\mathrm{u})} > k_0^{(\mathrm{u})}$ and $\Big(\big(k_{0,j}^{(\mathrm{u})} - 1\big)/n\Big)^{\chi} < a_{0,j}^{(\mathrm{i})} \leqslant \big(k_{0,j}^{(\mathrm{u})}/n\big)^{\chi}$, it is enough to bound $(\ell/n)^{\chi} - ((\ell-1)/n)^{\chi}$ for $k_0^{(\mathrm{u})} + 1 \leqslant \ell \leqslant n$. For such ℓ, we use the fact that $k_0^{(\mathrm{u})} > n(\log\log n)^{-1}$ on the event $\mathcal{H}_{1,0}$ and the mean value theorem to obtain

$$\mathbb{1}[\mathcal{H}_{1,0}]\left[\left(\frac{\ell}{n}\right)^{\chi} - \left(\frac{\ell-1}{n}\right)^{\chi}\right] \leqslant \mathbb{1}[\mathcal{H}_{1,0}]\frac{\chi}{n}\left(\frac{n}{\ell-1}\right)^{1-\chi} \leqslant \frac{\chi(\log\log n)^{1-\chi}}{n}. \tag{5.10}$$

Next, we couple the type-L child of $k_0^{(\mathrm{u})} \in V\big(G_n, k_0^{(\mathrm{u})}\big)$ and $0 \in V(\mathcal{T}_{\mathbf{x},n}, 0)$ so that $k_{0,1}^{(\mathrm{u})} = k_{0,1}^{(\mathrm{i})}$ and $\big(k_{0,1}^{(\mathrm{u})}/n\big)^{\chi} \approx a_{0,1}^{(\mathrm{i})}$. Independently from $a_0^{(\mathrm{i})}$, let $U_{0,1} \sim \mathrm{U}[0, 1]$ and $a_{0,1}^{(\mathrm{i})} = U_{0,1}a_0^{(\mathrm{i})}$. Recalling $v[2] = k_{0,1}^{(\mathrm{u})}$ in the BFS, it follows from Definitions 4 and 8 that we can define $k_{0,1}^{(\mathrm{u})}$ to satisfy

$$S_{n,v[2]-1}[1] \leqslant U_{0,1}S_{n,v[1]-1}[1] < S_{n,v[2]}[1],$$

or equivalently,

$$S_{n,v[2]-1}[1] \leqslant \frac{a_{0,1}^{(\mathrm{i})}}{a_0^{(\mathrm{i})}}S_{n,v[1]-1}[1] < S_{n,v[2]}[1], \tag{5.11}$$

and take $k_{0,1}^{(\mathrm{u})} = k_{0,1}^{(\mathrm{i})}$. Now we show that under this coupling, $a_{0,1}^{(\mathrm{i})} \approx \big(k_{0,1}^{(\mathrm{u})}/n\big)^{\chi}$ on the 'good' event $\mathcal{H}_{1,0} \cap \mathcal{H}_{1,1} \cap F_{1,1}$. Observe that $S_{n,v[2]}[1] = \Omega\big((\log\log n)^{-3\chi}\big)$ on the good event, because for n large enough,

$$\begin{aligned} S_{n,v[2]}[1] &\geqslant U_{0,1}S_{n,v[1]-1}[1] \geqslant (\log\log n)^{-2\chi}\left[\left(\frac{v[1]-1}{n}\right)^{\chi} - C^{\star}n^{-\frac{\chi}{12}}\right] \\ &\geqslant (\log\log n)^{-2\chi}\big[(\log\log n)^{-\chi} - 2C^{\star}n^{-\frac{\chi}{12}}\big]. \end{aligned}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

Since $S_{n,j}[1]$ increases with j, the last calculation implies that $|S_{n,j}[1]-(j/n)^{\chi}|\leqslant C^{\star}n^{-\frac{\chi}{12}}$ for $j=k_{0,1}^{(\mathrm{u})}, k_{0,1}^{(\mathrm{u})}-1$. Furthermore, a little calculation shows that on the event $\mathcal{H}_{1,0}\cap F_{1,1}$, there is a constant $c:=c(x_1,\mu)$ such that

$$\left|\left(S_{n,v[1]-1}[1]/a_0^{(\mathrm{i})}\right)-1\right|=cn^{-\frac{\chi}{12}}(\log\log n)^{\chi}.$$

Swapping $\left(S_{n,v[1]-1}[1]/a_0^{(\mathrm{i})}\right)$, $S_{n,v[2]}[1]$, and $S_{n,v[2]-1}[1]$ in (5.11) for one, $(v[2]/n)^{\chi}$, and $((v[2]-1)/n)^{\chi}$ at the costs above, we conclude that there exists $\widehat{C}:=\widehat{C}(x_1,\mu)$ such that on the good event, $\left|a_{0,1}^{(\mathrm{i})}-\left(k_{0,1}^{(\mathrm{u})}/n\right)^{\chi}\right|\leqslant\widehat{C}n^{-\frac{\chi}{12}}(\log\log n)^{\chi}$. Hence, we can take $C_1:=\widehat{C}\vee\chi$ for the event $\mathcal{H}_{1,2}$ and apply Lemma 9 to obtain

$$\begin{aligned}\mathbb{P}_{\mathbf{x}}\left[\bigcap_{j=1}^{3}F_{1,j}\cap\mathcal{H}_{1,2}^{c}\cap\mathcal{H}_{1,1}\cap\mathcal{H}_{1,0}\right]&=\mathbb{P}_{\mathbf{x}}\left[\left\{\mathbf{Y}_{\mathrm{Be}}^{(v[1],n)}\neq\mathbf{V}_{\mathrm{Po}}^{(v[1],n)}\right\}\cap\bigcap_{j=1}^{3}F_{1,j}\cap\bigcap_{j=0}^{1}\mathcal{H}_{1,j}\right]\\&\leqslant\mathbb{P}_{\mathbf{x}}\left[\left\{\mathbf{Y}_{\mathrm{Be}}^{(v[1],n)}\neq\mathbf{V}_{\mathrm{Po}}^{(v[1],n)}\right\}\cap\bigcap_{j=1}^{3}F_{1,j}\cap\mathcal{H}_{1,0}\right]\\&\leqslant Cn^{-\gamma}(\log\log n)^{1-\chi},\end{aligned}\tag{5.12}$$

where $C:=C(x_1,\mu,\kappa)$ is as in Lemma 9. The lemma is proved by applying (5.8), (5.9), and (5.12) to (5.7). □

Before proving Lemma 7, we require a final result that says that under the graph coupling, the event that vertex $k_0^{(\mathrm{u})}$ has a low degree occurs w.h.p. The next lemma is proved in [24] using Chebyshev's inequality. Recall that $A_n:=A_{2/3,n}$ is the event in (2.1) with $\alpha=2/3$.

Lemma 10. *Assume that $x_i\in(0,\kappa]$ for $i\geqslant 2$ and $\mathbf{x}\in A_n$. Let $\mathcal{H}_{1,i}$, $0,\ldots,3$, be as in (5.2). There is a coupling of $\left(G_n,k_0^{(\mathrm{u})}\right)$ and $(\mathcal{T}_{\mathbf{x},n},0)$, with a positive integer p and a positive constant $C:=C(p,\kappa)$ such that*

$$\mathbb{P}_{\mathbf{x}}\left[\bigcap_{i=0}^{2}\mathcal{H}_{1,i}\cap\mathcal{H}_{1,3}^{c}\right]\leqslant C(\log n)^{-\frac{p}{r}}(\log\log n)^{\frac{p}{\mu+1}}\quad\textit{for all } n\geqslant 3.$$

We now complete the proof of Lemma 7 by applying Lemmas 8 and 10.

Proof of Lemma 7. The lemma follows from bounding the right-hand side of

$$\mathbb{P}_{\mathbf{x}}\left[\left(\bigcap_{i=0}^{3}\mathcal{H}_{1,i}\right)^{c}\right]=\mathbb{P}_{\mathbf{x}}\left[\mathcal{H}_{1,0}^{c}\right]+\sum_{i=1}^{3}\mathbb{P}_{\mathbf{x}}\left[\bigcap_{j=0}^{i-1}\mathcal{H}_{1,j}\cap\mathcal{H}_{1,i}^{c}\right],$$

under the coupling in the proof of Lemma 8. To bound $\mathbb{P}_{\mathbf{x}}\left[\mathcal{H}_{1,0}^{c}\right]$ and $\mathbb{P}_{\mathbf{x}}\left[\mathcal{H}_{1,0}\cap\mathcal{H}_{1,1}^{c}\right]$, recall that $a_0^{(\mathrm{i})}=U_0^{\chi}$ and $a_{0,1}^{(\mathrm{i})}=U_{0,1}a_0^{(\mathrm{i})}$, where U_0 and $U_{0,1}$ are independent standard uniform variables. Hence,

$$\begin{aligned}\mathbb{P}_{\mathbf{x}}\left[\mathcal{H}_{1,0}\cap\mathcal{H}_{1,1}^{c}\right]&=\mathbb{E}\left[\mathbb{1}\left[\mathcal{H}_{1,0}\right]\mathbb{P}_{\mathbf{x}}\left[U_{0,1}\leqslant\left(a_0^{(\mathrm{i})}\right)^{-1}(\log\log n)^{-2\chi}\,|U_0\right]\right]\\&\leqslant\mathbb{P}_{\mathbf{x}}\left[U_{0,1}\leqslant(\log\log n)^{-\chi}\right]\\&=(\log\log n)^{-\chi},\end{aligned}\tag{5.13}$$

and $\mathbb{P}_{\mathbf{x}}\left[\mathcal{H}_{1,0}^{c}\right]=\mathbb{P}_{\mathbf{x}}\left[U_0\leqslant(\log\log n)^{-1}\right]=(\log\log n)^{-1}$. Bounding the remaining probabilities using Lemmas 8 and 10 completes the proof. □

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

6. Proof of the general case

Recall the definitions of $k_{\bar{w}}^{(\mathrm{u})}$, $a_{\bar{w}}^{(\mathrm{i})}$, and $k_{\bar{w}}^{(\mathrm{i})}$ in the beginning of Section 3.2 and Definition 8. In this section, we inductively couple the uniformly rooted Pólya urn graph $\big(G_n, k_0^{(\mathrm{u})}\big)$ and the intermediate Pólya point tree $(\mathcal{T}_{\mathbf{x},n}, 0)$ in Definition 8 so that w.h.p., $\Big(B_r\big(G_n, k_0^{(\mathrm{u})}\big), k_0^{(\mathrm{u})}\Big) \cong (B_r(\mathcal{T}_{\mathbf{x},n}, 0), 0)$, and for $k_{\bar{w}}^{(\mathrm{u})} \in V\Big(B_r\big(G_n, k_0^{(\mathrm{u})}\big)\Big)$ and $\bar{w} \in V(B_r(\mathcal{T}_{\mathbf{x},n}, 0)$, their ages are close enough $\Big(\big(k_{\bar{w}}^{(\mathrm{u})}/n\big)^{\chi} \approx a_{\bar{w}}^{(\mathrm{i})}\Big)$ and their PA labels and arrival times match $\big(k_{\bar{w}}^{(\mathrm{i})} = k_{\bar{w}}^{(\mathrm{u})}\big)$. Given a positive integer r, let $L[q] := (0, 1, \ldots, 1)$ and $|L[q]| = q+1$ for $2 \leqslant q \leqslant r$, so that $L[q]$ and $k_{L[q]}^{(\mathrm{u})}$ are the type-L children in $\partial\mathfrak{B}_q := V(B_q(\mathcal{T}_{\mathbf{x},n}, 0)) \setminus V(B_{q-1}(\mathcal{T}_{\mathbf{x},n}, 0))$ and $\partial\mathcal{B}_q := V\Big(B_q\big(G_n, k_0^{(\mathrm{u})}\big)\Big) \setminus V\Big(B_{q-1}\big(G_n, k_0^{(\mathrm{u})}\big)\Big)$. Let $\chi/12 := \beta_1 > \beta_2 > \cdots > \beta_r > 0$, so that $n^{-\beta_q} > n^{-\beta_{q-1}}(\log\log n)^{q\chi}$ for n large enough. To ensure that we can couple the $(q+1)$-neighbourhoods for $1 \leqslant q \leqslant r-1$, we define the coupling events analogous to $\mathcal{H}_{1,i}$, $i = 1, 2, 3$, in (5.2):

