

**Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.**

Preprint

ISSN 2198-5855

**On the structure of genealogical trees associated with
explosive Crump–Mode–Jagers branching processes**

Tejas Iyer¹, Bas Lodewijks²

submitted: November 29, 2023

¹ Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: tejas.iyer@wias-berlin.de

² University of Augsburg
Universitätsstraße 14
86159 Augsburg
Germany
E-Mail: bas.lodewijks@uni-a.de

No. 3060
Berlin 2023



2020 *Mathematics Subject Classification.* 60J80, 90B15, 05C80.

Key words and phrases. Explosive Crump–Mode–Jagers branching processes, super-linear preferential attachment, preferential attachment with fitness, random recursive trees, condensation.

The research of TI has been funded by Deutsche Forschungsgemeinschaft (DFG) through DFG Project no. 443759178. TI also acknowledges funding from the 2023 WIAS Young Scientist grant, as well as previous funding from the LMS Early Career Fellowship ECF-1920-55, which helped support this research. BL has received funding from the European Union's Horizon 2022 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 101108569 and has been supported by the grant GrHyDy ANR-20-CE40-0002.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: <http://www.wias-berlin.de/>

On the structure of genealogical trees associated with explosive Crump–Mode–Jagers branching processes

Tejas Iyer, Bas Lodewijks

Abstract

We study the structure of genealogical trees associated with explosive Crump–Mode–Jagers branching processes (stopped at the explosion time), proving criteria for the associated tree to contain a node of infinite degree (a *star*) or an infinite path. Next, we provide uniqueness criteria under which with probability 1 there exists exactly one of a unique star or a unique infinite path. Under the latter uniqueness criteria we also provide an example where, with strictly positive probability less than 1, there exists a unique star in the model. We thus illustrate that this probability is not restricted to being 0 or 1. Moreover, we provide structure theorems when there is a star, where we prove that certain trees appear as sub-trees in the tree infinitely often. We apply our results to a general discrete evolving tree model, named *explosive recursive trees with fitness*. As a particular case, we study a family of *super-linear preferential attachment models with fitness*. For these models, we derive phase transitions in the model parameters in three different examples, leading to either exactly one star with probability 1 or one infinite path with probability 1, with every node having finite degree. Furthermore, we highlight examples where sub-trees T of *arbitrary* size can appear infinitely often; behaviour that is markedly distinct from super-linear preferential attachment models studied in the literature so far.

Contents

| | | |
|----------|--|----------|
| 1 | Introduction | 3 |
| 1.1 | Random recursive trees with fitness | 4 |
| 1.2 | Overview of our contribution and structure | 6 |
| 1.2.1 | Structure of the paper | 6 |
| 1.3 | Notation and preliminaries | 7 |
| 2 | Statements of main results | 9 |
| 2.1 | Global assumptions | 9 |
| 2.2 | Sufficient criteria for a star | 10 |
| 2.3 | Sufficient criteria for an infinite path and structural results in the star regime | 11 |
| 2.4 | Uniqueness conditions related to the existence of a star or an infinite path | 13 |
| 2.5 | Proof techniques and relation to existing literature | 14 |

| | | |
|----------|--|-----------|
| 3 | Examples of applications and an open problem | 14 |
| 3.1 | The structure of explosive recursive trees with fitness | 15 |
| 3.1.1 | The star/path transition in explosive recursive trees with fitness | 16 |
| 3.1.2 | Sub-trees in explosive recursive trees with fitness when there is a star | 17 |
| 3.2 | Phase transitions in specific models of explosive recursive trees with fitness | 18 |
| 3.2.1 | Connection to existing literature: super-linear preferential attachment | 18 |
| 3.2.2 | Assumptions for the fitness function and vertex-weight distribution | 19 |
| 3.2.3 | Phase transitions for super-linear preferential attachment models with fitness | 21 |
| 3.2.4 | Discussion related to Theorem 3.21: the super-linear case | 23 |
| 3.3 | Proof techniques | 26 |
| 3.4 | Open problem | 26 |
| 4 | Proofs of main results | 26 |
| 4.1 | Sufficient criteria for a star | 26 |
| 4.1.1 | Proof of Theorem 2.5 | 33 |
| 4.2 | Sufficient criteria for an infinite path and structural results in the star regime | 33 |
| 4.2.1 | Proof of Theorem 2.8 | 34 |
| 4.2.2 | Proof of Theorem 2.10 | 35 |
| 4.3 | Uniqueness conditions related to the existence of a star or an infinite path | 36 |
| 4.3.1 | Proof of Theorem 2.12 | 37 |
| 5 | Proofs for applications of main results | 39 |
| 5.1 | Preliminary results and tools | 39 |
| 5.2 | Structure theorems related to explosive recursive trees | 40 |
| 5.2.1 | Proof of Theorem 3.4 | 40 |
| 5.2.2 | Proof of Theorem 3.7 | 42 |
| 5.3 | Possibility for both a star and infinite path to occur with positive probability | 44 |
| 5.3.1 | Proof of Theorem 2.15 | 45 |
| 6 | Examples of phase transitions in specific models of recursive trees with fitness | 47 |
| 6.1 | Phase transitions for the emergence of an infinite star or path: proof of Theorem 3.16 | 48 |
| 6.1.1 | Conditions for an infinite star: Item 1 of Theorem 3.4. | 49 |

| | | |
|----------|--|-----------|
| 6.1.2 | Conditions for an infinite path, Assumption 2.6. | 50 |
| 6.1.3 | Proofs of Lemmas 6.3 and 6.4 | 52 |
| 6.2 | Sub-tree counts in specific models of recursive trees with fitness | 56 |
| 6.2.1 | Proof of Corollary 3.8 | 57 |
| 6.2.2 | Proof of Theorem 3.21 | 58 |
| A | Additional results for other fitness functions | 61 |
| B | Verifying conditions for other inter-birth time distributions | 69 |
| B.1 | Conditions for an infinite star, Assumption 2.2. | 70 |
| B.2 | Conditions for an infinite path, Assumption 2.6. | 74 |
| C | Proof of Lemma 6.6 | 74 |

1 Introduction

Given a population of an entity moving towards explosion, that is, the emergence of infinitely many individuals in finite time, what can be said about the genealogical tree associated with the population at the time of explosion? On the one hand, an infinite path in the tree may be interpreted as an infinite line of evolution, with infinitely many ‘variants’ contributing to the explosion; on the other, a node of infinite degree (which we often call a *star*) may be interpreted, informally, as the emergence of a ‘dominant variant’. This is the goal of the present investigation, where as a simplified model of an evolving population, we use *Crump–Mode–Jagers* branching processes.

In a CMJ branching process (named after [17, 35]), an ancestral *root* individual produces offspring according to a collection of points on the non-negative real line. Each individual ‘born’ produces offspring according to an identically distributed collection of points, translated by their birth time (see Section 1.3 for a more formal description). One is generally interested in properties of the population as a function of time. Classical work from the 1970s and ‘80s related to this model generally deals with the *Malthusian case*, which, informally, refers to the fact that the population grows exponentially in time. These include strong laws of large numbers for *characteristics* associated with the process [56], properties of birth times in the k th generation [43], an $x \log x$ theorem [59], and numerous other results, for example [10, 57, 39, 37, 38]; see also the classical books [5, 36]. A number of more recent results are concerned with *asymptotic fluctuations* associated with the process in the Malthusian case, see, for example, [31, 41, 32, 44]. Other results are motivated by applications of these processes, including M/G/1 queues [28], vaccination and epidemic modelling [6, 7, 49], and numerous applications to random graphs, see Section 1.1 below.

Far fewer results exist for CMJ branching processes when a Malthusian parameter does not exist. In a particular case of reinforced branching process, a *condensation* phase transition can occur, where non-exponential growth occurs due to individuals having random weights

that influence their offspring distribution. In this case, a ‘small’ numbers of individuals with large weight produce larger and larger families, which in turn lead a rate of growth faster than exponential [19]. In more extreme circumstances, individuals produce larger and larger families so quickly that the process explodes in finite time. General criteria for explosion have been provided in terms of the solution of a functional fixed point equation by Komjáthy [45], who also extended the necessary and sufficient criteria for explosion in branching random walks in [1]. We refer the reader to [45] for a more comprehensive overview of the literature related to explosion in branching processes. In [11], the authors provide sufficient criteria for *local explosion* in closely related growth-fragmentation processes. Meanwhile, more, sufficient criteria for explosion in CMJ processes are in preparation in [33].

1.1 Random recursive trees with fitness

Aside from the applications outlined above, CMJ branching processes are often involved in the analysis of random graph models, more often, random trees. As far as the authors are aware, direct applications date back to Pittel [61], in providing a new proof for the limiting behaviour of the heights of *random recursive trees* and *affine preferential attachment trees*, but the technique of using continuous-time embeddings to analyse discrete combinatorial processes is more classical, going back to works of Arthreya and Karlin [3, 4].

Later, works by, for example [62, 12, 30, 58, 25] showed that CMJ branching processes can be applied to a large number of growing tree models. A natural framework of evolving trees, which corresponds to genealogical trees of CMJ branching processes and encompasses many existing models of recursive trees (which we refer to as random recursive trees with fitness [34]), posits that nodes v arrive one at a time, and are assigned a random i.i.d. *weight* W_v sampled from a measure μ on an arbitrary measure space (S, \mathcal{S}) (see Definition 3.1). Newly arriving nodes then connect, with edges directed outwards from the target nodes, with probability proportional to a general, measurable fitness function $f : \mathbb{N}_0 \times S \rightarrow [0, \infty)$ that incorporates information about the current out-degree of the target, and its weight. A natural quantity of interest in this model is the proportion of nodes N_k (at the n th time-step) having out-degree k . This model may be roughly classified according to the following conjectured phases [34]:

- 1 The *non-condensation phase*: There exists $\lambda > 0$ such that $\sum_{j=1}^{\infty} \mathbb{E} \left[\prod_{i=0}^{j-1} \frac{f(i, W)}{f(i, W) + \lambda} \right] = 1$. In this case, if p_k denotes the limit of \tilde{N}_k , we have $\sum_{k=0}^{\infty} p_k = 1$. In other words, all of the mass of edges is distributed around nodes of *microscopic* degrees.
- 2 The *condensation phase* (see [15, 19, 21, 18]): We have $\sum_{j=1}^{\infty} \mathbb{E} \left[\prod_{i=0}^{j-1} \frac{f(i, W)}{f(i, W) + \lambda} \right] < 1$, for any $\lambda > 0$ such that the sum converges. In this case, $0 < \sum_{k=0}^{\infty} p_k < 1$, so that a positive fraction of ‘mass’ is lost from the empirical measure to a sub-linear number of nodes of ‘large degrees’.
- 3 The *extreme-condensation phase*: For any $\lambda > 0$ we have $\sum_{j=1}^{\infty} \mathbb{E} \left[\prod_{i=0}^{j-1} \frac{f(i, W)}{f(i, W) + \lambda} \right] = \infty$. In this case, $\sum_{k=0}^{\infty} p_k = 0$, so that all of the ‘mass’ of edges is concentrated in nodes of ‘large’ degrees.

Part of the goal of this article is to investigate the behaviour of the third phase above, in the

explosive case that

$$\sum_{i=0}^{\infty} \frac{1}{f(i, w)} < \infty, \quad \text{for } \mu\text{-almost all } w \in S.$$

Note that this implies that $\prod_{i=0}^{\infty} \frac{f(i, W)}{f(i, W) + \lambda} > 0$ almost surely, so that the condition in Phase 3 is satisfied.

In [58], Oliveira and Spencer showed that, in the case $f(i, w) = i^p$, with $1 + \frac{1}{k} < p \leq 1 + \frac{1}{k-1}$ for some $k \in \mathbb{N}$ (equality only for $k \geq 2$), the infinite tree associated with this process is somehow extreme: it contains a single node of infinite degree, connected to infinitely many children with an associated sub-tree of size at most k , and only finitely many with an associated sub-tree of size $k + 1$ or larger. This paper uses the fact that the associated CMJ branching process is *explosive*, a technique also exploited in similar works related to ‘balls-in-bins’ processes [55]. Related work by Arthreya [2, Theorem 2.1] attempts to prove that, according to a certain summability condition, either every node in the infinite tree has finite degree, or with positive probability there exists a single node such that all but finitely many new nodes connect to this node. However, there is a mistake here, in that [2, Theorem 2.1b] should really state: the probability that there exists a single node such that all but finitely many new-coming nodes connect to this node is zero (indeed, note that [2, Corollary 2.2] directly contradicts [58]). Nevertheless, the associated summability condition and result in this paper is interesting, and motivates the question of whether there is a critical condition that guarantees the existence of a node of infinite degree in the infinite tree, or every node having finite degree, cf. Theorem 3.4, below.

A related question is whether or not, in the associated *recursive tree with fitness* model, the index associated with the node of maximal degree is fixed after some finite time, or changes infinitely often, that is, whether or not there is a *persistent hub*. A unique node of infinite degree in the infinite tree associated with the model thus implies the existence of such a hub. In a slightly different model of evolving graphs, when $f(i, w) = g(i)$ with g being a *concave* sub-linear function, one of the results of Dereich and Mörters [20] shows that a persistent hub emerges if and only if $\sum_{i=1}^{\infty} g(i)^{-2} < \infty$. In the recursive tree model described above, Galashin [23] proved that, if $f(i, w) = g(i)$, with g *convex* and unbounded, a persistent hub always appears. This has been extended to a much wider range of functions g , independent of the weight w , by Banerjee and Bhamidi in [8].

When weights are added, however, in the sense that $f(i, w)$ may depend on w , a different picture emerges. Suppose that w takes values in $[0, \infty)$. In the case $f(i, w) = w(i + 1)$, under a particular set of conditions leading to the *condensation phase* (Item 2 above), in [19] the authors show that there is no persistent hub, and the size of the node of maximal degree grows sub-linearly in the size of the tree. In the case $f(i, w) = i + w$ or $f(i, w) = w$, when the weights w are sampled according to certain classes of distributions, in [53, 63] and [52, 51], respectively, the authors provide critical criteria depending on the parameters of the weight distribution, for the existence, or non-existence, of a persistent hub. A number of other particular models of so called *preferential attachment with fitness* have been studied, see, for example, [26, 42, 22], and related works regarding local weak limits of preferential attachment type models [9, 50, 24].

1.2 Overview of our contribution and structure

In this paper we provide general sufficient conditions for the genealogical tree associated with an explosive CMJ branching process to contain a node of infinite degree or an infinite path at the explosion time (in Theorems 2.5 and 2.8, respectively). When there is a node of infinite degree, we provide criteria for one to see a fixed tree as a sub-tree of a child of that node either infinitely often, or finitely often, in Theorem 2.10. We also prove uniqueness criteria in Theorem 2.12, under which there almost surely exists a unique node of infinite degree or a unique infinite path. Under the conditions of the uniqueness theorem, we provide a counterexample in Theorem 2.15, where the events that there is a node of infinite degree, or an infinite path, both have positive probability, less than 1. Finally, in Theorems 3.4, 3.7, and Corollary 3.8 we apply our results to the recursive tree with fitness model and prove phase transitions in three particular models in Theorems 3.16 and 3.21. We encourage the reader more interested in this discrete model to refer to these results first.

The question of whether the genealogical tree of an explosive CMJ branching process contains an infinite path or a node of infinite degree has not been investigated in this level of generality previously. Our techniques involve significant improvements of those of [58] (see also Sections 2.5 and 3.3), and thus allow us to greatly extend the picture associated with the general recursive tree with fitness model. Intriguingly, our results show that when there is a unique node of infinite degree in the infinite tree associated with the model, in many particular cases there exist children of the node of infinite degree that have arbitrarily large, but finite, degree (or even an arbitrarily large, but finite, number of descendants). Previous comments in the literature seem to indicate that it was believed that when there is an infinite degree node, the degrees of all other nodes are bounded, see [8, Section 3]. Finally, we remark that the phase transitions related to the emergence of a node of infinite degree are reminiscent of a different notion of *condensation* in conditioned Bienaymé-Galton-Watson trees [40].

1.2.1 Structure of the paper

The paper is structured in the following way. Below, in Section 1.3, we introduce a formal description of the model and the notation we use in this paper. Section 2 states the main results, which are most general and require certain assumptions on the inter-birth time distribution. Section 3 then discusses the particular example of exponentially distributed inter-birth times and how this relates to a family of discrete tree models coined *recursive trees with fitness*. Here, we derive sufficient conditions such that the assumptions used for the main results are satisfied. Moreover, when considering certain sub-families of recursive trees with fitness, we prove more precise results in terms of phase diagrams for the existence of either unique infinite-degree nodes or unique infinite paths. As mentioned above, we encourage the reader more interested in results related to the discrete recursive tree model (which is also less abstract), to refer to the results of Section 3.1 first, before reading the section below.

Aside for a few exceptions, Section 4 proves the main results of Section 2, Section 5 proves the most general results of Section 3, and Section 6 proves the particular examples of Section 3. Finally, we consider a number of other models in Appendix A and B, showing that the assumptions subject to which the main results hold are valid more broadly.

1.3 Notation and preliminaries

In this paper, we consider properties of the genealogical trees associated with Crump-Mode-Jagers branching processes; a tree-valued stochastic process $(\mathcal{T}_t)_{t \geq 0}$ which one may regard as the genealogical tree representing a population evolving over time. The goal then, is to define a state space of individuals, in this setting, the infinite *Ulam-Harris* tree of potential individuals associated with a common ancestor, *birth times* $\mathcal{B}(u)$ associated with individuals, which themselves are encoded by a random function (X, W) , and then define \mathcal{T}_t as the set of individuals born up to time t . Note that the notation we use is slightly different to the common notation regarding CMJ branching processes, see Remark 1.1 further below.

First, we generally use $\mathbb{N} := \{1, 2, \dots\}$, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and for $n \in \mathbb{N}$ we let $[n] := \{1, \dots, n\}$. We consider *individuals* in the process as being labelled by elements of the infinite *Ulam-Harris* tree $\mathcal{U}_\infty := \bigcup_{n \geq 0} \mathbb{N}^n$; where $\mathbb{N}^0 := \{\emptyset\}$ contains a single element \emptyset which we call the *root*. We denote elements $u \in \mathcal{U}_\infty$ as a tuple, so that, if $u = (u_1, \dots, u_k) \in \mathbb{N}^k$, $k \geq 1$, we write $u = u_1 \cdots u_k$. An individual $u = u_1 u_2 \cdots u_k$ is to be interpreted recursively as the u_k th child of the individual $u_1 \cdots u_{k-1}$; for example, $1, 2, \dots$ represent the offspring of \emptyset . Suppose that $(\Omega, \Sigma, \mathbb{P})$ is a complete probability space and (S, \mathcal{S}) is a measure space. We also equip \mathcal{U}_∞ with the sigma algebra generated by singleton sets. Then, we fix a random mappings $X : \Omega \times \mathcal{U}_\infty \rightarrow [0, \infty]$, $W : \Omega \times \mathcal{U}_\infty \rightarrow S$, and define $(X, W) : \Omega \times \mathcal{U}_\infty \rightarrow [0, \infty] \times S$, so that $u \mapsto (X(u), W_u)$. In general, for $u \in \mathcal{U}_\infty$ and $j \in \mathbb{N}$, one interprets W_u as a ‘weight’ associated with u , and $X(u_j)$ the waiting time between the birth of the $(j - 1)$ th and j th *child* of u .

We introduce some notation related to elements $u \in \mathcal{U}_\infty$: we use $|\cdot|$ to measure the *length* of a tuple u , so that, if $u = \emptyset$ we set $|u| = 0$, whilst if $u = u_1 \cdots u_k$ then $|u| = k$. If, for some $x \in \mathcal{U}_\infty$, we have $x = uv$, we say u is a *ancestor* of x . We introduce a notation to refer to ancestors: given $\ell \leq |u|$, we set $u_{|\ell} := u_1 \cdots u_\ell$. It will be helpful to equip \mathcal{U}_∞ with the lexicographic total order \leq_L : given elements u, v we say $u \leq_L v$ if either u is a ancestor of v , or $u_\ell < v_\ell$ where $\ell = \min \{i \in \mathbb{N} : u_i \neq v_i\}$. We say a subset $T \subset \mathcal{U}_\infty$ is a *tree* if, given that $u \in T$, we also have $u_{|\ell} \in T$, for each $\ell \leq |u|$. Note that any such trees can be viewed as graphs in the natural way, connecting nodes to their children.

Now, we use the values of X to associate *birth times* $\mathcal{B}(u)$ to individuals $u \in \mathcal{U}_\infty$. In particular, we define $\mathcal{B} : \Omega \times \mathcal{U}_\infty \rightarrow [0, \infty]$ recursively as follows:

$$\mathcal{B}(\emptyset) := 0 \quad \text{and for } u \in \mathcal{U}_\infty, i \in \mathbb{N}, \quad \mathcal{B}(ui) := \mathcal{B}(u) + \sum_{j=1}^i X(u_j).$$

Consequentially, a value of $X(ui) = \infty$ indicates that the individual u has stopped producing offspring, and does not produce i children or more.

Finally, we set $\mathcal{T}_t = \{x \in \mathcal{U}_\infty : \mathcal{B}(x) \leq t\}$ and identify for each $t \geq 0$, \mathcal{T}_t as the *genealogical tree* of individuals with birth time at most t . Again we emphasise that one may think of this, intuitively, as the set of all individuals, originating from a common ancestor, that have been born by time t . More formally, we identify the process with $(\tilde{\mathcal{T}}_t)_{t \geq 0}$; a measurable mapping $\Omega \times [0, \infty) \times \mathcal{U}_\infty \rightarrow [0, \infty]$. Then, if $u \in \mathcal{T}_t$, we set $\tilde{\mathcal{T}}_t(u) = \mathcal{B}(u)$, and otherwise, set $\tilde{\mathcal{T}}_t(u) = \infty$. In addition, we set $\mathcal{W}_t = \{(x, \mathcal{B}(x), W_x) : x \in \mathcal{T}_t\}$, so that $(\mathcal{W}_t)_{t \geq 0}$ also incorporates information about the random ‘weights’ of individuals in the tree \mathcal{T}_t . We let $(\mathcal{F}_t)_{t \geq 0}$ and $(\mathcal{W}_t)_{t \geq 0}$ denote the filtrations generated by $(\tilde{\mathcal{T}}_t)_{t \geq 0}$ and $(\mathcal{W}_t)_{t \geq 0}$, respectively; and by taking their completions if necessary, assume that both \mathcal{F}_t and \mathcal{W}_t are complete. By

abuse of notation, we use the symbol \mathcal{T}_t to refer to $\widetilde{\mathcal{T}}_t$, the set \mathcal{T}_t , and the graph associated with \mathcal{T}_t where the vertex set is \mathcal{T}_t and edges connect elements to their children. For a given choice of X, W , we say $(\mathcal{T}_t)_{t \geq 0}$ is the genealogical tree process associated with an (X, W) -Crump-Mode-Jagers branching process; often, we refer to $(\mathcal{T}_t)_{t \geq 0}$ directly as an (X, W) -Crump-Mode-Jagers branching process, viewed as a stochastic process in t , adapted to its natural filtration $(\mathcal{F}_t)_{t \geq 0}$.¹

For $u \in \mathcal{U}_\infty$, we let $\mathcal{P}_i(u)$ denote the time, after the *birth* of u , required for u to produce i offspring. That is,

$$\mathcal{P}_i(u) := \sum_{j=1}^i X(uj) \quad \text{and} \quad \mathcal{P}(u) := \mathcal{P}_\infty(u) = \sum_{j=1}^{\infty} X(uj).$$

Note that, as a result of this definition, for any $u, v \in \mathcal{U}_\infty$ with $u = u_1 \cdots u_k$ and $v = v_1 \cdots v_\ell$, if we set $uv_0 := u$ we have

$$\mathcal{B}(uv) - \mathcal{B}(u) = \sum_{j=0}^{\ell-1} (\mathcal{P}_{v_{j+1}}(uv_{|j})). \quad (1.1)$$

It will also be beneficial to extend the notation \mathcal{P} to arbitrary trees T : for $T \subseteq \mathcal{U}_\infty$, we define

$$\mathcal{P}_T(u) := \inf \{t > 0 : \mathcal{B}(uv) - \mathcal{B}(u) < t \text{ for all } v \in T\}. \quad (1.2)$$

Thus, with the above notation $\mathcal{P}_i(u) = \mathcal{P}_{[i]}(u)$. We also set $\mathcal{P}_T := \mathcal{P}_T(\emptyset)$, $\mathcal{P}_i = \mathcal{P}_i(\emptyset)$, and $\mathcal{P} := \mathcal{P}(\emptyset)$.

We generally assume a dependence between the values $(\mathcal{P}_i(u))_{i \in \mathbb{N} \cup \{\infty\}}$ and W_u . However, for brevity of notation, we often do not explicitly indicate this dependence. We use the notation \mathcal{P}_i and \mathcal{P} to denote generic copies of random variables distributed like $\mathcal{P}_i(u)$, and $\mathcal{P}(u)$ respectively.

For each $u \in \mathcal{U}_\infty$ it will be helpful to have a map $\xi^{(u)} : \Omega \times [0, \infty] \rightarrow \mathbb{N}$ indicating the number of children u has produced, more precisely, we define $\xi^{(u)}(t)$ such that

$$\xi^{(u)}(t) = \begin{cases} \sum_{i=1}^{\infty} \mathbf{1}_{\{\mathcal{P}_i(u) \leq s\}} & \text{if } t = \mathcal{B}(u) + s, s \in [0, \infty]; \\ 0 & \text{otherwise.} \end{cases}$$

With regards to the process $(\mathcal{T}_t)_{t \geq 0}$, we define the stopping times $(\tau_k)_{k \in \mathbb{N}_0}$ such that

$$\tau_k := \inf \{t \geq 0 : |\mathcal{T}_t| \geq k\},$$

where we adopt the convention that the infimum of the empty set is ∞ . One readily verifies that $(|\mathcal{T}_t|)_{t \geq 0}$ is right-continuous, and thus $|\mathcal{T}_{\tau_k}| \geq k$. For each $k \in \mathbb{N}$ we define the tree \mathcal{T}_k as the tree consisting of the first k individuals in \mathcal{T}_{τ_k} ordered by birth time, breaking ties lexicographically. We call $\tau_\infty := \lim_{k \rightarrow \infty} \tau_k$ the *explosion time* of the process. We also define the tree $\mathcal{T}_\infty := \bigcup_{k=1}^{\infty} \mathcal{T}_k$. Note that it may be the case that $|\mathcal{T}_\infty| < \infty$; in this case $\tau_\infty = \infty$ and

$$\mathcal{T}_\infty = \{x \in \mathcal{U}_\infty : \mathcal{B}(x) < \infty\},$$

¹Note that distinct functions (X, W) , (X', W') may lead to the same tree, if, for example $X(ui) = X'(ui) = \infty$, but $X(u(i+1)) \neq X'(u(i+1))$, but this is only a formal technicality, which we can overcome by viewing (X, W) as an appropriate *equivalence class* of functions.

and the set on the right-hand side is finite. If $|\mathcal{T}_\infty| < \infty$, we say *extinction* occurs, otherwise *survival* occurs. For a non-negative real-valued random variable Z and $\lambda > 0$ we let $\mathcal{M}_\lambda(Z), \mathcal{L}_\lambda(Z)$ denote the associated moment generating function and Laplace transform, respectively, i.e.

$$\mathcal{M}_\lambda(Z) := \mathbb{E} [e^{\lambda Z}] \quad \text{and} \quad \mathcal{L}_\lambda(Z) := \mathbb{E} [e^{-\lambda Z}].$$

Moreover, if the random variable Z has, additionally, some dependence on a random variable $W \in \mathcal{S}$, we write

$$\mathcal{M}_\lambda(Z; W) := \mathbb{E} [e^{\lambda Z} | W] \quad \text{and} \quad \mathcal{L}_\lambda(Z; W) := \mathbb{E} [e^{-\lambda Z} | W].$$

In addition, for real valued random variables Z_1 and Z_2 , we say $Z_1 \leq_S Z_2$, if, for each $a \in \mathbb{R}$

$$\mathbb{P}(Z_1 > a) \leq \mathbb{P}(Z_2 > a).$$

Finally, for $r \geq 0$ we use $\text{Exp}(r)$ to denote the exponential distribution with parameter r .

Remark 1.1. With the more commonly used notation for CMJ branching processes, one assigns a point process (denoted $\xi^{(u)}$) to each $u \in \mathcal{U}_\infty$, and refers to the points $\sigma_1^{(u)} \leq \sigma_2^{(u)}, \dots$ associated with this point process (in the notation used here $\mathcal{B}(u1), \mathcal{B}(u2), \dots$). We do not use this framework here, because, this requires one to be able to write the measure $\xi^{(u)} = \sum_{i=1}^{\infty} \delta_{\sigma_i}$, which requires one to impose σ -finiteness assumptions on the point process (see, for example, [48, Corollary 6.5]). This σ -finiteness is implied by the classical Malthusian condition, but, in this general setting, we believe it is easier to have a framework where one can directly refer to the points $\mathcal{B}(u1), \mathcal{B}(u2), \dots$

2 Statements of main results

In this paper, we are interested in properties of the infinite tree \mathcal{T}_∞ , in particular the question of whether or not this tree contains an infinite path or a node of infinite degree. This section deals with results in their most general form: Section 2.1 states some global assumptions imposed throughout the paper, Section 2.2 deals with criteria for a node of infinite degree (or star), Section 2.3 deals with criteria for an infinite path, and structural properties of the tree \mathcal{T}_∞ when there is a star, and Section 2.4 deals with uniqueness properties, providing criteria for a the appearance of a unique star or unique infinite path (but not both) to appear almost surely. In Theorem 2.15 we also show that in the regime where there is, almost surely, exactly one of a unique star or infinite path, either may appear with positive probability. Finally, we provide an overview of the proof techniques used, and the relation to existing literature in Section 2.5.

2.1 Global assumptions

In general in this paper, we assume that the values of $(X(uj))_{j \in \mathbb{N}}$ depend on W_u . We also assume that the sequences of random variables

$$((X(uj))_{j \in \mathbb{N}}, W_u) \quad \text{are i.i.d. for different } u \in \mathcal{U}_\infty; \quad (2.1)$$

although we expect that some of our techniques may carry over to a more general setting. For a given $w \in \mathcal{S}$, we let $(X_w(uj))_{j \in \mathbb{N}}$ denote a sequence $(X(uj))_{j \in \mathbb{N}}$, conditionally on the

weight $W_u = w$. Another common assumption we use throughout is the following: for any given $w \in S$ and any $u \in \mathcal{U}_\infty$, the sequence of random variables

$$(X_w(ui))_{i \in \mathbb{N}} \text{ is mutually independent.} \quad (2.2)$$

The existence of an explosive (X, W) -CMJ process satisfying (2.1) is a straightforward application of (for example) the Kolmogorov extension theorem (indeed, this is the classical definition of a CMJ process).

We also generally assume that the event $\{|\mathcal{T}_\infty| = \infty\}$ has positive probability, and

$$\mathbb{P}(\tau_\infty < \infty \mid |\mathcal{T}_\infty| = \infty) = 1. \quad (2.3)$$

That is, the process is almost surely explosive, when conditioned on survival. In general, we say an event A occurs *almost surely on survival* if $\mathbb{P}(A \mid |\mathcal{T}_\infty| = \infty) = 1$. In all statements in this paper referring to an “explosive (X, W) -CMJ process”, we assume it satisfies (2.1) and (2.3).

In this paper, we also rely on the following well-known fact in graph theory.

Lemma 2.1 (König’s Lemma). *Any infinite tree contains a node of infinite degree or an infinite path.*

2.2 Sufficient criteria for a star

In this subsection, we provide sufficient criteria for the infinite tree \mathcal{T}_∞ to contain an infinite star. Our main assumptions are as follows.

Assumption 2.2. We have the following conditions.

- 1 There exist non-negative real-valued random variables $(Y_n)_{n \in \mathbb{N}_0}$ with finite mean such that, for any $w \in S$,

$$\sum_{i=n+1}^{\infty} X_w(i) \leq_S Y_n. \quad (2.4)$$

- 2 If we let $\mu_n := \mathbb{E}[Y_n]$, then we also have, for some $c \in (0, \infty)$,

$$\limsup_{n \rightarrow \infty} \mathcal{M}_{c\mu_n^{-1}}(Y_n) < \infty. \quad (2.5)$$

Moreover, we assume that $(\mu_n)_{n \in \mathbb{N}_0}$ is non-increasing in n , with $\lim_{n \rightarrow \infty} \mu_n = 0$.

- 3 For each $n \in \mathbb{N}$ and any given $w \in S$, the random variables $(X_w(i))_{i \in \mathbb{N}}$ are mutually independent.

- 4 For each $n \in \mathbb{N}$,

$$\sum_{i=n+1}^{\infty} X(i) > 0 \text{ almost surely,} \quad (2.6)$$

and additionally, we have

$$\mathbb{E}[\xi(0)] = \mathbb{E}[\sup\{k : X(k) = 0\}] < 1. \quad (2.7)$$

5 With c as appearing in Equation (2.5),

$$\sum_{i=1}^{\infty} \mathbb{E} \left[\mathcal{L}_{c\mu_i^{-1}}(\mathcal{P}_i(\emptyset); W) \right] < \infty. \quad (2.8)$$

Remark 2.3. Note that under Condition 1 of Assumption 2.2, we have $\mathbb{P}(|\mathcal{T}_\infty| = \infty) = 1$. ◀

Remark 2.4. In Assumption 2.2 we can consider Conditions 1 and 3 as a *uniform explosivity* condition: it implies that, for any $u \in \mathcal{U}_\infty$ and any $\varepsilon > 0$,

$$\lim_{L \rightarrow \infty} \mathbb{P} \left(\sum_{j=L+1}^{\infty} X_j(u) \geq \varepsilon \mid X_1(u), \dots, X_L(u) \right) = 0,$$

with the convergence uniform in $X_1(u), \dots, X_L(u)$; a fact that is crucial for Lemma 4.6, and hence the proof of Theorem 2.5, to hold. Condition 4 is there as a technical assumption, used, for example, in Proposition 4.4 and Lemma 4.5. It ensures that $\tau_k < \tau_\infty$ for each $k \in \mathbb{N}$. Indeed, if, for example, $\mathbb{E}[\xi(0)] > 1$, the tree consisting of all the individuals born *instantaneously* at time 0 is the genealogical tree of a supercritical Bienaymé-Galton-Watson branching process. Hence, with positive probability this tree is infinitely large, and thus there may be no node of infinite degree in this infinite tree. Condition 2 is used to prove the Chernoff type concentration bound in Lemma 4.1 which, when combined with the summability condition in Condition 5, leads to the proof of the crucial Proposition 4.3. ◀

The conditions of Assumption 2.2 allow us to formulate the following theorem.

Theorem 2.5 (Infinite star). *Under Assumption 2.2, almost surely the infinite tree \mathcal{T}_∞ contains a node of infinite degree (i.e. an infinite star).*

The proof of Theorem 2.5 appears in Section 4.1.

2.3 Sufficient criteria for an infinite path and structural results in the star regime

In this subsection, we provide sufficient criteria for \mathcal{T}_∞ to contain an infinite path and whether or not \mathcal{T}_∞ contains infinitely many copies of a fixed tree T . We first state the following assumption.

Assumption 2.6. There exists a collection of numbers $\{\nu_n^w \in [0, \infty) : w \in S, n \in \mathbb{N}\}$, such that for any $w \in S$,

$$\sum_{i=1}^{\infty} \mathbb{P}(\mathcal{P} < \nu_i^w) = \infty, \quad (2.9)$$

and

$$\liminf_{i \rightarrow \infty} \mathbb{P} \left(\sum_{j=i+1}^{\infty} X_w(j) \geq \nu_i^w \right) > 0. \quad (2.10)$$

Remark 2.7. Assumption 2.6 intuitively states that, conditionally on the weight w of the root \emptyset , infinitely many children i of \emptyset produce an infinite offspring within time ν_i^w . On the other hand, the root takes at least ν_i^w amount of time after the birth of its i th child to produce an infinite offspring, with a probability that is bounded from below, uniformly in i . Hence, infinitely many children i explode before their parent \emptyset . ◀

We can then formulate the following theorem.

Theorem 2.8 (Infinite path). *Under Assumption 2.6, the tree \mathcal{T}_∞ contains an infinite path almost surely on survival.*

The proof of Theorem 2.8 appears in Section 4.2.

Similar criteria to those related to the criteria for an infinite path allow us to also determine results related to the *structure* of \mathcal{T}_∞ in the sense that, when we know that \mathcal{T}_∞ contains an infinite star, certain sub-structures appear infinitely often; others only finitely often. We define $\mathcal{T}_\infty(\downarrow u) := \{v \in \mathcal{T}_\infty : v = uw, w \in \mathcal{U}_\infty\}$ as the sub-tree in \mathcal{T}_∞ rooted at u . For a fixed tree T containing \emptyset and $u \in \mathcal{U}_\infty$, we define $uT := \{uv : v \in T\}$. We say such a tree T is a sub-tree rooted at $u \in \mathcal{T}_\infty$, if, $uT \subseteq \mathcal{T}_\infty(\downarrow u)$. Recalling Equation (1.2) we then have the following set of assumptions.

Assumption 2.9. Let $(\mathcal{T}_t)_{t \geq 0}$ be an explosive (X, W) -CMJ process. For a given finite tree $T \subseteq \mathcal{U}_\infty$ containing \emptyset we have the following conditions.

- 1 There exists a collection of numbers $\{\nu_n^w \in [0, \infty) : w \in S, n \in \mathbb{N}_0\}$, such that for any $w \in S$,

$$\sum_{i=1}^{\infty} \mathbb{P}(\mathcal{P}_T < \nu_i^w) = \infty, \quad (2.11)$$

and

$$\liminf_{i \rightarrow \infty} \mathbb{P}\left(\sum_{j=i+1}^{\infty} X_w(j) \geq \nu_i^w\right) > 0. \quad (2.12)$$

- 2 For any $w \in S$ and with $(\tilde{X}_w(i), i \in \mathbb{N}) \sim (X_w(i), i \in \mathbb{N})$, independent of the process $(\mathcal{T}_t)_{t \geq 0}$,

$$\sum_{i=1}^{\infty} \mathbb{P}\left(\mathcal{P}_T < \sum_{j=i+1}^{\infty} \tilde{X}_w(j)\right) < \infty. \quad (2.13)$$

We can then formulate the following theorem.

Theorem 2.10 (Sub-tree count). *Let $(\mathcal{T}_t)_{t \geq 0}$ be an explosive (X, W) -CMJ process. Then, for any finite tree $T \subseteq \mathcal{U}_\infty$ containing \emptyset :*

- 1 *If Condition 1 of Assumption 2.9 is satisfied, almost surely, if $u \in \mathcal{T}_\infty$ has infinite degree, T appears infinitely often as a sub-tree rooted at a child of u .*
- 2 *If Condition 2 of Assumption 2.9 is satisfied, almost surely, if $u \in \mathcal{T}_\infty$ has infinite degree, T appears only finitely often as a sub-tree of a child of u . In particular, if Assumption 2.2 is satisfied, almost surely, the tree T appears only finitely often as a sub-tree of \mathcal{T}_∞ .*

The proof of Theorem 2.10 appears in Section 4.2.

2.4 Uniqueness conditions related to the existence of a star or an infinite path

In many cases, we expect \mathcal{T}_∞ to contain *exactly one* node of infinite degree or *exactly one* infinite path and also, often expect *co-existence* of an infinite path and node of infinite degree to be impossible. This leads us to the following assumption.

Assumption 2.11. We have the following conditions.

- 1 Condition 4 of Assumption 2.2 is satisfied.
- 2 The distribution of $\mathcal{P}(\emptyset)$ contains no atom on $[0, \infty)$.
- 3 For some $\varepsilon > 0$ and for each $i \in \mathbb{N}$, the distribution of $\mathcal{B}(i)$ contains no atom on $[0, \varepsilon)$.

We then have the following result.

Theorem 2.12 (Unique infinite star or path). *Let $(\mathcal{T}_t)_{t \geq 0}$ be an explosive (X, W) -CMJ process. Then:*

- 1 *If Conditions 1 and 2 of Assumption 2.11 are satisfied, \mathcal{T}_∞ contains at most 1 node of infinite degree, almost surely.*
- 2 *If Conditions 1 and 3 of Assumption 2.11 are satisfied, \mathcal{T}_∞ contains at most 1 infinite path, almost surely.*
- 3 *If all the conditions of Assumption 2.11 are satisfied, almost surely, on survival, \mathcal{T}_∞ contains exactly one of the following: a node of infinite degree, an infinite path.*

The proof of Theorem 2.12 appears in Section 4.3.

Remark 2.13. Though assumed to hold throughout, Theorem 2.12 can be proved without assuming (2.2). ◀

Remark 2.14. A case when the tree \mathcal{T}_∞ has more than one infinite path with positive probability is when *time 0 explosion* can occur, i.e. when $\mathbb{E}[\xi(0)] > 1$, so that $\tau_\infty = 0$ with positive probability, and \mathcal{T}_∞ is the genealogical tree of a supercritical Bienaymé-Galton-Watson branching process. This case is ruled out by Item 2 of Assumption 2.11, which, in particular, implies that $\xi(0) = 0$ almost surely. ◀

Given Item 3 of Theorem 2.12, one might expect the event that \mathcal{T}_∞ contains a node of infinite degree to occur with probability 0 or 1: for example, perhaps one might expect this to event to ‘look like’ a tail event, measurable with respect to the tail sigma algebra of an appropriate filtration. The following theorem shows that this is not actually the case in full generality.

Theorem 2.15. *There exists an explosive (X, W) -CMJ process satisfying the conditions of Assumption 2.11 such that*

$$\mathbb{P}(\mathcal{T}_\infty \text{ contains a node of infinite degree}) \in (0, 1). \quad (2.14)$$

Note that, applying Item 3 of Theorem 2.12, Equation (2.14) also implies that for such an explosive (X, W) -CMJ process,

$$\mathbb{P}(\mathcal{T}_\infty \text{ contains an infinite path}) \in (0, 1).$$

We note that, unlike the other theorems stated in this section, the proof of Theorem 2.15 appears in Section 5, in particular in Section 5.3.

2.5 Proof techniques and relation to existing literature

As mentioned in the introduction, Theorem 2.5 was proved in the case that the $(X(i), i \in \mathbb{N})$ are independent, with $X(i) \sim \text{Exp}(i^p)$, $p > 1$ in [58]. The technique used in that paper was to show that the number of nodes $u \in \mathcal{U}_\infty$ that are k -fertile in the tree \mathcal{T}_∞ (that is, contain a sub-tree of size at least $k + 1$), is almost surely finite for all $k > 1/(p - 1)$ [58, Lemma 5.1], and then deduce an infinite path cannot exist, almost surely. Applying Lemma 2.1 then yields the desired conclusion. An immediate generalisation of these techniques to the more general setting considered here does not allow one to prove Theorem 2.5. Indeed, as we will see in Theorem 3.21, Theorem 2.5 applies to cases where the number of k -fertile nodes is *almost surely infinite* for any $k \in \mathbb{N}$. Instead, we use a different approach to prove Theorem 2.5. By a first moment method and appropriate concentration bounds (Lemma 4.1), we show that the expected number of nodes $a \in \mathcal{U}_\infty$, with a high enough initial index a_1 , that explodes before all of its ancestors (that is, produces infinitely many offspring before any of its direct ancestors does), is finite (see Proposition 4.3). Combining this with a coupling argument in Proposition 4.4 (a significant generalisation of [58, Lemma 5.3]), we show that the expected number of nodes that explodes before all of their ancestors is finite. Finally, the *uniform explosivity* assumption (see Remark 2.4) allows one to deduce that the explosion time of the process τ_∞ is the infimum of the explosion times $\mathcal{B}(u) + \mathcal{P}(u)$ of individuals $u \in \mathcal{U}_\infty$. Using the aforementioned first moment arguments, we can show that this infimum coincides with an infimum over a finite set. Hence, at τ_∞ there exists at least one node of infinite degree.

The proofs of Theorems 2.8 and 2.10 use a different approach: by Borel-Cantelli arguments, we can show that *before* the explosion time associated with an individual, either infinitely many children explode themselves (leading to an infinite path, cf. Theorem 2.8), or otherwise certain finite sub-trees appear infinitely often when there is a star (cf. Theorem 2.10). The uniqueness conditions appearing in Theorem 2.12 are reminiscent of similar uniqueness conditions appearing, for example, previously in [58] (using the fact that the associated distribution of $\mathcal{P}(\emptyset)$ is smooth). However, this requires novelty when proving the existence of a unique infinite path in the level of generality we consider (see Lemma 4.8).

3 Examples of applications and an open problem

In this section, we provide applications of our main results, Theorem 2.5, Theorem 2.8, and Theorem 2.12 in the case that the inter-birth times $(X_w(i))_{i \in \mathbb{N}}$ are exponentially distributed. If $X_w(i)$ has an exponential distribution with parameter $f(i, w)$, say, the memory-less property and the property of minima of exponential distributions show that \mathcal{T}_∞ may be interpreted as the limiting infinite tree in a model of (W, f) -recursive trees with fitness. In this model, evolving in discrete time, nodes arrive one at a time, are assigned i.i.d. weights, and connect to

an existing node sampled with probability proportional to its ‘fitness function’. In this model, we are not only able to provide phase transitions related to the emergence of an infinite path, but also apply Theorem 2.10 to provide criteria for the emergence of a particular sub-tree infinitely often. This is the content of Section 3.1.

In Section 3.2 we consider more concrete cases, when the weights W are real-valued, closely connected to super-linear preferential attachment models. In Section 3.2.1 we provide some background for the analysis of such models, in the context of existing literature. Then, Theorem 3.16 in Section 3.2 provides a classification of the phases where one sees a unique node of infinite degree or a unique infinite path, proving phase transitions for three different examples. These results apply not only to the case that the values $(X_w(i))_{i \in \mathbb{N}}$ have exponential distributions, but other distribution types (see Remarks 3.6 and 3.17); which may be of interest in applications. In Theorem 3.21 we are able to characterise the sub-trees of children of the star that can emerge in this model. We discuss implications of these results in the particular ‘super-linear degree’ example in Section 3.2.4, which, in particular, allows us to produce phase diagrams in Figures 1 and 2.

Finally, we discuss the proof techniques involved in Section 3.3 and state an open problem in Section 3.4.

3.1 The structure of explosive recursive trees with fitness

Suppose the values of $(X_w(i), i \in \mathbb{N})$ are exponentially distributed and independent. The properties related to the exponential distribution yield that the sequence of trees $(\mathcal{T}_i)_{i \in \mathbb{N}}$ associated with an explosive (X, W) -CMJ branching process are identical in law to a sequence of *recursive trees with fitness* which we define below. First, we define the *fitness function* $f : \mathbb{N}_0 \times S \rightarrow [0, \infty)$ to be a measurable function such that $f(i, w)$ is the rate of the exponential random variable $X_w(i + 1)$.

In this section, we generally consider trees as being rooted with edges directed away from the root, and hence the number of ‘children’ of a node corresponds to its *out-degree*. More precisely, given a vertex labelled v in a directed tree T we let $\deg^+(v, T)$ denote its out-degree in T . We now define the recursive tree with fitness model.

Definition 3.1 (Recursive tree with fitness). Suppose that $(W_i)_{i \in \mathbb{N}}$ are i.i.d. copies of a random variable W that takes values in S , and let $f : \mathbb{N}_0 \times S \rightarrow [0, \infty)$ denote a fitness function. A (W, f) -recursive tree with fitness is the sequence of random trees $(\mathcal{T}_i)_{i \in \mathbb{N}}$ such that: \mathcal{T}_0 consists of a single node 0 with weight W_0 and for $n \geq 1$, \mathcal{T}_n is updated recursively from \mathcal{T}_{n-1} as follows:

- 1 Sample a vertex $j \in \mathcal{T}_{n-1}$ with probability proportional to its fitness, i.e., with probability

$$\frac{f(\deg^+(j, \mathcal{T}_{n-1}), W_j)}{\sum_{j=0}^{n-1} f(\deg^+(j, \mathcal{T}_{n-1}), W_j)}.$$

- 2 Connect j with an edge directed outwards to a new vertex n with weight W_n .

Remark 3.2. Due to the equivalence in law with trees associated with an (X, W) -CMJ branching process, by abuse of notation we refer to a sequence of recursive tree with fitness by $(\mathcal{T}_i)_{i \in \mathbb{N}}$, despite the fact that the vertex set of these trees is \mathbb{N}_0 rather than \mathcal{U}_∞ . ◀

Remark 3.3. The correspondence between recursive trees with fitness and the trees $(\mathcal{T}_i)_{i \in \mathbb{N}}$ associated with an (X, W) -CMJ process, when $X_w(i) \sim \text{Exp}(f(i+1, w))$ holds for all $i \in \mathbb{N}, w \in S$, is a consequence of the memory-less property and the fact that the minimum of exponential random variable is also exponentially distributed, with a rate given by the sum of the rates of the corresponding variables; see for example [34, Section 2.1]. The use of such continuous-time embeddings to analyse combinatorial processes was pioneered by Arthreya and Karlin [4]. This correspondence allows us to translate our main results to the infinite recursive tree, which we also denote by \mathcal{T}_∞ . ◀

3.1.1 The star/path transition in explosive recursive trees with fitness

Our first result pertains to the existence of a node of infinite degree or an infinite path in recursive trees with fitness. To this end, we make the following assumption:

$$\exists w^* \in S : \forall w \in S, j \in \mathbb{N} : f(j, w) \geq f(j, w^*) \quad \text{and} \quad \sum_{j=0}^{\infty} \frac{1}{f(j, w^*)} < \infty. \quad (w^*)$$

That is, there exists a minimiser $w^* \in S$ that, uniformly in $j \in \mathbb{N}$, minimises $f(j, \cdot)$, and the reciprocals of $f(j, w^*)$ are summable. Moreover, we define

$$\mu_n^w := \sum_{i=n}^{\infty} \frac{1}{f(i, w)}, \quad w \in S, n \in \mathbb{N}, \quad \text{and set } \mu_n := \mu_n^{w^*}. \quad (3.1)$$

Note that $\mu_n^w \leq \mu_n < \infty$ for all $w \in S$ by (w^*) .

Theorem 3.4 (Star/path in explosive recursive trees). *Let $(\mathcal{T}_i)_{i \in \mathbb{N}}$ be a (W, f) -recursive tree with fitness and assume f satisfies (w^*) . Then,*

1 If, for some $c < 1$, we have

$$\sum_{n=1}^{\infty} \mathbb{E} \left[\prod_{i=0}^{\infty} \frac{f(i, W)}{f(i, W) + c\mu_n^{-1}} \right] < \infty, \quad (3.2)$$

the tree \mathcal{T}_∞ contains a unique node of infinite degree, and no infinite path.

2 If either for some $c > 1$ and all $w \in S$, we have

$$\sum_{n=1}^{\infty} \mathbb{E} \left[\prod_{i=0}^{\infty} \frac{f(i, W)}{f(i, W) + c(\mu_n^w)^{-1} \log n} \right] = \infty, \quad (3.3)$$

or, as a weaker condition, Equation (2.9) is satisfied with $\nu_n^w := d\mu_n^w$ for $d < 1$, the tree \mathcal{T}_∞ contains a unique infinite path, and no node of infinite degree.

The proof of Theorem 3.4 appears in Section 5, in particular Section 5.2.1.

Remark 3.5. It turns out that Equation (3.3) is a sufficient condition for Equation (2.9) to be satisfied with $\nu_n^w := d\mu_n^w$ for $d < 1$. However, we include it as a general comparison to Equation (3.2). ◀

Remark 3.6. Analogues of Theorem 3.4 extend to more general distributions of $X_w(i)$, other than exponential distributions. In particular, we can apply the same techniques used to prove Item 2 of Theorem 3.4 whenever $(X_w(i), i \in \mathbb{N})$ are independent and $\text{Var}(X_w(i)) \leq K \mathbb{E}[X_w(i)]^2$ for some $K > 0$, possibly depending on w . In this case the expected value in (3.3) is replaced by the Laplace transform $\mathbb{E}[\mathcal{L}_{c(\mu_n^w)^{-1} \log n}(\mathcal{P}(\emptyset); W)]$. See the proof in Section 5.2.1 for more details. ◀

3.1.2 Sub-trees in explosive recursive trees with fitness when there is a star

In Section 2.3, we described a tree T as a finite subset of the Ulam-Harris tree containing the root, with a natural directed edge structure induced by parents being connected to children. We apply the same notion here, upon identifying labels of elements of \mathcal{T}_∞ with the Ulam-Harris labelling. For a tree $T \subseteq \mathcal{U}_\infty$ and $u \in \mathcal{T}_\infty$ (when we label the elements of \mathcal{T}_∞ with the Ulam-Harris labelling), we say that T appears as a sub-tree, rooted at u , in \mathcal{T}_∞ if $uT \subseteq \mathcal{T}_\infty$. Because the presence of ‘earlier siblings’ in a copy of a tree T can influence the probability of a tree emerging² it is convenient to also assume that T is *sibling closed*, where we define as follows. If $u = u_1 \cdots u_m \in T$ then $u = u_1 \cdots u_{m-1} \ell \in T$ for each $\ell \in [u_m]$.

The occurrence of a sibling-closed tree T in \mathcal{T}_∞ may also depend on the order in which the vertices in T appear, which can vary in such a way that they preserve the lexicographic ordering. An *ordering* of a tree T with $|T| = k + 1$ vertices, for some $k \in \mathbb{N}$, is a permutation $O : T \rightarrow \{0, 1, \dots, k\}$, such that $O(u) \leq O(v)$ if and only if $u \leq_L v$. Given an ordering O , we generally refer to the vertices of a tree T with $k + 1$ vertices as $\{v_0, \dots, v_k\}$, where $v_i := O^{-1}(i)$ for each i . Given a sibling-closed tree T , we let $\mathcal{O}(T)$ denote the set of all orderings of T . For a given ordering O and $j \leq k$, we let $O|_j$ denote the (also sibling-closed) tree on the vertex set $\{v_0, \dots, v_j\}$; note that this is well defined because O preserves the order \leq_L . Also note that each $O|_j$ inherits the natural directed edge structure from T . For a given vertex v_i , with $i \leq j$, $\text{deg}^+(v_i, O|_j)$ denotes its out-degree in $O|_j$.

We then have the following theorem.

Theorem 3.7 (Sub-tree counts). *Fix $k \in \mathbb{N}$, and let $(\mathcal{T}_i)_{i \in \mathbb{N}}$ be a (W, f) -recursive tree with fitness such that f satisfies (w^*) and so that \mathcal{T}_∞ contains a unique star. Moreover, assume that for each $w \in S$ we have $\mu_n^w \geq c_1(w) \mu_n$, where $0 < c_1(w) \leq 1$. Let T be a sibling-closed tree, with $|T| = k + 1$. The tree \mathcal{T}_∞ contains T as a sub-tree infinitely often if and only if*

$$\sum_{n=1}^{\infty} \sum_{O \in \mathcal{O}(T)} \mathbb{E} \left[\prod_{j=0}^k \frac{\prod_{\ell=0}^{\text{deg}^+(v_j, T)-1} f(\ell, W_{v_j})}{\sum_{i=0}^j f(\text{deg}^+(v_i, O|_j), W_{v_i}) \mathbf{1}_{\text{deg}^+(v_i, O|_j) < \text{deg}^+(v_i, T)} + \mu_n^{-1}} \right] = \infty. \quad (3.4)$$

Corollary 3.8. *Fix $k \in \mathbb{N}$ and assume that, for each $i \in [k]$ we have $\mathbb{E}[f(i, W)^k] < \infty$. Let T be a sibling-closed tree with $k + 1$ vertices. Then, under the assumptions of Theorem 3.7 the tree \mathcal{T}_∞ contains T as a sub-tree infinitely often if and only if $\sum_{n=1}^{\infty} \mu_n^k = \infty$.*

The proof of Theorem 3.7 appears in Section 5.2.2. On the other hand, the proof of Corollary 3.8 appears separately, in Section 6.2.1.

²For example, if the tree T corresponds to a path, it is intuitively less likely to have a path emerge where every node of the path is the first child of its parent rather than a path where some nodes are born later but, by random chance, produce children faster.

3.2 Phase transitions in specific models of explosive recursive trees with fitness

In this section we investigate three particular cases of the results presented in Section 3.1, where we are able to prove *phase transitions* for the structure of the infinite limiting tree \mathcal{T}_∞ in terms of the *fitness function* and the *vertex-weight distribution*. We assume that the vertex-weights are non-negative and real-valued, i.e. they take values in $S = [0, \infty)$. We also assume that the fitness function $f(i, W)$ grows *faster than linear* in the degree (i.e. its first argument). These cases are thus examples of *super-linear preferential attachment with fitness*.

3.2.1 Connection to existing literature: super-linear preferential attachment

As alluded to in the introduction, Section 1.1, a model of recursive trees (with fitness) that has received substantial attention the last two decades are preferential attachment models. Such models are thought to serve as a good explanation of the formation *real-world* networks due the *preferential attachment paradigm*, which suggests that networks are constructed by adding vertices and edges successively, in such a way that new vertices prefer to be connected to existing vertices with large degree. In particular, many of such models intrinsically give rise to properties also found in many real-world networks (i.e. the scale-free property and (ultra)small-world property), rather than such properties being imposed on the model. We refer to [29] and the references therein for an extensive overview of the literature on such models and their applications.

Super-linear preferential attachment is a particular type of preferential attachment where new vertices connect to existing vertices with out-degree i with a probability proportional to $f(i)$, for some fitness function $f : \mathbb{N}_0 \rightarrow (0, \infty)$ such that $\sum_{i=0}^{\infty} f(i)^{-1} < \infty$. Most often, as in e.g. [16, 58, 64], the case $f(i) = (i + 1)^p$ for some $p > 1$ is studied, though there are also choices for f that satisfy the summability condition such that $f(i)i^{-p} \rightarrow 0$ as $i \rightarrow \infty$ for any $p > 1$. We coin these functions *barely super-linear*. Though Pólya urn models with barely super-linear fitness functions have been studied previously [27], as far as the authors are aware this is the first case such fitness functions are treated for preferential attachment models.

Super-linear preferential attachment models are suggested to possibly explain the formation of real-world networks such as the Internet, where these networks are in a ‘preasymptotic regime’ (are of relatively small size) where the explosive nature of the model cannot be observed yet, based on statistical parameter estimation, simulations, and non-rigorous analysis [46, 47, 60].

The inclusion of vertex-weights allows for a more heterogeneous and hence more realistic model, where different vertices may behave differently (in distribution), even when their out-degree is the same, as also discussed in the introduction. The presence of vertex-weights often leads to rich behaviour where phase transitions based on the vertex-weight distribution can be observed (see the introduction for examples), which we show in this section to be case for the examples we consider here, too.

We study a number of examples for which we can apply the results in Section 3.2. We state the assumptions for the fitness function f and the vertex-weight distribution, after which we present the results related to Theorems 3.4 and 3.7. We conclude the section with a discussion of these results in Section 3.2.4, where we also provide some interesting phase diagrams in Figures 1 and 2, and with some open problems in Section 3.4.

3.2.2 Assumptions for the fitness function and vertex-weight distribution

When stating the particular assumptions for the fitness function f and the vertex-weight distribution, it is helpful to use the notions of *slowly-varying* and *regularly-varying* functions, which we recall in the following definition.

Definition 3.9. A measurable function $L : [0, \infty) \rightarrow [0, \infty)$ is said to be *slowly varying* if for any $a > 0$ we have

$$\lim_{x \rightarrow \infty} \frac{L(ax)}{L(x)} = 1.$$

We say a measurable function $g : [0, \infty) \rightarrow [0, \infty)$ is *regularly varying* with exponent $\beta \in \mathbb{R}$ if $g(x) = x^\beta L(x)$, where $L : [0, \infty) \rightarrow [0, \infty)$ is slowly varying. Finally, we say that a *random variable* W is regularly varying with exponent $z < 0$ if the tail distribution $\mathbb{P}(W \geq x)$ is a regularly-varying function (in x) with exponent $z < 0$.

We then assume that the fitness function satisfies the following assumption.

Assumption 3.10 (Fitness function). The fitness function f is such that Equation (w^*) is satisfied with $w^* = 0$. Furthermore, there exists $s : \mathbb{N}_0 \rightarrow (0, \infty)$, which we call the *degree function*, and continuous functions $g : [0, \infty) \rightarrow (0, \infty)$ and $h : [0, \infty) \rightarrow [0, \infty)$, which we call the *weight functions*, such that f satisfies

$$f(i, w) := g(w)s(i) + h(w), \quad i \in \mathbb{N}_0, w \geq 0.$$

We then distinguish the following two cases, based on the weight functions.

- **Additive weights.** $g \equiv 1$ and h is regularly varying with exponent 1.
- **Mixed weights.** g and h are regularly varying with exponents 1 and $\gamma \geq 0$, respectively.

Remark 3.11. The assumption that $w^* = 0$ is not necessary, but used to simplify notation and computations. The results presented here follow equivalently for $w^* > 0$ as well. ◀

Remark 3.12. The function g and h are regularly varying with exponents 1 and $\gamma \geq 0$ in the mixed case; $g \equiv 1$ and h is regularly varying with exponent 1 in the additive case. The choice of the exponents is due to the fact that, when the vertex-weights are regularly varying with exponent $-(\alpha - 1) < 0$, then $g(W)$ and $h(W)$ are random variables that are regularly varying with exponents $-(\alpha - 1)$ and $-(\alpha - 1)/\gamma$, respectively (see Lemma 6.6 for details). Hence, changing the exponent of, for example, the regularly-varying function g to $\zeta \neq 1$ in the mixed case, is equivalent to changing the exponent of the regularly-varying random variable W from $-(\alpha - 1)$ to $-(\alpha - 1)/\zeta$ and changing the exponent of the regularly-varying function h from γ to γ/ζ . As such, we take $\zeta = 1$ without loss of generality. The function h is regularly varying with exponent 1 in the additive case without loss of generality for the same reason. ◀

Depending on the precise form of the degree function, the model behaviour markedly differs. We assume the *degree function* s satisfies the following assumption.

Assumption 3.13 (Degree function). The degree function $s : \mathbb{N}_0 \rightarrow (0, \infty)$ satisfies one of the following cases.

- **Super-linear preferential attachment.** s is regularly varying with exponent $p > 1$.
- **Barely super-linear preferential attachment.** s is regularly varying with exponent 1, such that $\sum_{i=0}^{\infty} s(i)^{-1} < \infty$.

As a particular example of the barely super-linear case, we consider

- **Barely super-linear log-stretched preferential attachment.** For some $\beta \in (0, 1)$,

$$s(i) = (i + 1) \exp((\log(i + 1))^\beta), \quad i \in \mathbb{N}_0. \quad (3.5)$$

Remark 3.14. We note that these choices for the fitness function f are not exhaustive, but do cover a wide range of examples. In particular, the weight types considered, i.e. additive or mixed weights, are common in the literature of linear preferential attachment with fitness (see e.g. [21, 26, 53, 19, 34, 22]). When the vertex-weights are constant almost surely, the additive and mixed cases all fall into the same classes of super-linear preferential attachment. The barely super-linear class has not been studied previously, as far as the authors are aware. ◀

Finally, we require several assumptions on the distribution of the vertex-weights. For different choices of the fitness function f , different assumptions are required, which are summarised in the following overview.

Assumption 3.15 (Vertex-weight distribution). The vertex-weights $(W_i)_{i \in \mathbb{N}}$ are i.i.d. and their tail distribution satisfies one (or more) of the following conditions.

- **Power law.** Let $\alpha > 1$. We have the following two conditions.

- 1 There exist $\bar{x} > 0$ and a slowly-varying function $\bar{\ell} : [0, \infty) \rightarrow [0, \infty)$ such that

$$\mathbb{P}(W \geq x) \leq \bar{\ell}(x)x^{-(\alpha-1)}, \quad x \geq \bar{x}. \quad (3.6)$$

- 2 There exist $\underline{x} > 0$ and a slowly-varying function $\underline{\ell} : [0, \infty) \rightarrow [0, \infty)$ such that

$$\mathbb{P}(W \geq x) \geq \underline{\ell}(x)x^{-(\alpha-1)}, \quad x \geq \underline{x}. \quad (3.7)$$

- **Log-stretched exponential.** Let $\nu > 1$. We have the following two conditions.

- 1 There exist $\bar{c}, \bar{x} > 0$ such that

$$\mathbb{P}(W \geq x) \leq e^{-\bar{c}(\log x)^\nu}, \quad x > \bar{x}. \quad (3.8)$$

- 2 There exist $\underline{c}, \underline{x} > 0$ such that

$$\mathbb{P}(W \geq x) \geq e^{-\underline{c}(\log x)^\nu}, \quad x \geq \underline{x}. \quad (3.9)$$

3.2.3 Phase transitions for super-linear preferential attachment models with fitness

Based on these assumptions for the fitness function, degree function, and vertex-weight distribution, we can formulate the following theorem that treats the appearance of a unique infinite-degree vertex or a unique infinite path in \mathcal{T}_∞ .