$$\begin{aligned}
\mathcal{H}_{q,1} &= \Big\{a_{L[q]}^{(\mathrm{i})} > (\log\log n)^{-\chi(q+1)}\Big\},\\
\mathcal{H}_{q,2} &= \Bigg\{\Big(B_q\big(G_n, k_0^{(\mathrm{u})}\big), k_0^{(\mathrm{u})}\Big) \cong (B_q(\mathcal{T}_{\mathbf{x},n}, 0), 0),\ \text{with } k_{\bar{v}}^{(\mathrm{u})} = k_{\bar{v}}^{(\mathrm{i})}\\
&\qquad \text{and } \left|a_{\bar{v}}^{(\mathrm{i})} - \left(\frac{k_{\bar{v}}^{(\mathrm{u})}}{n}\right)^{\chi}\right| \leqslant C_q n^{-\beta_q} \text{ for } \bar{v} \in V(B_q(\mathcal{T}_{\mathbf{x},n}, 0))\Bigg\},\\
\mathcal{H}_{q,3} &= \Big\{R_{\bar{v}}^{(\mathrm{i})} < (\log n)^{1/r} \text{ for } \bar{v} \in V(B_{q-1}(\mathcal{T}_{\mathbf{x},n}, 0))\Big\},
\end{aligned} \tag{6.1}$$

where $R_{\bar{v}}^{(\mathrm{i})}$ is as in (4.1) and $C_q := C_q(x_1, \mu)$ will be chosen in the proof of Lemma 12 below. Let $A_n := A_{2/3,n}$ be the event in (2.1), and let $\mathbb{P}_{\mathbf{x}}$ indicate the conditioning on a specific realisation of the fitness sequence $\mathbf{x}$. The next lemma is the key result of this section; it essentially states that if we can couple $\big(G_n, k_0^{(\mathrm{u})}\big)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$ so that $\Big(B_q\big(G_n, k_0^{(\mathrm{u})}\big), k_0^{(\mathrm{u})}\Big) \cong (B_q(\mathcal{T}_{\mathbf{x},n}, 0), 0)$ w.h.p., then we can achieve this for the $(q+1)$-neighbourhoods too.

Lemma 11. *Let $\mathcal{H}_{q,i}$, $1 \leqslant q \leqslant r-1$, $i = 1, 2, 3$, be as in (5.2) and (6.1). Assume that $x_i \in (0, \kappa]$ for $i \geqslant 2$ and $\mathbf{x} \in A_n$. Given $r < \infty$ and $1 \leqslant q \leqslant r-1$, if there is a coupling of $\big(G_n, k_0^{(\mathrm{u})}\big)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$, with a positive constant $C := C(x_1, \mu, \kappa, q)$ such that*

$$\mathbb{P}_{\mathbf{x}}\left[\left(\bigcap_{i=1}^{3} \mathcal{H}_{q,i}\right)^{c}\right] \leqslant C(\log\log n)^{-\chi} \quad \textit{for all } n \geqslant 3,$$

then there is a coupling of $\big(G_n, k_0^{(\mathrm{u})}\big)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$, with a positive constant $C' := C'(x_1, \mu, \kappa, q)$ such that

$$\mathbb{P}_{\mathbf{x}}\left[\left(\bigcap_{i=1}^{3} \mathcal{H}_{q+1,i}\right)^{c}\right] \leqslant C'(\log\log n)^{-\chi} \quad \textit{for all } n \geqslant 3.$$

Since Lemma 7 implies that such a graph coupling exists for $r = 1$, combining Lemmas 7 and 11 yields the following corollary, which is instrumental in proving Theorem 1.

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

Corollary 2. *Retaining the assumptions and the notation in Lemma 11, there is a coupling of* $(G_n, k_0^{(\mathrm{u})})$ *and* $(\mathcal{T}_{\mathbf{x},n}, 0)$*, with a positive constant* $C := C(x_1, \mu, \kappa, r)$ *such that*

$$\mathbb{P}_{\mathbf{x}}\left[\left(\bigcap_{i=1}^{3}\mathcal{H}_{r,i}\right)^c\right] \leqslant C(\log\log n)^{-\chi} \quad \textit{for all } n \geqslant 3.$$

The proof of Lemma 11 is in the same vein as that of Lemma 7. Fixing $1 \leqslant q \leqslant r-1$, we may assume that the two graphs are already coupled so that the event $\bigcap_{i=1}^{3}\mathcal{H}_{q,i}$ has occurred. On the event $\mathcal{H}_{q,1}$, it is clear from the definitions of $a_{L[q+1]}^{(\mathrm{i})}$ and $\mathcal{H}_{q+1,1}$ that $\mathcal{H}_{q+1,1}$ occurs w.h.p., while on the event $\mathcal{H}_{q,1} \cap \mathcal{H}_{q,3}$, we can easily show that $\mathcal{H}_{q+1,3}$ occurs w.h.p. So for most of the proof, we handle the event $\mathcal{H}_{q+1,2}$. To couple the graphs so that $\mathcal{H}_{q+1,2}$ occurs w.h.p., we consider the vertices of $\partial\mathcal{B}_q$ and $\partial\mathfrak{B}_q$ in the breadth-first order, and couple $k_{\bar{w}}^{(\mathrm{u})} \in \partial\mathcal{B}_q$ and $\bar{w} \in \partial\mathfrak{B}_q$ so that w.h.p., the numbers of children they have are the same and not too large, and the children's ages are close enough. Recall the definition of $v[t]$ in (3.1). If $v[t] = k_{\bar{w}}^{(\mathrm{u})}$, the associated events are as follows:

$$\begin{aligned}\mathcal{K}_{t,1} = \Bigg\{ & R_{\bar{w}}^{(\mathrm{u})} = R_{\bar{w}}^{(\mathrm{i})}, \text{ and for } 1 \leqslant j \leqslant 1 + R_{\bar{w}}^{(\mathrm{i})}, \\ & \qquad k_{\bar{w},j}^{(\mathrm{i})} = k_{\bar{w},j}^{(\mathrm{u})} \text{ and } \left| a_{\bar{w},j}^{(\mathrm{i})} - \left(\frac{k_{\bar{w},j}^{(\mathrm{u})}}{n}\right)^{\chi} \right| \leqslant C_{q+1} n^{-\beta_{q+1}} \Bigg\}, \\ \mathcal{K}_{t,2} = \Big\{ & R_{\bar{w}}^{(\mathrm{i})} < (\log n)^{1/r} \Big\},\end{aligned} \tag{6.2}$$

where $R_{\bar{w}}^{(\mathrm{u})}$ and $R_{\bar{w}}^{(\mathrm{i})}$ are as in (3.4) and (4.1), and C_{q+1} and β_{q+1} are as in the event $\mathcal{H}_{q+1,2}$. Below we only consider the coupling of the type-L children in detail, because the type-R case can be proved similarly. Observe that for n large enough, there must be a type-L child in $\partial\mathcal{B}_q$ on the event $\bigcap_{j=1}^{3}\mathcal{H}_{q,j}$. Define $v[\tau[q]] = k_{L[q]}^{(\mathrm{u})}$, so that

$$\tau[1] = 2 \quad \text{and} \quad \tau[q] = \left|V\left(B_{q-1}\left(G_n, k_0^{(\mathrm{u})}\right)\right)\right| + 1 \quad \text{for } 2 \leqslant q \leqslant r-1$$

are the exploration times of the type-L children. Noting that

$$\left(\mathcal{A}_{\tau[q]-1}, \mathcal{P}_{\tau[q]-1}, \mathcal{N}_{\tau[q]-1}\right) = \left(\partial\mathcal{B}_q, V\left(B_{q-1}\left(G_n, k_0^{(\mathrm{u})}\right)\right), V(G_n) \setminus V\left(B_q\left(G_n, k_0^{(\mathrm{u})}\right)\right)\right),$$

we let

$$\left(\left(\mathcal{Z}_j[\tau[q]], \widetilde{\mathcal{Z}}_j[\tau[q]]\right), j \in \mathcal{A}_{\tau[q]-1} \cup \mathcal{N}_{\tau[q]-1}\right) \quad \text{and} \quad \left(S_{n,j}[\tau[q]], 1 \leqslant j \leqslant n\right)$$

be as in (3.7), (3.8), and (3.10). For convenience, we also define

$$\zeta_q := \mathcal{Z}_{v[\tau[q]]}[\tau[q]], \quad \text{so that } \zeta_q \sim \text{Gamma}(x_{v[\tau[q]]} + 1, 1). \tag{6.3}$$

We use $(S_{n,j}[\tau[q]], 1 \leqslant j \leqslant n)$ to generate the children of vertex $v[\tau[q]]$, and let $\left(a_{L[q],j}^{(\mathrm{i})}, 2 \leqslant j \leqslant 1 + R_{L[q]}^{(\mathrm{i})}\right)$ be the points of the mixed Poisson process on $\left(a_{L[q]}^{(\mathrm{i})}, 1\right]$ with intensity

$$\frac{\zeta_q}{\mu\left(a_{L[q]}^{(\mathrm{i})}\right)^{1/\mu}} y^{1/\mu - 1} dy, \tag{6.4}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

and $a^{(\mathrm{i})}_{L[q],1} \sim \mathrm{U}\left[0,\, a^{(\mathrm{i})}_{L[q]}\right]$. In the sequel, we develop lemmas analogous to Lemmas 8 and 10 to show that on the event $\bigcap_{j=1}^{3} \mathcal{H}_{q,j}$, we can couple the vertices $v[\tau[q]] \in \partial\mathcal{B}_q$ and $\tau[q] \in \partial\mathfrak{B}_q$ so that the events $\mathcal{K}_{\tau[q],1}$ and $\mathcal{K}_{\tau[q],2}$ occur w.h.p. As in the 1-neighbourhood case, the difficult part is proving the claim for $\mathcal{K}_{\tau[q],1}$. From now on we write

$$\overline{\mathcal{H}}_{q,l} := \bigcap_{j=1}^{3} \mathcal{H}_{q,j} \cap \left\{|\partial\mathcal{B}_q| = l\right\} \quad \textit{for all } l \geqslant 1. \tag{6.5}$$

Lemma 12. *Retaining the assumptions and the notation in Lemma 11, let β_q, $\mathcal{K}_{\tau[q],1}$, and $\overline{\mathcal{H}}_{q,l}$ be as in (6.1), (6.2), and (6.5). Then there is a coupling of $\left(G_n, k^{(\mathrm{u})}_0\right)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$, with a positive constant $C := C(x_1, \mu, \kappa, q)$ such that*

$$\mathbb{P}_{\mathbf{x}}\left(\overline{\mathcal{H}}_{q,l} \cap \mathcal{H}_{q+1,1} \cap \mathcal{K}^c_{\tau[q],1}\right) \leqslant Cn^{-d}(\log\log n)^{q+1}(\log n)^{q/r} \quad \textit{for all } n \geqslant 3 \text{ and} l \geqslant 1,$$

where $0 < \gamma < \chi/12$ and $d := \min\left\{\chi/3 - 4\gamma,\, \gamma,\, 1-\chi,\, \beta_q\right\}$.