Theorem 3.16. *Let $(\mathcal{T}_i)_{i \in \mathbb{N}}$ be a (W, f) -recursive tree with fitness, where the fitness function f and degree function s satisfy one of the cases in Assumptions 3.10 and 3.13, respectively, and the vertex-weight distribution satisfies one of the cases in Assumption 3.15. The tree \mathcal{T}_∞ either contains a unique vertex with infinite degree and no infinite path almost surely, or contains a unique infinite path and no vertex with infinite degree almost surely, when the following conditions are met, based on the fitness function, degree function, and vertex-weight assumptions:*

| Fitness | Degree | Star | Path |
|----------|---------------|--|--|
| Mixed | Super-linear | (3.6) & $(p-1)(\alpha-1) > (\gamma - \frac{\gamma-1}{p}) \vee 1$ | (3.7) & $(p-1)(\alpha-1) < (\gamma - \frac{\gamma-1}{p}) \vee 1$ |
| Additive | Super-linear | (3.6) & $p(\alpha-1) > 1$ | (3.7) & $p(\alpha-1) < 1$ |
| Mixed | Log-stretched | (3.8) & $\beta\nu > 1$ | (3.9) & $\beta\nu < 1$ |

Table 1: The first column represents the form of the fitness function, as in Assumption 3.10; the second represents the form of the degree function, as in Assumption 3.13. The third and fourth column, respectively, list the required assumptions on the vertex-weight distribution, as in Assumption 3.15, together with the choices of the parameters that lead to either a unique node of infinite degree or a unique infinite path.

The proof of Theorem 3.16 appears in Section 6.1.

Remark 3.17. Though Theorem 3.16 is presented for exponentially distributed inter-birth times, we expect the same results to hold for a large family of distributions, such that $X_w(u, j)$ has mean $1/f(j-1, w)$ for each $u \in \mathcal{U}_\infty, j \in \mathbb{N}$, and $w \in [0, \infty)$. In particular, we discuss the extension of Theorem 3.16 to the following examples Appendix B:

- 1 **Gamma distribution:** For each $u \in \mathcal{U}_\infty, j \in \mathbb{N}$, and $w \in [0, \infty)$, the inter-birth time $X_w(u, j)$ follows a $\text{Gamma}(k, kf(j-1, w))$ distribution, for some $k > 0$.
- 2 **Beta distribution:** For each $u \in \mathcal{U}_\infty, j \in \mathbb{N}$, and $w \in [0, \infty)$, the inter-birth time $X_w(u, j)$ equals $\frac{\alpha+\beta}{\alpha} \frac{1}{f(j-1, w)} B(u, j)$, where $(B(u, j))_{u \in \mathcal{U}_\infty, j \in \mathbb{N}}$ is a sequence of i.i.d. copies of a $\text{Beta}(\alpha, \beta)$ random variable, for some $\alpha \geq 1$ and $\beta \leq (0, 1]$.
- 3 **Rayleigh distribution:** For each $u \in \mathcal{U}_\infty, j \in \mathbb{N}$, and $w \in [0, \infty)$, the inter-birth time $X_w(u, j)$ follows a $\text{Rayleigh}(\sqrt{2/\pi}/f(j-1, w))$ distribution. ◀

Remark 3.18. In Section A we discuss three more examples of barely super-linear preferential attachment with fitness: the **log-stretched** case with **additive** fitness, as well as a **poly-log** case with either additive or **mixed** fitness. Again, the extension of the results to the other inter-birth time distributions as in Remark 3.17 apply here, too.

We do not include these results here, as we cannot prove a complete phase diagram, i.e. for certain parameter choices we cannot prove the appearance of a unique infinite-degree vertex nor a unique infinite path. ◀

Remark 3.19. When the vertex-weights are almost surely bounded (or, in particular, a deterministic constant), it follows from the above theorem that in all cases, a unique vertex

with infinite degree emerges in \mathcal{T}_∞ almost surely. Indeed, in such a case the vertex-weight distribution would satisfy, the upper bounds in (3.6) or (3.8) for any $\alpha > 1, \nu > 1$. As such, we can take $\alpha \rightarrow \infty, \nu \rightarrow \infty$ to conclude the claim. In the case $s(i) = (i + 1)^p$ with $p > 1$, this recovers the results of Oliveira and Spencer [58, Theorem 1.1]. ◀

Remark 3.20. It is interesting to note that, though two different techniques with distinct assumptions are used to prove Theorems 2.5 and 2.8 (for the existence of an infinite star or path in \mathcal{T}_∞), the application of these two general results in Theorem 3.16 allows us to obtain a complete phase diagram for the three examples discussed here. ◀

When the infinite tree \mathcal{T}_∞ contains a unique vertex with infinite degree almost surely, we can also quantify Theorem 3.21, in the sense that, depending on assumptions on the fitness function f and vertex-weight distribution, we can determine almost surely whether or not \mathcal{T}_∞ contains an infinite number of copies of which kinds of sub-trees. We remark that we can do this in a relatively general manner, subject to the assumption that \mathcal{T}_∞ contains a unique vertex with infinite degree. That is, for any degree function s that satisfies the (barely) super-linear cases in Assumption 3.13 the results below apply.

To this end, we define for a finite tree T and constant $z > 0$,

$$G_1 = G_1(T, z) := \sum_{v \in T} \deg^+(v, T) \mathbf{1}_{\deg^+(v, T) > z}, \quad G_2 = G_2(T, z) := \sum_{v \in T} \mathbf{1}_{\deg^+(v, T) > z}. \quad (3.10)$$

We can then have the following result.

Theorem 3.21. *Let $(\mathcal{T}_i)_{i \in \mathbb{N}}$ be a (W, f) -recursive tree with fitness, where we assume that f satisfies Assumption 3.10, that the degree function s satisfies Assumption 3.13, and that the vertex-weight distribution satisfies (3.6) and (3.7) for some slowly-varying functions $\underline{\ell}, \bar{\ell}$, respectively, and some $\alpha > 1$. Furthermore, we assume that \mathcal{T}_∞ contains a unique vertex with infinite degree, almost surely. Fix $k \in \mathbb{N}$ and let T be a sibling-closed tree of size $k + 1$. Then, we have the following.*

- **Super-linear case:** We take $\gamma = 1$ for additive weights, $\gamma \geq 0$ for mixed weights, and set $z = (\alpha - 1)/(\gamma \vee 1)$. The tree T almost surely appears infinitely often as a sub-tree of \mathcal{T}_∞ when

$$p < 1 + \frac{1}{k - (G_1(T, z) - zG_2(T, z))}. \quad (3.11)$$

The tree T almost surely appears finitely often as a sub-tree of \mathcal{T}_∞ when

$$p > 1 + \frac{1}{k - (G_1(T, z) - zG_2(T, z))}. \quad (3.12)$$

- **Barely super-linear case:** The tree T appears as a sub-tree of \mathcal{T}_∞ infinitely often, almost surely.

The proof of Theorem 3.21 appears in Section 6.2, in particular, Section 6.2.2.

Remark 3.22. In fact, we have a strengthening of Theorem 3.21 which applies to any (W, f) -recursive tree with fitness (in particular without the assumption that W is real-valued), as long as, for any $j \in \mathbb{N}_0$, $f(j, W)$ satisfies (3.6) and (3.7) for some slowly-varying functions $\bar{\ell}_j, \underline{\ell}_j$, respectively, and an exponent $z > 0$, and for each $k \in \mathbb{N}_0$ there exists $i_k \in \{0, \dots, k\}$ such that $\inf_{i \leq k} f(i, w) = f(i_k, w)$ for all $w \in S$. See Proposition 6.5 for more details. ◀

Remark 3.23. When the vertex-weight distribution satisfies $\mathbb{E}[W^a] < \infty$ for all $a > 0$, it directly follows from the assumptions on the fitness function f in Assumption 3.10 and Corollary 3.8 that the results in Theorem 3.21 extend to this case, where we set $G_1(T, z) = G_2(T, z) = 0$.

Furthermore, we stress that for *all* barely super-linear cases and for certain super-linear cases (see the upcoming discussion), sub-trees T of arbitrary size appear infinitely often as a sub-tree of \mathcal{T}_∞ , almost surely. This is markedly different when compared to the super-linear preferential attachment model $f(i, w) = (i + 1)^p$ ($p > 1$) studied by Oliveira and Spencer in [58], where only sub-trees T with size at most $\lceil (p - 1)^{-1} \rceil$ can appear infinitely often almost surely. ◀

3.2.4 Discussion related to Theorem 3.21: the super-linear case

Let us provide some intuition for Theorem 3.21 by discussing two particular examples.

Super-linear degree, mixed weights. We let $f(i, w) = (w + 1)(i + 1)^p$ for some $p > 1$. That is, we consider the mixed weights case for the fitness function f , with $g(x) := x + 1$, $h \equiv 0$, and $\gamma = 0$, as in Assumption 3.10, and the super-linear case for the degree function s , i.e. $s(i) := (i + 1)^p$, $p > 1$, as in Assumption 3.13. We require that \mathcal{T}_∞ contains a unique vertex with infinite degree almost surely, so that, by Theorem 3.16, we assume that the vertex-weight distribution satisfies (3.6) and that $(p - 1)(\alpha - 1) > 1$. Then, additionally assume that the vertex-weight distribution satisfies (3.7) with the same $\alpha > 1$ but potentially with a different slowly-varying function. Now, Theorem 3.21 states that a tree T of size $k + 1$, for some $k \in \mathbb{N}$, appears infinitely often as a sub-tree of \mathcal{T}_∞ almost surely when

$$p < 1 + \frac{1}{k - (G_1(T, \alpha - 1) - (\alpha - 1)G_2(T, \alpha - 1))}.$$

First, noting that the sum of all degrees equals $|T| - 1 = k$, we observe that

$$G_1(T, \alpha - 1) - (\alpha - 1)G_2(T, \alpha - 1) \in [0, k), \quad (3.13)$$

for any choice of T and α , so that the upper bound yields a restriction on p . We omit the arguments of G_1 and G_2 from here on out for ease of writing. Combining our two assumptions we then require that

$$\frac{1}{\alpha - 1} < p - 1 < \frac{1}{k - (G_1 - (\alpha - 1)G_2)},$$

and we can only find $p > 1$ that satisfy both inequality when

$$k - (G_1 - (\alpha - 1)G_2) < \alpha - 1. \quad (3.14)$$

Now, if $k < \alpha - 1$, there is no vertex in T with an out-degree larger than $\alpha - 1$, so that $G_1 = G_2 = 0$ and the inequality is satisfied. For $k \geq \alpha - 1$ we distinguish two cases. (i) There is no vertex in T with a degree larger than $\alpha - 1$. It again follows that $G_1 = G_2 = 0$, so that the inequality in (3.14) is *not* satisfied; (ii) There exists at least one vertex with degree larger than $\alpha - 1$. Then, $k - G_1 > 0$ since k equals the sum of all degrees, whilst $(\alpha - 1)(1 - G_2) \leq 0$ (since $G_2 \geq 1$), so that (3.14) is not satisfied.

We thus conclude that, when $(p - 1)(\alpha - 1) > 1$, only trees T with size $k + 1$, where $k < \alpha - 1$, appear infinitely often. In particular, we do not require any assumptions on the *structure* of

such trees T ; only their size is relevant. The reversed inequality in Theorem 3.21 can be analysed in a similar manner to derive the phase diagram in Figure 1.

Super-linear degree, additive weights. We let $f(i, w) = (i + 1)^p + w$ for some $p > 1$. That is, we consider the additive weight case for the fitness function f , with $h(x) = x$, as in Assumption 3.10, and the super-linear case for the degree function s , i.e. $s(i) = (i + 1)^p$, $p > 1$, as in Assumption 3.13. We make the same assumptions as in the first example, except that we now require $p(\alpha - 1) > 1$. First, when p is so large that $(p - 1)(\alpha - 1) > 1$, we can derive the same conclusions as in the first example. When p is such that

$$\frac{1}{\alpha - 1} \vee 1 < p < 1 + \frac{1}{\alpha - 1}, \quad (3.15)$$

different behaviour can be observed. Here, as we illustrate with the following particular family of trees, we observe the peculiar behaviour that the *structure* of a tree T plays a role in terms of whether it appears infinitely or finitely often as a sub-tree in \mathcal{T}_∞ .

Let T be an m -ary tree of size $k + 1$ for some $m, k \in \mathbb{N}$ such that $k = \ell m$ for some $\ell \in \mathbb{N}$, i.e. a tree where the ℓ internal vertices (non-leaf vertices) have out-degree m (note that stars are the particular case $\ell = 1$). We observe that, by distinguishing the two cases $m > \alpha - 1$ and $m \leq \alpha - 1$ (where in the latter case $G_1 = G_2 = 0$),

$$G_1(T, \alpha - 1) - (\alpha - 1)G_2(T, \alpha - 1) = \max\{k - (\alpha - 1)\ell, 0\}.$$

As a result, recalling $k = \ell m$, an m -ary tree T appears infinitely often as a sub-tree of \mathcal{T}_∞ , almost surely, when

$$\frac{1}{\alpha - 1} \vee 1 < p < 1 + \frac{1}{\ell \min\{\alpha - 1, m\}}.$$

Observe that for $\ell = 1$, i.e. stars of size $k + 1 = m + 1$, these inequalities are satisfied by any p that satisfies (3.15), so that a star of *any* size appears infinitely often as a sub-tree of \mathcal{T}_∞ , almost surely, when p satisfies (3.15). For $\ell \geq 2$, these inequalities can be satisfied only when $\alpha \leq 2$ and $\ell < 1/(2 - \alpha)$ or when $\alpha > 2$.

Again, the reversed inequality in Theorem 3.21 can be analysed in a similar manner to derive the phase diagram in Figure 2, where some of the phases for ternary trees ($m = 3$) are shown.

We thus observe that it is possible for two trees T_1, T_2 of sizes $k_1 + 1 < k_2 + 1$, respectively, to appear finitely and infinitely often as a sub-tree of \mathcal{T}_∞ , almost surely. This behaviour is completely opposite to the behaviour of the first example (or when $(p - 1)(\alpha - 1) > 1$ in this example).

As a final remark, any tree of any size appears infinitely often as a sub-tree of \mathcal{T}_∞ in the *barely sub-linear* case discussed in Theorem 3.21. As discussed in Section 2.5, this behaviour has not been observed in explosive tree models studied so far.

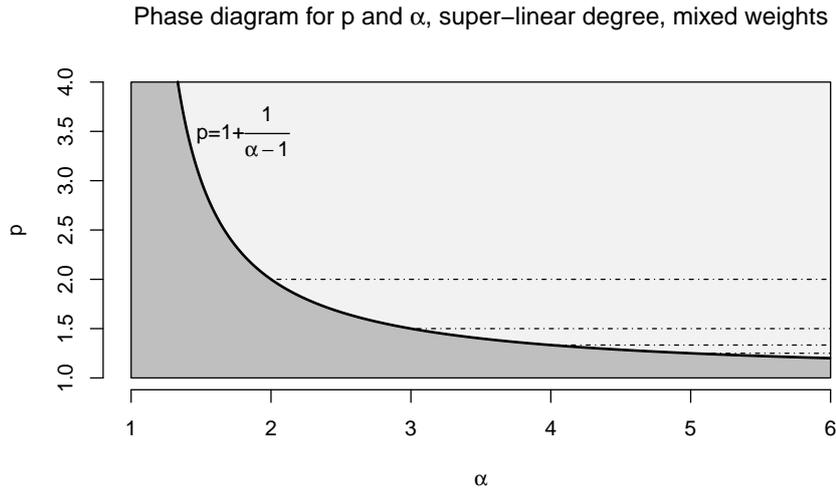


Figure 1: The phase diagram for the case $f(i, w) = (w + 1)(i + 1)^p$. Below the curve $(p - 1)(\alpha - 1) = 1$ (the dark shaded region), the tree \mathcal{T}_∞ contains a unique infinite path, whereas above the curve (the light shaded region) the tree contains a unique vertex of infinite degree. Moreover, in the light shaded region, below the k th horizontal dotted line, corresponding to $1 + 1/k$, (with $k = 1, 2, 3, 4$ visible here), any tree of size $k + 1$ appears as a sub-tree of a child of the node of infinite degree infinitely often, whereas above the k th horizontal line, in the light shaded region, trees of size $k + 1$ appear only finitely often as a sub-tree of a child of the node of infinite degree.

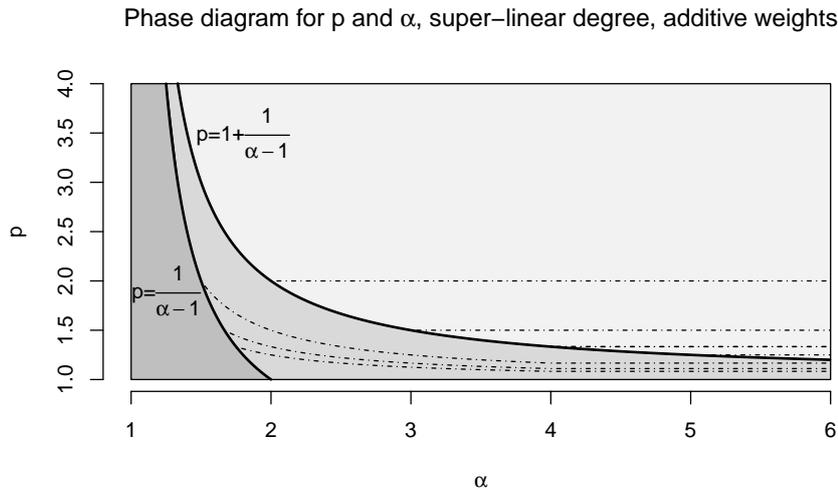


Figure 2: The phase diagram for the case $f(i, w) = (i + 1)^p + w$. Below the curve $p(\alpha - 1) = 1$ (the dark shaded region), the tree \mathcal{T}_∞ contains a unique infinite path, whereas above the curve (the lighter shaded regions) the tree contains a unique vertex of infinite degree. In the area above the line $(p - 1)(\alpha - 1) = 1$ (lightest shaded region), below the k th horizontal dotted line, corresponding to $1 + 1/k$, (with $k = 1, 2, 3, 4$ visible here), any tree of size $k + 1$ appears as a sub-tree of a child of the node of infinite degree infinitely often, whereas above the k th horizontal line, in the lightest shaded region, trees of size $k + 1$ appear only finitely often as a sub-tree of a child of the node of infinite degree. In the area in between the curves $p(\alpha - 1) = 1$ and $(p - 1)(\alpha - 1) = 1$ (the semi-dark region) a star of *any* size appears as a sub-tree of a child of the node of infinite degree infinitely often. Additionally, below the ℓ th curve in this region, corresponding to $(p - 1)\ell \min\{\alpha - 1, 3\} = 1$, (with $\ell = 2, 3, 4$ visible here) ternary trees of size $3\ell + 1$ appear as a sub-tree of the node of infinite degree infinitely often, whereas above the ℓ th curve, such trees only appear finitely often.

3.3 Proof techniques

The proofs of the results in Section 3 generally apply the results of Section 2. However, we are unaware of previous proofs of Lemma 5.1, used to prove Theorem 3.4. Moreover, exploiting the memory-less property of the exponential distribution allows the derivation of a necessary and sufficient condition for the emergence of a structure infinitely often when \mathcal{T}_∞ contains an infinite star, as in Theorem 3.7. The proof of this theorem, we believe, is more elegant than the approach used to prove [58, Theorem 1.2] for the particular case $f(i, w) = (i+1)^p, p > 1$. The assumptions made in Section 3.2 allow us to deduce a variety of phase-transitions in applied models (cf. Theorem 3.16), which we believe extend to a fairly general family of distributions (see Remark 3.17).

We prove the most general results, as presented in Section 3.1, in Section 5, and prove the results for the examples presented in Section 3.2 in Section 6.

3.4 Open problem

It is unclear whether or not *any* explosive (X, W) -CMJ process for which the vertex-weights are almost surely constant *always* yields an infinite star almost surely. In other words, when assuming that the $(X(i))_{i \in \mathbb{N}}$ are mutually independent and positive, we have the following open problem.

Open problem 3.24. *Consider an explosive (X, W) -CMJ process $(\mathcal{T}_t)_{t \geq 0}$, such that the sequence $(X(i))_{i \in \mathbb{N}}$ are mutually independent and positive. Is it the case that almost surely \mathcal{T}_∞ contains a unique vertex with infinite degree?*

Remark 3.25. In the case of the recursive tree with fitness, if $X(i) \sim \text{Exp}(g(i))$, with g unbounded, and convex, we believe that the affirmative of Open problem 3.24 holds, by an argument using a combination of the result of Galashin [23, Theorem 1] and Proposition 4.4.

◀

Remark 3.26. Note that the counter-example in Theorem 2.15 relies on the dependence of the $(X(i))_{i \in \mathbb{N}}$ on the weights. Thus, if \mathcal{T}_∞ *does* almost surely contain a unique node of infinite degree in Open problem 3.24, one may interpret this, informally, as saying that ‘new nodes are unable to out-compete older nodes, without the influence of a random weight’.

◀

4 Proofs of main results

This section is dedicated to proving the most general results, as presented in Section 2. We prove the existence of an infinite star (cf. Theorem 2.5) in Section 4.1, prove the existence of an infinite path (cf. Theorem 2.8) and the structural result of sub-trees in the star regime (cf. Theorem 2.10) in Section 4.2, and finally prove the uniqueness properties (cf. Theorem 2.12) in Section 4.3.

4.1 Sufficient criteria for a star

This section is dedicated to the proof of Theorem 2.5. We first have the following lemma:

Lemma 4.1. Fix $k, n \in \mathbb{N}$, let $(a_1, \dots, a_k) \in \mathbb{N}^k$. Let $(Z_i)_{i \in [k]}$ be independent random variables, and let $(X(\ell))_{\ell \geq n+1}$ be as in (2.1), satisfying Conditions 1 and 2 of Assumption 2.2. Then, there exists $n_0 > 0$ and $C = C(n_0) > 0$ such that, for all $n \geq n_0$,

$$\mathbb{P} \left(\sum_{i=1}^k Z_i \leq \sum_{\ell=n+1}^{\infty} X(\ell) \right) \leq \mathbb{P} \left(\sum_{i=1}^k Z_i \leq Y_n \right) \leq C \prod_{j=1}^k \mathcal{L}_{c\mu_n^{-1}}(Z_i), \quad (4.1)$$

where Y_n is as in (2.4) and independent of the $(Z_i)_{i \in [k]}$.

Proof. First note that the first inequality in (4.1) is an immediate consequence of (2.4). Then, for any $\lambda > 0$,

$$\mathbb{P} \left(\sum_{i=1}^k Z_i \leq Y_n \right) = \mathbb{P} \left(\exp \left(\lambda Y_n - \lambda \sum_{i=1}^k Z_i \right) \geq 1 \right) \leq \mathcal{M}_{Y_n}(\lambda) \prod_{i=1}^k \mathcal{L}_{\lambda}(Z_i),$$

where the last inequality uses Markov's inequality and the independence of the random variables. Next, setting $\lambda := c\mu_n^{-1}$ and recalling that, by (2.5), we can set

$$C := \limsup_{n \rightarrow \infty} \mathcal{M}_{Y_n}(c\mu_n^{-1}) < \infty,$$

we deduce the result. \square

We now introduce the following terminology, used in the remainder of the section, which, although not strictly needed, we believe makes the proofs conceptually easier to understand. For $a, b \in \mathcal{U}_{\infty}$ we say that

“ a has at least k children before b explodes”

if $\mathcal{B}(a) + \mathcal{P}_k(a) < \mathcal{B}(b) + \mathcal{P}(b)$. We say that

“ a explodes before all of its ancestors”

if, for each $\ell < |a|$, we have $\mathcal{B}(a) + \mathcal{P}(a) < \mathcal{B}(a_{|\ell}) + \mathcal{P}(a_{|\ell})$. Finally, for $a \in \mathcal{U}_{\infty}$ with $|a| \geq 1$, we say that $a = a_1 \cdots a_m$ is a_1 -conservative if, for each $j \in \{2, \dots, m\}$, we have $a_j \leq a_1$. (Note that this implies that any a such that $|a| = 1$ is a_1 -conservative.)

Lemma 4.2. Under Assumption 2.2, there exist $\eta < 1$ and $K = K(\eta) > 0$ such that for all $a_1 > K(\eta)$, all integers $m \in \mathbb{N}$, and some constant $C > 0$,

$$\sum_{\substack{a: |a|=m \\ a \text{ is } a_1\text{-conservative}}} \mathbb{P}(a \text{ has at least } a_1 \text{ children before } \emptyset \text{ explodes}) \leq C \eta^{m-1} \mathbb{E} \left[\mathcal{L}_{c\mu_{a_1}^{-1}}(\mathcal{P}_{a_1}; W) \right].$$

Proof. Suppose that $a = a_1 \cdots a_m \in \mathcal{U}_{\infty}$. For a to have at least a_1 children before the explosion of \emptyset , in particular, each of the births corresponding to the ancestors of a need to occur (leading to a term as in Equation (1.1) with $u = \emptyset$ and $v = a$). Thus, for a_1 sufficiently large, by (1.1) and Lemma 4.1, we have

$$\begin{aligned} \mathbb{P}(a \text{ has at least } a_1 \text{ children before } \emptyset \text{ explodes}) &= \mathbb{P}(\mathcal{B}(a) + \mathcal{P}_{a_1}(a) < \mathcal{P}(\emptyset)) \quad (4.2) \\ &= \mathbb{P} \left(\left(\sum_{j=1}^{m-1} \mathcal{P}_{a_{j+1}}(a_{|j}) \right) + \mathcal{P}_{a_1}(a) \leq \sum_{k=a_1+1}^{\infty} X(k) \right) \\ &\stackrel{(4.1)}{\leq} C \prod_{j=1}^{m-1} \mathbb{E} \left[\mathcal{L}_{c\mu_{a_1}^{-1}}(\mathcal{P}_{a_{j+1}}; W) \right] \times \mathbb{E} \left[\mathcal{L}_{c\mu_{a_1}^{-1}}(\mathcal{P}_{a_1}; W) \right], \end{aligned}$$

where the last line follows from the fact that, by (2.1), for each $u \in \mathcal{U}_\infty$ the sequence $(\mathcal{P}_j(u))_{j \in \mathbb{N}}$ is independent and distributed like $(\mathcal{P}_j(\emptyset))_{j \in \mathbb{N}}$. When we sum over the possible conservative sequences a that are a_1 -conservative, each a_j takes values between 1 and a_1 , for $j = 2, \dots, m$. Thus,

$$\begin{aligned} & \sum_{\substack{a: |a|=m \\ a \text{ } a_1\text{-conservative}}} \mathbb{P}(a \text{ has at least } a_1 \text{ children before } \emptyset \text{ explodes}) \\ & \leq \sum_{a_2=1}^{a_1} \sum_{a_3=1}^{a_1} \cdots \sum_{a_m=1}^{a_1} C \prod_{j=1}^{m-1} \mathbb{E} \left[\mathcal{L}_{c\mu_{a_1}^{-1}}(\mathcal{P}_{a_{j+1}}; W) \right] \times \mathbb{E} \left[\mathcal{L}_{c\mu_{a_1}^{-1}}(\mathcal{P}_{a_1}; W) \right] \\ & = C \left(\sum_{\ell=1}^{a_1} \mathbb{E} \left[\mathcal{L}_{c\mu_{a_1}^{-1}}(\mathcal{P}_\ell; W) \right] \right)^{m-1} \mathbb{E} \left[\mathcal{L}_{c\mu_{a_1}^{-1}}(\mathcal{P}_{a_1}; W) \right]. \end{aligned} \quad (4.3)$$

We now need only show that for a_1 sufficiently large,

$$\sum_{\ell=1}^{a_1} \mathbb{E} \left[\mathcal{L}_{c\mu_{a_1}^{-1}}(\mathcal{P}_\ell; W) \right] < \eta.$$

Indeed, since $\sum_{i=\ell}^{\infty} \mathbb{E} \left[\mathcal{L}_{c\mu_\ell^{-1}}(\mathcal{P}_i; W) \right] < \infty$ by (2.8) in Assumption 2.2, there exists $L = L(\eta) > 0$ such that, for all $a_1 > L$,

$$\sum_{\ell=L}^{a_1} \mathbb{E} \left[\mathcal{L}_{c\mu_{a_1}^{-1}}(\mathcal{P}_\ell; W) \right] < \sum_{\ell=L}^{\infty} \mathbb{E} \left[\mathcal{L}_{c\mu_\ell^{-1}}(\mathcal{P}_\ell; W) \right] < \frac{\eta}{2}, \quad (4.4)$$

where the inequality uses the fact that $c\mu_n^{-1}$ is non-decreasing in n . On the other hand, since $\lim_{n \rightarrow \infty} c\mu_n^{-1} = \infty$, by bounded convergence (bounding the integrand by 1) we have

$$\lim_{n \rightarrow \infty} \sum_{\ell=1}^{L-1} \mathbb{E} \left[\mathcal{L}_{c\mu_n^{-1}}(\mathcal{P}_\ell; W) \right] = 0.$$

As a result, for some $K \geq L$ sufficiently large and for all $a_1 > K$, we arrive at

$$\sum_{\ell=1}^{L-1} \mathbb{E} \left[\mathcal{L}_{c\mu_\ell^{-1}}(\mathcal{P}_\ell; W) \right] < \frac{\eta}{2}. \quad (4.5)$$

Combining Equations (4.4) and (4.5) in (4.3), we conclude the proof. \square

The above lemma provides an upper bound for the probability of the event that a vertex a explodes before the root of the tree, in the case that a is a_1 -conservative. However, when a does not satisfy this condition, we can view a as a concatenation of a number of conservative sequences. That is, we write $a = \bar{b}_1 \cdots \bar{b}_\ell$, where $\bar{b}_i = b_{i1} \cdots b_{im_i}$ for each $i \in [\ell]$ and for some $\ell \in \mathbb{N}$, $(m_i)_{i \in [\ell]} \in \mathbb{N}^\ell$, and $(b_{i,j})_{i \in [\ell], j \in [m_i]}$, such that \bar{b}_i is b_{i1} -conservative for each $i \in [\ell]$. By the independence of birth processes of distinct individuals (or in fact, the independence of disjoint sub-trees) by Equation (2.2), we are able to apply Lemma 4.2 to each conservative sequence in the concatenation to arrive at a bound for the expected number of individuals that explode before all its ancestors.

Proposition 4.3. *Under Assumption 2.2, there exists $K' > 0$ sufficiently large, such that*

$$\mathbb{E} [\#\{a \in \mathcal{U}_\infty : a_1 > K', a \text{ explodes before all its ancestors}\}] < \infty.$$

Proof. As explained before the proposition statement, we think of sequences $a \in \mathcal{U}_\infty$ as a concatenation of conservative sequences. Let $a = a_1 \dots a_m$ be a sequence of length $m \in \mathbb{N}$, and assume that there exist $k \in [m]$ and indices $I_1 < I_2 < \dots < I_k$ such that $a_1 =: a_{I_1} < a_{I_2} < \dots < a_{I_k}$. That is, the I_j are the indices of the running maxima of the sequence a . For brevity of notation, we also set $I_{k+1} := m + 1$, and set $a_0 = \emptyset$.

To show that a explodes before all its ancestors, we think of a as a concatenation of the conservative sequences $a_{I_j} \dots a_{I_{j+1}-1}$, with $j \in [k]$. By the fact that, for each $u \in \mathcal{U}_\infty$, the sequence $(\mathcal{P}_i(u))_{i \in \mathbb{N}}$ is independent and distributed like $(\mathcal{P}_i(\emptyset))_{i \in \mathbb{N}}$, each of these conservative sequences can be seen as corresponding to an a_{I_j} -conservative individual rooted at $a_{I_{j-1}}$, $j \in [k]$. We can thus apply Lemma 4.2 to all these concatenated sequences.

Since, by definition we have $a_{I_{\ell+1}} > a_{I_\ell}$, applying a similar logic to (4.2) we have the following inclusion:

$$\begin{aligned} E_a &:= \{a \text{ explodes before any of its ancestors explodes}\} \\ &= \bigcap_{\ell=1}^k \{a_1 \dots a_{I_{\ell+1}-1} \text{ gives birth to at least } a_{I_\ell} \text{ children before } a_1 \dots a_{I_{\ell-1}} \text{ explodes}\} \\ &= \bigcap_{\ell=1}^k \left\{ \left(\sum_{j=I_\ell}^{I_{\ell+1}-2} \mathcal{P}_{a_{j+1}}(a_{|j}) \right) + \mathcal{P}_{a_{I_\ell}}(a_{|I_{\ell+1}-1}) \leq \sum_{i=a_{I_\ell}+1}^{\infty} X(a_1 \dots a_{I_{\ell-1}i}) \right\} =: \bigcap_{\ell=1}^k E_{a,\ell}. \end{aligned}$$

Now, note that the events $(E_{a,\ell}, \ell \in [k])$ are not independent, since, for a given ℓ , the term $\mathcal{P}_{a_{I_\ell}}(a_{|I_{\ell+1}-1})$ appearing in $E_{a,\ell}$ may be correlated with the term $\sum_{i=a_{I_{\ell+1}}+1}^{\infty} X(a_1 \dots a_{I_{\ell+1}-1}i)$ appearing in $E_{a,\ell+1}$. However, by the third condition of Assumption 2.2, these events are conditionally independent, given the weights of a and all its ancestors, $W_\emptyset, W_{a_1}, W_{a_1 a_2}, \dots, W_a$. Thus,

$$\begin{aligned} &\mathbb{P}(E_a \mid W_\emptyset, W_{a_1}, W_{a_1 a_2}, \dots, W_a) \\ &= \prod_{\ell=1}^k \mathbb{P} \left(\left(\sum_{j=I_\ell}^{I_{\ell+1}-2} \mathcal{P}_{a_{j+1}}(a_{|j}) \right) + \mathcal{P}_{a_{I_\ell}}(a_{|I_{\ell+1}-1}) \leq \sum_{i=a_{I_\ell}+1}^{\infty} X(a_1 \dots a_{I_{\ell-1}i}) \mid W_\emptyset, \dots, W_a \right) \\ &\stackrel{(2.4)}{\leq} \prod_{\ell=1}^k \mathbb{P} \left(\left(\sum_{j=I_\ell}^{I_{\ell+1}-2} \mathcal{P}_{a_{j+1}}(a_{|j}) \right) + \mathcal{P}_{a_{I_\ell}}(a_{|I_{\ell+1}-1}) \leq Y_{a_{I_\ell}}^{(a_1 \dots a_{I_{\ell-1}})} \mid W_\emptyset, W_{a_1}, \dots, W_a \right) \\ &:= \prod_{\ell=1}^k \mathbb{P} \left(\tilde{E}_{a,\ell} \mid W_\emptyset, W_{a_1}, W_{a_1 a_2}, \dots, W_a \right), \end{aligned}$$

where each $Y_{a_{I_\ell}}^{(a_1 \dots a_{I_{\ell-1}})}$ is independent and distributed like $Y_{a_{I_\ell}}$. Now, each of the terms $\tilde{E}_{a,\ell}$ are independent, as they depend on different weights. Hence, so are each of the terms appearing in the above product, so that

$$\mathbb{P}(E_a) \leq \prod_{\ell=1}^k \mathbb{P} \left(\tilde{E}_{a,\ell} \right). \quad (4.6)$$

We now let $d_j := I_{j+1} - I_j - 1$ for $j \in [k-1]$ and $d_k := m - I_k$ denote the number of entries between the running maxima in the sequence a . We can then define, for $(d_j)_{j \in [k]} \in \mathbb{N}_0^k$ (and with the convention that $[0]$ is the empty set),

$$\begin{aligned} &\mathcal{P}_k(a_{I_1}, a_{I_2}, \dots, a_{I_k}, d_1, \dots, d_k) \\ &:= \{a \in \mathcal{U}_\infty : \text{For all } j \in \{1, \dots, k\} \text{ and all } i \in [d_j], a_{I_j+i} \in [a_{I_j}]\} \end{aligned}$$

as the set of all sequences a with running maxima $a_1 = a_{I_1}, \dots, a_{I_k}$ and d_j many entries between the j^{th} and $(j + 1)^{\text{th}}$ maximum. For ease of writing, we omit the arguments of \mathcal{P}_k . We then write the expected value of the number of individuals $a \in \mathcal{U}_\infty$ that explode before all their ancestors such that $a_1 > K'$ as

$$\sum_{\substack{a \in \mathcal{U}_\infty \\ a_1 > K'}} \mathbb{P}(E_a) = \sum_{m=1}^{\infty} \sum_{\substack{a: |a|=m \\ a_1 > K'}} \mathbb{P}(E_a) \leq \sum_{m=1}^{\infty} \sum_{k=1}^m \sum_{a_{I_k} > \dots > a_{I_1} > K'} \sum_{\substack{(d_\ell)_{\ell \in [k]} \in \mathbb{N}_0^k \\ \sum_{\ell=1}^k d_\ell = m-k}} \sum_{a \in \mathcal{P}_k} \prod_{\ell=1}^k \mathbb{P}(\tilde{E}_{a,\ell}).$$

In the first step, we introduce a sum over all sequence lengths m . In the second step, we furthermore sum over the number of running maxima k , the values of the running maxima a_{I_1}, \dots, a_{I_k} , the number of entries d_ℓ between each maxima I_ℓ and $I_{\ell+1}$ (or between I_k and m if $\ell = m$), and all sequences $a \in \mathcal{P}_k$ that admit such running maxima and inter-maxima lengths. Moreover, we use (4.6) to bound $\mathbb{P}(E_a)$ from above, now that we know the number of running maxima in a .

We can now take the sum over $a \in \mathcal{P}_k$ into the product, due to the fact that we can decompose each sequence $a \in \mathcal{P}_k(a_{I_1}, \dots, a_{I_k}, d_1, \dots, d_k)$ into a concatenation of sequences $a^{(1)} \dots a^{(k)}$, with $a^{(\ell)} := a_{I_\ell} \dots a_{I_{\ell+1}-1} \in \mathcal{P}_1(a_{I_\ell}, d_\ell)$ for each $\ell \in [k]$. This yields, for a_{I_1}, \dots, a_{I_k} and d_1, \dots, d_k fixed,

$$\sum_{a \in \mathcal{P}_k} \prod_{\ell=1}^k \mathbb{P}(\tilde{E}_{a,\ell}) = \prod_{\ell=1}^k \sum_{\substack{a^{(\ell)}: |a^{(\ell)}|=d_\ell+1 \\ a^{(\ell)} \text{ } a_{I_\ell}\text{-conservative}}} \mathbb{P}(\tilde{E}_{a,\ell}).$$

We can then directly apply Lemma 4.2 to each of the sums in the product to obtain, for some $\eta < 1$ and with C sufficiently large, the upper bound

$$\prod_{\ell=1}^k \sum_{\substack{a^{(\ell)}: |a^{(\ell)}|=d_\ell \\ a^{(\ell)} \text{ } a_{I_\ell}\text{-conservative}}} \mathbb{P}(\tilde{E}_{a,\ell}) \leq \prod_{\ell=1}^k C \eta^{d_\ell} \mathbb{E}[\mathcal{L}_{c\mu_{a_{I_\ell}}^{-1}}(\mathcal{P}_{a_{I_\ell}}; W)] = C^k \eta^{\sum_{\ell=1}^k d_\ell} \prod_{\ell=1}^k \mathbb{E}[\mathcal{L}_{c\mu_{a_{I_\ell}}^{-1}}(\mathcal{P}_{a_{I_\ell}}; W)].$$

We substitute this in (4.6) to arrive at

$$\begin{aligned} & \sum_{m=1}^{\infty} \sum_{k=1}^m \sum_{a_{I_k} > \dots > a_{I_1} > K'} \sum_{\substack{(d_j)_{j \in [k]} \in \mathbb{N}_0^k \\ \sum_{\ell=1}^k d_\ell = m-k}} C^k \eta^{\sum_{\ell=1}^k d_\ell} \prod_{\ell=1}^k \mathbb{E}[\mathcal{L}_{c\mu_{a_{I_\ell}}^{-1}}(\mathcal{P}_{a_{I_\ell}}; W)] \\ & \leq \sum_{m=1}^{\infty} \sum_{k=1}^m \binom{m-1}{k-1} C^k \eta^{m-k} \sum_{a_{I_1} > K'} \dots \sum_{a_{I_k} > K'} \prod_{\ell=1}^k \mathbb{E}[\mathcal{L}_{c\mu_{a_{I_\ell}}^{-1}}(\mathcal{P}_{a_{I_\ell}}; W)] \\ & = \sum_{m=1}^{\infty} \sum_{k=1}^m \binom{m-1}{k-1} C^k \eta^{m-k} \left(\sum_{a > K'} \mathbb{E}[\mathcal{L}_{c\mu_a^{-1}}(\mathcal{P}_a; W)] \right)^k. \end{aligned}$$

By (4.4) we can bound the innermost sum from above by $\eta/2$ when K' is sufficiently large, so that we obtain the upper bound

$$\sum_{m=1}^{\infty} \sum_{k=1}^m \binom{m-1}{k-1} C^k \eta^{m-k} = \frac{C}{2} \sum_{m=1}^{\infty} (1 + C/2)^{m-1} \eta^m < \infty,$$

where the last step follows when $\eta < (1 + C/2)^{-1}$, which holds by choosing K' sufficiently large, as follows from (4.4) and (4.5). \square

The following proposition requires the following notation:

$$\mathcal{U}_L := \{u \in \mathcal{U}_\infty : u_i \leq L, i \in [|u|]\} \cup \{\emptyset\}.$$

Following [58], we say elements of \mathcal{U}_L are *L-moderate*. We also let \mathcal{U}_L^c denote the set complement of \mathcal{U}_L .

Proposition 4.4. *Let $(\mathcal{T}_t)_{t \geq 0}$ be a (X, W) -CMJ branching process satisfying Condition 4 of Assumption 2.2. Then, almost surely, for all $t \in (0, \infty)$, we have*

$$|\{u \in \mathcal{T}_t : u \text{ is } L\text{-moderate}\}| < \infty.$$

Proof. First, let $(X^{(L)}, W^{(L)})$ denote the distribution of an auxiliary CMJ branching process, where, if the symbol ' \sim ' denotes equality in distribution, we have $W^{(L)} \sim W$ and

$$(X^{(L)}(j), W^{(L)}) \sim \begin{cases} (X(j), W) & \text{if } j \leq L, \\ (\infty, W) & \text{otherwise.} \end{cases}$$

In other words, a $(X^{(L)}, W^{(L)})$ -CMJ process is truncated to ensure that no node produces more than L children. Moreover, if $\mathcal{B}_L(u)$ denotes the distribution of the random variable $\mathcal{B}(u)$ under the distribution of the $(X^{(L)}, W^{(L)})$ -CMJ process, this definition ensures that if $u \in \mathcal{U}_L$ then $\mathcal{B}_L(u) \sim \mathcal{B}(u)$. Now, note that, if $(\mathcal{T}_t^{(L)})_{t \geq 0}$ denotes an $(X^{(L)}, W^{(L)})$ -CMJ process, for each $t \in (0, \infty)$ we have $\mathbb{E}_{(X^{(L)}, W^{(L)})} [\xi(t)] \leq L$, and, by (2.7), also $\mathbb{E}_{(X^{(L)}, W^{(L)})} [\xi(0)] < 1$. Therefore, by [45, Theorem 3.1(b)], $(\mathcal{T}_t^{(L)})_{t \geq 0}$ is *conservative*, i.e. almost surely, for each $t \in (0, \infty)$,

$$|\mathcal{T}_t^{(L)}| < \infty. \tag{4.7}$$

We now construct a coupling of a (X, W) -CMJ branching process $(\overline{\mathcal{T}}_t)_{t \geq 0}$ with a $(X^{(L)}, W^{(L)})$ -CMJ process $(\overline{\mathcal{T}}_t^{(L)})_{t \geq 0}$. Note that, by definition, for each L -moderate $u \in \mathcal{U}_L$, we have $X^{(L)}(u) \sim X(u)$. We then construct $(\overline{\mathcal{T}}_t)_{t \geq 0}$ in the natural way from the random variables defining $(\overline{\mathcal{T}}_t^{(L)})_{t \geq 0}$: for each $u \in \mathcal{U}_L$, we set $(X(u), W_u) = (X^{(L)}(u), W_u^{(L)})$, and for each $u \in \mathcal{U}_L^c$, with $|u| = m \geq 1$, sample $X(u)$ independently, conditionally on the weight $W_{u_{|m-1}}$. One readily verifies that this coupling has the correct marginal distributions. Moreover, on this coupling, almost surely, for each $t \geq 0$, we have

$$|\{u \in \overline{\mathcal{T}}_t : u \text{ is } L\text{-moderate}\}| = |\overline{\mathcal{T}}_t^{(L)}| \stackrel{(4.7)}{<} \infty,$$

as desired □

Recall that, for a (X, W) -Crump-Mode-Jagers branching process $(\mathcal{T}_t)_{t \geq 0}$, we have

$$\tau_\infty = \lim_{k \rightarrow \infty} \tau_k = \inf \{t > 0 : |\mathcal{T}_t| = \infty\}.$$

Recall also that we have $\mathcal{T}_\infty = \bigcup_{k=1}^\infty \mathcal{T}_k = \bigcup_{k=1}^\infty \mathcal{T}_{\tau_k}$.

Lemma 4.5. *Let $(\mathcal{T}_t)_{t \geq 0}$ be a (X, W) -CMJ branching process satisfying Condition 4 of Assumption 2.2. Then,*

$$\mathcal{T}_\infty = \{u \in \mathcal{U}_\infty : \mathcal{B}(u) < \tau_\infty\} \subseteq \mathcal{T}_{\tau_\infty}. \tag{4.8}$$

Proof. The second inclusion is clear, hence we just prove the first equality. If $|\mathcal{T}_\infty| < \infty$, in other words, the process becomes extinct, then $\tau_\infty = \infty$ and the equality is clear. Otherwise, on the event $\{|\mathcal{T}_\infty| = \infty\}$, since by (2.3) we have $\tau_\infty < \infty$, it suffices to show that for any $k \in \mathbb{N}$, $\tau_k < \tau_\infty$. First, we order elements of \mathcal{U}_∞ , $u^{(1)}, u^{(2)} \dots$, according to birth time, breaking ties with the lexicographic ordering. Suppose that $\tau_k = \tau_\infty < \infty$ for some $k \in \mathbb{N}$. Then, since $|\mathcal{T}_\infty| = \infty$, the set

$$\Lambda = \{u \in \mathcal{U}_\infty : \mathcal{B}(u) = \tau_\infty = \tau_k\}$$

is infinite and each element of Λ is a descendant (though not necessarily a child) of one of $u^{(1)}, \dots, u^{(k-1)}$. Now, there are two (not mutually exclusive) cases: either there exists a minimal element $u' \in \Lambda$ such that Λ contains infinitely many elements of the form $u'v \in \Lambda$, with $v \in \mathcal{U}_\infty$, or, one of $u^{(1)}, \dots, u^{(k-1)}$ has infinitely many children in Λ . The latter case occurs with probability 0, because, by Equation (2.6), for each $j \in [k-1]$ and $n \in \mathbb{N}$, we have $\sum_{i=n+1}^\infty X(u^{(j)}i) > 0$ almost surely. Hence, a single individual cannot produce infinitely many children instantaneously. But now, for the prior case, the size of the collection

$$\{u'v : v \in \mathcal{U}_\infty\},$$

is the total progeny of a Bienaymé-Galton-Watson branching process with offspring distribution $\xi^{(u')}(0) \sim \xi(0)$, and by (2.7), this is finite almost surely. We deduce that, on $\{|\mathcal{T}_\infty| = \infty\}$, we have $\tau_k < \tau_\infty$ almost surely. \square

Lemma 4.6. *Let $(\mathcal{F}_t)_{t \geq 0}$ be a (X, W) -CMJ branching process that satisfies Assumption 2.2. Almost surely,*

$$\tau_\infty = \inf_{u \in \mathcal{T}_\infty} \{\mathcal{B}(u) + \mathcal{P}(u)\}.$$

Proof. First, as a shorthand in this proof, we define $\tau_\infty^+ := \inf_{u \in \mathcal{T}_\infty} \{\mathcal{B}(u) + \mathcal{P}(u)\}$, so that we need only show that $\tau_\infty = \tau_\infty^+$ almost surely. Note that, for each $u \in \mathcal{U}_\infty$, we have $\tau_\infty \leq \mathcal{B}(u) + \mathcal{P}(u)$ and hence $\tau_\infty \leq \tau_\infty^+$. Moreover, Assumption 2.2 guarantees that $\mathcal{P}(\emptyset) < \infty$ almost surely, hence, in particular, $\tau_\infty \leq \mathcal{P}(\emptyset) < \infty$. Thus, by Proposition 4.4 and Lemma 4.5, for any $L-1 \in \mathbb{N}$, $\mathcal{T}_\infty \subseteq \mathcal{F}_{\tau_\infty}$ contains only finitely many $(L-1)$ -moderate elements. Since \mathcal{T}_∞ is almost surely infinite (because of Condition 1 of Assumption 2.2), by the pigeonhole principle, there must be infinitely many elements of the form $u^*L \in \mathcal{T}_\infty$, with $u^* \in \mathcal{U}_\infty$. Now, for such u^* and for any $\varepsilon > 0$,

$$\begin{aligned} & \lim_{L \rightarrow \infty} \mathbb{P}(\mathcal{B}(u^*) + \mathcal{P}(u^*) - \mathcal{B}(u^*L) > \varepsilon) \\ &= \lim_{L \rightarrow \infty} \mathbb{P}\left(\sum_{j=L+1}^\infty X(u^*j) > \varepsilon\right) \leq \lim_{L \rightarrow \infty} \mathbb{P}(Y_L > \varepsilon) = 0. \end{aligned}$$

Formally, to find these elements u^*L we must condition on the sigma-algebra generated by the ‘information’ in \mathcal{T}_∞ . Thus, suppose that $\mathscr{W}(\mathcal{T}_\infty) := \bigcup_{k \in \mathbb{N}} \mathscr{W}_{\tau_k}$ denotes the sigma algebra generated by the random variables $\{(\mathcal{B}(u), W_u) : u \in \mathcal{T}_\infty\}$. Clearly, τ_∞ and \mathcal{T}_∞ are $\mathscr{W}(\mathcal{T}_\infty)$ -measurable. Now, if (conditioning on this sigma algebra) there exists $u^* \in \mathcal{T}_\infty$ such that for each $j \in \mathbb{N}$ we have $u^*j \in \mathcal{T}_\infty$, then $\mathcal{B}(u^*j) \leq \tau_\infty$ for each $j \in \mathbb{N}$. Hence, $\mathcal{B}(u^*) + \mathcal{P}(u^*) = \tau_\infty$, and we are done. Otherwise, for any $L \in \mathbb{N}$, we can guarantee the existence of $u^* \in \mathcal{T}_\infty$ and a final $J_L \geq L$ such that $u^*J_L \in \mathcal{T}_\infty$ but $u^*(J_L + 1) \notin \mathcal{T}_\infty$. Then, for any $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P}(\tau_\infty^+ - \tau_\infty > \varepsilon | \mathscr{W}(\mathcal{T}_\infty)) &\leq \mathbb{P}(\mathcal{B}(u^*) + \mathcal{P}(u^*) - \tau_\infty > \varepsilon | \mathscr{W}(\mathcal{T}_\infty)) \\ &\leq \mathbb{P}\left(\sum_{j=J_L+1}^\infty X(u^*j) > \varepsilon \middle| \mathscr{W}(\mathcal{T}_\infty)\right) \leq \mathbb{P}(Y_L \geq \varepsilon). \end{aligned}$$

Taking limits $L \rightarrow \infty$ and then $\varepsilon \rightarrow 0$, we deduce that, $\tau_\infty^+ = \tau_\infty$ almost surely, as required. \square

4.1.1 Proof of Theorem 2.5

Proof of Theorem 2.5. Suppose that K' is taken as in Proposition 4.3. Note that we may view any $w \in \mathcal{U}_\infty$ as a concatenation $w = uv$, where $u \in \mathcal{U}_{K'}$ is K' -moderate, and $v = v_1 \cdots v_k$, where $v_1 > K'$ (here we also allow v to be empty, so that K' -moderate nodes w may also be interpreted as a concatenation). Now, note that (on $\mathcal{B}(u) < \infty$) the birth times $\mathcal{B}(uv) - \mathcal{B}(u) \sim \mathcal{B}(v)$, and thus, by arguments analogous to those appearing in Proposition 4.3, for any $u \in \mathcal{U}_\infty$ (in particular for $u \in \mathcal{U}_{K'}$),

$$\mathbb{E} \left[\left| \left\{ a = a_1 \cdots a_m \in \mathcal{U}_\infty : a_1 > K', ua \text{ explodes before } ua_{|m-1}, ua_{|m-2}, \dots, u \right\} \right| \right] < \infty. \quad (4.9)$$

Now, since $\tau_\infty < \infty$ almost surely, we infer from Proposition 4.4 with $L = K'$, that $|\{u \in \mathcal{U}_{K'} : \mathcal{B}(u) \leq \tau_\infty\}| < \infty$ almost surely. Therefore, by (4.9), the set

$$S := \{u \in \mathcal{U}_\infty : \mathcal{B}(u) \leq \tau_\infty, u \text{ explodes before all of its ancestors}\}.$$

is finite almost surely. By the definition of S and the fact that the infimum of a finite set is attained by (at least) one of the elements, by Lemma 4.6, almost surely there exists $u^* \in S$ such that

$$\mathcal{B}(u^*) + \mathcal{P}(u^*) = \inf_{v \in S} \{\mathcal{B}(v) + \mathcal{P}(v)\} = \inf_{u \in \mathcal{U}_\infty} \{\mathcal{B}(u) + \mathcal{P}(u)\} = \tau_\infty.$$

This implies that u^* has infinite degree in $\mathcal{T}_{\tau_\infty}$. Moreover, since by Condition 1 of Assumption 2.2 we have $\sum_{i=n+1}^\infty X(u^*i) > 0$ almost surely, it follows that for each $i \in \mathbb{N}$ we have $\mathcal{B}(u^*i) < \tau_\infty$ almost surely. Therefore, by the equality in (4.8), u^* has infinite degree in \mathcal{T}_∞ as well. \square

4.2 Sufficient criteria for an infinite path and structural results in the star regime

To prove Theorem 2.8, we first state and prove the following lemma.

Lemma 4.7. *Let $(\mathcal{T}_t)_{t \geq 0}$ be a (X, W) -CMJ branching process. Under Assumption 2.6,*

$$\mathbb{P} \left(\bigcap_{u \in \mathcal{U}_\infty} \bigcap_{j=1}^\infty \bigcup_{i=j}^\infty \{\mathcal{B}(ui) + \mathcal{P}(ui) < \mathcal{B}(u) + \mathcal{P}(u)\} \right) = 1. \quad (4.10)$$

Proof. We first fix $u \in \mathcal{U}_\infty$ and condition on the random variables $\mathcal{B}(u)$ and W_u . We then sample each of the values of $\mathcal{P}(ui)$, then the values of $X_{W_u}(ui)$, for $i \in \mathbb{N}$. Note that for each $u \in \mathcal{U}_\infty$, the random variables $(\mathcal{P}(ui))_{i \in \mathbb{N}}$ are i.i.d. and distributed like $\mathcal{P}(i)$. Thus, by (2.9) and the converse of the Borel-Cantelli lemma, on sampling each $\mathcal{P}(ui)$, for any $w \in S$ with probability 1, we have $\mathcal{P}(ui) < \nu_i^w$ infinitely often. As a result, conditionally on the value of the weight W_u , with probability 1 we have $\mathcal{P}(ui) < \nu_i^{W_u}$ for infinitely many $i \in \mathbb{N}$. Conditioning on this event, let i_1, i_2, \dots , denote indices such that, for each $\ell \in \mathbb{N}$, we have $\mathcal{P}(ui_\ell) < \nu_{i_\ell}^{W_u}$

almost surely. Then, by (2.10), there exists $I_0 \in \mathbb{N}$, such that, for some $\delta = \delta(I_0) > 0$ and for all $i_\ell \geq I_0$,

$$\mathbb{P} \left(\sum_{k=i_\ell+1}^{\infty} X_{W_u}(uk) \geq \nu_{i_\ell}^{W_u} \mid W_u \right) \geq \delta.$$

As a result, for some $\phi(i_\ell) > i_\ell$ sufficiently large,

$$\mathbb{P} \left(\sum_{k=i_\ell+1}^{\phi(i_\ell)} X_{W_u}(uk) \geq \nu_{i_j}^{W_u} \mid W_u \right) \geq \frac{\delta}{2}.$$

We now pass to a sub-sequence $(i_{\ell_n})_{n \in \mathbb{N}}$ of $(i_\ell)_{\ell \in \mathbb{N}}$, such that $i_{\ell_1} > I_0$, and $i_{\ell_{n+1}} = \inf\{i_\ell : \ell \in \mathbb{N}, i_\ell > \phi(i_{\ell_n})\}$. Then, by Equation (2.2), conditionally on W_u , the random variables

$$\left(\sum_{k=i_{\ell_n}+1}^{i_{\ell_{n+1}}} X_{W_u}(uk) \right)_{n \in \mathbb{N}}$$

are independent of each other. Again applying the converse of the Borel-Cantelli lemma, we obtain

$$\mathbb{P} \left(\bigcap_{j=1}^{\infty} \bigcup_{n=j}^{\infty} \left\{ \sum_{k=i_{\ell_n}+1}^{i_{\ell_{n+1}}} X_{W_u}(uk) \geq \nu_{i_{\ell_n}}^{W_u} \right\} \mid \mathcal{B}(u), W_u \right) = 1.$$

But for every index $i^* \in (i_{\ell_n})_{n \in \mathbb{N}}$, we have by the definition of the sequence $(i_\ell)_{\ell \in \mathbb{N}}$,

$$\mathcal{P}(ui^*) < \nu_{i^*}^{W_u} \leq \sum_{i=i^*+1}^{\infty} X_{W_u}(ui).$$

Adding $\mathcal{B}(u) + \sum_{i=1}^{i^*} X_{W_u}(ui)$ (which equals $\mathcal{B}(ui^*)$) to both sides, we deduce that, almost surely, conditionally on $\mathcal{B}(u)$ and W_u ,

$$\mathcal{B}(ui^*) + \mathcal{P}(ui^*) < \mathcal{B}(u) + \mathcal{P}(u).$$

Then, by taking expectations over $\mathcal{B}(u)$ and W_u , we have

$$\mathbb{P} \left(\bigcap_{j=1}^{\infty} \bigcup_{i=j}^{\infty} \{ \mathcal{B}(ui) + \mathcal{P}(ui) < \mathcal{B}(u) + \mathcal{P}(u) \} \right) = 1. \quad (4.11)$$

But now, since \mathcal{U}_∞ is countable, we deduce (4.10). \square

4.2.1 Proof of Theorem 2.8

Proof of Theorem 2.8. On $\{|\mathcal{T}_\infty| = \infty\}$, let us first assume that \mathcal{T}_∞ does not contain a node of infinite degree. It then immediately follows from König's Lemma (Lemma 2.1) that \mathcal{T}_∞ contains an infinite path. We then assume, on $\{|\mathcal{T}_\infty| = \infty\}$, that there exists $u^* \in \mathcal{T}_\infty$ such that u^* has infinite degree, i.e. $\mathcal{B}(u^*) + \mathcal{P}(u^*) = \tau_\infty$. But then, by Equation (4.10) from Lemma 4.7, there must be a child of u^* , u^*j say, such that $\mathcal{B}(u^*j) + \mathcal{P}(u^*j) < \mathcal{B}(u^*) + \mathcal{P}(u^*) = \tau_\infty$, a contradiction. \square

4.2.2 Proof of Theorem 2.10

Proof of Theorem 2.10. For each $u \in \mathcal{U}_\infty$, we set

$$(\mathcal{F}_t^{(u)})_{t \geq 0} := \left\{ uv \in \mathcal{U}_\infty : \sum_{j=0}^{\ell-1} (\mathcal{P}_{v_{j+1}}(uv|_j) \leq t) \right\},$$

where we recall that if $v = v_1 \cdots v_m$ then $uv|_j = uv_1 \cdots v_j$. Note that, if $\mathcal{B}(u) < \infty$, by (1.1) we have $\mathcal{F}_t^{(u)} = \{uv : uv \in \mathcal{F}_{t+\mathcal{B}(u)}\}$. However, the definition we use allows us to define $(\mathcal{F}_t^{(u)})_{t \geq 0}$ even when $\mathcal{B}(u) = \infty$. Note that, by (2.1), $(\mathcal{F}_t^{(u)})_{t \geq 0} \sim (\mathcal{F}_t)_{t \geq 0}$, and for each $u \in \mathcal{U}_\infty$, the $((\mathcal{F}_t^{(ui)})_{t \geq 0})_{i \in \mathbb{N}}$ are i.i.d. Now, upon sampling each of the random variables $(X(ui))_{i \in \mathbb{N}}$, (regardless of whether $\mathcal{B}(u) < \infty$ or not), recalling the notation from (1.2), note that if we have

$$\mathcal{P}_T(ui) < \sum_{k=i+1}^{\infty} X(uk),$$

then, if $\mathcal{B}(u) + \mathcal{P}(u) < \infty$, we have $\mathcal{B}(ui) + \mathcal{P}_T(ui) < \mathcal{B}(u) + \mathcal{P}(u)$. Now, by exploiting an almost identical argument to the proof of Lemma 4.7 (as in Equations (2.9) and (2.10)), combining Equations (2.11) and (2.12) from Condition 1 of Assumption 2.9 allows us to deduce that

$$\mathbb{P} \left(\bigcap_{u \in \mathcal{U}_\infty} \bigcap_{j=1}^{\infty} \bigcup_{i=j}^{\infty} \left\{ \mathcal{P}_T(ui) < \sum_{k=i+1}^{\infty} X(uk) \right\} \right) = 1.$$

But then, this implies that if a node $u^* \in \mathcal{U}_\infty$ has infinite degree in \mathcal{T}_∞ , there exist infinitely many indices i such that $\mathcal{B}(u^*i) + \mathcal{P}_T(u^*i) < \mathcal{B}(u^*) + \mathcal{P}(u^*)$, and hence, by Lemma 4.5, that $(u^*i)T \subseteq \mathcal{T}_\infty$. This proves the first statement.

For the second statement, by Equation (2.13) we have

$$\sum_{i=1}^{\infty} \mathbb{P} \left(\mathcal{P}_T(ui) < \sum_{k=i+1}^{\infty} X(uk) \mid W_u \right) < \infty,$$

almost surely. Thus, by a conditional analogue of the Borel-Cantelli lemma,

$$\mathbb{P} \left(\bigcap_{j=1}^{\infty} \bigcup_{i=j}^{\infty} \left\{ \mathcal{P}_T(ui) < \sum_{k=i+1}^{\infty} X(uk) \right\} \mid W_u \right) = 0,$$

almost surely. Taking expectations over W_u and a union bound over $u \in \mathcal{U}_\infty$, we deduce that

$$\mathbb{P} \left(\bigcup_{u \in \mathcal{U}_\infty} \bigcap_{j=1}^{\infty} \bigcup_{i=j}^{\infty} \left\{ \mathcal{P}_T(ui) < \sum_{k=i+1}^{\infty} X(uk) \right\} \right) = 0.$$

This implies that, if u^* is such that $\mathcal{B}(u^*) + \mathcal{P}(u^*) = \tau_\infty$, almost surely, there exist only finitely many indices i such that T appears as a sub-tree of u^*i in \mathcal{T}_∞ .

The final statement in Item 3 follows from the proof of Theorem 2.5, Equation 4.1.1 in particular, which states that \mathcal{T}_∞ contains a finite number of stars u^* in \mathcal{T}_∞ almost surely under Assumption 2.2. \square

4.3 Uniqueness conditions related to the existence of a star or an infinite path

In this section, it is useful to introduce some extra notation. First, recall that, given $u \in \mathcal{U}_\infty$, the random variable $\mathcal{B}(u) + \mathcal{P}(u)$ determines the time at which u ‘explodes’, in the sense that, if $\sigma = \mathcal{B}(u) + \mathcal{P}(u)$, the individual u has infinite degree in \mathcal{T}_σ . We also define the random variable $\tau_{\text{path}}(u)$ as the amount of time taken after the ‘birth’ of u , for there to be an infinite path containing u in the process $(\mathcal{T}_t)_{t \geq 0}$. To make this more precise, we extend the notation for elements $v \in \mathcal{U}_\infty$ to $v \in \mathbb{N}^{\mathbb{N}}$: for $v = v_1 v_2 \cdots \in \mathbb{N}^\infty$ and $\ell \in \mathbb{N}$, we set $v|_\ell = v_1 v_2 \cdots v_\ell \in \mathcal{U}_\infty$. Then, for $u \in \mathcal{U}_\infty$ we set

$$\tau_{\text{path}}(u) := \inf_{t \geq 0} \left\{ \exists v \in \mathbb{N}^\infty : \sum_{j=0}^{i-1} (\mathcal{P}_{v_{j+1}}(u v|_j)) \leq t \text{ for all } i \in \mathbb{N} \right\}.$$
³

Thus, by this definition, if $\sigma = \mathcal{B}(u) + \tau_{\text{path}}(u)$ then \mathcal{T}_σ contains an infinite path passing through u .

The approach we use in this section is surprisingly simple, and reminiscent of the approach used, for example, in [58] to show a unique node has infinite degree: we show that, for each $u \in \mathcal{U}_\infty$, the random variables $\tau_{\text{path}}(u)$ or $\mathcal{P}(u)$ have distributions that contain no atoms on $[0, \infty)$. We use this to show that for any pair $u, v \in \mathcal{U}_\infty$, (which have ‘independent’ sub-trees), the probability that both have infinite degrees, or lie on infinite paths simultaneously is 0. As \mathcal{U}_∞ is countable, we can readily take a union bound over all these pairs, and deduce the result. Condition 2 of Assumption 2.11 (and (2.1)) already provides this property to $\mathcal{P}(u)$. For $\tau_{\text{path}}(u)$ we use the following result.