Similar to Lemma 8, the key step is to couple the type-R children of vertex $k^{(\mathrm{u})}_{L[q]}$. Let the Bernoulli sequence $\mathbf{Y}^{(v[\tau[q]],n)}_{\mathrm{Be}}$ be as in Definition 7, which by Lemma 5 encodes the type-R children of vertex $k^{(\mathrm{u})}_{L[q]}$. Additionally, define

$$M_{L[q]} := \min\left\{j \in [n] : (j/n)^{\chi} \geqslant a^{(\mathrm{i})}_{L[q]}\right\}. \tag{6.6}$$

We want to use the bins $\left(a^{(\mathrm{i})}_{L[q]},\, ((j/n)^{\chi})^{n}_{j=M_{L[q]}}\right)$ to construct the discretised mixed Poisson process that encodes the ages and the PA labels, i.e. $\left(\left(a^{(\mathrm{i})}_{L[q],j},\, k^{(\mathrm{i})}_{L[q],j}\right),\, 2 \leqslant j \leqslant 1 + R^{(\mathrm{i})}_{L[q]}\right)$. However, it is possible that $M_{L[q]} \neq v[\tau[q]] + 1$, in which case the numbers of Bernoulli and Poisson variables do not match and a modification of $\mathbf{Y}^{(v[\tau[q]],n)}_{\mathrm{Be}}$ is needed. If $M_{L[q]} \leqslant v[\tau[q]]$, define $Y_{j\to v[\tau[q]]}$, $M_{L[q]} \leqslant j \leqslant v[\tau[q]]$, as Bernoulli variables with means $P_{j\to v[\tau[q]]} := 0$, and concatenate the vectors $\left(Y_{j\to v[\tau[q]]},\, M_{L[q]} \leqslant j \leqslant v[\tau[q]]\right)$ and $\mathbf{Y}^{(v[\tau[q]],n)}_{\mathrm{Be}}$. This corresponds to the fact that vertex j cannot send an outgoing edge to vertex $v[\tau[q]]$ in $\left(G_n, k^{(\mathrm{u})}_0\right)$. If $M_{L[q]} \geqslant v[\tau[q]] + 1$, let $\mathbf{Y}^{(v[\tau[q]],n)}_{\mathrm{Be}}$ be as in Definition 7. Saving notation, we redefine

$$\mathbf{Y}^{(v[\tau[q]],n)}_{\mathrm{Be}} := \left(Y_{\widetilde{M}_{L[q]}\to v[\tau[q]]},\, Y_{(\widetilde{M}_{L[q]}+1)\to v[\tau[q]]},\, \ldots,\, Y_{n\to v[\tau[q]]}\right), \tag{6.7}$$

where

$$\widetilde{M}_{L[q]} := \min\left\{M_{L[q]},\, v[\tau[q]] + 1\right\}. \tag{6.8}$$

We also make the following adjustment so that the upcoming Poisson random vector is a discretisation of the mixed Poisson point process on $\left(a^{(\mathrm{i})}_{L[q]},\, 1\right]$, and is still comparable to the modified $\mathbf{Y}^{(v[\tau[q]],n)}_{\mathrm{Be}}$. When $M_{L[q]} \leqslant v[\tau[q]] + 1$, define the Poisson means

$$\lambda^{[\tau[q]]}_{M_{L[q]}} := \int_{a^{(\mathrm{i})}_{L[q]}}^{\left(\frac{M_{L[q]}}{n}\right)^{\chi}} \frac{\zeta_q}{\mu\left(a^{(\mathrm{i})}_{L[q]}\right)^{1/\mu}} y^{1/\mu-1}\,dy, \qquad \lambda^{[\tau[q]]}_{j} := \int_{\left(\frac{j-1}{n}\right)^{\chi}}^{\left(\frac{j}{n}\right)^{\chi}} \frac{\zeta_q}{\mu\left(a^{(\mathrm{i})}_{L[q]}\right)^{1/\mu}} y^{1/\mu-1}\,dy \tag{6.9}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

for $M_{L[q]}+1 \leqslant j \leqslant n$. When $M_{L[q]} \geqslant v[\tau[q]]+2$, we let

$$\lambda_j^{[\tau[q]]} := 0 \quad \text{for } v[\tau[q]]+1 \leqslant j < M_{L[q]}, \tag{6.10}$$

in addition to (6.9), so that there are no Poisson points outside the interval $\left(a^{(\mathrm{i})}_{L[q]}, 1\right]$.

Definition 10. Given $k^{(\mathrm{u})}_{L[q]}$, $a^{(\mathrm{i})}_{L[q]}$, and ζ_q defined in (6.3), let $\widetilde{M}_{L[q]}$ be as in (6.8); for $\widetilde{M}_{L[q]}+1 \leqslant j \leqslant n$, let $V_{j\to v[\tau[q]]}$ be conditionally independent Poisson random variables, each with parameters $\lambda_j^{[\tau[q]]}$ given in (6.9) and (6.10). Define this (mixed) Poisson sequence by the random vector

$$\mathbf{V}^{(v[\tau[q]],n)}_{\mathrm{Po}} := \left(V_{\widetilde{M}_{L[q]}\to v[\tau[q]]}, V_{(\widetilde{M}_{L[q]}+1)\to v[\tau[q]]}, \ldots, V_{n\to v[\tau[q]]}\right).$$

We proceed to define the events analogous to $F_{1,j}$, $j=1, 2, 3$, in (5.5). On these events, $P_{j\to v[\tau[q]]}$ in (3.11) and $\lambda_j^{[\tau[q]]}$ are close enough for most j, and so we can couple the point processes as before. Let $\phi(n)=\Omega(n^\chi)$, $0<\gamma<\chi/12$, $\mathcal{Z}_j[\tau[q]]$, $B_j[\tau[q]]$, and $S_{n,j}[\tau[q]]$ be as in (3.7), (3.9), and (3.10), and let $C^\star_q := C^\star_q(x_1, \mu)$ be a positive constant that we specify later. Define

$$\begin{aligned} F_{\tau[q],1} &= \left\{ \max_{\lceil\phi(n)\rceil \leqslant j \leqslant n} \left| S_{n,j}[\tau[q]] - \left(\frac{j}{n}\right)^{\chi} \right| \leqslant C^\star_q n^{-\frac{\chi}{12}} \right\}, \\ F_{\tau[q],2} &= \bigcap_{j=\lceil\phi(n)\rceil; j\notin\mathcal{P}_{\tau[q]-1}}^{n} \left\{ \left| B_j[\tau[q]] - \frac{\mathcal{Z}_j[\tau[q]]}{(\mu+1)j} \right| \leqslant \frac{\mathcal{Z}_j[\tau[q]] n^{-\gamma}}{(\mu+1)j} \right\}, \\ F_{\tau[q],3} &= \bigcap_{j=\lceil\phi(n)\rceil; j\notin\mathcal{P}_{\tau[q]-1}}^{n} \left\{ \mathcal{Z}_j[\tau[q]] \leqslant j^{1/2} \right\}. \end{aligned} \tag{6.11}$$

The following analogue of Lemma 9 is the main ingredient for proving Lemma 12.

Lemma 13. *Retaining the assumptions and the notation in Lemma 12, let* $\mathbf{Y}^{(v[\tau[q]],n)}_{\mathrm{Be}}$, $\mathbf{V}^{(v[\tau[q]],n)}_{\mathrm{Po}}$, *and* $F_{\tau[q],i}$, $i=1, 2, 3$, *be as in Definition 7, Definition 10, and (6.11). Then we can couple the random vectors so that there is a positive constant* $C=C(x_1, \mu, \kappa, q)$ *such that*

$$\mathbb{P}_{\mathbf{x}}\left[\left\{\mathbf{Y}^{(v[\tau[q]],n)}_{\mathrm{Be}} \neq \mathbf{V}^{(v[\tau[q]],n)}_{\mathrm{Po}}\right\} \cap \bigcap_{i=1}^{3} F_{\tau[q],i} \cap \overline{\mathcal{H}}_{q,l}\right] \leqslant Cn^{-d}(\log\log n)^{q+1}(\log n)^{q/r}$$

for all $n \geqslant 3$, *where* $d := \min\{\beta_q, \gamma, 1-\chi\}$.

Sketch of proof. We only highlight the main steps here, leaving the details to [24]. As in the proof of Lemma 9, we construct a Bernoulli and a Poisson sequence using suitable means $\widehat{P}_{j\to v[\tau[q]]}$, $\widetilde{M}_{L[q]} \leqslant j \leqslant n$, with $\widetilde{M}_{L[q]}$ as in (6.8). Then we couple the four processes. However, this time we need to handle the cases where we couple a Bernoulli variable with mean zero and a Poisson variable with positive mean, or vice versa. We take care of these cases by choosing the appropriate $\widehat{P}_{j\to v[\tau[q]]}$. First we observe that on the event $\left(\bigcap_{j=1}^{3} F_{\tau[q],j}\right) \cap \left(\bigcap_{j=1}^{3} \mathcal{H}_{q,j}\right)$, $P_{j\to v[\tau[q]]}$ is close enough to

$$\widehat{P}_{j\to v[\tau[q]]} := \left(\frac{v[\tau[q]]}{j}\right)^{\chi} \frac{\zeta_q}{(\mu+1)v[\tau[q]]} \leqslant 1 \tag{6.12}$$

TABLE 1. Combinations of means for $\widetilde{M}_{L[q]} \leqslant j \leqslant \max\{M_{L[q]}, k^{(\mathrm{u})}_{L[q]}+1\}$. Note that $j = k^{(\mathrm{u})}_{L[q]}+1$ when $M_{L[q]} = k^{(\mathrm{u})}_{L[q]}+1$, and in that case, the coupling is the same as that of Lemma 9.

	$M_{L[q]} \leqslant k^{(\mathrm{u})}_{L[q]}$	$M_{L[q]} \geqslant k^{(\mathrm{u})}_{L[q]}+2$	$M_{L[q]} = k^{(\mathrm{u})}_{L[q]}+1$
$P_{j\to v[\tau[q]]}$	0	as in (3.11)	as in (3.11)
$\lambda_j^{[\tau[q]]}$	as in (6.9)	0	as in (6.9)
$\widehat{P}_{j\to v[\tau[q]]}$	0	as in (6.12)	as in (6.12)

for $j \in \{v[\tau[q]]+1, \ldots, n\} \setminus V\Big(B_q\big(G_n, k_0^{(\mathrm{u})}\big)\Big)$, whereas for $\max\{M_{L[q]}, v[\tau[q]]+1\} \leqslant j \leqslant n$, we have

$$\lambda_j^{[\tau[q]]} \approx \widehat{P}_{j\to v[\tau[q]]}$$

because $a^{(\mathrm{i})}_{L[q]} \approx (v[\tau[q]]/n)^{\chi}$ on the event $\mathcal{H}_{q,2}$. Hence, we can couple $Y_{j\to v[\tau[q]]}$ and $V_{j\to v[\tau[q]]}$ for $j \in \{\max\{M_{L[q]}, v[\tau[q]]+1\}, \ldots, n\} \setminus V\Big(B_q\big(G_n, k_0^{(\mathrm{u})}\big)\Big)$ as in Lemma 9. When $j \in V\Big(B_q\big(G_n, k_0^{(\mathrm{u})}\big)\Big)$, the Bernoulli variable $Y_{j\to v[\tau[q]]}$ has mean zero, and we couple it to a Bernoulli variable with mean (6.12) as in Lemma 9. By a little computation,

$$|\partial\mathcal{B}_j| < 1 + j(\log n)^{j/r}, \qquad \Big|V\Big(B_j\big(G_n, k_0^{(\mathrm{u})}\big)\Big)\Big| < 1 + j + j^2(\log n)^{j/r} \tag{6.13}$$

for $1 \leqslant j \leqslant q$ on the event $\mathcal{H}_{q,3}$. Because the $\widehat{P}_{j\to v[\tau[q]]}$ are sufficiently small and there are at most $O((\log n)^{q/r})$ such pairs of Bernoulli variables, we can use a union bound to show that the probability that $Y_{j\to v[\tau[q]]} \neq V_{j\to v[\tau[q]]}$ for any such j tends to zero as $n \to \infty$. We now consider the coupling of $Y_{j\to v[\tau[q]]}$ and $V_{j\to v[\tau[q]]}$ for $\widetilde{M}_{L[q]} \leqslant j \leqslant \max\Big\{M_{L[q]}, k^{(\mathrm{u})}_{L[q]}+1\Big\}$. In Table 1 below, we give the possible combinations of the Bernoulli and Poisson means, and our choice of intermediate means. As indicated in the table, we choose $\widehat{P}_{j\to v[\tau[q]]} > 0$ whenever $P_{j\to v[\tau[q]]} > 0$. When $M_{L[q]} \leqslant v[\tau[q]]$, we couple $V_{j\to v[\tau[q]]}$ and a Poisson variable with mean zero, while when $M_{L[q]} \geqslant v[\tau[q]]+2$, $V_{j\to v[\tau[q]]} := 0$ by construction, and it is coupled with a Poisson variable with mean (6.12). However, the number of these pairs is small because $M_{L[q]} \approx v[\tau[q]]+1$ when $a^{(\mathrm{i})}_{L[q]} \approx (v[\tau[q]]/n)^{\chi}$. Consequently, another union bound argument shows that the probability that any of these couplings fail tends to zero as $n \to \infty$. □

We are now ready to prove Lemma 12.