Lemma 4.8. *Let $(\mathcal{T}_t)_{t \geq 0}$ be an explosive (X, W) -CMJ process. Under Condition 3 of Assumption 2.11, the distribution of $\tau_{\text{path}}(\emptyset)$ contains no atom on $[0, \infty)$.*

Proof. We argue by contradiction and suppose that $\tau_{\text{path}}(\emptyset)$ contains an atom. Set

$$\tilde{a} := \inf \{ t \in [0, \infty) : \mathbb{P}(\tau_{\text{path}}(\emptyset) = t) > 0 \}.$$

Now, exploiting Condition 3 of Assumption 2.11 and the definition of \tilde{a} , let a be such that $\mathbb{P}(\tau_{\text{path}}(\emptyset) = a) > 0$, and $a - \tilde{a} < \varepsilon$ (with ε as in Condition 3). Let \mathcal{G}_1 denote the sigma algebra generated by $(\mathcal{B}(i))_{i \in \mathbb{N}}$. The ancestral node \emptyset contains an infinite path precisely when one of its children lies on an infinite path. Thus,

$$\begin{aligned} \mathbb{P}(\tau_{\text{path}}(\emptyset) = a \mid \mathcal{G}_1) &= \mathbb{P} \left(\bigcup_{i=1}^{\infty} \{ \tau_{\text{path}}(i) = a - \mathcal{B}(i) \} \mid \mathcal{G}_1 \right) \\ &\leq \sum_{i=1}^{\infty} \mathbb{P} \left(\tau_{\text{path}}(i) = a - \mathcal{B}(i) \mid \mathcal{G}_1 \right). \end{aligned} \tag{4.12}$$

³Note that, since the definition of $\tau_{\text{path}}(u)$ includes the uncountable set \mathbb{N}^∞ , it is not immediately clear that this is measurable. This is the reason that we assume that $(\Omega, \bar{\Sigma}, \mathbb{P})$ and \mathcal{F}_t are complete: complete sigma algebras with respect to probability measures are closed under the Souslin operation [14, Theorem 1.10.5, page 38], which makes a number of uncountable unions measurable. In particular, for each $t \geq 0$, we can write $\{ \tau_{\text{path}}(u) \leq t \} = \bigcup_{v \in \mathbb{N}^\infty} \bigcap_{i \in \mathbb{N}} \{ \sum_{j=0}^{i-1} (\mathcal{P}_{v_{j+1}}(u v|_j)) \leq t \} \in \bar{\Sigma}$, and similarly, we know for each $t \geq 0$ that $\{ \tau_{\text{path}}(\emptyset) \leq t \} \in \mathcal{F}_t$.

Now, as the event $\{\tau_{\text{path}}(i) = a - \mathcal{B}(i)\}$ depends on \mathcal{G}_1 only via the sigma algebra generated by $\mathcal{B}(i)$, we may re-write the summands on the right hand side as

$$\begin{aligned} \mathbb{P}\left(\tau_{\text{path}}(i) = a - \mathcal{B}(i) \mid \mathcal{B}(i)\right) &= \mathbb{P}\left(\tau_{\text{path}}(i) = a - \mathcal{B}(i) \mid \mathcal{B}(i)\right) \mathbf{1}_{\{\mathcal{B}(i) < \varepsilon\}} \\ &\quad + \mathbb{P}\left(\tau_{\text{path}}(i) = a - \mathcal{B}(i) \mid \mathcal{B}(i)\right) \mathbf{1}_{\{\mathcal{B}(i) \geq \varepsilon\}}. \end{aligned}$$

Now, note that (by (2.1)), as $\tau_{\text{path}}(i)$ is identically distributed to $\tau_{\text{path}}(\emptyset)$, the distribution of $\tau_{\text{path}}(i)$ has no atom smaller than $\tilde{a} > a - \varepsilon$. As a result, on the event $\mathcal{B}(i) \geq \varepsilon$,

$$\mathbb{P}\left(\tau_{\text{path}}(i) = a - \mathcal{B}(i) \mid \mathcal{B}(i)\right) = 0, \quad \text{almost surely.}$$

Combining this with (4.12), this leads to the inequality

$$\mathbb{P}\left(\tau_{\text{path}}(\emptyset) = a \mid \mathcal{G}_1\right) \leq \sum_{i=1}^{\infty} \mathbb{P}\left(\tau_{\text{path}}(i) = a - \mathcal{B}(i) \mid \mathcal{B}(i)\right) \mathbf{1}_{\{\mathcal{B}(i) < \varepsilon\}},$$

and since $\mathbb{P}\left(\tau_{\text{path}}(\emptyset) = a\right) > 0$, it must be the case that for some i ,

$$0 < \mathbb{P}\left(\tau_{\text{path}}(i) = a - \mathcal{B}(i), \mathcal{B}(i) < \varepsilon\right) = \mathbb{E}\left[\mathbb{P}\left(\tau_{\text{path}}(i) = a - \mathcal{B}(i) \mid \mathcal{B}(i)\right) \mathbf{1}_{\{\mathcal{B}(i) < \varepsilon\}}\right].$$

But now, when we integrate over the possible values of $\mathcal{B}(i)$ on the right-hand side, this implies that

$$\left\{\mathcal{B}(i) \in [0, \varepsilon) : \mathbb{P}\left(\tau_{\text{path}}(i) = a - \mathcal{B}(i) \mid \mathcal{B}(i)\right) > 0\right\}$$

is a set of positive measure with respect to \mathbb{P} . But then this implies that, with respect to the distribution induced by $\mathcal{B}(i)$, the set

$$\{x \in [0, \varepsilon) : \mathbb{P}\left(\tau_{\text{path}}(i) = a - x\right) > 0\}, \quad (4.13)$$

has positive measure. But now, as the set of atoms of the distribution of $\tau_{\text{path}}(i)$ (as with any finite measure) must always be countable, (4.13) must be countable. As the distribution of $\mathcal{B}(i)$ contains no atoms on $[0, \varepsilon)$ by Condition 3 in Assumption 2.11, countable subsets of $[0, \varepsilon)$ are null sets with respect to the distribution induced by $\mathcal{B}(i)$; this implies that (4.13) must be a null set, a contradiction. \square

4.3.1 Proof of Theorem 2.12

Proof of Theorem 2.12. The first and second statements of the theorem are clearly satisfied on $\{|\mathcal{T}_\infty| < \infty\}$. Now note that, for \mathcal{T}_∞ to have two nodes of infinite degree, there must be two elements $u, v \in \mathcal{U}_\infty$ such that

$$\mathcal{B}(u) + \mathcal{P}(u) = \mathcal{B}(v) + \mathcal{P}(v) = \tau_\infty.$$

By Equation (2.3), $\tau_\infty < \infty$ almost surely on $\{|\mathcal{T}_\infty| = \infty\}$, hence the left-hand side must also be finite. But now, note that, as long as v is not an ancestor of u , (which is the case if $v >_L u$ with respect to the lexicographical ordering), then $\mathcal{P}(v)$ is independent of $\mathcal{B}(u), \mathcal{B}(v)$, and

$\mathcal{P}(u)$. Moreover, by Condition 2 of Assumption 2.11, $\mathcal{P}(v)$ has no atom on $[0, \infty)$. Therefore, for $v >_L u$ we have,

$$\mathbb{P} \left(\mathcal{P}(v) = \mathcal{P}(u) + \mathcal{B}(v) - \mathcal{B}(u) \mid \mathcal{B}(u), \mathcal{B}(v), \mathcal{P}(v) \right) \mathbf{1}_{\{\mathcal{B}(u), \mathcal{B}(v), \mathcal{P}(u) < \infty\}} = 0 \quad \text{almost surely,}$$

and hence, for any $u, v \in \mathcal{U}_\infty$ with $u \neq v$,

$$\mathbb{P}(\mathcal{B}(u) + \mathcal{P}(u) = \mathcal{B}(v) + \mathcal{P}(v), \mathcal{B}(u) + \mathcal{P}(u) < \infty) = 0.$$

But then, since \mathcal{U}_∞ is countable, taking a union bound, we deduce

$$\begin{aligned} & \mathbb{P}(\mathcal{T}_\infty \text{ has two nodes of infinite degree}) \\ & \leq \mathbb{P} \left(\bigcup_{u, v \in \mathcal{U}_\infty, u \neq v} \{\mathcal{B}(u) + \mathcal{P}(u) = \mathcal{B}(v) + \mathcal{P}(v), \mathcal{B}(u) + \mathcal{P}(u) < \infty\} \right) \quad (4.14) \\ & \leq \sum_{u, v \in \mathcal{U}_\infty, u \neq v} \mathbb{P}(\mathcal{B}(u) + \mathcal{P}(u) = \mathcal{B}(v) + \mathcal{P}(v), \mathcal{B}(u) + \mathcal{P}(u) < \infty) = 0. \end{aligned}$$

This proves the first statement of the theorem. In a similar manner, using Condition 3 of Assumption 2.11 and Lemma 4.8, we have

$$\begin{aligned} & \mathbb{P}(\mathcal{T}_\infty \text{ has two infinite paths}) \\ & \leq \mathbb{P} \left(\bigcup_{u, v \in \mathcal{U}_\infty, u \neq v} \{\mathcal{B}(u) + \tau_{\text{path}}(u) = \mathcal{B}(v) + \tau_{\text{path}}(v), \mathcal{B}(u) + \tau_{\text{path}}(u) < \infty\} \right) \quad (4.15) \\ & \leq \sum_{u, v \in \mathcal{U}_\infty, u \neq v} \mathbb{P}(\mathcal{B}(u) + \tau_{\text{path}}(u) = \mathcal{B}(v) + \tau_{\text{path}}(v), \mathcal{B}(u) + \tau_{\text{path}}(u) < \infty) = 0, \end{aligned}$$

proving the second statement. Finally, for the third statement we need only prove that a node of infinite degree and an infinite path cannot co-exist. Note that there may be u such that $\mathcal{P}(u)$ and $\tau_{\text{path}}(\emptyset)$ are correlated, for example, if \emptyset is a parent of u . But, noting $\mathbb{N}^k = \{u \in \mathcal{U}_\infty : |u| = k\}$,

$$\begin{aligned} & \{\tau_{\text{path}}(\emptyset) = \mathcal{B}(u) + \mathcal{P}(u), \tau_{\text{path}}(\emptyset) < \infty\} \\ & \subseteq \bigcup_{v \in \mathbb{N}^{|u|+1}} \{\mathcal{B}(v) + \tau_{\text{path}}(v) = \mathcal{B}(u) + \mathcal{P}(u), \mathcal{B}(v) + \tau_{\text{path}}(v) < \infty\}. \end{aligned}$$

Now, for $v \in \mathbb{N}^{|u|+1}$, $\tau_{\text{path}}(v)$ is independent of u , and therefore, for each $v \in \mathbb{N}^{|u|+1}$ we have

$$\begin{aligned} & \mathbb{P}(\mathcal{B}(v) + \tau_{\text{path}}(v) = \mathcal{B}(u) + \mathcal{P}(u) \mid \mathcal{B}(v), \mathcal{B}(u), \mathcal{P}(u)) \mathbf{1}_{\{\mathcal{B}(v), \mathcal{B}(u), \mathcal{P}(u) < \infty\}} \\ & = \mathbb{P}(\tau_{\text{path}}(v) = \mathcal{B}(u) + \mathcal{P}(u) - \mathcal{B}(v) \mid \mathcal{B}(v), \mathcal{B}(u), \mathcal{P}(u)) \mathbf{1}_{\{\mathcal{B}(v), \mathcal{B}(u), \mathcal{P}(u) < \infty\}} = 0, \end{aligned}$$

almost surely. Therefore, again using a union bound, we have

$$\begin{aligned} & \mathbb{P}(\mathcal{T}_\infty \text{ has an infinite path and a node of infinite degree, } |\mathcal{T}_\infty| = \infty) \quad (4.16) \\ & \leq \mathbb{P} \left(\bigcup_{u \in \mathcal{U}_\infty} \{\mathcal{B}(u) + \mathcal{P}(u) = \tau_{\text{path}}(\emptyset), \mathcal{B}(u) + \tau_{\text{path}}(u) < \infty\} \right) \\ & \leq \mathbb{P} \left(\bigcup_{u \in \mathcal{U}_\infty} \bigcup_{v \in \mathbb{N}^{|u|+1}} \{\mathcal{B}(u) + \mathcal{P}(u) = \mathcal{B}(v) + \tau_{\text{path}}(v), \mathcal{B}(u) + \tau_{\text{path}}(u) < \infty\} \right) \\ & \leq \sum_{u \in \mathcal{U}_\infty} \sum_{v \in \mathbb{N}^{|u|+1}} \mathbb{P}(\mathcal{B}(u) + \mathcal{P}(u) = \mathcal{B}(v) + \tau_{\text{path}}(v), \mathcal{B}(u) + \tau_{\text{path}}(u) < \infty) = 0. \end{aligned}$$

We deduce the final statement by using a union bound, and combining Equations (4.14), (4.15), and (4.16) with König's lemma (Lemma 2.1). \square

5 Proofs for applications of main results

In this section we use the general results of Section 2 to prove the main theorems in Section 3.1. In particular, we show that the conditions stated in Theorem 3.4 are sufficient and that the condition in Theorem 3.7 is necessary and sufficient to apply the results of Section 2 to exponentially distributed inter-birth times.

5.1 Preliminary results and tools

We first collect some useful lemmas that we use throughout the proofs in this section.

Lemma 5.1. *Let $(X(i))_{i \in \mathbb{N}}$ be independent random variables and fix $\varepsilon > 0$ and $(b_n)_{n \in \mathbb{N}} \in (0, \infty)^{\mathbb{N}}$. Then,*

$$\mathbb{P} \left(\sum_{i=1}^{\infty} X_i < b_n \right) \geq \prod_{i=1}^{\infty} \mathcal{L}_{(1+\varepsilon) \log(n) b_n^{-1}}(X_i) - \frac{1}{n^{1+\varepsilon}}.$$

Proof. Let $Y \sim \text{Exp}(b_n^{-1}(1 + \varepsilon) \log n)$ be given, independent of each of the X_i . Then,

$$\mathbb{P} \left(\sum_{i=1}^{\infty} X_i < b_n \right) \geq \mathbb{P} \left(\sum_{i=1}^{\infty} X_i < Y, Y \leq b_n \right) \geq \mathbb{P} \left(\sum_{i=1}^{\infty} X_i < Y \right) - \mathbb{P}(Y \geq b_n).$$

Evaluating both probabilities by using that Y is an exponential random variable, we obtain

$$\mathbb{E} \left[e^{-(1+\varepsilon) \log(n) b_n^{-1} \sum_{i=1}^{\infty} X_i} \right] - e^{-(1+\varepsilon) \log n} = \prod_{i=1}^{\infty} \mathcal{L}_{(1+\varepsilon) \log(n) b_n^{-1}}(X_i) - \frac{1}{n^{1+\varepsilon}},$$

as desired. □

We believe that the following lemmas are more well-known, but we provide a proof of the first for completeness. Note, as is clear from the proof, that (5.1) is a special case of the more general Equation (4.1) from Lemma 4.1.

Lemma 5.2. *Let Z_0, \dots, Z_k be independent exponential random variables with parameters r_0, \dots, r_{k-1} , respectively, and fix $\lambda > 0$. Then,*

$$\mathbb{P} \left(\sum_{j=0}^k Z_j \leq \lambda \right) \geq \prod_{j=0}^k \frac{r_j}{r_j + (k+1)\lambda^{-1}}.$$

Moreover, let $(X_{w^*}(j))_{j \in \mathbb{N}}$ be independent, exponential random variables with rates $(r_j)_{j \in \mathbb{N}}$, satisfying $\mu_i := \sum_{j=i+1}^{\infty} \mathbb{E}[X_{w^*}(j)] < \infty$ for $i \in \mathbb{N}$. For any $c < 1$ and for any $i \in \mathbb{N}$,

$$\mathbb{P} \left(\sum_{j=0}^k Z_j \leq \sum_{j=i+1}^{\infty} X_{w^*}(j) \right) \leq \frac{1}{1-c} \prod_{j=0}^k \frac{r_j}{r_j + c\mu_i^{-1}}. \quad (5.1)$$

Proof. For the first inequality, we may, without loss of generality take $\lambda = 1$, since for an exponential random variable $X \sim \text{Exp}(r)$, X/λ is again exponentially distributed with parameter λr . Then,

$$\begin{aligned} \mathbb{P}\left(\sum_{j=0}^k Z_j \leq 1\right) &\geq \mathbb{P}(Z_j \leq 1/(k+1) \text{ for all } j \in \{0, \dots, k\}) \\ &= \prod_{j=0}^k (1 - e^{-r_j/(k+1)}) \geq \prod_{j=0}^k \frac{r_j}{r_j + (k+1)}, \end{aligned}$$

where the right-hand side uses the inequality $1 - e^{-x} \geq \frac{x}{1+x}$, for all $x \geq 0$. Meanwhile, the upper bound in (5.1) uses a standard Chernoff bound. First, note that for $c < 1$,

$$c\mu_i^{-1} = c \left(\sum_{j=i+1}^{\infty} \frac{1}{r_j} \right)^{-1} < r_j, \quad \text{for all } j \geq i+1. \quad (5.2)$$

Furthermore, for each $i \in \mathbb{N}$,

$$\prod_{j=i+1}^{\infty} \frac{r_j - c\mu_i^{-1}}{r_j} = \prod_{j=i+1}^{\infty} \left(1 - \frac{c\mu_i^{-1}}{r_j}\right) > 1 - c\mu_i^{-1} \sum_{j=i+1}^{\infty} \frac{1}{r_j} = 1 - c, \quad (5.3)$$

where the last step uses the inequality $\prod_i (1 - p_i) \geq 1 - \sum_i p_i$, for $p_i \in [0, 1]$. By the moment generating function of exponential random variables, we have

$$\mathbb{E}\left[e^{c\mu_i^{-1} \sum_{j=i+1}^{\infty} X_{w^*(j)}}\right] = \prod_{j=i+1}^{\infty} \frac{r_j}{r_j - c\mu_i^{-1}} \stackrel{(5.3)}{\leq} \frac{1}{1-c}, \quad (5.4)$$

where the exponential moments exist by (5.2). Finally, this yields

$$\mathbb{P}\left(\sum_{j=0}^k Z_j \leq \sum_{j=i+1}^{\infty} X_{w^*(j)}\right) \leq \mathbb{P}\left(e^{c\mu_i^{-1}(\sum_{j=i+1}^{\infty} X_{w^*(j)} - \sum_{j=0}^k Z_j)} \geq 1\right) \leq \frac{1}{1-c} \prod_{j=0}^k \frac{r_j}{r_j + c\mu_i^{-1}},$$

as desired. \square

Lemma 5.3 (Paley-Zygmund Inequality). *Let Z be a non-negative random variable with finite variance, and let $d \in (0, 1)$. Then,*

$$\mathbb{P}(Z \geq d\mathbb{E}[Z]) \geq (1-d)^2 \frac{\mathbb{E}[Z]^2}{\text{Var}(Z) + \mathbb{E}[Z]^2}.$$

5.2 Structure theorems related to explosive recursive trees

5.2.1 Proof of Theorem 3.4

Proof of Theorem 3.4. Recall that for fixed u, W_u , the random variables $(X_{W_u}(wi))_{i \in \mathbb{N}}$ are independent, with each $X_{W_u}(wi) \sim \text{Exp}(f(i-1, W_u))$, conditionally on W_u . Since the exponential distribution is a smooth distribution, one readily verifies that the conditions of Assumption 2.11 are met; hence the associated tree \mathcal{T}_{∞} contains either a unique node of infinite

degree, or a unique infinite path. Now, to prove Item 1, note that Conditions 3 and 4 of Assumption 2.2 are immediately satisfied. Moreover, since f satisfies (w^*) , Equation (2.4) and thus Condition 1 is satisfied by setting

$$Y_n := \sum_{i=n+1}^{\infty} \tilde{X}_{w^*}(i), \quad (5.5)$$

where each $\tilde{X}_{w^*}(i) \sim \text{Exp}(f(i-1, w^*))$. By choosing $c < 1$ according to Equation (3.2) and by following the calculations in Equations (5.3) and (5.4) (with $r_j := f(j-1, w^*)$), we see that Condition 2 is satisfied. Finally, note that, uniformly in w , we have

$$\prod_{i=n}^{\infty} \frac{f(i, w)}{f(i, w) + c\mu_n^{-1}} > 0.$$

Indeed, taking logarithms and using the inequality

$$-\log(1-x) \leq x + \frac{x^2}{2(1-x)}, \quad 0 < x < 1,$$

we obtain

$$\begin{aligned} -\log\left(\prod_{i=n}^{\infty} \frac{f(i, w)}{f(i, w) + c\mu_n^{-1}}\right) &= -\sum_{i=n}^{\infty} \log\left(1 - \frac{c\mu_n^{-1}}{f(i, w) + c\mu_n^{-1}}\right) \\ &\leq \sum_{i=n}^{\infty} \left(\frac{c\mu_n^{-1}}{f(i, w^*) + c\mu_n^{-1}} + \frac{c^2\mu_n^{-2}}{2f(i, w^*)(f(i, w^*) + c\mu_n^{-1})}\right) < \frac{3}{2}, \end{aligned}$$

which implies that

$$\prod_{i=n}^{\infty} \frac{f(i, w)}{f(i, w) + c\mu_n^{-1}} > e^{-3/2}.$$

Thus, by Equation (3.2),

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{E}[\mathcal{L}_{c\mu_n^{-1}}(\mathcal{P}_n(\emptyset); W)] &= \sum_{n=1}^{\infty} \mathbb{E}\left[\prod_{i=0}^{n-1} \frac{f(i, W)}{f(i, W) + c\mu_n^{-1}}\right] \\ &< e^{3/2} \sum_{n=1}^{\infty} \mathbb{E}\left[\prod_{i=0}^{\infty} \frac{f(i, W)}{f(i, W) + c\mu_n^{-1}}\right] < \infty, \end{aligned}$$

so that Condition 5 is satisfied, where the final step uses the assumption in Item 1. We can now apply Theorems 2.5 and 2.12 to obtain the desired result.

For Item 2, if there exists $c > 1$ satisfying Equation (3.3), we write $c = d^{-1}(1+\varepsilon)$, for some $0 < d < 1$ and $\varepsilon > 0$. Otherwise, assume (2.9) is satisfied with $\nu_n^w = d\mu_n^w = d\mathbb{E}[\sum_{i=n+1}^{\infty} X_w(i)]$ for some $d \in (0, 1)$. First note that, since the $X_w(i)$ are mutually independent exponential random variables,

$$\text{Var}\left(\sum_{i=n+1}^{\infty} X_w(i)\right) = \sum_{i=n}^{\infty} \frac{1}{f(i, w)^2} \leq (\mu_n^w)^2. \quad (5.6)$$

By the Paley-Zygmund inequality, it thus follows that

$$\mathbb{P} \left(\sum_{i=n+1}^{\infty} X_w(i) \geq d\mu_n^w \right) \geq (1-d)^2 \frac{(\mu_n^w)^2}{(\mu_n^w)^2 + \sum_{i=n}^{\infty} f(i, w)^{-2}} \stackrel{(5.6)}{\geq} \frac{(1-d)^2}{2}, \quad (5.7)$$

which implies (2.10) in either case. We now need only show that Equation (3.3) implies that (2.9) is satisfied with this choice of $(\nu_n^w)_{n \in \mathbb{N}}$. We apply Lemma 5.1 with $b_n := d\mu_n^w$ and $\varepsilon > 0$ so that $c = d^{-1}(1 + \varepsilon)$, and by conditioning on the vertex-weight W , to obtain

$$\mathbb{P}(\mathcal{P} < d\mu_n^w) \geq \mathbb{E} \left[\prod_{i=0}^{\infty} \frac{f(i, W)}{f(i, W) + c(\mu_n^w)^{-1} \log n} \right] - \frac{1}{n^{1+\varepsilon}}.$$

We thus obtain that (2.9) is satisfied when Equation (3.3) holds. \square

5.2.2 Proof of Theorem 3.7

Proof of Theorem 3.7. In the proof we seek to apply Theorem 2.10. We first apply Item 1 of Theorem 2.10 to show that, if (3.4) is satisfied, T appears as a sub-tree of \mathcal{T}_∞ infinitely often. That is, we show that Condition 1 of Assumption 2.9 is satisfied when assuming Equation (3.4) holds. As we assume that Equation 2.2 holds as well, this implies Item 1 of Theorem 2.10.

In a similar manner to the proof of Theorem 3.4 above, we choose $\nu_n^w = d\mu_n^w$, for $d < 1$. This way, Equation (2.12) follows from Equation (5.7), using the Paley-Zygmund inequality. To show (2.11), we first condition on an ordering $O \in \mathcal{O}$ in which the vertices of T appear. In particular, given such an ordering, by extending the definitions of the stopping times associated with the (X, W) -CMJ process, we can define the stopping times

$$\tau_{O|_j} := \inf \{t \geq 0 : O|_j \subseteq \mathcal{F}_t\}.$$

Now, with $\mathcal{O}(T)$ the set of possible orderings of the vertices of T , we wish to show that

$$\sum_{n=1}^{\infty} \mathbb{P}(T \subseteq \mathcal{F}_{d\mu_n^w}) = \sum_{n=1}^{\infty} \sum_{O \in \mathcal{O}(T)} \mathbb{P}(0 = \tau_{O|_0} \leq \tau_{O|_1} \cdots \leq \tau_{O|_k} \leq d\mu_n^w) = \infty. \quad (5.8)$$

We make the following observations:

- (I) The probability of seeing an ordering $O|_j, j = 0, \dots, k$, depends on the minima of the exponential random variables associated with the next vertex in the tree to appear. Suppose that the vertices of T are $\{v_0, \dots, v_k\}$, labelled by order of appearance, and that $v_j^+ \in \{v_0, \dots, v_{j-1}\}$ is the *parent* of v_j for each $j \in [k]$. Then, conditionally on the weights of the nodes in T , W_{v_0}, \dots, W_{v_k} , the ordering $(O|_j)_{j=0, \dots, k}$ occurs with probability

$$\mathbb{P} \left(O \mid W_{v_0}, \dots, W_{v_k} \right) = \prod_{j=0}^k \frac{f(\deg^+(v_j^+), W_{v_j^+})}{\sum_{i=0}^j f(\deg^+(v_i, O|_j), W_{v_i}) \mathbf{1}_{\deg^+(v_i, O|_j) < \deg^+(v_i, T)}}.$$

The j th product in the above equation denotes the probability that the next exponential clock associated with the process, amongst those yet to ring, corresponds to v_j^+ , and hence respects the ordering O .

(II) We may write

$$\prod_{j=0}^k f(\deg^+(v_j^+, O_{|j}), W_{v_j^+}) = \prod_{j=0}^k \prod_{\ell=0}^{\deg^+(v_j, T)-1} f(\ell, W_{v_j}).$$

(III) By properties of the exponential distribution, the distribution of the minimum of a collection of independent exponential random variables is independent of the index that attains that minimum. Thus, conditionally on seeing the ordering O , the waiting times $(\tau_{O_{|j+1}} - \tau_{O_{|j}})_{j=0, \dots, k}$ are independent, with

$$\tau_{O_{|j+1}} - \tau_{O_{|j+1}} \sim \text{Exp} \left(\sum_{i=0}^j f(\deg^+(i, O_{|j}), W_i) \mathbf{1}_{\deg^+(i, O_{|j}) < \deg^+(i, T)} \right), \quad j = 0, \dots, k.$$

Thus, by Lemma 5.2, we have

$$\begin{aligned} & \mathbb{P} \left(0 = \tau_{O_{|0}} \leq \tau_{O_{|1}} \leq \dots \leq \tau_{O_{|k}} \leq d\mu_n^w \mid W_{v_0}, \dots, W_{v_k}, O \right) \\ & \geq \left(\frac{d}{k+1} \right)^{k+1} \prod_{j=0}^k \frac{\sum_{i=0}^j f(\deg^+(i, O_{|j}), W_i) \mathbf{1}_{\deg^+(i, O_{|j}) < \deg^+(i, T)}}{\sum_{i=0}^j f(\deg^+(i, O_{|j}), W_i) \mathbf{1}_{\deg^+(i, O_{|j}) < \deg^+(i, T)} + (\mu_n^w)^{-1}} \\ & \geq \left(\frac{dc_1(w)}{k+1} \right)^{k+1} \prod_{j=0}^k \frac{\sum_{i=0}^j f(\deg^+(i, O_{|j}), W_i) \mathbf{1}_{\deg^+(i, O_{|j}) < \deg^+(i, T)}}{\sum_{i=0}^j f(\deg^+(i, O_{|j}), W_i) \mathbf{1}_{\deg^+(i, O_{|j}) < \deg^+(i, T)} + \mu_n^{-1}}, \end{aligned}$$

where the last line follows from our assumption that $\mu_n^w \geq c_1(w)\mu_n$, with $c_1(w) < 1$.

Combining Items (I)-(III) and taking expectations over the weights W_{v_0}, \dots, W_{v_k} , we deduce that Equation (5.8) holds if Equation (3.4) is satisfied.

For the converse statement, if Equation (3.4) is not satisfied and the sum instead converges, we now apply Item 2 of Theorem 2.10 in a similar manner. In particular, we now wish to show that, for any $w \in S$,

$$\sum_{i=1}^{\infty} \mathbb{P} \left(T \subseteq \mathcal{T}_{\sum_{j=i+1}^{\infty} \tilde{X}_w(j)} \right) = \sum_{i=1}^{\infty} \sum_{O \in \mathcal{O}(T)} \mathbb{P} \left(0 = \tau_{O_{|0}} \leq \tau_{O_{|1}} \leq \dots \leq \tau_{O_{|k}} \leq \sum_{j=i+1}^{\infty} \tilde{X}_w(j) \right) < \infty. \quad (5.9)$$

But now, for each j we have $\tilde{X}_w(j) \leq_S \tilde{X}_{w^*}(j)$, where $\tilde{X}_{w^*}(j) \sim X_{w^*}(j)$, since we assume that Equation (w^*) holds. Thus, using the same observations as above but instead looking for an upper bound in Item (III), we apply Equation (5.1) in Lemma 5.2 with $c < 1$ fixed, to see

that

$$\begin{aligned}
& \mathbb{P} \left(0 = \tau_{O_{l_0}} \leq \tau_{O_{l_1}} \cdots \leq \tau_{O_{l_k}} \leq \sum_{j=i+1}^{\infty} \tilde{X}_w(j) \mid W_{v_0}, \dots, W_{v_k}, O \right) \\
& \leq \mathbb{P} \left(0 = \tau_{O_{l_0}} \leq \tau_{O_{l_1}} \cdots \leq \tau_{O_{l_k}} \leq \sum_{j=i+1}^{\infty} \tilde{X}_{w^*}(j) \mid W_{v_0}, \dots, W_{v_k}, O \right) \\
& \stackrel{(5.1)}{\leq} \frac{1}{1-c} \prod_{j=0}^k \frac{\sum_{i=0}^j f(\deg^+(i, O_{l_j}), W_i) \mathbf{1}_{\deg^+(i, O_{l_j}) < \deg^+(i, T)}}{\sum_{i=0}^j f(\deg^+(i, O_{l_j}), W_i) \mathbf{1}_{\deg^+(i, O_{l_j}) < \deg^+(i, T)} + c\mu_i^{-1}} \\
& \leq \frac{1}{(1-c)c^{k+1}} \prod_{j=0}^k \frac{\sum_{i=0}^j f(\deg^+(i, O_{l_j}), W_i) \mathbf{1}_{\deg^+(i, O_{l_j}) < \deg^+(i, T)}}{\sum_{i=0}^j f(\deg^+(i, O_{l_j}), W_i) \mathbf{1}_{\deg^+(i, O_{l_j}) < \deg^+(i, T)} + \mu_i^{-1}}.
\end{aligned} \tag{5.10}$$

Again, combining Equation (5.10) with Items (I) and (II) and taking expectations over the weights W_{v_0}, \dots, W_{v_k} , we deduce that Equation (5.9) holds if Equation (3.4) is not satisfied. Hence, Condition 2 of Assumption 2.9 holds, so that Item 2 in Theorem 2.10 yields the desired result. \square

5.3 Possibility for both a star and infinite path to occur with positive probability

The goal of this section is to produce the counter-example in Theorem 2.15. We first show, in a similar spirit to [45, Claim 2.3], that infinite paths, when they appear in finite time, can appear arbitrarily fast.