Proof of Lemma 12. Recall the definition of $\overline{\mathcal{H}}_{q,l}$ in (6.5). We bound the right-hand side of

$$\begin{aligned}
&\mathbb{P}_{\mathbf{x}}\Big[\overline{\mathcal{H}}_{q,l} \cap \mathcal{H}_{q+1,1} \cap \mathcal{K}^c_{\tau[q],1}\Big] \\
&\leqslant \mathbb{P}_{\mathbf{x}}\Bigg[\overline{\mathcal{H}}_{q,l} \cap \bigcap_{j=1}^{3} F_{\tau[q],j} \cap \mathcal{H}_{q+1,1} \cap \mathcal{K}^c_{\tau[q],1}\Bigg] + \mathbb{P}_{\mathbf{x}}\Bigg[\overline{\mathcal{H}}_{q,l} \cap \Bigg(\bigcap_{j=1}^{3} F_{\tau[q],j}\Bigg)^c\Bigg] \\
&\leqslant \mathbb{P}_{\mathbf{x}}\Bigg[\overline{\mathcal{H}}_{q,l} \cap \bigcap_{j=1}^{3} F_{\tau[q],j} \cap \mathcal{H}_{q+1,1} \cap \mathcal{K}^c_{\tau[q],1}\Bigg] + \sum_{j=1}^{3} \mathbb{P}_{\mathbf{x}}\Big[\overline{\mathcal{H}}_{q,l} \cap F^c_{\tau[q],j}\Big],
\end{aligned} \tag{6.14}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

starting from the last three terms. Using (6.13), a computation similar to that for Lemma 2 shows that there are positive constants $C_q^\star = C_q^\star(x_1, \mu)$ and $c^\star = c^\star(x_1, \mu, q)$ such that

$$\mathbb{P}_{\mathbf{x}}\left[\overline{\mathcal{H}}_{q,l} \cap \left\{\max_{\lceil\phi(n)\rceil \leqslant j \leqslant n}\left|S_{n,j}[\tau[q]] - \left(\frac{j}{n}\right)^{\chi}\right| \leqslant C_q^\star n^{-\frac{\chi}{12}}\right\}\right] \leqslant c^\star n^{-\frac{\chi}{6}}, \tag{6.15}$$

which bounds $\mathbb{P}_{\mathbf{x}}\left[\overline{\mathcal{H}}_{q,l} \cap F^c_{\tau[q],1}\right]$. Combining (6.13) and the argument of Lemma 3 also yields

$$\mathbb{P}_{\mathbf{x}}\left[\overline{\mathcal{H}}_{q,l} \cap F^c_{\tau[q],3}\right] \leqslant C\kappa^4 n^{-\chi} \quad \text{and} \quad \mathbb{P}_{\mathbf{x}}\left[\overline{\mathcal{H}}_{q,l} \cap F^c_{\tau[q],2}\right] \leqslant C' n^{4\gamma - \frac{\chi}{3}}, \tag{6.16}$$

for some positive constants $C := C(q)$, $C' = C'(x_1, \gamma, \mu, q)$, and $0 < \gamma < \chi/12$. Next, we bound the first term under a suitable coupling of the vertices $k^{(\mathrm{u})}_{L[q]} \in \partial\mathcal{B}_q$ and $L[q] \in \partial\mathfrak{B}_q$, starting from their type-R children. Assume that

$$\left(\left(k^{(\mathrm{u})}_{\bar{w}}, k^{(\mathrm{i})}_{\bar{w}}, a^{(\mathrm{i})}_{\bar{w}}\right), \bar{w} \in V\left(B_q(\mathcal{T}_{\mathbf{x},n}, 0)\right)\right), \quad \left(\left(R^{(\mathrm{u})}_{\bar{w}}, R^{(\mathrm{i})}_{\bar{w}}\right), \bar{w} \in V(B_{q-1}(\mathcal{T}_{\mathbf{x},n}, 0))\right),$$
$$\left(\left(\mathcal{Z}_j[\tau[q]], \widetilde{\mathcal{Z}}_j[\tau[q]]\right), j \in [n] \setminus V(B_{q-1}(\mathcal{T}_{\mathbf{x},n}, 0))\right)$$

are coupled so that $\mathrm{E}_q := \bigcap_{j=1}^3 \mathcal{H}_{q,j} \cap \bigcap_{j=1}^3 F_{\tau[q],j}$ occurs. In view of Definitions 8 and 7, it follows from Lemma 13 that on the event E_q, there is a coupling such that $R^{(\mathrm{u})}_{L[q]} = R^{(\mathrm{i})}_{L[q]}$ and $k^{(\mathrm{u})}_{L[q],j} = k^{(\mathrm{i})}_{L[q],j}$ for $2 \leqslant j \leqslant 1 + R^{(\mathrm{i})}_{L[q]}$ w.h.p. Note that $k_{L[q],j} \geqslant M_{L[q]}$ when $\mathbf{Y}^{(\nu[\tau[q]],n)}_{\mathrm{Be}}$ and $\mathbf{V}^{(\nu[\tau[q]],n)}_{\mathrm{Po}}$ are coupled. Thus, using $M_{L[q]} \geqslant n(a^{(\mathrm{i})}_{L[q]})^{1/\chi}$, a little calculation shows that there is a constant $\widehat{c} := \widehat{c}(\mu)$ such that

$$\mathbb{1}\left[\bigcap_{i=1}^{2} \mathcal{H}_{q,i}\right]\left[\left(\frac{j}{n}\right)^{\chi} - \left(\frac{j-1}{n}\right)^{\chi}\right] \leqslant \frac{\widehat{c}(\log\log n)^{(1+q)(1-\chi)}}{n}, \qquad M_{L[q]} \leqslant j \leqslant n,$$

implying that the ages of the type-R children are close enough. We proceed to couple the type-L child of $k^{(\mathrm{u})}_{L[q]} \in \partial\mathcal{B}_q$ and $L[q] \in \partial\mathfrak{B}_q$ on the event $\mathrm{E}_q \cap \mathcal{H}_{q+1}$. Independently from all the variables generated so far, let $U_{L[q],1} \sim \mathrm{U}[0, 1]$. Set

$$a^{(\mathrm{i})}_{L[q],1} := a^{(\mathrm{i})}_{L[q+1]} = U_{L[q],1} a^{(\mathrm{i})}_{L[q]},$$

so that $a^{(\mathrm{i})}_{L[q],1}$ is the age of vertex $(L[q], 1) \in \partial\mathfrak{B}_{q+1}$. Temporarily defining $f := k^{(\mathrm{u})}_{L[q]}$ and $g := k^{(\mathrm{u})}_{L[q],1}$, we define g to satisfy

$$S_{n,g-1}[\tau[q]] \leqslant U_{L[q],1} S_{n,f-1}[\tau[q]] < S_{n,g}[\tau[q]]. \tag{6.17}$$

In light of Definition 8 and Lemma 6, $k^{(\mathrm{i})}_{L[q],1} = g$. To establish $a^{(\mathrm{i})}_{L[q],1} \approx (g/n)^{\chi}$, we first show that we can substitute $S_{n,j}[\tau[q]]$ with $(j/n)^{\chi}$ for $j = g-1, g$ at a small enough cost. A straightforward computation shows that on the event $\mathrm{E}_q \cap \mathcal{H}_{q+1,1}$,

$$U_{L[q],1} \geqslant (\log\log n)^{-(q+2)\chi} \quad \text{and} \quad f \geqslant n(\log\log n)^{-(q+1)} - C_q n^{1-\beta_q/\chi},$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

where C_q is the constant in the event $\mathcal{H}_{q,2}$. Consequently, there is a constant $C := C(x_1, \mu, q)$ such that on the event $\mathrm{E}_q \cap \mathcal{H}_{q+1,1}$,

$$S_{n,g}[\tau[q]] \geqslant U_{L[q],1} S_{n,f-1}[\tau[q]] \geqslant (\log\log n)^{-(q+2)\chi} \left[\left(\frac{f-1}{n}\right)^{\chi} - C_q^{\star} n^{-\frac{\chi}{12}}\right] \\ \geqslant C(\log\log n)^{-(2q+3)\chi}. \tag{6.18}$$

This implies that $g = \Omega(n^{\chi})$, and so $|S_{n,j}[\tau[q]] - (j/n)^{\chi}| \leqslant C_q^{\star} n^{-\chi/12}$ for $j = g-1, g$. Additionally, a direct calculation yields

$$\left|\left(S_{n,f-1}[\tau[q]]/a^{(\mathrm{i})}_{L[q]}\right) - 1\right| \leqslant \widetilde{C} n^{-\beta_q} (\log\log n)^{(q+1)\chi}$$

for some $\widetilde{C} := \widetilde{C}(x_1, \mu, q)$. By replacing $S_{n,j}[\tau[q]], j = g-1, g$, in (6.17) with $a^{(\mathrm{i})}_{L[q]}$ and $(j/n)^{\chi}$ at the costs above, and using $\beta_{q+1} < \beta_q$, we conclude that, on the event $\mathrm{E}_q \cap \mathcal{H}_{q+1,1}$, there is a constant $\widehat{C} := \widehat{C}(x_1, \mu, q)$ such that $|a^{(\mathrm{i})}_{L[q],1} - (g/n)^{\chi}| \leqslant \widehat{C} n^{-\beta_{q+1}}$. Now, pick $C_{q+1} := \widehat{C} \vee \widehat{c}$ for the event $\mathcal{K}_{\tau[q],1}$ and $\mathcal{H}_{q+1,2}$. By Lemma 13, there is a positive constant $C := C(x_1, \mu, \kappa, q)$ such that

$$\mathbb{P}_{\mathbf{x}}\left[\overline{\mathcal{H}}_{q,l} \cap \bigcap_{j=1}^{3} F_{\tau[q],j} \cap \mathcal{H}_{q+1,1} \cap \mathcal{K}^c_{\tau[q],1}\right] \leqslant \mathbb{P}_{\mathbf{x}}\left[\left\{\mathbf{Y}^{(v[\tau[q]],n)}_{\mathrm{Be}} \neq \mathbf{V}^{(v[\tau[q]],n)}_{\mathrm{Po}}\right\} \cap \bigcap_{j=1}^{3} F_{\tau[q],j} \cap \overline{\mathcal{H}}_{q,l}\right] \\ \leqslant C n^{-d} (\log\log n)^{q+1} (\log n)^{q/r}, \tag{6.19}$$

where $d = \min\{\beta_q, \gamma, 1-\chi\}$. The lemma follows from applying (6.15), (6.16), and (6.19) to (6.14). □

The following analogue of Lemma 10 shows that under the graph coupling, w.h.p. $v[\tau[q]]$ has at most $(\log n)^{1/r}$ children. We omit the proof as it is similar to that of Lemma 10.

Lemma 14. *Retaining the assumptions and the notation in Lemma 11, let $\mathcal{K}_{\tau[q],j}$ and $\overline{\mathcal{H}}_{q,j}$ be as in (6.2) and (6.5). Then, given positive integers $r < \infty$ and $q < r$, there is a coupling of $\left(G_n, k^{(\mathrm{u})}_0\right)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$, with an integer $p > 0$ and a constant $C := C(\kappa, p) > 0$ such that*

$$\mathbb{P}_{\mathbf{x}}\left[\overline{\mathcal{H}}_{q,l} \cap \mathcal{K}_{q,1} \cap \mathcal{K}^c_{q,2}\right] \leqslant C(\log n)^{-\frac{p}{r}} (\log\log n)^{\frac{p(q+1)}{\mu+1}} \quad \textit{for all } n \geqslant 3.$$

We now apply Lemmas 12 and 14 to prove Lemma 11.

Proof of Lemma 11. We begin by stating the type-R analogues of Lemmas 12 and 14. Suppose that $\left(G_n, k^{(\mathrm{u})}_0\right)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$ are coupled so that for a type-R child $v[t] = k^{(\mathrm{u})}_{\bar{w}} \in \partial\mathcal{B}_q$, the event

$$\mathcal{H}_{q+1,1} \cap \mathcal{J}_{q,l,t} := \mathcal{H}_{q+1,1} \cap \overline{\mathcal{H}}_{q,l} \cap \bigcap_{\{s<t\,:\, v[s]\in\partial\mathcal{B}_q\}} (\mathcal{K}_{s,1} \cap \mathcal{K}_{s,2}) \tag{6.20}$$

has occurred for some $l \geqslant 2$, where $\mathcal{H}_{q+1,1}$, $\overline{\mathcal{H}}_{q,l}$, $\mathcal{K}_{s,1}$, and $\mathcal{K}_{s,2}$ are as in (6.1), (6.5), and (6.2). Let $\left(\left(\mathcal{Z}_j[t], \widetilde{\mathcal{Z}}_{j-1}[t]\right), j \in \mathcal{A}_{t-1} \cup \mathcal{N}_{t-1}\right)$ and $\left(S_{n,j}[t], 1 \leqslant j \leqslant n\right)$ be as in (3.7), (3.8), and (3.10). We use these variables to generate the type-R children of $k^{(\mathrm{u})}_{\bar{w}}$, and to sample the points

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

$\left(a^{(\mathrm{i})}_{\bar{w},i},\ 1 \leqslant i \leqslant R^{(\mathrm{i})}_{\bar{w}}\right)$ of the mixed Poisson process with intensity (6.4), where ζ_q and $a^{(\mathrm{i})}_{L[q]}$ are replaced with $\mathcal{Z}_{k^{(\mathrm{i})}_{\bar{w}}}[t] \sim \mathrm{Gamma}\left(x_{k^{(\mathrm{i})}_{\bar{w}}}, 1\right)$ and $a^{(\mathrm{i})}_{\bar{w}}$. Note that $v[t] > k^{(\mathrm{u})}_{L[q]}$, and on the event $\mathcal{J}_{q,l,t}$, the number of discovered vertices up to time t can be bounded as

$$|\mathcal{A}_{t-1} \cup \mathcal{P}_{t-1}| \leqslant 2 + q + (q+1)^2(\log n)^{(q+1)/r},$$

as implied by (6.13). Hence, we can proceed similarly as for Lemmas 12 and 13. In particular, there is a coupling of $\left(G_n, k^{(\mathrm{u})}_0\right)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$, with positive constants $C := C(x_1, \mu, \kappa, q)$ and $c := c(\kappa, p)$ such that for $n \geqslant 3$,

$$\begin{aligned} \mathbb{P}_{\mathbf{x}}\left[\mathcal{K}^c_{t,1} \cap \mathcal{J}_{q,l,t}\right] &\leqslant Cn^{-d}(\log\log n)^{q+1}(\log n)^{\frac{q+1}{r}}, \\ \mathbb{P}_{\mathbf{x}}\left[\mathcal{K}^c_{t,2} \cap \mathcal{K}_{t,1} \cap \mathcal{J}_{q,l,t}\right] &\leqslant c(\log n)^{-\frac{p}{r}}(\log\log n)^{\frac{p(q+1)}{\mu+1}}, \end{aligned} \tag{6.21}$$

where $d = \min\{\chi/3 - 4\gamma, 1-\chi, \gamma, \beta_q\}$. Define $\widetilde{\mathcal{H}}_q := \bigcap_{j=1}^3 \mathcal{H}_{q,j}$. To bound $\mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{H}}^c_{q+1}\right]$ using Lemma 12, Lemma 14, and (6.21), we use

$$\begin{aligned} \mathbb{P}_{\mathbf{x}}\left[\left(\widetilde{\mathcal{H}}_{q+1}\right)^c\right] &\leqslant \mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{H}}^c_{q+1} \cap \widetilde{\mathcal{H}}_q\right] + \mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{H}}^c_q\right] \\ &= \mathbb{P}_{\mathbf{x}}\left[\left(\bigcap_{j=2}^3 \mathcal{H}_{q+1,j}\right)^c \cap \mathcal{H}_{q+1,1} \cap \widetilde{\mathcal{H}}_q\right] + \mathbb{P}_{\mathbf{x}}\left[\mathcal{H}^c_{q+1,1} \cap \widetilde{\mathcal{H}}_q\right] + \mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{H}}^c_q\right] \\ &= \mathbb{P}_{\mathbf{x}}\left[\left(\bigcap_{\{s\,:\,v[s]\in\partial\mathcal{B}_q\}} (\mathcal{K}_{s,1} \cap \mathcal{K}_{s,2})\right)^c \cap \mathcal{H}_{q+1,1} \cap \widetilde{\mathcal{H}}_q\right] \qquad (6.22) \\ &\qquad + \mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{H}}_q \cap \mathcal{H}^c_{q+1,1}\right] + \mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{H}}^c_q\right], \qquad (6.23) \end{aligned}$$

where the last equality follows from

$$\left(\bigcap_{j=2}^3 \mathcal{H}_{q+1,j}\right)^c \cap \widetilde{\mathcal{H}}_q = \left(\bigcap_{\{s\,:\,v[s]\in\partial\mathcal{B}_q\}} (\mathcal{K}_{s,1} \cap \mathcal{K}_{s,2})\right)^c \cap \widetilde{\mathcal{H}}_q.$$

The lemma is proved once we show that the probabilities in (6.22) and (6.23) are of order at most $(\log\log n)^{-\chi}$. By assumption, there is a constant $C := C(x_1, \mu, \kappa, q)$ such that $\mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{H}}^c_q\right] \leqslant C(\log\log n)^{-\chi}$. Recall that $a^{(\mathrm{i})}_{L[q],1} = U_{L[q],1} a^{(\mathrm{i})}_{L[q]}$, where $U_{L[q],1} \sim \mathrm{U}[0, 1]$ is independent of $a^{(\mathrm{i})}_{L[q]}$. Thus, similarly to (5.13),

$$\mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{H}}_q \cap \mathcal{H}^c_{q+1,1}\right] \leqslant (\log\log n)^{-\chi}.$$

For (6.22), define

$$\widetilde{\mathcal{K}}_{q,s,j} = \mathcal{K}_{\left|V\left(B_q\left(G_n, k^{(\mathrm{u})}_0\right)\right)\right| + s, j}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

for $j=1, 2$ and $1 \leqslant s \leqslant |\partial\mathcal{B}_q|$. Note that $|\partial\mathcal{B}_q| \leqslant c[n, q] := \lfloor 1 + q(\log n)^{\frac{q}{r}} \rfloor$ on the event $\mathcal{H}_{q,3}$, and so

$$\mathbb{P}_{\mathbf{x}}\left[\left(\bigcap_{\{s\,:\, v[s]\in\partial\mathcal{B}_q\}} (\mathcal{K}_{s,1}\cap\mathcal{K}_{s,2})\right)^c \cap \mathcal{H}_{q+1,1}\cap\widetilde{\mathcal{H}}_q\right]$$
$$= \mathbb{P}_{\mathbf{x}}\left[\bigcup_{l=1}^{c[n,q]}\left\{\{|\partial\mathcal{B}_q| = l\}\cap\left(\bigcap_{s=1}^{l}\left(\widetilde{\mathcal{K}}_{q,s,1}\cap\widetilde{\mathcal{K}}_{q,s,2}\right)\right)^c\right\}\cap\mathcal{H}_{q+1,1}\cap\widetilde{\mathcal{H}}_q\right].$$

By a union bound, and recalling the definition of $\overline{\mathcal{H}}_{q,l}$ in (6.5), the probability above is at most

$$\sum_{l=1}^{c[n,q]}\left\{\mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{K}}^c_{q,1,2}\cap\widetilde{\mathcal{K}}_{q,1,1}\cap\mathcal{H}_{q+1,1}\cap\overline{\mathcal{H}}_{q,l}\right] + \mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{K}}^c_{q,1,1}\cap\mathcal{H}_{q+1,1}\cap\overline{\mathcal{H}}_{q,l}\right]\right.$$
$$\left. + \sum_{s=2}^{l}\mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{K}}^c_{q,s,2}\cap\widetilde{\mathcal{K}}_{q,s,1}\cap\mathcal{H}_{q+1,1}\cap\mathcal{J}_{q,l,s}\right] + \mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{K}}^c_{q,s,1}\cap\mathcal{H}_{q+1,1}\cap\mathcal{J}_{q,l,s}\right]\right\}. \tag{6.24}$$

We bound (6.24) using (6.21), Lemma 12, and Lemma 14, where we choose $p > 2(r-1)$ in Lemma 14 so that

$$\sum_{l=1}^{c[n,q]}\sum_{s=2}^{l}\mathbb{P}_{\mathbf{x}}\left[\widetilde{\mathcal{K}}^c_{q,s,2}\cap\widetilde{\mathcal{K}}_{q,s,1}\cap\mathcal{J}_{q,l,s}\right]\to 0$$

as $n\to\infty$. It follows that there is a constant $C := C(x_1, \mu, \kappa, q)$ such that (6.24) is bounded by $C(\log\log n)^{-\chi}$. □

7. Completion of the local weak limit proof

Equipped with Corollary 2, we can prove Theorem 1.

Proof of Theorem 1. In view of (1.1), it is enough to establish (1.5) of Theorem 1. Let $G'_n \sim \mathrm{PA}(\pi, X_1)_n$ (Definition 1), let k_0 be its randomly chosen vertex, let $(\mathcal{T}_{\mathbf{X},n}, 0)$ be the intermediate Pólya point tree $(\mathcal{T}_{\mathbf{x},n}, 0)$ randomised over $\mathbf{X}$, and let $(\mathcal{T}, 0)$ be the π-Pólya point tree in Definition 3. By the triangle inequality for the total variation distance, for any $r < \infty$ we have

$$d_{\mathrm{TV}}\left(\mathscr{L}\left(\left(B_r\left(G'_n, k_0\right), k_0\right)\right), \mathscr{L}((B_r(\mathcal{T}, 0), 0))\right)$$
$$\leqslant d_{\mathrm{TV}}\left(\mathscr{L}\left(\left(B_r\left(G'_n, k_0\right), k_0\right)\right), \mathscr{L}\left((B_r(\mathcal{T}_{\mathbf{X},n}, 0), 0)\right)\right)$$
$$+ d_{\mathrm{TV}}\left(\mathscr{L}\left((B_r(\mathcal{T}_{\mathbf{X},n}, 0), 0)\right), \mathscr{L}((B_r(\mathcal{T}, 0), 0))\right).$$

Let $A_n := A_{2/3,n}$ be as in (2.1). Applying Jensen's inequality to the total variation distance, the above is bounded by

$$\mathbb{E}\left[d_{\mathrm{TV}}\left(\mathscr{L}\left(\left(B_r\left(G'_n, k_0\right), k_0\right)|\mathbf{X}\right), \mathscr{L}\left((B_r(\mathcal{T}_{\mathbf{X},n}, 0), 0)|\mathbf{X}\right)\right)\right]$$
$$+ d_{\mathrm{TV}}\left(\mathscr{L}\left((B_r(\mathcal{T}_{\mathbf{X},n}, 0), 0)\right), \mathscr{L}((B_r(\mathcal{T}, 0), 0))\right)$$
$$\leqslant \mathbb{E}\left[\mathbb{1}[A_n] d_{\mathrm{TV}}\left(\mathscr{L}\left(\left(B_r\left(G'_n, k_0\right), k_0\right)|\mathbf{X}\right), \mathscr{L}\left((B_r(\mathcal{T}_{\mathbf{X},n}, 0), 0)|\mathbf{X}\right)\right)\right]$$
$$+ \mathbb{P}\left[A_n^c\right] + d_{\mathrm{TV}}\left(\mathscr{L}\left(\left(B_r\left(\mathcal{T}_{\mathbf{X},n}, 0\right), 0\right)\right), \mathscr{L}((B_r(\mathcal{T}, 0), 0))\right).$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

We prove that each term on the right-hand side is of order at most $(\log\log n)^{-\chi}$, starting from the expectation. Let $\mathcal{H}_{r,j}, j=1, 2, 3$, be as in (6.1). For the $(\mathbf{x}, n)$-Pólya urn tree G_n with $\mathbf{x}\in A_n$, Corollary 2 implies that there is a coupling of $\big(G_n, k_0^{(\mathrm{u})}\big)$ and $(\mathcal{T}_{\mathbf{x},n}, 0)$, with a positive constant $C := C(x_1, \mu, \kappa, r)$ such that

$$\mathbb{P}_{\mathbf{x}}\left[B_r(\mathcal{T}_{\mathbf{x},n}, 0), 0 \ncong \left(B_r\left(G_n, k_0^{(\mathrm{u})}\right), k_0^{(\mathrm{u})}\right)\right] \leqslant \mathbb{P}_{\mathbf{x}}\left[\left(\bigcap_{j=1}^{3}\mathcal{H}_{r,j}\right)^c\right] \leqslant C(\log\log n)^{-\chi}; \tag{7.1}$$

following from the definition (1.4) of the total variation distance, the expectation is at most $C(\log\log n)^{-\chi}$. The probability $\mathbb{P}[A_n^c]$ can be bounded using Lemma 1 (with $\alpha=2/3$). We now couple $(B_r(\mathcal{T}, 0), 0)$ and $(B_r(\mathcal{T}_{\mathbf{X},n}, 0), 0)$ to bound the last term. For this coupling, denote by E the event that the PA labels in $B_r(\mathcal{T}_{\mathbf{X},n}, 0)$ are distinct, that is, $k_{\bar{w}}^{(\mathrm{i})} \neq k_{\bar{v}}^{(\mathrm{i})}$ for any $\bar{w}\neq\bar{v}$. We first construct $B_r(\mathcal{T}_{\mathbf{X},n}, 0)$; then on the event E, we set $B_r(\mathcal{T}, 0)$ as $B_r(\mathcal{T}_{\mathbf{X},n}, 0)$, inheriting the Ulam–Harris labels, ages, types, and fitness from the latter. If the PA labels are not distinct, we generate $B_r(\mathcal{T}, 0)$ independently from $B_r(\mathcal{T}_{\mathbf{X},n}, 0)$. Under this coupling,

$$\mathbb{P}[(B_r(\mathcal{T}_{\mathbf{X},n}, 0), 0)\ncong(B_r(\mathcal{T}, 0), 0)] = \mathbb{P}[\{(B_r(\mathcal{T}_{\mathbf{X},n}, 0), 0)\ncong(B_r(\mathcal{T}, 0), 0)\}\cap \mathrm{E}^c]. \tag{7.2}$$