Lemma 5.4. *Let $(\mathcal{F}_t)_{t \geq 0}$ be an explosive (X, W) -CMJ process, and suppose that $\mathcal{B}(1) > 0$ almost surely. It is either the case that $\mathbb{P}(\tau_{\text{path}}(\emptyset) < \infty) = 0$, or, for any $\varepsilon > 0$,*

$$\mathbb{P}(\tau_{\text{path}}(\emptyset) < \varepsilon) > 0.$$

Proof. Suppose that $\mathbb{P}(\tau_{\text{path}}(\emptyset) < \infty) > 0$. Then, we define

$$\varepsilon^* := \inf \{ \varepsilon \geq 0 : \mathbb{P}(\tau_{\text{path}}(\emptyset) < \varepsilon) > 0 \}.$$

By assumption, we have $\varepsilon^* < \infty$, and moreover, $\mathbb{P}(\tau_{\text{path}}(\emptyset) < \varepsilon^*) = 0$ (by monotone convergence). Suppose that $\varepsilon^* > 0$. As an infinite path from \emptyset must pass through the first generation, we have

$$\mathbb{P}(\tau_{\text{path}}(\emptyset) < \varepsilon^*) = \mathbb{P} \left(\bigcup_{i \in \mathbb{N}} \{ \mathcal{B}(i) + \tau_{\text{path}}(i) < \varepsilon^* \} \right). \tag{5.11}$$

Since for each i we have $\mathcal{B}(i) \geq \mathcal{B}(1)$, we must have $\mathbb{P}(\mathcal{B}_1 + \tau_{\text{path}} < \varepsilon^*) > 0$; where $\mathcal{B}_1 \sim \mathcal{B}(1) \sim X(1)$ and $\tau_{\text{path}} \sim \tau_{\text{path}}(\emptyset)$, and $\mathcal{B}_1, \tau_{\text{path}}$ are independent. Indeed, otherwise a union bound would show that the right-hand side of (5.11) equals zero. But then, it must be the case that

$$\left\{ \mathcal{B}_1 \in (0, \varepsilon^*) : \mathbb{P} \left(\tau_{\text{path}} < \varepsilon^* - \mathcal{B}_1 \mid \mathcal{B}_1 \right) > 0 \right\}$$

must be a set of positive \mathbb{P} -measure. Again, as in (the proof of) Lemma 4.8, this implies that with respect to the distribution induced by \mathcal{B}_1 , the set

$$\{x \in (0, \varepsilon^*) : \mathbb{P}(\tau_{\text{path}} < \varepsilon^* - x) > 0\}$$

has positive measure. In particular, there exists an $x \in (0, \varepsilon^*)$ such that $\mathbb{P}(\tau_{\text{path}} < \varepsilon^* - x) > 0$, a contradiction. We deduce that $\varepsilon^* = 0$. \square

5.3.1 Proof of Theorem 2.15

Proof of Theorem 2.15. We define an explosive (X, W) -CMJ process with weights $W_u = (R_u, I_u) \in (0, \infty) \times \{0, 1\}$ for $u \in \mathcal{U}_\infty$. Here R_u and I_u are independent and distributed such that $\mathbb{P}(I_u = 1) = \mathbb{P}(I_u = 0) = 1/2$, and R_u has a smooth distribution such that, for some constant $\alpha > 1$, for all $x \in [1, \infty)$,

$$\mathbb{P}(R_u > x) \geq x^{-(\alpha-1)}. \quad (5.12)$$

We also fix $p > 1$ such that $(\alpha - 1)(p - 1) < 1$. We can think of the variable I_u as an indication of the *type* of a node u , effecting the distribution of $(X_{W_u}(uj))_{j \in \mathbb{N}}$. We then define the (X, W) -CMJ process $(\mathcal{T}_t)_{t \geq 0}$ so that, conditionally on W_u , the random variables $(X_{W_u}(ui))_{i \in \mathbb{N}}$ are mutually independent, and for each $u \in \mathcal{U}_\infty, i \in \mathbb{N}$,

$$X_{W_u}(ui) \sim \begin{cases} \text{Exp}(R_u i^p), & \text{if } W_u = (R_u, 0); \\ \text{Exp}(1), & \text{if } W_u = (R_u, 1), i = 1; \\ s_i, & \text{if } W_u = (R_u, 1), i > 1; \end{cases}$$

where the $(s_i)_{i \geq 2}$ are constants with each $s_i > 0$ defined to satisfy the following: if for each $k \geq 1$ we set

$$s_k := \sum_{j=k+1}^{\infty} s_j < \infty, \quad \text{then we have} \quad \mathbb{E}[e^{-(R_k \vee 1)s_k}] > 1 - 2^{-k}; \quad (5.13)$$

where $R_k \sim R_\emptyset$ is the first element of the weight associated with the individual $k \in \mathcal{U}_\infty$. We stress that this condition is satisfied when we choose the constants s_i sufficiently small. Note that under this construction, the conditions of Assumption 2.11 are satisfied: clearly, in general $\mathcal{B}(1) \sim X(1)$ contains no atom on $[0, \infty)$, so Condition 3 is satisfied. Moreover, for any $z \geq 0$, by the conditional independence of $(X_{W_u}(ui))_{i \in \mathbb{N}}$,

$$\mathbb{P}\left(\sum_{i=1}^{\infty} X_{W_u}(ui) = z \mid W_u\right) = \mathbb{P}\left(X_{W_u}(u1) = z - \sum_{i=2}^{\infty} X_{W_u}(ui) \mid W_u\right) = 0,$$

since $X_{W_u}(u1)$ always has an exponential (hence smooth) distribution. Note also, that

$$\mathbb{P}(\tau_\infty < \infty) = 1, \quad \text{and} \quad \mathbb{P}(|\mathcal{T}_\infty| = \infty) = 1.$$

Now, we have the following claim:

Claim 5.4.1. *For all $u \in \mathcal{U}_\infty$,*

$$\mathbb{P}\left(\bigcap_{j=1}^{\infty} \bigcup_{i=j}^{\infty} \{\mathcal{B}(ui) + \mathcal{P}(ui) < \mathcal{B}(u) + \mathcal{P}(u), I_{ui} = 0\} \mid I_u = 0\right) = 1. \quad (5.14)$$

Given Claim 5.4.1, in a similar manner to the proof of Theorem 2.8 (which uses (4.10)), we can exploit Equation (5.14) to deduce that $\mathbb{P}(\tau_{\text{path}}(\emptyset) < \infty \mid I_\emptyset = 0) = 1$. In particular, we can ensure that $\tau_{\text{path}}(\emptyset) \leq \mathcal{P}(\emptyset)$. Indeed, Equation (5.14) guarantees that there exists a child i such that $\mathcal{B}(i) + \mathcal{P}(i) \leq \mathcal{P}(\emptyset)$; this child, in turn, by (5.14) has a child of infinite degree by time $\mathcal{P}(\emptyset)$. Iterating this argument, we have an infinite path by time $\mathcal{P}(\emptyset)$.

Now, it cannot be the case that there exists $u \in \mathcal{U}_\infty$ such that $I_u = 0$ and $\tau_\infty = \mathcal{B}(u) + \mathcal{P}(u)$. Indeed, assume such a u does exist. Then, by the above argument it would follow that $\tau_{\text{path}}(u) = \mathcal{B}(u) + \mathcal{P}(u) = \tau_\infty$. Hence u is contained in an infinite path *and* has infinite degree in \mathcal{T}_∞ , contradicting Theorem 2.12. On the other hand, for any u such that $I_u = 1$, by construction, we have $\mathcal{P}(u) \geq \varsigma_1$. However, by Lemma 5.4 we deduce that, since $\mathbb{P}(\tau_{\text{path}}(\emptyset) < \infty) > 0$ by the above argument, we have $\mathbb{P}(\tau_{\text{path}}(\emptyset) < \varsigma_1) > 0$. It thus follows that, on $\{\tau_{\text{path}}(\emptyset) < \varsigma_1\}$ the tree \mathcal{T}_∞ contains an infinite path almost surely.

Now, suppose that $I_\emptyset = 1$, an event that also occurs with probability $1/2$. Then, for each child $k \in \mathbb{N}$ (in the first generation),

$$\begin{aligned} \mathbb{P}(\mathcal{B}(k1) < \mathcal{P}(\emptyset) \mid I_\emptyset = 1) &= \mathbb{P}\left(X(k1) < \sum_{j=k+1}^{\infty} X(j) \mid I_\emptyset = 1\right) \\ &= \mathbb{P}(X(k1) < \varsigma_k) \leq 1 - \mathbb{E}\left[e^{-(R_k \vee 1)\varsigma_k}\right], \end{aligned}$$

where the last inequality follows from the fact that $X(k1)$ is exponentially distributed, either with parameter 1 or R_k , and the probability is maximised if we choose the maximum of the two. But then, by (5.13) and a union bound,

$$\mathbb{P}\left(\bigcup_{k \in \mathbb{N}} \mathcal{B}(k1) < \mathcal{P}(\emptyset) \mid I_\emptyset = 1\right) < \sum_{k=1}^{\infty} 2^{-k} = 1.$$

It follows that, with positive probability, \mathcal{T}_∞ consists of a single infinite star and hence contains no infinite path. \square

We conclude with the proof of Claim 5.4.1.

Proof of Claim 5.4.1. By using a similar Borel-Cantelli argument as in the proof of Equation (4.11) of Lemma 4.7 (as in Equations (2.9) and (2.10)), it suffices to prove the following: there exists a collection of numbers $\{\nu_n^r \in (0, \infty) : r \in (0, \infty), n \in \mathbb{N}\}$, such that for any $r \in (0, \infty)$,

$$\sum_{n=1}^{\infty} \mathbb{P}(\mathcal{P} < \nu_n^r, I_n = 0) = \infty,$$

and

$$\liminf_{n \rightarrow \infty} \mathbb{P}\left(\sum_{i=n+1}^{\infty} X_{(r,0)}(i) \geq \nu_n^r\right) > 0. \quad (5.15)$$

Here, we set

$$\nu_n^r := d\mathbb{E}\left[\sum_{i=n+1}^{\infty} X_{(r,0)}(i)\right] = d \sum_{i=n+1}^{\infty} \frac{1}{ri^p},$$

for some $d < 1$. Then, note that we may deduce Equation (5.15) by a similar application of the Paley-Zygmund inequality as in Equation (5.7). Moreover, applying Lemma 5.1 with $b_n = \nu_n^r$ and $\varepsilon > 0$ fixed and using the independence of R_n and I_n , for $n \in \mathbb{N}$, we have

$$\begin{aligned} \mathbb{P}(\mathcal{P} < \nu_n^r, I_n = 0) &= \frac{1}{2} \mathbb{P} \left(\sum_{j=1}^{\infty} X_{(R_n, 0)}(j) < \nu_n^r \right) \\ &\geq \frac{1}{2} \left(\mathbb{E} \left[\prod_{j=1}^{\infty} \frac{R_n j^p}{R_n j^p + (1 + \varepsilon)(\nu_n^r)^{-1} \log(n)} \right] - \frac{1}{n^{1+\varepsilon}} \right). \end{aligned}$$

It thus suffices to show that

$$\sum_{n=1}^{\infty} \mathbb{E} \left[\prod_{j=1}^{\infty} \frac{R_n j^p}{R_n j^p + (1 + \varepsilon)(\nu_n^r)^{-1} \log(n)} \right] = \infty. \quad (5.16)$$

To this end, we note that we may bound

$$(1 + \varepsilon)(\nu_n^r)^{-1} \log(n) = \frac{(1 + \varepsilon)r}{d} \log(n) \left(\sum_{j=n+1}^{\infty} \frac{1}{j^p} \right)^{-1} < C_1 \log(n) n^{p-1},$$

where we bound the sum from below by an integral and where $C_1 > 0$ is a constant. For n sufficiently large such that $(1 + \varepsilon)dr \log(n) n^{p-1} > 1$, we now bound

$$\begin{aligned} \mathbb{E} \left[\prod_{j=1}^{\infty} \frac{R_n j^p}{R_n j^p + (1 + \varepsilon)(\nu_n^r)^{-1} \log(n)} \right] &\geq \mathbb{E} \left[\prod_{j=1}^{\infty} \frac{R_n j^p}{R_n j^p + C_1 \log(n) n^{(p-1)}} \mathbf{1}_{R_n \geq C_1 \log(n) n^{(p-1)}} \right] \\ &\geq \left(\prod_{j=1}^{\infty} \frac{j^p}{j^p + 1} \right) \mathbb{P} (R_n \geq C_1 \log(n) n^{p-1}) \\ &\geq C_2 \left(\prod_{j=1}^{\infty} \frac{j^p}{j^p + 1} \right) (\log(n) n^{p-1})^{-(\alpha-1)}, \end{aligned}$$

for some constant $C_2 > 0$ and where we use (5.12) in the final step. Since $p > 1$, $\sum_{j=1}^{\infty} j^{-p} < \infty$, which implies that the infinite product on the right-hand side is strictly larger than zero. Since $(\alpha - 1)(p - 1) < 1$, the lower bound on the right-hand side is not summable in i , so that we obtain (5.16), which concludes the proof. \square

6 Examples of phase transitions in specific models of recursive trees with fitness

The aim of this section is to prove Theorems 3.16 and 3.21. The previous section provided conditions for the emergence of a unique vertex with infinite degree or a unique infinite path to appear almost surely in the case the inter-birth times are exponential random variables whose rate depends on some fitness function f . In this section we turn these conditions into phase transitions for three specific examples. We start by proving Theorem 3.16 in Section 6.1 and then prove Theorem 3.21 in Section 6.2.

We also note that, in this section, we use the commonly applied 'big O', 'little o' and 'big theta' notation: we say $f(x) = \mathcal{O}(g(x))$ if there exists positive constants M, x_0 such that, for all $x \geq x_0$, we have $|f(x)| \leq Mg(x)$; we say $f(x) = o(g(x))$ if $\lim_{x \rightarrow \infty} |f(x)/g(x)| = 0$; and finally, we say $f(x) = \Theta(g(x))$ if $f(x) = \mathcal{O}(g(x))$ and $g(x) = \mathcal{O}(f(x))$.

6.1 Phase transitions for the emergence of an infinite star or path: proof of Theorem 3.16

To prove Theorem 3.16, it suffices to check the conditions subject to which Theorem 3.4 holds. In the remainder of this section, it is be useful to have the following preliminary results from [13] that we use throughout.

Lemma 6.1 ([13, Propositions 1.3.6, 1.5.7, 1.5.9a, and 1.5.10]). *Let ℓ be a slowly-varying function and r_1, r_2 be regularly-varying functions with exponents $\alpha_1, \alpha_2 \in \mathbb{R}$, respectively, and fix $a \in \mathbb{R}$. Then,*

- (i) *For $a > 0$, it holds that $\lim_{x \rightarrow \infty} \ell(x)x^a = \infty$ and $\lim_{x \rightarrow \infty} \ell(x)x^{-a} = 0$.*
- (ii) *The function $r_1^a := (r_1(\cdot))^a$ is regularly varying with exponent $\alpha_1 a$.*
- (iii) *If $\alpha_2 > 0$, then $r_1(r_2(\cdot))$ is regularly varying with exponent $\alpha_1 \alpha_2$.*
- (iv) *The function $r_1 + r_2$ is regularly varying with exponent $\max\{\alpha_1, \alpha_2\}$.*
- (v) *The function $r_1 r_2$ is regularly varying with exponent $\alpha_1 + \alpha_2$.*
- (vi) *If $\alpha_1 < -1$, then*

$$\int_x^\infty r_1(t) dt = (-\alpha_1 - 1 + o(1)) x r_1(x), \quad \text{as } x \rightarrow \infty.$$

When $\alpha_1 = -1$, the result remains true in the sense that the integral is slowly varying in x , and that it is $o(x r_1(x))$.

We generally use Lemma 6.1 without always referring to it explicitly in the remainder of this section. We mainly use it in the following manner. First, Lemma 6.1 shows that regularly-varying functions are closed under elementary operations. Next, Item (vi) shows that, if $s(n)$ is regularly varying with exponent $p > 0$, then $\sum_{i=n}^\infty s(i)^{-1}$ is regularly varying with exponent $p - 1$. In addition, Item (i) shows that slowly-varying functions are often "negligible" in the sense that, for a regularly-varying function f with exponent $\gamma \neq 0$, we can write $f(x) = x^{\gamma+o(1)}$. Moreover, it shows that for any regularly-varying function s with exponent $p > 1$ that $\sum_{i=1}^\infty s(i)^{-1} < \infty$, whilst for any regularly-varying function s with exponent $p < 1$ we have $\sum_{i=1}^\infty s(i)^{-1} = \infty$.

The following approximations for μ_n^w , as defined in (3.1), are also used heavily.

Lemma 6.2. *Consider the two choices for the fitness function f (mixed or additive weights), as in Assumption 3.10, as well as the two choices for the degree function s (super-linear*

and barely super-linear), as in Assumption 3.13. For the *super-linear* case and for each fixed $w \in [0, \infty)$,

$$\mu_n^w = \frac{1 + o(1)}{g(w)(p-1)} ns(n)^{-1}.$$

For the *barely super-linear* case there exists a slowly-varying function L such that for each $w \in [0, \infty)$,

$$\mu_n^w = \frac{1 + o(1)}{g(w)} L(n).$$

In particular, for the barely super-linear *log-stretched* case, it follows that

$$L(n) = \beta^{-1} (\log n)^{1-\beta} e^{-(\log n)^\beta}.$$

Proof. We fix $w \in [0, \infty)$ throughout and generally use that

$$\mu_n^w = \sum_{i=n}^{\infty} \frac{1}{f(i, w)} = \sum_{i=n}^{\infty} \frac{1}{g(w)s(i) + h(w)} = \frac{1 + o(1)}{g(w)} \sum_{i=n}^{\infty} \frac{1}{s(i)},$$

where we recall that $g \equiv 1$ in the additive case, and we can omit the term $h(w)$ in the fraction at the cost of a $o(1)$ term. Taking s as in the *super-linear* case, since s is assumed to be regularly varying with exponent $p > 1$, by the integral test and using Lemma 6.1, we thus find that

$$\mu_n^w = \frac{1 + o(1)}{g(w)(p-1)} ns(n)^{-1}.$$

With s as in the barely super-linear case, it follows from (vi) in Lemma 6.1 and an integral test that

$$L(n) := \int_n^{\infty} \frac{1}{s(x)} dx.$$

For the *log-stretched* case, we have with s as in (3.5),

$$L(n) = \int_{n+1}^{\infty} x^{-1} \exp(-(\log x)^\beta) dx.$$

Using a change of variables $y = (\log x)^\beta$, we obtain, with $\Gamma(s, x)$, $s, x > 0$, the upper incomplete gamma function,⁴

$$L(n) = \beta^{-1} \Gamma(1/\beta, (\log(n+1))^\beta) = (\beta^{-1} + o(1)) (\log n)^{1-\beta} e^{-(\log n)^\beta},$$

which concludes the proof. □

6.1.1 Conditions for an infinite star: Item 1 of Theorem 3.4.

To provide conditions for an infinite star to appear almost surely, we apply Theorem 3.4 by proving the condition in Item 1 is satisfied in certain cases. To this end, we state the following lemma.

⁴The upper incomplete gamma function is defined, for $s, x > 0$ by $\Gamma(s, x) := \int_x^{\infty} t^{s-1} e^{-t} dt$.

Lemma 6.3. Equation (3.2) in Item 1 of Theorem 3.4 is satisfied when the following conditions are met, based on the assumptions for the fitness type, degree function s , and vertex-weight distribution:

| Weight | Degree | Star |
|----------|---------------|--|
| Mixed | Super-linear | (3.6) & $(p-1)(\alpha-1) > (\gamma - \frac{\gamma-1}{p}) \vee 1$ |
| Additive | Super-linear | (3.6) & $p(\alpha-1) > 1$ |
| Mixed | Log-stretched | (3.8) & $\beta\nu > 1$ |

Table 2: The first column represents the form of the fitness function, as in Assumption 3.10; the second represents the form of the degree function, as in Assumption 3.13. The third column lists the required assumptions on the vertex-weight distribution, as in Assumption 3.15, together with the choices of the parameters that lead to a unique node of infinite degree.

We split the proof of Lemma 6.3 based on the different choices for the fitness function f and degree function s , as in Assumptions 3.10 and 3.13, respectively. As a general approach, we bound, for some sequence $(k_n)_{n \in \mathbb{N}} \in [0, \infty)^{\mathbb{N}}$,

$$\mathbb{E} \left[\prod_{j=0}^{\infty} \frac{f(j, W)}{f(j, W) + c\mu_n^{-1}} \right] \leq \prod_{j=0}^{\infty} \frac{\sup_{w \leq k_n} f(j, w)}{\sup_{w \leq k_n} f(j, w) + c\mu_n^{-1}} + \mathbb{P}(W \geq k_n).$$

We now use that, by Assumption 3.10, for $x \geq 0$ and $j \in \mathbb{N}$,

$$\sup_{w \leq x} f(j-1, w) \leq \left(\sup_{w \leq x} g(w) \right) s(j-1) + \sup_{w \leq x} h(w) =: \bar{g}(x) s(j-1) + \bar{h}(x), \quad (6.1)$$

where the suprema are well-defined as we assume g and h to be continuous. Since g and h are regularly varying with exponents 1 and γ in the **mixed** case and $g \equiv 1$ and h is regularly varying with exponent 1 in the **additive** case, it follows from [13, Theorem 1.5.3] that $\bar{g}(x) = (1 + o(1))g(x)$ and $\bar{h}(x) = (1 + o(1))h(x)$; in particular, they are regularly varying with the same exponent (and $\bar{g} = g \equiv 1$ in the additive case). As such, using the weight functions \bar{g} and \bar{h} falls in the same family of weight functions as g and h . The advantage of using \bar{g} and \bar{h} is that these are monotone increasing. However, by this argument we can thus use g and h instead and assume, without loss of generality, that they are monotone increasing. This yields

$$\mathbb{E} \left[\prod_{j=0}^{\infty} \frac{f(j, W)}{f(j, W) + c\mu_n^{-1}} \right] \leq \exp \left(-c\mu_n^{-1} \sum_{j=0}^{\infty} \frac{1}{f(j, k_n) + c\mu_n^{-1}} \right) + \mathbb{P}(W \geq k_n). \quad (6.2)$$

Using Assumption 3.15, one can choose k_n to grow with n sufficiently fast so that the second term on the right-hand side is summable in n . It then remains to show that the exponential term is sufficiently small for this choice of k_n as well, to obtain that the left hand side is summable in n .

6.1.2 Conditions for an infinite path, Assumption 2.6.

To prove the appearance of an infinite path, we use Item 2 of Theorem 3.4. In particular we verify Condition (2.9) with $\nu_n^w := d\mu_n^w$ for some constant $d \in (0, 1)$ and μ_n^w as in (3.1). Notably, the approach we use to do so works for *any* inter-birth distribution satisfying the variance condition of Remark 3.6; it is not tailored to the exponential distribution.

We thus have the following lemma.

Lemma 6.4. *The condition in (2.9) satisfied when the following conditions are met, based on the assumptions for the fitness type, degree function s , and vertex-weight distribution:*

| Weight | Degree | Path |
|----------|---------------|--|
| Mixed | Super-linear | (3.7) & $(p-1)(\alpha-1) < (\gamma - \frac{\gamma-1}{p}) \vee 1$ |
| Additive | Super-linear | (3.7) & $p(\alpha-1) < 1$ |
| Mixed | Log-stretched | (3.9) & $\beta\nu < 1$ |

Table 3: The first column represents the form of the fitness function, as in Assumption 3.10; the second represents the form of the degree function, as in Assumption 3.13. The third column lists the required assumptions on the vertex-weight distribution, as in Assumption 3.15, together with the choices of the parameters that lead to a unique infinite path.

We observe that combining Lemmas 6.3 and 6.4 proves Theorem 3.16.

As a general approach to proving Lemma 6.4, we start with a general lower bound for each of the terms in the sum in (2.9), and then show that this lower bound is not summable for each of the specific cases dealt with in Lemma 6.4. We first introduce the event $\{W \geq k_n\}$ for some sequence $(k_n)_{n \in \mathbb{N}}$ such that the tail probability $\mathbb{P}(W \geq k_n)$ is not summable. By a similar argument as in (6.1), but applied to $\inf_{w \geq x} f(j-1, w)$ (which is monotone increasing in x), we may assume, without loss of generality, that f is monotone increasing in its second argument.

Then, on the event $\{W \geq k_n\}$, we stochastically dominate each inter-birth time $X_W(j)$ by $\tilde{X}_{k_n}(j)$, which is distributed as an exponential random variable with rate $f(j-1, k_n)$, and is independent of both W and $X_W(j)$, for each $j \in \mathbb{N}$. It follows that \mathcal{P} stochastically dominates $\tilde{\mathcal{P}} = \sum_{j=1}^{\infty} \tilde{X}_{k_n}(j)$, which again is independent of W and \mathcal{P} . We thus obtain the lower bound

$$\mathbb{P}(\mathcal{P} < \nu_n^w) \geq \mathbb{P}(\{\mathcal{P} < \nu_n^w\} \cap \{W \geq k_n\}) \geq \mathbb{P}(\tilde{\mathcal{P}} < \nu_n^w) \mathbb{P}(W \geq k_n).$$

Since we choose k_n such that the second probability on the right-hand side is not summable, it suffices to show that the first probability is uniformly bounded from below in n . By Markov's inequality and the choice of ν_n^w , we obtain the lower bound

$$\begin{aligned} \mathbb{P}(\tilde{\mathcal{P}} < \nu_n^w) &\geq \left(1 - \frac{1}{d\mu_n^w} \mathbb{E}[\tilde{\mathcal{P}}]\right) \mathbb{P}(W \geq k_n) \\ &= \left(1 - \frac{1}{d\mu_n^w} \sum_{j=0}^{\infty} \frac{1}{f(j, w)}\right) \mathbb{P}(W \geq k_n). \end{aligned} \tag{6.3}$$

With k_n such that the probability on the right-hand side is not summable and since $d \in (0, 1)$ is arbitrary, it thus suffices to prove that

$$\limsup_{n \rightarrow \infty} \frac{1}{\mu_n^w} \sum_{j=0}^{\infty} \frac{1}{f(j, k_n)} < 1. \tag{6.4}$$

If this holds, it follows that the condition in (2.9) is satisfied. We thus prove (6.4) for the cases in Lemma 6.4.

6.1.3 Proofs of Lemmas 6.3 and 6.4

We now proceed to state the proofs of Lemmas 6.3 and 6.4. We do this case by case, and we combine the proofs of the related statements in both lemmas. We use the approaches outlined after the statements of both lemmas.

Proof of Lemmas 6.3 and 6.4, s super-linear, mixed weights. We first prove the claim from Lemma 6.3. We distinguish between $\gamma \leq 1$ and $\gamma > 1$, and we start with the former case. Observe that, when $\gamma \leq 1$, it follows that $\max\{\gamma - (\gamma - 1)/p, 1\} = 1$. Hence, we can take $\varepsilon > 0$ small enough such that $(1 - \varepsilon)(p - 1)(\alpha - 1) > 1 = \max\{\gamma - (\gamma - 1)/p, 1\}$. We set $\beta := (1 - \varepsilon)(p - 1)$ and $k_n := n^\beta$. We use (3.6) to deduce that $\mathbb{P}(W \geq k_n) \leq \bar{\ell}(k_n)n^{-(\alpha-1)\beta}$, which, as $\beta(\alpha - 1) > 1$, is summable. For the first term on the right-hand side of (6.2), we write

$$\exp\left(-\sum_{j=0}^{\infty} \frac{c\mu_n^{-1}}{f(j, k_n) + c\mu_n^{-1}}\right) = \exp\left(-\frac{c\mu_n^{-1}}{g(n^\beta)} \sum_{j=0}^{\infty} \frac{1}{s(j) + h(n^\beta)/g(n^\beta) + c\mu_n^{-1}/g(n^\beta)}\right). \quad (6.5)$$

As $\gamma \leq 1$, it follows that $\beta\gamma < p - 1$. Recalling that, by Lemma 6.2, μ_n^{-1} is regularly varying with exponent $p - 1$, it follows that

$$a_n := \frac{h(n^\beta) + c\mu_n^{-1}}{g(n^\beta)}$$

is regularly varying with exponent $\varepsilon(p - 1)$. We now write $s(x) = \ell(x)x^p$ for some slowly-varying function ℓ . Moreover, since $p > 1$, we have $j^p + a_n \leq (j + a_n^{1/p})^p$. As a result, for any $\eta > 0$ there exists $J = J(\eta) \in \mathbb{N}$ such that for all $j \geq J$,

$$s(j) + a_n = \ell(j)j^p + a_n \leq j^\eta(j^p + a_n) \leq (j + a_n^{1/p})^{p+\eta}. \quad (6.6)$$

We can thus bound the sum in (6.5) from below by

$$\sum_{j=J}^{\infty} \frac{1}{s(j) + a_n} \geq \sum_{j=J}^{\infty} (j + a_n^{1/p})^{-(p+\eta)} \geq \sum_{j=J+[a_n^{1/p}]}^{\infty} j^{-(p+\eta)} = \Theta(a_n^{-(p+\eta-1)/p}), \quad (6.7)$$

where the final step uses an integral test. Using this in (6.5), we thus obtain, for some constants $C, C' > 0$, the upper bound

$$\exp\left(-C \frac{\mu_n^{-1} a_n^{-(p+\eta-1)/p}}{g(n^\beta)}\right) = \exp\left(-C' a_n^{(1-\eta)/p}\right).$$

As the term in the exponential varies regularly with exponent $\varepsilon(p - 1)(1 - \eta)/p > 0$, the exponential term is summable. We thus conclude that (3.2) is satisfied, which concludes the proof for $\gamma \leq 1$.

We then consider the case $\gamma > 1$. We first use our assumption, and fix $\varepsilon > 0$ small such that

$$\frac{(1 - \varepsilon)(p - 1)(\alpha - 1)}{\gamma - (\gamma - 1)/p} > 1, \quad (6.8)$$

which is possible since $\max\{\gamma - (\gamma - 1)/p, 1\} = \gamma - (\gamma - 1)/p$ when $\gamma, p > 1$. We then set $\beta := (1 - \varepsilon)(p - 1)/(\gamma - (\gamma - 1)/p)$. By the same argument as before, since $\beta(\alpha - 1) > 1$, the

second term on the right-hand side of (6.2) is summable. To bound the first term, we again write the upper bound

$$\exp\left(-c\mu_n^{-1}\sum_{j=0}^{\infty}\frac{1}{f(j,n^\beta)+c\mu_n^{-1}}\right)=\exp\left(-\frac{c\mu_n^{-1}}{g(n^\beta)}\sum_{j=0}^{\infty}\frac{1}{s(j)+(h(n^\beta)+c\mu_n^{-1})/g(n^\beta)}\right). \tag{6.9}$$

Note that the exponent associated with the regularly-varying function $h(n^\beta)$ is $\beta\gamma = (1 - \varepsilon)(p - 1)\gamma/(\gamma - (\gamma - 1)/p) > p - 1$ for ε sufficiently small (since we assume that $\gamma > 1$). Thus, making ε smaller if necessary (which means that (6.8) still holds), recalling that, by Lemma 6.2, μ_n is regularly varying with exponent $-(p - 1)$ in the super-linear case, it follows that, if $a_n := (h(n^\beta) + c\mu_n^{-1})/g(n^\beta)$, $a_n^{1/p}$ is regularly varying with exponent

$$\beta(\gamma - 1)/p = \frac{(1 - \varepsilon)(p - 1)(\gamma - 1)}{\gamma(p - 1) + 1}.$$

As a result, using the same approach as in (6.7), we can bound the sum appearing in the exponent in the right-hand side of (6.9) from below by

$$\sum_{j=J}^{\infty}\frac{1}{s(j)+a_n}\geq\sum_{j=J}^{\infty}(j+a_n^{1/p})^{-(p+\eta)}\geq\sum_{j=J+[a_n^{1/p}]}^{\infty}j^{-(p+\eta)}=\Theta(a_n^{-(p+\eta-1)/p}).$$

Combining this lower bound with (6.9), for some constant $C''' > 0$, we obtain the upper bound

$$\exp\left(-C'''\frac{\mu_n^{-1}a_n^{-(p+\eta-1)}}{g(n^\beta)}\right).$$

Since we can choose η arbitrarily small, balancing the exponent of μ_n^{-1} (which is $p - 1$) with the exponents of $a_n^{-(p-1)/p}$ and $g(n^\beta)$ it follows that the fraction is regularly varying in n with an exponent that is positive if

$$p - 1 > (1 - \varepsilon)\left(\frac{(p - 1)^2(\gamma - 1)}{\gamma(p - 1) + 1} + \frac{p(p - 1)}{\gamma(p - 1) + 1}\right) = (1 - \varepsilon)(p - 1).$$

As this inequality is clearly always satisfied for any $\varepsilon > 0$ we obtain that the left-hand side of (6.2) is summable (which implies (3.2)), as desired.