From the definition of $\mathcal{H}_{r,2}$ in (6.1), we can use (7.1) and Lemma 1 to bound the second term in (7.2) by $\mathbb{P}[\mathrm{E}^c]\leqslant\mathbb{P}\left[\left(\bigcap_{j=1}^{3}\mathcal{H}_{r,j}\right)^c\right]\leqslant C(\log\log n)^{-\chi}$. So, once again by (1.4),

$$d_{\mathrm{TV}}\left(\mathscr{L}\big((B_r(\mathcal{T}_{\mathbf{X},n}, 0), 0)\big), \mathscr{L}((B_r(\mathcal{T}, 0), 0))\right)\leqslant C(\log\log n)^{-\chi},$$

which concludes the proof. □

8. Remarks on the limiting distributions of the degree statistics

In this section, we discuss the connection of Corollary 1 and Theorem 2 to [4, 7, 20, 25, 30]. To this end, we state the probability mass function of the limiting distributions of the degrees of the uniformly chosen vertex k_0 in $\mathrm{PA}(\pi, X_1)_n$ and its type-L neighbour $k_{L[1]}$. These results are direct consequences of Corollary 1 and Theorem 2 (see [24] for a detailed proof). We also show that both probability mass functions also exhibit power-law behaviour, which follows from a dominated convergence argument (see [25]). Below we write $a_n\sim b_n$ to indicate $\lim_{n\to\infty} a_n/b_n = 1$. We start with the degree of the uniformly chosen vertex.

Proposition 1. *Retaining the assumptions and the notation in Theorem 2, let ξ_0 be as in the theorem. The probability mass function of the distribution of ξ_0 is*

$$p_\pi(j) = (\mu+1)\int_0^\infty \frac{\Gamma(x+j-1)\Gamma(x+\mu+1)}{\Gamma(x)\Gamma(x+\mu+j+1)}d\pi(x), \qquad j\geqslant 1. \tag{8.1}$$

Furthermore, if $\mathbb{E}X_2^{\mu+1}<\infty$, then as $j\to\infty$,

$$p_\pi(j)\sim C_\pi j^{-(\mu+2)}, \qquad C_\pi := (\mu+1)\int_0^\infty \frac{\Gamma(x+\mu+1)}{\Gamma(x)}d\pi(x). \tag{8.2}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

Remark 4. We use Proposition 1 to relate Theorem 2 to some known results.

(i) When π is a point mass at 1, $(p_\pi(j), j \geqslant 1)$ is the probability mass function of $\mathrm{Geo}_1(\sqrt{U})$ (also known as the Yule–Simon distribution), which is a mixture of geometric distributions on the positive integers, with the parameter sampled according to the distribution of $\sqrt{U}$, where $U \sim \mathrm{U}(0, 1)$. For such a fitness sequence and a model that allows for self-loops, [7] established that the limiting empirical degree distribution is $(p_\pi(j), j \geqslant 1)$; furthermore, using Stein's method, [30, Theorem 6.1] showed that the total variation distance is at most $n^{-1} \log n$.

(ii) In the multiple-edge setting, and assuming only the fitness has a finite mean, [25] used stochastic approximation to obtain the a.s. limit of the empirical degree distribution. Theorem 2 and Proposition 1 are special cases of [25, Theorems 2.4 and 2.6], but the distributional representation and the rate in Theorem 2 are new.

In the following, we give the limiting probability mass function of the degree of the type-L neighbour $k_{L[1]}$ of the uniformly chosen vertex k_0, which shows that the distribution also exhibits power-law behaviour. When the PA tree has constant initial attractiveness, the proposition is a special case of [4, Lemma 5.2] and [20, Lemma 5.9].

Proposition 2. *Retaining the assumptions and the notation in Theorem 1, let $R_{L[1]}$ be as in the theorem. The probability mass function of the random variable $R_{L[1]} + 2$ in the theorem is given by*

$$q_\pi(j) = \mu(\mu+1)(j-1) \int_0^\infty \frac{\Gamma(x+j-1)\Gamma(x+\mu+1)}{\Gamma(x+1)\Gamma(x+\mu+j+1)} d\pi(x), \qquad j \geqslant 2. \tag{8.3}$$

Furthermore, if $\mathbb{E}X_2^\mu < \infty$, then as $j \to \infty$,

$$q_\pi(j) \sim C_\pi j^{-(\mu+1)}, \qquad C_\pi = \mu(\mu+1) \int_0^\infty \frac{\Gamma(x+\mu+1)}{\Gamma(x+1)} d\pi(x).$$

Comparing Propositions 1 and 2, we see that in the limit, the degree of vertex $k_{L[1]}$ has a heavier tail than the degree of k_0. This is due to the fact that $k_{L[1]}$ has received an incoming edge from k_0, and so $k_{L[1]}$ is more likely to have a higher degree than k_0.

9. Proofs for the approximation results

Here we prove the lemmas in Section 2. The arguments are adapted from [33] and have a similar flavour to the arguments of [6], both of which studied different PA models. For the proofs below, we recall that $\phi(n) = \Omega(n^\chi)$, where χ is as in (1.2).

Proof of Lemma 1. Given $p > 2$, choose $1/2 + 1/p < \alpha < 1$. Let $A_{\alpha,n}$ be as in (2.1), let $T^\star_m := \sum_{i=2}^m X_i$, and let C_p be the positive constant given in Lemma 15 below, which bounds the moment of a sum of variables in terms of the moments of the summands. Then

$$\begin{aligned}\mathbb{P}\big[A^c_{\alpha,n}\big] &= \mathbb{P}\Bigg[\bigcup_{j=\lceil\phi(n)\rceil}^\infty \big\{\big|T^\star_j - (j-1)\mu\big| > j^\alpha\big\}\Bigg] \\ &\leqslant \sum_{j=\lceil\phi(n)\rceil}^\infty \mathbb{P}\big[\big|T^\star_j - (j-1)\mu\big| > j^\alpha\big] \quad \text{(by a union bound)}\end{aligned}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

$$\leqslant \sum_{j=\lceil\phi(n)\rceil}^{\infty} \mathbb{E}\big[\big|T_j^\star-(j-1)\mu\big|^p\big]j^{-\alpha p} \quad \text{(by Chebyshev's inequality)}$$
$$\leqslant C_p\mathbb{E}\big[|X_2-\mu|^p\big]\sum_{j=\lceil\phi(n)\rceil}^{\infty} j^{-p(\alpha-1/2)} \quad \text{(by Lemma 15)}$$
$$\leqslant C_p\mathbb{E}\big[|X_2-\mu|^p\big]\int_{\lceil\phi(n)\rceil-1}^{\infty} y^{-p(\alpha-1/2)}dy$$
$$= C_p[p(\alpha-1/2)-1]^{-1}\mathbb{E}\big[|X_2-\mu|^p\big](\lceil\phi(n)\rceil-1)^{1-p(\alpha-1/2)},$$

where $p(\alpha-1/2)>1$ and $\mathbb{E}[|X_2-\mu|^p]<\infty$. The lemma follows from the fact that $\phi(n)=\Omega(n^\chi)$. □

The next lemma can be found in [34, Item 16, p.60], where [13] is credited.

Lemma 15. *Let $Y_1,\ldots,Y_n$ be independent random variables such that for $i=1,\ldots,n$, $\mathbb{E}Y_i=0$ and $\mathbb{E}|Y_i|^p<\infty$ for some $p\geqslant 2$. Let $W_n:=\sum_{i=1}^n Y_i$; then*

$$\mathbb{E}\big[|W_n|^p\big]\leqslant C_pn^{p/2-1}\sum_{i=1}^n \mathbb{E}\big[|Y_i|^p\big],$$

where

$$C_p:=\frac{1}{2}p(p-1)\max\left(1,2^{p-3}\right)\left[1+\frac{2}{p}K_{2m}^{(p-2)/2m}\right],$$

and the integer m satisfies the condition $2m\leqslant p\leqslant 2m+2$ and $K_{2m}=\sum_{i=1}^m \frac{i^{2m-1}}{(i-1)!}$.

Keeping the notation $T_i:=\sum_{h=1}^i x_h$ and $\phi(n)=\Omega(n^\chi)$, we now prove Lemma 2 under the assumption that the event $A_{\alpha,n}$ holds for the realisation of the fitness sequence $\mathbf{x}$ (written as $\mathbf{x}\in A_{\alpha,n}$). We use the subscript $\mathbf{x}$ in $\mathbb{P}_{\mathbf{x}}$ and $\mathbb{E}_{\mathbf{x}}$ to indicate the conditioning on $\mathbf{X}=\mathbf{x}$. The first step is to derive an expression for $\mathbb{E}_{\mathbf{x}}[S_{n,k}]$ that holds for any realisation of the fitness sequence, where we modify a moment formula used in proving [36, Proposition 1.3].

Lemma 16. *Let T_i be as above, and let B_i, $S_{n,i}$ be as in (1.8) and (1.9). Then for $1\leqslant k<n$ and a positive integer p,*

$$\mathbb{E}_{\mathbf{x}}\big[S_{n,k}^p\big]=\left[\prod_{h=0}^{p-1}\frac{T_k+k+h}{T_n+n-1+h}\right]\prod_{j=0}^{p-1}\prod_{i=k+1}^{n-1}\left[1+\frac{1}{T_i+i-1+j}\right]. \tag{9.1}$$

Proof. Since $(B_i,\ 1\leqslant i\leqslant n)$ are independent beta random variables, we use the moment formula of the beta distribution to show that for $p\geqslant 1$,

$$\mathbb{E}_{\mathbf{x}}\big[S_{n,k}^p\big]=\prod_{i=k+1}^n \mathbb{E}\left[(1-B_i)^p\right]=\prod_{i=k+1}^n\prod_{j=0}^{p-1}\frac{T_{i-1}+i-1+j}{T_i+i-1+j},$$
$$=\prod_{j=0}^{p-1}\left\{\frac{(T_k+k+j)}{(T_n+n-1+j)}\,\frac{(T_n+n-1+j)}{(T_k+k+j)}\prod_{i=k+1}^n\frac{T_{i-1}+i-1+j}{T_i+i-1+j}\right\}.$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

Noting that T_k+k+j and $T_n+n-1+j$ in the second product above cancel with $(T_n+n-1+j)/(T_k+k+j)$, we can rewrite the final term as

$$\begin{aligned}\mathbb{E}_{\mathbf{x}}\big[S_{n,k}^p\big] &= \left[\prod_{h=0}^{p-1}\frac{T_k+k+h}{T_n+n-1+h}\right]\prod_{i=k+1}^{n-1}\prod_{j=0}^{p-1}\frac{T_i+i+j}{T_i+i-1+j}\\ &= \left[\prod_{h=0}^{p-1}\frac{T_k+k+h}{T_n+n-1+h}\right]\prod_{j=0}^{p-1}\prod_{i=k+1}^{n-1}\left[1+\frac{1}{T_i+i-1+j}\right],\end{aligned}$$

hence concluding the proof. □

Taking $k=1$ in (9.1) recovers the original formula of [36], where T_i here is A_i in [36]. Next, we use the moment formula to show that when $\mathbf{x}\in A_{\alpha,n}$, the difference between the mean of $S_{n,k}$ and $(k/n)^\chi$ is small enough for large n and $k\geqslant\lceil\phi(n)\rceil$.

Lemma 17. *Given $1/2<\alpha<1$ and a positive integer n, assume that $\mathbf{x}\in A_{\alpha,n}$. Then there is a positive constant $C:=C(x_1,\mu,\alpha)$ such that for all $\lceil\phi(n)\rceil\leqslant k\leqslant n$,*

$$\left|\mathbb{E}_{\mathbf{x}}[S_{n,k}]-\left(\frac{k}{n}\right)^\chi\right|\leqslant Cn^{\chi(\alpha-1)}. \tag{9.2}$$

Proof. We first prove the upper bound for $\mathbb{E}_{\mathbf{x}}[S_{n,k}]$, using the techniques for proving [33, Lemma 4.4]. Applying the formula (9.1) (with $p=1$), for $\mathbf{x}\in A_{\alpha,n}$ and $k\geqslant\lceil\phi(n)\rceil$, we obtain

$$\mathbb{E}_{\mathbf{x}}[S_{n,k}]\leqslant\frac{k(\mu+1)+k^\alpha+b}{n(\mu+1)-n^\alpha+b-1}\prod_{i=k+1}^{n-1}\left[1+\frac{1}{i\mu-i^\alpha+i+b-1}\right], \tag{9.3}$$

where $b:=x_1-\mu$. We rewrite the first term on the right-hand side of (9.3) as follows:

$$\begin{aligned}\left(\frac{k}{n}\right)\frac{\mu+1+k^{-1+\alpha}+k^{-1}b}{\mu+1-n^{-1+\alpha}+n^{-1}(b-1)} &= \left(\frac{k}{n}\right)\left[1+\frac{k^{\alpha-1}+n^{\alpha-1}+(k^{-1}-n^{-1})b+n^{-1}}{\mu+1-n^{\alpha-1}+n^{-1}(b-1)}\right]\\ &\leqslant\frac{k}{n}\big(1+\overline{C}k^{\alpha-1}\big)\\ &\leqslant\frac{k}{n}\big(1+\overline{C}n^{\chi(\alpha-1)}\big),\end{aligned}$$

where $\overline{C}:=\overline{C}(x_1,\mu,\alpha)$ is some positive constant. To bound the product term on the right-hand side of (9.3), we take the logarithm and bound

$$\left|\sum_{i=k+1}^{n-1}\log\left(1+\frac{1}{i(\mu+1)-i^\alpha+b-1}\right)-\frac{1}{\mu+1}\log\left(\frac{n}{k}\right)\right|.$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

By the triangle inequality, we have

$$\left|\sum_{i=k+1}^{n-1}\log\left(1+\frac{1}{i\mu-i^{\alpha}+i+b-1}\right)-\frac{1}{\mu+1}\log\left(\frac{n}{k}\right)\right|$$

$$\leqslant\left|\sum_{i=k+1}^{n-1}\log\left(1+\frac{1}{i\mu-i^{\alpha}+i+b-1}\right)-\frac{1}{i(\mu+1)-i^{\alpha}+b-1}\right| \tag{9.4}$$

$$+\left|\sum_{j=k+1}^{n-1}\frac{1}{j(\mu+1)-j^{\alpha}+b-1}-\frac{1}{\mu+1}\log\left(\frac{n}{k}\right)\right|. \tag{9.5}$$

We use the fact that $y-\log(1+y)\leqslant y^2$ for $y\geqslant 0$ to bound (9.4). Letting $y_i=(i(\mu+1)-i^{\alpha}+b-1)^{-1}$, this implies that (9.4) is bounded by $\sum_{i=k+1}^{n}y_i^2$ for $k\geqslant\lceil\phi(n)\rceil$ and n large enough, and by an integral comparison,

$$\sum_{i=k+1}^{n}y_i^2=O(n^{-\chi}).$$

For (9.5), we have

$$\left|\sum_{i=k+1}^{n-1}\frac{1}{i(\mu+1)-i^{\alpha}+b-1}-\frac{1}{(\mu+1)}\log\left(\frac{n}{k}\right)\right|$$

$$=\left|\sum_{i=k+1}^{n-1}\left(\frac{1}{i(\mu+1)-i^{\alpha}+b-1}-\frac{1}{(\mu+1)i}\right)+O\big(k^{-1}\big)\right|$$

$$\leqslant\sum_{i=k+1}^{n-1}\left|\frac{i^{\alpha}-b+1}{i(\mu+1)(i(\mu+1)-i^{\alpha}+b-1)}\right|+O\big(k^{-1}\big)$$

$$\leqslant C'\sum_{i=k+1}^{n-1}i^{-2+\alpha}+O\big(k^{-1}\big)$$

$$\leqslant C'(1-\alpha)^{-1}[\phi(n)]^{\alpha-1}+O\big(n^{-\chi}\big),$$

where $C':=C'(x_1,\mu,\alpha)$ is some positive constant. Combining the bounds above, a little calculation shows that there are positive constants $\overline{C}:=\overline{C}(x_1,\mu,\alpha)$ and $\widetilde{C}:=\widetilde{C}(x_1,\mu,\alpha)$ such that for $\mathbf{x}\in A_{\alpha,n}$ and $\lceil\phi(n)\rceil\leqslant k\leqslant n$,

$$\mathbb{E}_{\mathbf{x}}[S_{n,k}]\leqslant(k/n)^{\chi}\left(1+\overline{C}n^{\chi(\alpha-1)}\right)\exp\left\{\widetilde{C}n^{\chi(\alpha-1)}\right\}.$$

Since $e^x=1+x+O(x^2)$ for x near zero, there is a positive constant $C:=C(x_1,\mu,\alpha)$ such that for $\lceil\phi(n)\rceil\leqslant k\leqslant n$ and n large enough,

$$\mathbb{E}_{\mathbf{x}}[S_{n,k}]\leqslant(k/n)^{\chi}\left(1+Cn^{\chi(\alpha-1)}\right)\leqslant(k/n)^{\chi}+Cn^{\chi(\alpha-1)},$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

hence proving the desired upper bound. The lower bound can be proved by noting that for $\lceil \phi(n) \rceil \leqslant k \leqslant n$,

$$\mathbb{E}_{\mathbf{x}}[S_{n,k}] \geqslant \frac{k(\mu+1) - k^{\alpha} + b}{n(\mu+1) + n^{\alpha} + b - 1} \prod_{i=k+1}^{n} \left[1 + \frac{1}{i\mu + i^{\alpha} + i + b - 1}\right],$$

and repeating the calculations above. The details are omitted. □

With Lemma 17 and a martingale argument, we can prove Lemma 2.

Proof of Lemma 2. Let $\hat{\delta}_n = Cn^{\chi(\alpha-1)/4}$, where $C := C(x_1, \mu, \alpha)$ is the positive constant in (9.2) of Lemma 17. Writing $K := \lceil \phi(n) \rceil$, define

$$\widetilde{E}_{n,\hat{\delta}_n} := \left\{ \max_{K \leqslant k \leqslant n} \left| S_{n,k} - \left(\frac{k}{n}\right)^{\chi} \right| \geqslant 2\hat{\delta}_n \right\}.$$

The lemma follows from bounding $\mathbb{P}_{\mathbf{x}}\left[\widetilde{E}_{n,\hat{\delta}_n}\right]$ under the assumption $\mathbf{x} \in A_{\alpha,n}$. By the triangle inequality, we have

$$\mathbb{P}_{\mathbf{x}}\left[\widetilde{E}_{n,\hat{\delta}_n}\right] \leqslant \mathbb{P}_{\mathbf{x}}\left[\max_{K \leqslant k \leqslant n} \left|S_{n,k} - \mathbb{E}_{\mathbf{x}}[S_{n,k}]\right| + \max_{K \leqslant j \leqslant n} \left|\mathbb{E}_{\mathbf{x}}[S_{n,j}] - (j/n)^{\chi}\right| \geqslant 2\hat{\delta}_n \right].$$

Applying Lemma 17 to bound the difference between $(j/n)^{\chi}$ and $\mathbb{E}_{\mathbf{x}}[S_{n,j}]$, we obtain

$$\mathbb{P}_{\mathbf{x}}\left[\widetilde{E}_{n,\hat{\delta}_n}\right] \leqslant \mathbb{P}_{\mathbf{x}}\left[\max_{K \leqslant k \leqslant n} \left|S_{n,k} - \mathbb{E}_{\mathbf{x}}[S_{n,k}]\right| \geqslant \hat{\delta}_n \right] \leqslant \mathbb{P}_{\mathbf{x}}\left[\max_{K \leqslant k \leqslant n} \left|S_{n,k}(\mathbb{E}_{\mathbf{x}}[S_{n,k}])^{-1} - 1\right| \geqslant \hat{\delta}_n \right],$$

where the second inequality is due to $\mathbb{E}_{\mathbf{x}}[S_{n,k}] \leqslant 1$. We bound the right-hand side of the above using a martingale argument. Recall that $\mathbb{E}_{\mathbf{x}}[S_{n,k}] = \prod_{j=k+1}^{n} \mathbb{E}[1 - B_j]$. Define $M_0 := 1$, and for $j = 1, \ldots, n-K$, let

$$M_j := \prod_{i=n-j+1}^{n} \frac{1 - B_i}{\mathbb{E}[1 - B_i]} = \frac{S_{n,n-j}}{\mathbb{E}[S_{n,n-j}]}.$$

Let $\mathcal{F}_j$ be the σ-algebra generated by $(B_i, n-j+1 \leqslant i \leqslant n)$ for $1 \leqslant j \leqslant n-K$, with $\mathcal{F}_0 = \varnothing$. It follows that $\big((M_j, \mathcal{F}_j), 0 \leqslant j \leqslant n-K\big)$ is a martingale with $\mathbb{E}[M_j] = 1$. Since $(M_j - 1)^2$ is a submartingale, Doob's inequality [14, Theorem 4.4.2, p.204] yields

$$\mathbb{P}_{\mathbf{x}}\left[\widetilde{E}_{n,\hat{\delta}_n}\right] \leqslant \mathbb{P}_{\mathbf{x}}\left[\max_{0 \leqslant j \leqslant n-K} \left|M_j - 1\right| \geqslant \hat{\delta}_n \right] \leqslant \hat{\delta}_n^{-2} \mathrm{Var}_{\mathbf{x}}(M_{n-K}), \tag{9.6}$$

where $\mathrm{Var}_{\mathbf{x}}(M_{n-K})$ is the variance conditional on $\mathbf{x}$. We use the formulas for the first and second moments of the beta distribution to bound the variance:

$$\begin{aligned} \mathrm{Var}_{\mathbf{x}}(M_{n-K}) = \mathbb{E}_{\mathbf{x}}\left[(M_{n-K})^2\right] - 1 &= \left[\prod_{j=K+1}^{n} \frac{(T_{j-1}+j)}{(T_{j-1}+j-1)} \frac{(T_j+j-1)}{(T_j+j)} \right] - 1 \\ &= \prod_{j=K+1}^{n} \left[1 + \frac{T_j - T_{j-1}}{(T_j+j)(T_{j-1}+j-1)}\right] - 1. \end{aligned}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

Below we allow the positive constant $C' = C'(x_1, \mu, \alpha)$ to vary from line to line. As $\left|\sum_{i=2}^{j} x_i - (j-1)\mu\right| \leqslant j^\alpha$ for all $K+1 \leqslant j \leqslant n$ when $\mathbf{x} \in A_{\alpha,n}$, a little computation yields

$$\begin{aligned}\mathrm{Var}_{\mathbf{x}}(M_{n-K}) &\leqslant \prod_{j=K+1}^{n} \left[1 + \frac{\mu + j^\alpha + (j-1)^\alpha}{\{(\mu+1)j - \mu - j^\alpha + x_1\}\{(\mu+1)j - 2\mu - j^\alpha + x_1 - 1\}}\right] - 1 \\ &\leqslant \prod_{j=K+1}^{n} \left(1 + C' j^{\alpha-2}\right) - 1 \\ &\leqslant C' K^{\alpha-1}. \end{aligned} \tag{9.7}$$

Applying (9.7) to (9.6) completes the proof. □

We conclude this section with the proof of Lemma 3, recalling that $\mathcal{Z}_j \sim \mathrm{Gamma}(x_j, 1)$ and $\widetilde{\mathcal{Z}}_j \sim \mathrm{Gamma}(T_j + j, 1)$.