We then prove the claim in Lemma 6.4. Again, we distinguish between the cases $\gamma \leq 1$ and $\gamma > 1$. Let $\varepsilon > 0$ be sufficiently small such that

$$\begin{cases} (p - 1)(\alpha - 1) < 1 - \varepsilon & \text{if } \gamma \leq 1, \\ (p - 1)(\alpha - 1)/(\gamma - (\gamma - 1)/p) < 1 - \varepsilon & \text{if } \gamma > 1. \end{cases} \tag{6.10}$$

In either case, we set $k_n = n^{(1-\varepsilon)/(\alpha-1)}$. Then $\mathbb{P}(W \geq k_n) \geq \ell(k_n)n^{-(1-\varepsilon)}$ by (3.7), which is not summable in n . Moreover, we have $\mu_n(w) = \ell(n)n^{-(p-1)}$ for ℓ some slowly-varying function (which may depend on w up to constants only) by Lemma 6.2. Now, when $\gamma \leq 1$, we have the upper bound

$$\sum_{j=0}^{\infty}\frac{1}{f(j,k_n)}=\sum_{j=0}^{\infty}\frac{1}{g(k_n)s(j)+h(k_n)}\leq\frac{1}{g(k_n)}\sum_{j=0}^{\infty}\frac{1}{s(j)}.$$

Since s varies regularly with exponent $p > 1$, it follows that the sum is finite. Since g is regularly varying with exponent 1 we can write $g(x) = x^{1+o(1)}$ so that we obtain the upper bound $n^{-(1-\varepsilon+o(1))/(\alpha-1)}$. As $(1-\varepsilon)/(\alpha-1) > p-1$, it follows that (6.4) is satisfied.

Otherwise, for $\gamma > 1$, by similar computations as in (6.6) and (6.7) and using that $j^p + x \geq \eta(j + x^{1/p})^p$ for j, x , sufficiently large and η small, we obtain, for $\eta > 0$ sufficiently small,

$$\begin{aligned} \sum_{j=0}^{\infty} \frac{1}{f(j, k_n)} &= \frac{1}{g(k_n)} \sum_{j=0}^{\infty} \frac{1}{s(j) + h(k_n)/g(k_n)} \\ &\leq \frac{1}{\eta g(k_n)} \sum_{j=\lfloor (h(k_n)/g(k_n))^{1/p} \rfloor}^{\infty} j^{-(p-\eta)} = \mathcal{O}\left(\frac{g(k_n)^{-(1+\eta)/p}}{h(k_n)^{(p-1-\eta)/p}}\right). \end{aligned}$$

By Lemma 6.1, the term $g(k_n)^{-(1-\eta)/p}/h(k_n)^{(p-1-\eta)/p}$ is regularly varying, with exponent

$$-\frac{1-\varepsilon}{\alpha-1} \left(\frac{1+\eta}{p} + \gamma \frac{p-1-\eta}{p} \right) = -\frac{1-\varepsilon}{\alpha-1} \left(\gamma - \frac{\gamma-1}{p} - \frac{(\gamma-1)\eta}{p} \right) < -(p-1),$$

for $\eta > 0$ sufficiently small, where the last step uses the second inequality in (6.10). Consequently, it follows that $g(k_n)^{-(1-\eta)/p}/h(k_n)^{(p-1-\eta)/p} = o(\mu_n^w)$ and we obtain (6.4). \square

Proof of Lemmas 6.3 and 6.4, s super-linear, additive weights. We first prove the claim in Lemma 6.3. We fix $\varepsilon \in (0, 1/p)$ small such that $p(\alpha-1)(1-\varepsilon) > 1$. Then, with $k_n := n^{(1-\varepsilon)p}$ and using (3.6), $\mathbb{P}(W \geq k_n) \leq \bar{\ell}(k_n)n^{-(\alpha-1)p(1-\varepsilon)}$, and since $(\alpha-1)p(1-\varepsilon) > 1$ we can deduce that the second term on the right-hand side of (6.2) is summable. For the first term, we bound $f(j, k_n)$ from above by $\sup_{j \leq n^{1-\varepsilon}} f(j, k_n)$ for all $j \leq n^{1-\varepsilon}$. With a similar argument as in (6.1), we may assume that f is monotone increasing in its first argument, since s is regularly-varying with exponent $p > 1$. This thus yields the upper bound

$$\exp\left(-\sum_{j=0}^{\infty} \frac{c\mu_n^{-1}}{f(j, k_n) + c\mu_n^{-1}}\right) \leq \exp\left(-\frac{c\mu_n^{-1}n^{1-\varepsilon}}{f(n^{1-\varepsilon}, k_n) + c\mu_n^{-1}}\right).$$

As $(1-\varepsilon)p > p-1$ using Lemma 6.2, we see that $f(n^{1-\varepsilon}, k_n) + c\mu_n^{-1} = s(n^{1-\varepsilon}) + h(n^{(1-\varepsilon)p}) + c\mu_n^{-1}$ is regularly varying with exponent $(1-\varepsilon)p$. It follows that the fraction is a regularly-varying function in n , with exponent $(p-1)\varepsilon > 0$. So, for some slowly-varying function L , we can write the exponential term as $\exp(-L(n)n^{(p-1)\varepsilon})$. We thus arrive at the desired conclusion that the exponential term is summable in n . This shows that the left-hand side of (6.2) is summable and hence that (3.2) is satisfied.

We then prove the claim in Lemma 6.4. We fix $\varepsilon > 0$ sufficiently small such that $p(\alpha-1) < 1-\varepsilon$ and let $k_n := n^{(1-\varepsilon)/(\alpha-1)}$. We then have $\mathbb{P}(W \geq k_n) \geq \underline{\ell}(k_n)n^{-(1-\varepsilon)}$ by (3.7), which is not summable in n . Then, for w fixed, $\mu_n^w = \ell(n)n^{-(p-1)}$, for some slowly-varying function ℓ (which may depend on w up to constants only), by Lemma 6.2. By calculations similar to those in (6.6) and (6.7), we have for any $\eta > 0$ and some large constant $C > 0$, for all n sufficiently large,

$$\sum_{j=0}^{\infty} \frac{1}{f(j, k_n)} = \sum_{j=0}^{\infty} \frac{1}{s(j) + h(k_n)} \leq C \sum_{j=\lfloor h(k_n)^{1/p} \rfloor}^{\infty} j^{-(p-\eta)} = \mathcal{O}(h(k_n)^{-(p-1-\eta)/p}).$$

Since h varies regularly with exponent 1, the last term is $\mathcal{O}(n^{-(p-1-\eta)/(1-\varepsilon)})$ since $p(\alpha-1) < 1-\varepsilon$. As we can choose η arbitrarily small, by the choice of k_n , it follows that (6.4) holds. \square

Proof of Lemmas 6.3 and 6.4, s log-stretched case, mixed weights. First, we prove the claim from Lemma 6.3. We assume that $\beta\nu > 1$ and recall that

$$f(i, w) = g(w)(i + 1)e^{(\log(i+1))^\beta} + h(w),$$

where g and h are regularly-varying functions with exponents 1 and $\gamma \geq 0$, respectively. Fix $\varepsilon > 0$ small so that $\beta\nu > 1 + \varepsilon$. We apply (6.2) with

$$k_n := \exp((\log n)^{(1+\varepsilon)/\nu}), \quad \text{and} \quad \mu_n = \frac{1 + o(1)}{g(0)\beta} (\log n)^{1-\beta} \exp(-(\log n)^\beta), \quad (6.11)$$

where the latter follows from Lemma 6.2. Now, using (3.8) we obtain

$$\mathbb{P}(W \geq k_n) \leq \exp(-\bar{c}(\log k_n)^\nu) = \exp(-\bar{c}(\log n)^{1+\varepsilon}) \leq n^{-(1+\varepsilon)}, \quad (6.12)$$

for n sufficiently large, which implies that the second term on the right-hand side of (6.2) is summable. For the first term, we define

$$I_n := \exp((\log n)^\beta). \quad (6.13)$$

Since g is regularly varying with exponent one, we can write $g(x) = \ell(x)x = x^{1+o(1)}$ for some slowly-varying function ℓ . As a result, using (6.11),

$$\mu_n^{-1}/g(k_n) = (g(w)\beta + o(1))(\log n)^{\beta-1} \exp((\log n)^\beta - (\log n)^{(1+\varepsilon)/\nu}(1 + o(1))). \quad (6.14)$$

We thus obtain that for all $j \geq I_n$, when n is large,

$$(j + 1) \exp((\log(j + 1))^\beta) \geq I_n \exp((\log I_n)^\beta) = \exp((\log n)^\beta + (\log n)^{\beta^2}) \geq \mu_n^{-1}/g(k_n).$$

Moreover, since $\beta\nu > 1 + \varepsilon$, it holds that $h(k_n) = o(\mu_n^{-1})$, irrespective of the regularly-varying exponent $\gamma \geq 0$ of h (since we can again write $h(x) = x^{\gamma+o(1)}$). As a result, we obtain the lower bound

$$\begin{aligned} & \sum_{j=0}^{\infty} \frac{1}{g(k_n)(j + 1) \exp((\log(j + 1))^\beta) + h(k_n) + \mu_n^{-1}} \\ & \geq \frac{1}{g(k_n)} \sum_{j=I_n}^{\infty} \frac{1}{3(j + 1) \exp((\log(j + 1))^\beta)} = \frac{\mu_{I_n}(1 + o(1))}{3g(k_n)/g(0)} \geq \frac{g(0)\mu_{I_n}}{4g(k_n)}, \end{aligned}$$

for n sufficiently large, where we recall the definition of μ_n from (3.1) (with $w^* = 0$) and use (the proof of) Lemma 6.2 in the final step. Substituting this bound into the sum in the exponent of the first term on the right-hand side of (6.2), we obtain the upper bound $\exp(-\frac{1}{4}cg(0)\mu_{I_n}\mu_n^{-1}/g(k_n))$. By the choice of I_n in (6.13) and μ_n as in (6.11),

$$\mu_{I_n} = \frac{1 + o(1)}{g(0)\beta} (\log I_n)^{1-\beta} \exp(-(\log I_n)^\beta) = \exp(-(\log n)^{\beta^2}(1 - o(1))).$$

By (6.14), we deduce that that the dominant term in $\exp(-\frac{c}{4C}\mu_{I_n}\mu_n^{-1}/g(k_n))$ is

$$\exp(-\exp(C'(\log n)^\beta))$$

for some $C' > 0$, since $\beta\nu > 1 + \varepsilon$. By a similar argument as in (6.12), this upper bound is summable in n , as desired.

We then prove the claim in Lemma 6.4. We fix $\varepsilon > 0$ sufficiently small such that $\beta\nu < 1 - \varepsilon$ and set $k_n := \exp((\log n)^{(1-\varepsilon)/\nu})$. It follows from (3.9) that, for n sufficiently large

$$\mathbb{P}(W \geq k_n) \geq e^{-c(\log n)^{1-\varepsilon}} \geq n^{-1},$$

which is not summable. We then have upper bound

$$\begin{aligned} \sum_{i=0}^{\infty} \frac{1}{f(i, k_n)} &= \frac{1}{g(k_n)} \sum_{i=0}^{\infty} \frac{1}{(i+1) \exp((\log(i+1))^\beta) + h(k_n)/g(k_n)} \\ &\leq \frac{1}{g(k_n)} \sum_{i=0}^{\infty} \frac{1}{(i+1) \exp((\log(i+1))^\beta)} \leq \frac{C_\beta}{g(k_n)}, \end{aligned} \tag{6.15}$$

for some constant $C_\beta > 0$, which follows since for any $\varepsilon > 0$ and i sufficiently large,

$$\frac{1}{(i+1) \exp((\log(i+1))^\beta)} \leq \frac{1}{(i+1)(\log(i+1))^{1+\varepsilon}},$$

so that the sum is indeed finite. Hence, since g is regularly varying with exponent one and we can thus write $g(x) = x^{1+o(1)}$, we obtain

$$\sum_{i=0}^{\infty} \frac{1}{f(i, k_n)} \leq C_\beta \exp(-(1+o(1))(\log n)^{(1-\varepsilon)/\nu}).$$

Since $\beta\nu < 1 - \varepsilon$, the $-(\log n)^{(1-\varepsilon)/\nu}$ term in the exponential function dominates the $(\log n)^\beta$ term in the exponential function in $(\mu_n^w)^{-1}$; it thus follows that (6.4) is satisfied. \square

6.2 Sub-tree counts in specific models of recursive trees with fitness

The aim of this subsection is to turn the conditions of Theorem 3.7 into more concrete conditions, which we then proceed to check for the different examples discussed in Section 3 to prove Theorem 3.21. We start by bounding the summands of (3.4) from below and above.

Lower bound. For a lower bound, we use any one fixed ordering $O \in \mathcal{O}(T)$, drop the indicators from the denominator, and introduce the indicator of the event $\{\sup_{i \leq k, j \leq k-1} f(i, W_j) < \mu_n^{-1}\}$. By bounding all terms $f(\deg^+(v_i, O_{|j}), W_{v_i})$ (except for $f(\ell, W_{v_j})$) in the denominator from above by μ_n^{-1} , we obtain the lower bound

$$\begin{aligned} &\mathbb{E} \left[\mathbf{1}_{\sup_{i \leq k, j \leq k-1} f(i, W_{v_j}) < \mu_n^{-1}} \prod_{j=0}^{k-1} \prod_{\ell=0}^{\deg^+(v_j, T)-1} \frac{f(\ell, W_{v_j})}{\sum_{i=0}^j f(\deg^+(v_i, O_{|j}), W_{v_i}) + \mu_n^{-1}} \right] \\ &\geq \mathbb{E} \left[\mathbf{1}_{\sup_{i \leq k, j \leq k-1} f(i, W_{v_j}) < \mu_n^{-1}} \prod_{j=0}^{k-1} \prod_{\ell=0}^{\deg^+(v_j, T)-1} \frac{f(\ell, W_{v_j})}{f(\ell, W_{v_j}) + k\mu_n^{-1}} \right], \end{aligned}$$

where we use that the sum in the denominator has at most $k - 1$ terms besides $f(\ell, W_{v_j})$. By the independence of the vertex-weights, setting

$$f^+(k, W) := \sup_{i \leq k} f(i, W), \quad \text{and} \quad f^-(k, W) := \inf_{i \leq k} f(i, W), \tag{6.16}$$

and using that the mapping $x \mapsto x + y$ is increasing in x , we obtain the lower bound

$$\prod_{j=0}^{k-1} \mathbb{E} \left[\mathbf{1}_{f^+(k,W) < \mu_n^{-1}} \left(\frac{f^-(k,W)}{f^-(k,W) + k\mu_n^{-1}} \right)^{\deg^+(v_j,T)} \right], \quad (6.17)$$

where we use that $\deg^+(v_j, T) \leq k$ for any vertex v_j and any tree T of size $k + 1$.

Upper bound. For an upper bound we use a similar approach. First, irrespective of the ordering $O \in \mathcal{O}(T)$, we can drop all terms from the denominator except $f(\ell, W_{v_j})$ and μ_n^{-1} and use independence of the vertex weights to obtain the upper bound

$$\begin{aligned} & \mathbb{E} \left[\prod_{j=0}^{k-1} \frac{\prod_{\ell=0}^{\deg^+(v_j,T)-1} f(\ell, W_{v_j})}{\sum_{i=0}^j f(\deg^+(v_i, O_{|j}), W_{v_i}) \mathbf{1}_{\deg^+(v_i, O_{|j}) < \deg^+(v_i, T)} + \mu_n^{-1}} \right] \\ & \leq \prod_{j=0}^{k-1} \mathbb{E} \left[\prod_{\ell=0}^{\deg^+(v_j,T)-1} \frac{f(\ell, W_{v_j})}{f(\ell, W_{v_j}) + \mu_n^{-1}} \right] \leq \prod_{j=0}^{k-1} \mathbb{E} \left[\left(\frac{f^+(k, W)}{f^+(k, W) + \mu_n^{-1}} \right)^{\deg^+(v_j, T)} \right], \end{aligned} \quad (6.18)$$

where now $f^+(k, w) := \sup_{i \leq k} f(i, w)$ and we again use that the out-degrees are at most k . As the terms in the expected value are at most one, we condition on the size of $f^+(k, W)$ to obtain the upper bound

$$\prod_{j=0}^{k-1} \left(\mathbb{E} \left[\mathbf{1}_{f^+(k,W) < \mu_n^{-1}} \left(\frac{f^+(k, W)}{f^+(k, W) + \mu_n^{-1}} \right)^{\deg^+(v_j, T)} \right] + \mathbb{P}(f^+(k, W) \geq \mu_n^{-1}) \right).$$

As a result, since $|\mathcal{O}(T)| < (k + 1)!$ for any tree of size $k + 1$, we can use the above in (3.4) to obtain that, for some constant $D = D(k) > 0$, the inner sum in (3.4) is at most

$$D \prod_{j=0}^{k-1} \left(\mathbb{E} \left[\mathbf{1}_{f^+(k,W) < \mu_n^{-1}} \left(\frac{f^+(k, W)}{f^+(k, W) + \mu_n^{-1}} \right)^{\deg^+(v_j, T)} \right] + \mathbb{P}(f^+(k, W) \geq \mu_n^{-1}) \right). \quad (6.19)$$

We now use the lower and upper bound to determine, based on certain conditions on the vertex-weights, whether the double sum in Theorem 3.7 is finite or infinite. In particular, we are now in a position to prove Corollary 3.8.

6.2.1 Proof of Corollary 3.8

Proof of Corollary 3.8. We start by using the lower bound in (6.17). For each expected value in the product, we have, since $f^-(k, W) \leq f^+(k, W)$, the lower bound

$$\begin{aligned} & \mathbb{E} \left[\mathbf{1}_{f^+(i,W) < \mu_n^{-1}} \left(\frac{f^-(k, W)}{(k + 1)\mu_n^{-1}} \right)^{\deg^+(v_j, T)} \right] \\ & \geq (k + 1)^{-k} \mu_n^{\deg^+(v_j, T)} \mathbb{E} \left[\mathbf{1}_{f^+(i,W) < \mu_n^{-1}} f^-(k, W)^{\deg^+(v_j, T)} \right]. \end{aligned}$$

As μ_n^{-1} diverges with n and by the moment condition, by monotone convergence we find that the expected value on the right-hand side equals $\mathbb{E} \left[f^-(k, W)^{\deg^+(v_j, T)} \right] (1 - o(1))$, which is

finite. As such, substituting this in (6.17), we arrive at the lower bound

$$C \prod_{j=0}^{k-1} \mu_n^{\deg^+(v_j, T)} = C \mu_n^k,$$

since the sum of the out-degrees equals $|T| - 1 = k$, and where $C > 0$ is a constant.

For an upper bound we use (6.18): omitting the terms $f^*(k, W)$ from the denominator, yields the upper bound

$$D \prod_{j=0}^{k-1} \left(\mu_n^{\deg^+(v_j, T)} \mathbb{E} \left[f^+(k, W)^{\deg^+(v_j, T)} \right] \right).$$

By the moment condition we again have that, irrespective of the choice of $k \in \mathbb{N}$ and the tree T , that expected values are finite. As such, we obtain an upper bound $D' \mu_n^k$ for some constant D' . Applying Theorem 3.7 then yields the desired result. \square

6.2.2 Proof of Theorem 3.21

Recall the quantities G_1 and G_2 from (3.10). We first have the following result, which deals with the case that $f(k, W)$ is a regularly varying random variable.

Proposition 6.5. *Assume that $f(k, W)$ satisfies (3.6) and (3.7) for some slowly-varying functions $\bar{\ell}_k, \underline{\ell}_k$, respectively, and an exponent $z > 0$, for any $k \in \mathbb{N}_0$ (where z is independent of k), and assume that for each $k \in \mathbb{N}_0$ there exists $i_k \in \{0, \dots, k\}$ such that $\inf_{i \leq k} f(i, w) = f(i_k, w)$ for all $w \in S$. If there exists $\eta_0 > 0$ such that*

$$\sum_{n=1}^{\infty} \mu_n^{k - (G_1(T, z) - zG_2(T, z)) - \eta} < \infty \quad \text{for all } \eta \in (0, \eta_0), \quad (6.20)$$

then the tree \mathcal{T}_∞ contains T as a sub-tree finitely often, almost surely. If there exists $\eta_0 > 0$ such that

$$\sum_{n=1}^{\infty} \mu_n^{k - (G_1(T, z) - zG_2(T, z)) + \eta} = \infty, \quad \text{for all } \eta \in (0, \eta_0), \quad (6.21)$$

then the tree \mathcal{T}_∞ contains T infinitely often, almost surely.

We finally state the following technical lemma, whose proof we defer to Section C in the Appendix.

Lemma 6.6. *Let r be a regularly-varying function with exponent $\rho > 0$ and let W be a random variable with a regularly-varying tail distribution with exponent $-(\zeta - 1) < 0$. Then, $r(W)$ is has a regularly-varying distribution with exponent $-(\zeta - 1)/\rho$.*

With this proposition and lemma at hand, we can prove the results in Theorem 3.21.

Proof of Theorem 3.21. We fix $k \in \mathbb{N}$ and let T be a sibling-closed tree of size $k + 1$. In all of the cases we discuss, we use the approximations of μ_n for the two main choices of the degree function s (super-linear and barely super-linear), as in Lemma 6.2, and distinguish the proof between these two main cases.

Super-linear case. For the super-linear case with additive weights, we have that $f(k, x)$ is regularly varying in x with exponent 1; for the super-linear case with mixed weights, $f(k, x)$ is regularly varying in x with exponent $\gamma \vee 1$, for any $k \in \mathbb{N}_0$. Applying (the proof of) Lemma 6.6 and assuming that the vertex-weight distribution satisfies (3.6) and (3.7) with exponent $\alpha - 1$, we obtain that $f(k, W)$ satisfies these tail distribution inequalities with exponent $z = \alpha - 1$ in the additive weight case and $z = (\alpha - 1)/(\gamma \vee 1)$ in the mixed weight case for any $k \in \mathbb{N}_0$ (where the slowly-varying functions may depend on k). In both cases, for any $k \in \mathbb{N}_0$ and $w \geq 0$,

$$\inf_{i \leq k} f(i, w) = g(w) \left(\inf_{i \leq k} s(i) \right) + h(w) = g(w) s(i_k) + h(w),$$

where $i_k = \arg \min_{i \leq k} s(i)$. As i_k does not depend on $w \geq 0$, the condition on $\inf_{i \leq k} f(i, w)$ in Proposition 6.5 is satisfied. Moreover, in both cases, μ_n is regularly varying with exponent $-(p - 1)$. Hence, the inequalities in (3.11) and (3.12) follow by applying (6.20) and (6.21), respectively.

Barely super-linear case. The proof follows directly from the fact that μ_n^a is not summable for any $a \geq 0$, as follows from Lemma 6.2, and that

$$k - (G_1(T, z) - zG_2(T, z)) = \sum_{v \in T} \deg^+(v, T) - \sum_{v \in T} (\deg^+(v, T) - z) \mathbf{1}_{\deg^+(v, T) > z} \geq 0,$$

which concludes the proof. \square

It remains to prove Proposition 6.5.

Proof of Proposition 6.5. We first recall that we assume that the random variables $(f(k, W))_{k \in \mathbb{N}_0}$ satisfy (3.6) and (3.7) for some slowly-varying functions $\bar{\ell}_k, \underline{\ell}_k$, respectively, and some exponent $z > 0$ independent of k . By the definition of f^+ and f^- in (6.16) and the assumption that for each $k \in \mathbb{N}_0$ there exists $i_k \in \{0, \dots, k\}$ such that $f^-(k, w) := \inf_{i \leq k} f(i, w) = f(i_k, w)$ for all $w \in S$, we have

$$\mathbb{P}(f(i_k, W) \geq x) = \mathbb{P}(f^-(k, W) \geq x) \leq \mathbb{P}(f^+(k, W) \geq x) \leq \sum_{i=0}^k \mathbb{P}(f(i, W) \geq x). \quad (6.22)$$

By the assumptions on the tail distribution of $f(i, W)$, and by Lemma 6.1, we thus find that we obtain an upper and lower bound for the tail probabilities of $f^-(k, W)$ and $f^+(k, W)$ that are regularly-varying with exponent $-z$. We now apply [54, Proposition 1.3.2] to $f^+(k, W)$, using the regularly-varying upper bound on its tail distribution (with $X = f^+(k, W)$, $\beta = \deg^+(v, T)$, and regularly-varying exponent $-z$) to determine that, for some slowly-varying function L_1 ,

$$\begin{aligned} & \mathbb{E} \left[\mathbf{1}_{f^+(k, W) < \mu_n^{-1}} f^+(k, W)^{\deg^+(v, T)} \right] \\ & \leq \begin{cases} \mathbb{E} \left[f^+(k, W)^{\deg^+(v, T)} \right] (1 - o(1)) & \text{if } \deg^+(v, T) < z, \\ L_1(\mu_n^{-1}) & \text{if } \deg^+(v, T) = z, \\ \frac{z + o(1)}{\deg^+(v, T) - z} L_1(\mu_n^{-1}) \mu_n^{z - \deg^+(v, T)} & \text{if } \deg^+(v, T) > z. \end{cases} \end{aligned} \quad (6.23)$$

We intend to use this in each of the product terms in the upper bound in (6.19) by omitting the term $f^+(k, W)$ in the denominator (which yields a further upper bound) and taking out a

factor $\mu_n^{\deg^+(v_j, T)}$ for each $j \leq k - 1$. We thus have a sum of two quantities in each product term, namely

$$\mu_n^{\deg^+(v_j, T)} \mathbb{E} \left[\mathbf{1}_{f^+(k, W) < \mu_n^{-1}} f^+(k, W)^{\deg^+(v_j, T)} \right] + \mathbb{P} (f^+(k, W) \geq \mu_n^{-1}). \quad (6.24)$$

When $\deg^+(v_j, T) < z$, the case distinction in (6.23) tells us that the first term is regularly varying in μ_n with exponent $\deg^+(v_j, T)$. Since, by (6.22) $f^+(k, W)$ satisfies (3.6) (for some slowly-varying function $\bar{\ell}_k$, constant $\bar{x}_k > 0$ and exponent $z > 0$), the second term is regularly varying in μ_n with exponent $z > \deg^+(v_j, T)$. Since μ_n tends to zero, it follows that, in this case, the first term in (6.24) is dominant.

When $\deg^+(v_j, T) = z$, the case distinction tells us that the first term is regularly varying in μ_n with exponent $\deg^+(v_j, T) = z$, and by again applying (3.6), so is $\mathbb{P} (f^+(k, W) \geq \mu_n^{-1})$. Finally, when $\deg^+(v_j, T) > z$, it is readily verified that both terms are equal, up to a multiplicative constant. Hence, we can bound the entire term in (6.24) from above by

$$L_2(\mu_n^{-1}) \mu_n^{\min\{z, \deg^+(v_j, T)\}}$$

for some slowly-varying function L_2 . Using this in each of the product terms in (6.19), we obtain the upper bound

$$\begin{aligned} & D \prod_{j=0}^{k-1} \left(\mathbb{E} \left[\mathbf{1}_{f^+(k, W) < \mu_n^{-1}} \left(\frac{f^+(k, W)}{f^+(k, W) + \mu_n^{-1}} \right)^{\deg^+(v_j, T)} \right] + \mathbb{P} (f^+(k, W) \geq \mu_n^{-1}) \right) \\ & \leq D \mu_n^k \prod_{j=0}^{k-1} (L_2(\mu_n^{-1}) \mu_n^{\min\{z - \deg^+(v_j, T), 0\}}), \end{aligned}$$

where we have taken the terms $\mu_n^{\deg^+(v_j, T)}$ outside of the product to obtain a term μ_n^k . We now recall that $G_2(T, z)$ counts the number of vertices v_j in T such that $\deg^+(v_j, T) > z$, and $G_1(T, z)$ counts the sum of all degrees of such vertices v_j (see (3.10)). As a result, for some slowly-varying function L_3 , using that products of slowly-varying functions are still slowly varying by Lemma 6.1, we obtain the upper bound

$$L_3(\mu_n^{-1}) \mu_n^{k - G_1(T, z)} \mathbb{P} (f^+(k, W) \geq \mu_n^{-1})^{G_2(T, z)},$$

Now, we use that $f^+(k, W)$ satisfies (3.6) and apply 6.1 to arrive at

$$L_4(\mu_n^{-1}) \mu_n^{k - (G_1(T, z) - z G_2(T, z))}, \quad (6.25)$$

for some slowly-varying function L_4 . Since we can bound $L_4(\mu_n^{-1}) \leq \mu_n^{-\eta}$ for any $\eta > 0$ and n sufficiently large, the desired result in (6.20) is obtained by applying Theorem 3.7.

To prove (6.21), we start from the lower bound in (6.17). Noting that the integrand is bounded from above by one, we have

$$\begin{aligned} & \mathbb{E} \left[\mathbf{1}_{f^+(k, W) < \mu_n^{-1}} \prod_{j=0}^{k-1} \prod_{\ell=0}^{\deg^+(v_j, T) - 1} \frac{f(\ell, W_{v_j})}{f(\ell, W_{v_j}) + k \mu_n^{-1}} \right] \\ & \geq \prod_{j=0}^{k-1} \mathbb{E} \left[\mathbf{1}_{f^-(k, W) < \mu_n^{-1}} \left(\frac{f^-(k, W)}{f^-(k, W) + k \mu_n^{-1}} \right)^{\deg^+(v_j, T)} \right] \\ & \quad - \mathbb{P} (f^-(k, W) < \mu_n^{-1}, f^+(k, W) \geq \mu_n^{-1})^k. \end{aligned} \quad (6.26)$$

We first tend to the product on the right-hand side. By bounding the term $f^-(k, W)$ in the denominator from above by μ_n^{-1} , we bound it from below by

$$(k+1)^{-k} \mu_n^k \prod_{j=0}^k \mathbb{E} \left[\mathbf{1}_{f^-(k, W) < \mu_n^{-1}} f^-(k, W)^{\deg^+(v_j, T)} \right]$$

We now use the regularly-varying lower bound on the tail distribution of $f^-(k, W)$, as derived from (6.22). Again applying [54, Proposition 1.3.2], we obtain a lower bound analogous to (6.23) (and by possibly changing the slowly-varying function). The steps in (6.24) through (6.25) are then identical, so that we obtain, for some slowly-varying function L_5 , the lower bound

$$(k+1)^{-k} L_5(\mu_n^{-1}) \mu_n^{k - (G_1(T, z) - zG_2(T, z))}.$$

We now substitute this in (6.26) with $C = (k+1)^{-k}$ and bound the probability from above by omitting the first event, to arrive at

$$\prod_{j=0}^{k-1} \mathbb{E} \left[\mathbf{1}_{f^+(k, W) < \mu_n^{-1}} \left(\frac{f^-(k, W)}{f^-(k, W) + k\mu_n^{-1}} \right)^{\deg^+(v_j, T)} \right] \geq CL_5(\mu_n^{-1}) \mu_n^{k - (G_1(T, z) - zG_2(T, z))} - \mathbb{P}(f^+(k, W) \geq \mu_n^{-1})^k. \quad (6.27)$$

As follows from (6.22), the probability $\mathbb{P}(f^+(k, W) \geq \mu_n^{-1})^k$ is bounded from above by a regularly-varying term (in μ_n) with exponent kz . By Lemma 6.1, the right-hand side is regularly-varying with exponent $\min\{k - (G_1(T, z) - zG_2(T, z)), kz\}$. In particular, it is positive for all n large if the first argument attains the minimum. To determine that this is the case, we observe, with a similar argument as in from (3.13), that the first argument of the minimum is at most k . Hence, when $z \geq 1$, it is immediate. When $z < 1$, we instead use that $G_1(T, z) = k$ (since now G_1 simply sums all degrees) and $G_2(T, z) = k + 1 - N$, where $N \geq 1$ is the number of leaves (i.e. nodes with out-degree zero) in T . We thus have

$$k(G_1(T, z) - zG_2(T, z)) \leq kz \iff z(k + 1 - N) \leq kz,$$

which holds since $N \geq 1$. As a result, we can bound the right-hand side of (6.27) from below, for some slowly-varying function L_6 , by

$$L_6(\mu_n^{-1}) \mu_n^{k - (G_1(T, z) - (\alpha-1)G_2(T, z)/z)}.$$

We finally use Lemma 6.1 to bound $L_6(\mu_n^{-1}) \geq \mu_n^\eta$ for any $\eta > 0$ and n sufficiently large and apply Theorem 3.7 to finally arrive at (6.21), which concludes the proof. \square

A Additional results for other fitness functions

In Theorem 3.16 we discussed three particular choices of the fitness function: a super-linear degree function s with either μ mixed or additive weight functions, and a log-stretched super-linear degree function s with mixed weight functions. For these choices, we were able to prove a complete phase diagram for the emergence of a unique vertex with infinite degree or a unique infinite path, and for sub-tree counts in the former case.

In this section, we discuss three additional choices of the fitness function: barely super-linear log-stretched degree function s with *additive* weight functions, and *poly-log* degree function

s with either *mixed* or *additive* weight functions. We present these cases here, since we are unable to prove a complete phase diagram; for each case there is a gap in the parameter choices where we cannot prove the emergence of a unique vertex with infinite degree nor a unique infinite path.

We summarise these choices in the following assumptions.