Proof of Lemma 3. Writing $T_j := \sum_{i=1}^{j} x_i$ and $Y_j \sim \mathrm{Gamma}(T_j + j - 1, 1)$, we first prove (2.4). Letting $E_{\varepsilon,j}$ be as in (2.3), we have

$$\begin{aligned}\mathbb{P}_{\mathbf{x}}\left[E_{\varepsilon,j}^c\right] &= \mathbb{P}_{\mathbf{x}}\left[\left|\frac{\mathcal{Z}_j}{\mathcal{Z}_j + \widetilde{\mathcal{Z}}_{j-1}} - \frac{\mathcal{Z}_j}{(\mu+1)j}\right| \geqslant \frac{\mathcal{Z}_j}{(\mu+1)j}\varepsilon\right] \\ &= \mathbb{P}_{\mathbf{x}}\left[\left|\frac{(\mu+1)j}{\mathcal{Z}_j + \widetilde{\mathcal{Z}}_{j-1}} - 1\right| \geqslant \varepsilon\right] \\ &\leqslant \mathbb{P}_{\mathbf{x}}\left[\left|\frac{\mathcal{Z}_j + \widetilde{\mathcal{Z}}_{j-1}}{(\mu+1)j} - 1\right| \geqslant \frac{\varepsilon}{1+\varepsilon}\right] \\ &= \mathbb{P}_{\mathbf{x}}\left[\left|\frac{Y_j}{(\mu+1)j} - 1\right| \geqslant \frac{\varepsilon}{1+\varepsilon}\right],\end{aligned}$$

and by Chebyshev's inequality,

$$\mathbb{P}_{\mathbf{x}}\left[E_{\varepsilon,j}^c\right] \leqslant \left(\frac{1+\varepsilon}{\varepsilon}\right)^4 \mathbb{E}_{\mathbf{x}}\left[\left(\frac{Y_j}{(\mu+1)j} - 1\right)^4\right].$$

Then (2.4) follows from bounding the moment above under the assumption $\mathbf{x} \in A_{\alpha,n}$, and applying a union bound. Let $a_j := T_j + j - 1$. Using the moment formula for the standard gamma distribution, a little calculation shows that the moment is equal to

$$\begin{aligned}&\frac{\mathbb{E}_{\mathbf{x}}\left[Y_j^4\right]}{(\mu+1)^4 j^4} - \frac{4\mathbb{E}_{\mathbf{x}}\left[Y_j^3\right]}{(\mu+1)^3 j^3} + \frac{6\mathbb{E}_{\mathbf{x}}\left[Y_j^2\right]}{(\mu+1)^2 j^2} - \frac{4\mathbb{E}_{\mathbf{x}}[Y_j]}{(\mu+1)j} + 1 \\ &= \frac{\prod_{k=0}^{3}(a_j+k)}{(\mu+1)^4 j^4} - \frac{4\prod_{k=0}^{2}(a_j+k)}{(\mu+1)^3 j^3} + \frac{6\prod_{k=0}^{1}(a_j+k)}{(\mu+1)^2 j^2} - \frac{4a_j}{(\mu+1)j} + 1.\end{aligned}$$

Noting that $|a_j - (\mu+1)j| \leqslant j^\alpha + x_1 + \mu + 1$ for $j \geqslant \phi(n)$, a direct computation shows that there is a positive constant $C := C(x_1, \mu, \alpha)$ such that

$$\mathbb{E}_{\mathbf{x}}\left[\left(\frac{Y_j}{(\mu+1)j} - 1\right)^4\right] \leqslant C j^{4\alpha-4}. \tag{9.8}$$

https://doi.org/10.1017/apr.2023.54 Published online by Cambridge University Press

We now prove (2.4) using (9.8). Let $C := C(x_1, \alpha, \mu)$ be a positive constant that may vary at each step of the calculation. Then

$$\mathbb{P}_{\mathbf{x}}\Bigg[\bigcup_{j=\lceil\phi(n)\rceil}^{n} E^c_{\varepsilon,j}\Bigg] \leqslant \sum_{j=\lceil\phi(n)\rceil}^{n} \mathbb{P}_{\mathbf{x}}[E^c_{\varepsilon,j}] \leqslant C\left(\frac{1+\varepsilon}{\varepsilon}\right)^4 \sum_{j=\lceil\phi(n)\rceil}^{n} j^{4\alpha-4}$$
$$\leqslant C\left(\frac{1+\varepsilon}{\varepsilon}\right)^4 \int_{\lceil\phi(n)\rceil-1}^{\infty} y^{4\alpha-4}dy \quad \leqslant C(1+\varepsilon)^4\varepsilon^{-4}n^{\chi(4\alpha-3)},$$

as required. Next, we use a union bound and Chebyshev's inequality to prove (2.5) as follows:

$$\mathbb{P}_{\mathbf{x}}\Bigg[\bigcup_{j=\lceil\phi(n)\rceil}^{n} \left\{\mathcal{Z}_j \geqslant j^{1/2}\right\}\Bigg] \leqslant \sum_{j=\lceil\phi(n)\rceil}^{n} \mathbb{E}_{\mathbf{x}}\big[\mathcal{Z}_j^4\big]j^{-2} = \sum_{j=\lceil\phi(n)\rceil}^{n} j^{-2}\prod_{\ell=0}^{3}(x_j+\ell).$$

If we further assume $x_2 \in (0, \kappa]$, then there are positive numbers C' and C'' such that

$$\mathbb{P}_{\mathbf{x}}\Bigg[\bigcup_{j=\lceil\phi(n)\rceil}^{n} \{\mathcal{Z}_j \geqslant j^{1/2}\}\Bigg] \leqslant C'\kappa^4 \sum_{j=\lceil\phi(n)\rceil}^{n} j^{-2} \leqslant C'\kappa^4 \int_{\phi(n)-1}^{\infty} y^{-2}dy \leqslant C''\kappa^4 n^{-\chi},$$

proving (2.6). □

Acknowledgements

The author is grateful to the referees and Nathan Ross for their careful reading of the manuscript and numerous helpful suggestions, and to Tejas Iyer for pointing out the connection to the work [18].

Funding information

This research was completed when the author was a Ph.D. student at the University of Melbourne. The author was supported by an Australian Government Research Training Program scholarship, and partially by ACEMS and the David Lachlan Hay Memorial Fund.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/apr.2023.54.

References

[1] ALDOUS, D. AND STEELE, J. M. (2004). The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures*, Springer, Berlin, pp. 1–72.

[2] BARABÁSI, A. L. AND ALBERT, R. (1999). Emergence of scaling in random networks. *Science* **286**, 509–512.

[3] BENJAMINI, I. AND SCHRAMM, O. (2001). Recurrence of distributional limits of finite planar graphs. *Electron. J. Prob.* **6**, article no. 23, 13 pp.

[4] BERGER, N., BORGS, C., CHAYES, J. T. AND SABERI, A. (2014). Asymptotic behavior and distributional limits of preferential attachment graphs. *Ann. Prob.* **42**, 1–40.
[5] BHAMIDI, S. (2007). Universal techniques to analyze preferential attachment trees: global and local analysis. In preparation.
[6] BLOEM-REDDY, B. AND ORBANZ, P. (2017). Preferential attachment and vertex arrival times. Preprint. Available at https://arxiv.org/abs/1710.02159.
[7] BOLLOBÁS, B., RIORDAN, O., SPENCER, J. AND TUSNÁDY, G. (2001). The degree sequence of a scale-free random graph process. *Random Structures Algorithms* **18**, 279–290.
[8] BORGS, C., CHAYES, J., DASKALAKIS, C. AND ROCH, S. (2007). First to market is not everything: an analysis of preferential attachment with fitness. In *STOC '07: Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, Association for Computing Machinery, New York, pp. 135–144.
[9] BOROVKOV, K. A. AND VATUTIN, V. A. (2006). On the asymptotic behaviour of random recursive trees in random environments. *Adv. Appl. Prob.* **38**, 1047–1070.
[10] BUBECK, S., MOSSEL, E. AND RÁCZ, M. (2015). On the influence of the seed graph in the preferential attachment model. *IEEE Trans. Network Sci. Eng.* **2**, 30–39.
[11] DEREICH, S. (2016). Preferential attachment with fitness: unfolding the condensate. *Electron. J. Prob.* **21**, article no. 3, 38 pp.
[12] DEREICH, S. AND ORTGIESE, M. (2014). Robust analysis of preferential attachment models with fitness. *Combinatorics Prob. Comput.* **23**, 386–411.
[13] DHARMADHIKARI, S. W. AND JOGDEO, K. (1969). Bounds on moments of certain random variables. *Ann. Math. Statist.* **40**, 1506–1509.
[14] DURRETT, R. (2019). *Probability: Theory and Examples*, 5th edn. Cambridge University Press.
[15] ERGÜN, G. AND RODGERS, G. J. (2002). Growing random networks with fitness. *Physica A* **303**, 261–272.
[16] GARAVAGLIA, A. (2019). Preferential attachment models for dynamic networks. Doctoral Thesis, Technische Universiteit Eindhoven.
[17] GARAVAGLIA, A., HAZRA, R. S., VAN DER HOFSTAD, R. AND RAY, R. (2022). Universality of the local limit of preferential attachment models. Preprint. Available at https://arxiv.org/abs/2212.05551
[18] GARAVAGLIA, A., VAN DER HOFSTAD, R. AND LITVAK, N. (2020). Local weak convergence for PageRank. *Ann. Appl. Prob.* **30**, 40–79.
[19] VAN DER HOFSTAD, R. (2017). *Random Graphs and Complex Networks*, Vol. 1. Cambridge University Press.
[20] VAN DER HOFSTAD, R. (2021). *Random Graphs and Complex Networks*, Vol. 2 (draft). Available at https://www.win.tue.nl/rhofstad/NotesRGCNII.pdf.
[21] HOLMGREN, C. AND JANSON, S. (2017). Fringe trees, Crump–Mode–Jagers branching processes and m-ary search trees. *Prob. Surveys* **14**, 53–154.
[22] IYER, T. (2023). Degree distributions in recursive trees with fitnesses. *Adv. Appl. Prob.* **55**, 407–443.
[23] KARP, R. M. (1990). The transitive closure of a random digraph. *Random Structures Algorithms* **1**, 73–93.
[24] LO, T. Y. Y. (2023). Supplementary material for 'Local weak limit of preferential attachment random trees with additive fitness'. Available at https://doi.org/10.1017/apr.2023.54.
[25] LODEWIJKS, B. AND ORTGIESE, M. (2020). A phase transition for preferential attachment models with additive fitness. *Electron. J. Prob.* **25**, article no. 146, 54 pp.
[26] LUKACS, E. (1955). A characterization of the gamma distribution. *Ann. Math. Statist.* **26**, 319–324.
[27] MAHMOUD, H. (2008). *Pólya Urn Models*. CRC Press, Boca Raton.
[28] MÓRI, T. F. (2005). The maximum degree of the Barabási–Albert random tree. *Combinatorics Prob. Comput.* **14**, 339–348.
[29] PAIN, M. AND SÉNIZERGUES, D. (2020). Correction terms for the height of weighted recursive trees. *Ann. Appl. Prob.* **32**, 3027–3059.
[30] PEKÖZ, E. A., RÖLLIN, A. AND ROSS, N. (2013). Total variation error bounds for geometric approximation. *Bernoulli* **19**, 610–632.
[31] PEKÖZ, E. A., RÖLLIN, A. AND ROSS, N. (2016). Generalized gamma approximation with rates for urns, walks and trees. *Ann. Prob.* **44**, 1776–1816.
[32] PEKÖZ, E. A., RÖLLIN, A. AND ROSS, N. (2017). Joint degree distributions of preferential attachment random graphs. *Adv. Appl. Prob.* **49**, 368–387.
[33] PEKÖZ, E. A., RÖLLIN, A. AND ROSS, N. (2019). Pólya urns with immigration at random times. *Bernoulli* **25**, 189–220.
[34] PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, Berlin, Heidelberg.
[35] RUDAS, A., TÓTH, B. AND VALKÓ, B. (2007). Random trees and general branching processes. *Random Structures Algorithms* **31**, 186–202.
[36] SÉNIZERGUES, D. (2021). Geometry of weighted recursive and affine preferential attachment trees. *Electron. J. Prob.* **26**, article no. 80, 56 pp.