Assumption A.1. Let $f(i, w) = g(w)s(i) + h(w)$, $i \in \mathbb{N}_0$, $w \geq 0$ satisfy (w^*) and one of the following three cases.

- **Barely super-linear log-stretched with additive weights:** s is as in (3.5), $g \equiv 1$, and $h(x) = \ell(x)x$ for some slowly-varying function $\ell : [0, \infty) \rightarrow [0, \infty)$, such that either

$$\exists \underline{a} \in [0, 1) : \lim_{x \rightarrow \infty} \frac{\log \log(1/\ell(x))}{\log \log x} = \underline{a}, \quad (\text{A.1})$$

or

$$\exists \bar{a} \in [0, 1) : \lim_{x \rightarrow \infty} \frac{\log \log(\ell(x))}{\log \log x} = \bar{a}. \quad (\text{A.2})$$

- **Barely super-linear poly-log:** For some $\sigma > 1$,

$$s(i) = (i + 2)(\log(i + 2))^\sigma, \quad i \in \mathbb{N}_0,$$

and g, h are as in the additive or mixed weight case of Assumption 3.10.

We then also introduce the following additional assumptions for the vertex-weight distribution.

Assumption A.2. The vertex-weights $(W_i)_{i \in \mathbb{N}}$ are i.i.d. and their tail distribution satisfies one of the following cases.

- **Power law with lower-order term.** Let $\tau \in (0, 1)$. We have the following two conditions.

- 1 There exist $\bar{c}, \bar{x} > 0$ and $\tau \in (0, 1)$ such that

$$\mathbb{P}(W \geq x) \leq x^{-1} e^{-\bar{c}(\log x)^\tau}, \quad x \geq \bar{x}. \quad (\text{A.3})$$

- 2 There exist $\underline{c}, \underline{x} > 0$ and $\tau' \in (0, 1)$ such that

$$\mathbb{P}(W \geq x) \geq x^{-1} e^{\underline{c}(\log x)^{\tau'}}, \quad x \geq \underline{x}. \quad (\text{A.4})$$

- **Stretched exponential.** Let $\kappa > 0$. We have the following two conditions.

- 1 There exist $\bar{c}, \bar{x} > 0$ such that

$$\mathbb{P}(W \geq x) \leq e^{-\bar{c}x^\kappa}, \quad x \geq \bar{x}. \quad (\text{A.5})$$

- 2 There exists $\underline{c}, \underline{x} > 0$ such that

$$\mathbb{P}(W \geq x) \geq e^{-\underline{c}x^\kappa}, \quad x \geq \underline{x}. \quad (\text{A.6})$$

We then have the following theorem, which is an equivalent result to Theorem 3.16.

Theorem A.3. *Let $(\mathcal{T}_i)_{i \in \mathbb{N}}$ be a (W, f) -recursive tree with fitness, where the fitness function f and degree function s satisfy one of the cases in Assumptions 3.10 and 3.13, respectively, and the vertex-weight distribution satisfies one of the cases in Assumption 3.15. The tree \mathcal{T}_∞ either contains a unique vertex with infinite degree and no infinite path almost surely, or contains a unique infinite path and no vertex with infinite degree almost surely, when the following conditions are met, based on the fitness function, degree function, and vertex-weight assumptions:*

| Weight | Degree | Star | Path |
|----------|---------------|--|--|
| Additive | Log-Stretched | (A.1),(A.3) & $(\tau \vee \beta) > (1 - \beta) \vee \underline{a}$ | (A.2),(A.4) & $\tau' > \beta \vee (1 - \beta), (\tau' \vee \beta) > \bar{a}$ |
| Mixed | Poly-log | (A.5) & $(\sigma - 1)\kappa > 1 + \kappa$ | (A.6) & $(\sigma - 1)\kappa < 1$ |
| Additive | Poly-log | (3.8) & $(\sigma - 1)(1 - 1/\nu) > 1$ | (3.7) & $\alpha < 2$ |

Table 4: The first column represents the form of the fitness function, as in Assumption 3.10; the second represents the form of the degree function, as in Assumption A.1. The third and fourth column, respectively, list the required assumptions on the vertex-weight distribution, as in Assumptions 3.15 and A.2, together with the choices of the parameters that lead to either a unique node of infinite degree or a unique infinite path.

The proof of Theorem A.3 follows from two lemmas, equivalent to Lemmas 6.3 and 6.4, which we state now.

Lemma A.4. *Equation (3.2) in Item 1 of Theorem 3.4 is satisfied when the following conditions are met, based on the assumptions for the fitness type, degree function s , and vertex-weight distribution:*

| Weight | Degree | Star |
|----------|---------------|--|
| Additive | Log-Stretched | (A.1),(A.3) & $(\tau \vee \beta) > (1 - \beta) \vee \underline{a}$ |
| Mixed | Poly-log | (A.5) & $(\sigma - 1)\kappa > 1 + \kappa$ |
| Additive | Poly-log | (3.8) & $(\sigma - 1)(1 - 1/\nu) > 1$ |

Table 5: The first column represents the form of the fitness function, as in Assumption 3.10; the second represents the form of the degree function, as in Assumption A.1. The third column lists the required assumptions on the vertex-weight distribution, as in Assumption A.2, together with the choices of the parameters that lead to a unique node of infinite degree.

Lemma A.5. *The condition in (2.9) is satisfied when the following conditions are met, based on the assumptions for the fitness type, degree function s , and vertex-weight distribution:*

| Weight | Degree | Path |
|----------|---------------|--|
| Additive | Log-Stretched | (A.2),(A.4) & $\tau' > \beta \vee (1 - \beta), (\tau' \vee \beta) > \bar{a}$ |
| Mixed | Poly-log | (A.6) & $(\sigma - 1)\kappa < 1$ |
| Additive | Poly-log | (3.7) & $\alpha < 2$ |

Table 6: The first column represents the form of the fitness function, as in Assumption 3.10; the second represents the form of the degree function, as in Assumption A.1. The third column lists the required assumptions on the vertex-weight distribution, as in Assumptions 3.15 and A.2, together with the choices of the parameters that lead to a unique infinite path.

It is clear that Theorem A.3 follows from Lemmas A.4 and A.5. Before we prove these two lemmas, we state the following result, which is an analogue of Lemma 6.2.

Lemma A.6. *Let f satisfy the [barely super-linear poly-log](#) case, as in [Assumption A.1](#). Then, for any $w \geq 0$,*

$$\mu_n^w = \frac{1 + o(1)}{g(w)(\sigma - 1)} (\log n)^{-(\sigma-1)}.$$

Proof. We use [Lemma 6.2](#), which tells us that $\mu_n^w = (1 + o(1))g(w)^{-1}L(n)$, where

$$L(n) := \int_{n+2}^{\infty} x^{-1}(\log x)^{-\sigma} dx.$$

Using a variable substitution $y = \log x$ and determining the integral yields the desired result. \square

The proofs of [Lemmas A.4](#) and [A.5](#) follow the same approach as those of [Lemmas 6.3](#) and [6.4](#).

Proof of [Lemmas A.4](#) and [A.5](#), s [log-stretched case](#), additive weights. To start, we prove the claim in [Lemma A.4](#). We assume that $\beta\tau > 1$ and recall that

$$f(i, w) = (i + 1)e^{(\log(i+1))^\beta} + h(w),$$

where h is a regularly-varying function with exponent 1. Fix $\varepsilon > 0$ sufficiently small so that $\beta\tau > 1 + \varepsilon$. We apply [\(6.2\)](#) with

$$k_n := n \exp(-(1 - \varepsilon)\bar{c}(\log n)^\tau), \quad \text{and} \quad \mu_n = \frac{1 + o(1)}{\beta} (\log n)^{1-\beta} \exp(-(\log n)^\beta), \quad (\text{A.7})$$

where the latter follows from [Lemma 6.2](#). Since $\tau \in (0, 1)$ and by [\(A.3\)](#), we have

$$\begin{aligned} \mathbb{P}(W > k_n) &\leq k_n^{-1} e^{-\bar{c}(\log k_n)^\tau} \\ &= \frac{1}{n} \exp\left(\left((1 - \varepsilon)\bar{c}(\log n)^\tau - \bar{c}(\log n - (1 - \varepsilon)\bar{c}(\log n)^\tau)^\tau\right)\right) \\ &= \frac{1}{n} \exp\left(-\varepsilon\bar{c}(\log n)^\tau(1 + o(1))\right), \end{aligned} \quad (\text{A.8})$$

for n sufficiently large, so that the second term on the right-hand side of [\(6.2\)](#) is summable. Next, we define

$$I_n := h(k_n) \exp\left(-(\log h(k_n))^\beta + \beta(\log h(k_n))^{2\beta-1}\right).$$

By Taylor's theorem, for $|x| < 1$, $(1 + x)^\beta = 1 + \beta x + \frac{\beta(\beta-1)}{2}x^2 + o(x^3)$. Factoring out the term $\log h(k_n)$ we can thus write

$$\begin{aligned} (\log I_n)^\beta &= \log h(k_n)^\beta \left(1 - (\log h(k_n))^{\beta-1} + \beta(\log h(k_n))^{2\beta-2}\right)^\beta \\ &= (\log h(k_n))^\beta - \beta(\log h(k_n))^{2\beta-1} + \frac{3}{2}\beta\left(\beta - \frac{1}{3}\right)(\log h(k_n))^{3\beta-2} \\ &\quad + o((\log h(k_n))^{3\beta-2}). \end{aligned} \quad (\text{A.9})$$

It thus follows that for $j \geq I_n$,

$$\begin{aligned} (j + 1) \exp\left((\log(j + 1))^\beta\right) &\geq I_n \exp\left((\log I_n)^\beta\right) \\ &= h(k_n) \exp\left(\frac{3\beta}{2}\left(\beta - \frac{1}{3}\right)(\log h(k_n))^{3\beta-2} + o((\log h(k_n))^{3\beta-2})\right). \end{aligned}$$

We claim that the exponential term is at least $1/2$ for n large. Indeed, for $\beta > 2/3$ the exponential term diverges with n , whilst for $\beta \in (0, 2/3]$ it converges to a constant that is at least 1. So, for all $j \geq I_n$ and n sufficiently large,

$$h(k_n) \leq 2(j+1) \exp((\log(j+1))^\beta).$$

Moreover, as μ_n is slowly varying and k_n and h are both regularly varying with exponents 1, we have $\mu_n^{-1} = o(h(k_n))$. We thus obtain the lower bound

$$\begin{aligned} \sum_{j=0}^{\infty} \frac{1}{(j+1) \exp(c(\log(j+1))^\beta) + h(k_n) + \mu_n^{-1}} &\geq \sum_{j=I_n}^{\infty} \frac{1}{5(j+1) \exp((\log(j+1))^\beta)} \\ &= \frac{\mu_{I_n}(1+o(1))}{5}, \end{aligned}$$

where we recall the definition of μ_n from (3.1) (with $w^* = 0$) and use (the proof of) Lemma 6.2 in the final step. Using this in the first term on the right-hand side of (6.2), we obtain the upper bound

$$\exp\left(-\frac{c+o(1)}{5}\mu_{I_n}\mu_n^{-1}\right). \quad (\text{A.10})$$

With I_n and μ_n as in (A.7) and using (A.9), we find

$$\begin{aligned} \mu_{I_n} &= \frac{1+o(1)}{\beta} (\log I_n)^{1-\beta} \exp(-(\log I_n)^\beta) \\ &= \frac{1+o(1)}{\beta} (\log n)^{1-\beta} \exp(-(\log h(k_n))^\beta + \beta(\log h(k_n))^{2\beta-1}(1+o(1))). \end{aligned} \quad (\text{A.11})$$

We now use that $h(x) = \ell(x)x$ for some slowly-varying function ℓ which satisfies (A.1) for some $a \in [0, \tau \vee \beta)$. We can then write, using that $h(x) = x^{1+o(1)}$ and $k_n = n^{1+o(1)}$, and for $|x| < 1$, that $(1+x)^\beta = 1 + \beta x + o(x^2)$,

$$\begin{aligned} &-(\log h(k_n))^\beta + \beta(\log h(k_n))^{2\beta-1}(1+o(1)) \\ &= -(\log k_n)^\beta (1 + \log(\ell(k_n))(\log k_n)^{-1})^\beta + \beta(\log n)^{2\beta-1}(1+o(1)) \\ &= -(\log k_n)^\beta - \beta(\log k_n)^{\beta-1} \log(\ell(k_n))(1+o(1)) + \beta(\log n)^{2\beta-1}(1+o(1)) \end{aligned} \quad (\text{A.12})$$

Now, again using the expression for $(1+x)^\beta$, observe that

$$(\log k_n)^\beta = (\log n)^\beta + (1-\varepsilon)\bar{c}\beta(\log n)^{\beta+\tau-1}(1+o(1)).$$

Hence we may write (A.12) as

$$-(\log n)^\beta + \left[(1-\varepsilon)\bar{c}\beta(\log n)^{\beta+\tau-1} + \beta(\log n)^{2\beta-1} - \beta(\log k_n)^{\beta-1} \log(\ell(k_n))\right](1+o(1)). \quad (\text{A.13})$$

Recall that we assume that ℓ is such that $\lim_{x \rightarrow \infty} \log \log(\ell(x)) / \log \log x = \bar{a}$ for some $\bar{a} \in [0, \tau \vee \beta)$ (as we assume that $\tau \vee \beta > \bar{a}$), that is, when $\log(\ell(x)) = (\log x)^{\bar{a}+o(1)}$, we can write

$$(\log k_n)^{\beta-1} \log(\ell(k_n)) = (\log k_n)^{\beta+\bar{a}-1+o(1)}.$$

Since $\bar{a} < \tau \vee \beta$, this term is negligible compared to the other two terms in the square brackets in (A.13), so that it can be included in the $o(1)$ term. We thus obtain, combining (A.11)

with (A.13),

$$\begin{aligned}\mu_{I_n} &= \\ &= \frac{1 + o(1)}{\beta} (\log n)^{1-\beta} \exp\left(-(\log n)^\beta + [(1-\varepsilon)\bar{c}\beta(\log n)^{\beta+\tau-1} + \beta(\log n)^{2\beta-1}](1+o(1))\right) \\ &= \mu_n \exp\left((C' + o(1))(\log n)^{\beta+(\tau\vee\beta)-1}\right).\end{aligned}$$

We can thus, finally, bound the first term on the right-hand side of (6.2) from above by substituting this in (A.10), which yields

$$\exp\left(-\frac{c + o(1)}{5} \exp\left((C' + o(1))(\log n)^{\beta+(\tau\vee\beta)-1}\right)\right),$$

which is summable since $\tau \vee \beta > 1 - \beta$ (and using an argument similar to that in (A.8)), as desired.

We then prove the claim in Lemma A.5. We set $k_n := n \exp(\underline{c}(\log n)^{\tau'})$. It follows from (A.4) that

$$\mathbb{P}(W \geq k_n) \geq n^{-1} \exp(-\underline{c}(\log n)^{\tau'}) \exp(\underline{c}(\log n + \underline{c}(\log n)^{\tau'})^{\tau'}) \geq n^{-1},$$

which is not summable. Then, we define

$$I_n := h(k_n) \exp(-(\log h(k_n))^\beta - (\log h(k_n))^{2\beta-1}), \quad (\text{A.14})$$

and write

$$\begin{aligned}\sum_{i=0}^{\infty} \frac{1}{f(i, k_n)} &= \sum_{i=0}^{\infty} \frac{1}{(i+1) \exp((\log(i+1))^\beta) + h(k_n)} \\ &\leq \sum_{i=0}^{I_n} \frac{1}{h(k_n)} + \sum_{i=I_n+1}^{\infty} \frac{1}{(i+1) \exp((\log(i+1))^\beta)}.\end{aligned}$$

We recall the definition of μ_n from (3.1) (with $w^* = 0$) and use (the proof of) Lemma 6.2 to deduce that the above equals

$$\exp(-(\log h(k_n))^\beta - (\log h(k_n))^{2\beta-1}) + (1 + o(1))\mu_{I_n}. \quad (\text{A.15})$$

We now need only show that this expression is $o(\mu_n^w)$, which clearly implies (6.4).

Let us start with the first term. By a similar sequence of computations as in (A.12), (A.13), and using the fact that, for $|x| < 1$ we have $(1+x)^\beta = 1 + \beta x + o(x^2)$, we obtain

$$\begin{aligned}&\exp(-(\log h(k_n))^\beta - (\log h(k_n))^{2\beta-1}) \\ &= \exp(-(\log k_n)^\beta - \beta(\log k_n)^{\beta-1} \log(\ell(k_n))(1+o(1)) - (\log n)^{2\beta-1}(1+o(1))) \\ &= \exp\left(-(\log n)^\beta - [\underline{c}\beta(\log n)^{\beta+\tau'-1} + (\log n)^{2\beta-1} + \beta(\log n)^{\beta-1} \log(\ell(k_n))](1+o(1))\right).\end{aligned} \quad (\text{A.16})$$

We recall that $h(x) = \ell(x)x$, where ℓ is slowly varying, such that

$$\lim_{x \rightarrow \infty} \log \log(1/\ell(x)) / \log \log x = \underline{a},$$

for some $\underline{a} \in [0, \tau' \vee \beta)$ (as by our assumption). That is, $\log(\ell(x)) = -(\log x)^{\underline{a}+o(1)}$. We thus have

$$(\log k_n)^{\beta-1} \log(\ell(k_n)) = -(\log k_n)^{\beta+\underline{a}-1+o(1)}. \quad (\text{A.17})$$

Since $\underline{a} < \tau' \vee \beta$, it follows that this term can be included in the $o(1)$ term in (A.16). We thus arrive at

$$\begin{aligned} \exp(-(\log h(k_n))^\beta) &= \exp\left(-(\log n)^\beta - [\underline{c}\beta(\log n)^{\beta+\tau'-1} + (\log n)^{2\beta-1}](1 + o(1))\right) \\ &= \mu_n^\beta (\log n)^{\beta-1} \exp\left(-[\underline{c}\beta(\log n)^{\beta+\tau'-1} + (\log n)^{2\beta-1}](1 + o(1))\right). \end{aligned}$$

Since we assume that $\tau' > \beta \vee (1 - \beta)$, we directly obtain that this is $o(\mu_n^w)$.

Recalling (A.15), it thus remains to prove that $\mu_{I_n} = o(\mu_n^w)$. By (A.16), (A.14), and (A.7), the desired result follows by showing that $\exp(-(\log I_n)^\beta) = o(\exp(-(\log n)^\beta))$ or, equivalently, $(\log I_n)^\beta - (\log n)^\beta$ diverges with n . First, we note that $h(x) = x^{1+o(1)}$ (as h is regularly varying with exponent 1) and $k_n = n^{1+o(1)}$. Again, using the approximation to $(1+x)^\beta$ in a similar manner as in (A.12),

$$\begin{aligned} (\log I_n)^\beta &= (\log h(k_n) - (\log h(k_n))^\beta - (\log h(k_n))^{2\beta-1})^\beta \\ &= (\log h(k_n))^\beta - \beta(\log h(k_n))^{2\beta-1}(1 + o(1)) \\ &= (\log k_n)^\beta + [\beta(\log k_n)^{\beta-1} \log(\ell(k_n)) - \beta(\log n)^{2\beta-1}](1 + o(1)) \\ &= (\log n)^\beta + [\underline{c}\beta(\log n)^{\beta+\tau'-1} + \beta(\log k_n)^{\beta-1} \log(\ell(k_n)) - \beta(\log n)^{2\beta-1}](1 + o(1)). \end{aligned}$$

By the same argument that leads to (A.17), we argue that we can include the second term in the square brackets within the $o(1)$ term as it is of lower order compared to either $(\log n)^{\beta+\tau'-1}$ or $(\log n)^{2\beta-1}$. The desired result thus follows since $\tau' > \beta \vee (1 - \beta)$. \square

Proof of Lemmas A.4 and A.5, s poly-log case, mixed weights. We first prove the claim in Lemma A.4. We assume that $(\sigma - 1)\kappa > 1 + \kappa$ and recall that

$$f(i, w) = g(w)(i + 2)(\log(i + 2))^\sigma + h(w),$$

where g and h are regularly-varying functions with exponents 1 and $\gamma \geq 0$, respectively. Fix $\varepsilon > 0$ sufficiently small so that $(\sigma - 1)\kappa > 1 + \varepsilon$. We apply (6.2) with

$$k_n := (\log n)^{(1+\varepsilon)/\kappa}, \quad \text{and} \quad \mu_n = \frac{1 + o(1)}{g(0)(\sigma - 1)} (\log n)^{-(\sigma-1)}, \quad (\text{A.18})$$

where the latter follows from Lemma 6.2. Now, using (A.5), we obtain

$$\mathbb{P}(W \geq k_n) \leq e^{-\bar{c}k_n^\kappa} = e^{-\bar{c}(\log n)^{1+\varepsilon}} \leq n^{-(1+\varepsilon)},$$

for n sufficiently large, so that the second term on the right-hand side of (6.2) is summable. Then, we define

$$I_n := (h(k_n) + \mu_n^{-1})/g(k_n). \quad (\text{A.19})$$

We write $g(x) = \ell_1(x)x$ and $h(x) = \ell_2(x)x^\gamma$ for some slowly-varying functions ℓ_1, ℓ_2 , and note that $g(x) = x^{1+o(1)}$ and $h(x) = x^{\gamma+o(1)}$. With k_n and μ_n as in (A.18), we thus have

$$I_n = (\log n)^{\max\{(\sigma-1)\kappa, (1+\varepsilon)\gamma\}/\kappa - (1+\varepsilon)/\kappa + o(1)}.$$

We note that I_n diverges with n since $(\sigma - 1)\kappa > 1 + \varepsilon$. It then directly follows for all n sufficiently large and all $j \geq I_n$, that

$$(j + 2)(\log(j + 2))^\sigma \geq j \geq I_n = (h(k_n) + \mu_n^{-1})/g(k_n).$$

As a result, for all n large we obtain the lower bound

$$\begin{aligned} \sum_{j=0}^{\infty} \frac{1}{g(k_n)(j+2)(\log(j+2))^\sigma + h(k_n) + \mu_n^{-1}} &\geq \frac{1}{g(k_n)} \sum_{j=I_n}^{\infty} \frac{1}{3(j+2)(\log(j+2))^\sigma} \\ &= \frac{\mu_{I_n}(1+o(1))}{3g(k_n)/g(0)}, \end{aligned}$$

where we recall the definition of μ_n from (3.1) (with $w^* = 0$) and use (the proof of) Lemma 6.2 in the final step. Substituting this bound into the first term on the right-hand side of (6.2), and using (A.18), we obtain the upper bound

$$\exp\left(-\frac{cg(0) + o(1)}{3} \frac{\mu_{I_n} \mu_n^{-1}}{g(k_n)}\right) = \exp\left(-(\log \log n)^{-(\sigma-1)} (\log n)^{(\sigma-1)-(1+\varepsilon)/\kappa+o(1)}\right),$$

(where we incorporate the constants into the $o(1)$ in the exponent of the $\log n$ term). This is summable when $(\sigma-1) - (1+\varepsilon)/\kappa > 1$ (again using an argument similar to that in (A.8)). Since ε is arbitrary, the desired result follows since $(\sigma-1)\kappa > 1 + \kappa$.

We then prove the claim in Lemma A.5. We fix $\varepsilon > 0$ sufficiently small such that $(\sigma-1)\kappa < 1 - \varepsilon$ and set $k_n := (\log n)^{(1-\varepsilon)/\kappa}$. It follows from (A.6) that

$$\mathbb{P}(W \geq k_n) \geq e^{-c(\log n)^{1-\varepsilon}} \geq n^{-1},$$

which is not summable in n . By a similar computation as in (6.15),

$$\sum_{i=0}^{\infty} \frac{1}{f(i, k_n)} = \frac{1}{g(k_n)} \sum_{i=0}^{\infty} \frac{1}{(i+2)(\log(i+2))^\sigma + h(k_n)/g(k_n)} \leq \frac{C_\sigma}{g(k_n)},$$

for some constant $C_\sigma > 0$. Since g varies regularly with exponent 1, we can write $g(x) = x^{1+o(1)}$, so that

$$\sum_{i=0}^{\infty} \frac{1}{f(i, k_n)} \leq (\log n)^{-(1-\varepsilon)/\kappa+o(1)}.$$

As a result, since $(\sigma-1)\kappa < 1 - \varepsilon$, this sum is $o(\mu_n^w)$. It follows that (6.4) is satisfied, which concludes the proof. \square

Proof of Lemmas A.4 and A.5, s barely super-linear poly-log case, additive weights. We first prove the claim in Lemma A.4. We assume that $(\sigma-1)(1-1/\nu) > 1$, and recall that

$$f(i, w) = (i+2)(\log(i+2))^\sigma + h(w),$$

where h is a regularly-varying function with exponent 1. Fix $\varepsilon > 0$ sufficiently small so that $\sigma(1-1/\nu) > 1 + \varepsilon$. We apply (6.2) with

$$k_n := \exp((\log n)^{(1+\varepsilon)/\nu}), \quad \text{and} \quad \mu_n = \frac{1+o(1)}{\sigma-1} (\log n)^{-(\sigma-1)}, \quad (\text{A.20})$$

where the latter follows from Lemma A.6. Now, using (3.8) we obtain

$$\mathbb{P}(W \geq k_n) \leq e^{-\bar{c}(\log k_n)^\nu} = e^{-\bar{c}(\log n)^{1+\varepsilon}} < n^{-(1+\varepsilon)},$$

for n sufficiently large, so that the second term on the right-hand side of (6.2) is summable. Then, we define

$$I_n := \frac{h(k_n)}{(\log h(k_n))^\sigma}.$$

It then follows that for all $j \geq I_n$ and n sufficiently large,

$$(j + 2)(\log(j + 2))^\sigma \geq I_n(\log I_n)^\sigma = h(k_n)(1 + o(1)) > h(k_n)/2.$$

Moreover, since $\mu_n^{-1} = o(h(k_n))$, irrespective of the values of $\sigma, \nu > 1$ and since $h(x) = x^{1+o(1)}$, we obtain for all n sufficiently large the lower bound

$$\begin{aligned} \sum_{j=0}^{\infty} \frac{1}{(j + 1)(\log(j + 1))^\sigma + h(k_n) + \mu_n^{-1}} &\geq \sum_{j=I_n}^{\infty} \frac{1}{5(j + 1)(\log(j + 1))^\sigma} \quad (\text{A.21}) \\ &= \frac{\mu_{I_n}(1 + o(1))}{5} \geq \frac{\mu_{I_n}}{6}, \end{aligned}$$

where we recall the definition of μ_n from (3.1) (with $w^* = 0$) and use (A.20) in the final step. We again use that $h(x) = x^{1+o(1)}$, which yields

$$\mu_{I_n} = \frac{1 + o(1)}{(\sigma - 1)} (\log I_n)^{-(\sigma-1)} = \frac{1 + o(1)}{(\sigma - 1)} (\log k_n)^{-(\sigma-1)} = \frac{1 + o(1)}{(\sigma - 1)} (\log n)^{-(1+\varepsilon)(\sigma-1)/\nu}.$$

Substituting this into the lower bound in (A.21), we may bound the sum in the exponent in the first term on the right-hand side of (6.2) to obtain the upper bound

$$\exp\left(-\frac{\varepsilon}{6}\mu_{I_n}\mu_n^{-1}\right) \leq \exp\left(-(\log n)^{(\sigma-1)(1-(1+\varepsilon)/\nu)+o(1)}\right).$$

By choosing ε sufficiently small, this upper bound is summable since $(\sigma - 1)(1 - 1/\nu) > 1$, which yields the desired result.

We then prove the claim in Lemma A.5. We fix $\alpha \in (1, 2)$, take $\varepsilon > 0$ sufficiently small such that $1 - \varepsilon > \alpha - 1$, and set $k_n := n^{(1-\varepsilon)/(\alpha-1)}$. We then have $\mathbb{P}(W \geq k_n) \geq \underline{\ell}(k_n)n^{-(1-\varepsilon)}$ by (3.7), which is not summable in n . We define $I_n := k_n/(\log k_n)^\sigma$ and bound

$$\sum_{i=0}^{\infty} \frac{1}{f(i, k_n)} \leq \sum_{i=0}^{I_n-1} \frac{1}{k_n} + \sum_{i=I_n}^{\infty} \frac{1}{(j + 2)(\log(j + 2))^\sigma} = \frac{1}{(\log k_n)^\sigma} + (1 + o(1))\mu_{I_n},$$

where we recall the definition of μ_n from (3.1) (with $w^* = 0$) and use (the proof of) Lemma 6.2. By the choice of k_n and I_n and with μ_n as in (A.20), we have

$$\frac{1}{\mu_n^w} \left(\frac{1}{(\log k_n)^\sigma} + (1 + o(1))\mu_{I_n} \right) = \mathcal{O}\left(\frac{1}{\log n}\right) + (1 + o(1)) \left(\frac{\log I_n}{\log n}\right)^{-(\sigma-1)} \rightarrow \left(\frac{\alpha - 1}{1 - \varepsilon}\right)^{\sigma-1}.$$

As a result, by choosing ε sufficiently small, since $\alpha - 1 < 1$, the right-hand side of the above is strictly smaller than one; it follows that (6.4) is satisfied. \square

B Verifying conditions for other inter-birth time distributions

In this section we consider the other birth-time distributions listed in Remark 3.17, and check the conditions of Assumptions 2.2 and 2.6 (it is clear that all distributions are continuous and thus satisfy Assumption 2.11). We do not repeat all the detailed calculations in the previous subsection. Rather, we show where calculations differ, and where similar or analogous arguments yield the desired results.

B.1 Conditions for an infinite star, Assumption 2.2.

We need to verify (2.5) and (2.8). We start with the former, and split between the different inter-birth distributions listed in Remark 3.17.

Gamma. Let the inter-birth times be gamma random variables, i.e. for $i \in \mathbb{N}, k > 0$, and $w \geq 0$, let $X_w(i) \sim \text{Gamma}(k, kf(i-1, w))$. Recall Y_n from (5.5) (with $w^* = 0$). Then,

$$\mathcal{M}_\lambda(Y_n) = \prod_{\ell=n}^{\infty} \left(\frac{kf(\ell, 0)}{kf(\ell, 0) - \lambda} \right)^k = \left(\prod_{\ell=n}^{\infty} \frac{kf(\ell, 0)}{kf(\ell, 0) - \lambda} \right)^k = \mathcal{M}_\lambda(\tilde{Y}_n)^k,$$

where

$$\tilde{Y}_n \stackrel{d}{=} \sum_{j=n+1}^{\infty} \text{Exp}(kf(j-1, 0)).$$

We can, by possibly changing the constant c so that $c < \min\{1, 1/k\}$, use the same computations as in (5.2) through (5.4) to derive the upper bound $\mathcal{M}_{c\mu_n^{-1}}(Y_n) \leq (1-c)^{-k}$ to conclude that Condition (2.5) of Assumption 2.2 holds.

For the next two examples, we use that

$$\sum_{\ell=n}^{\infty} \frac{1}{f(\ell, 0)^2} \leq \left(\sum_{\ell=n}^{\infty} \frac{1}{f(\ell, 0)} \right)^2 = \mu_n^{-2}.$$

Beta. Let the inter-birth times be distributed as follows. For any $i \in \mathbb{N}$ and $w \geq 0$, let

$$X_w(i) \stackrel{d}{=} \frac{\alpha + \beta}{\alpha} \frac{1}{f(i-1, w)} B_i, \tag{B.1}$$

where $(B_i)_{i \in \mathbb{N}}$ is a sequence of i.i.d. copies of a $\text{Beta}(\alpha, \beta)$ random variables, for some $\alpha \geq 1, \beta \in (0, 1]$. Recall Y_n from (5.5). Then, for $\lambda > 0$,

$$\mathbb{E} [e^{\lambda X_0(j)}] = \sum_{k=0}^{\infty} \prod_{\ell=0}^{k-1} \frac{\alpha + \ell}{\alpha + \beta + \ell} \left(\frac{\alpha + \beta}{\alpha} \frac{\lambda}{f(j-1, 0)} \right)^k \frac{1}{k!},$$

where we set the empty product $\prod_{\ell=0}^{-1}$ equal to one. As the terms in the product are at most one, we directly obtain the upper bound

$$\begin{aligned} \mathcal{M}_{\mu_n^{-1}}(Y_n) &= \prod_{j=n+1}^{\infty} \mathbb{E} [e^{\mu_n^{-1} X_0(j)}] \\ &\leq \prod_{j=n+1}^{\infty} \exp \left(\frac{\alpha + \beta}{\alpha} \frac{\mu_n^{-1}}{f(j-1, 0)} \right) = \exp \left(\frac{\alpha + \beta}{\alpha} \mu_n^{-1} \sum_{j=n}^{\infty} \frac{1}{f(j, 0)} \right). \end{aligned}$$

By the definition of μ_n^{-1} , it follows that this upper bound equals $e^{(\alpha+\beta)/\alpha}$, so that (2.5) is satisfied.

Rayleigh. Let the inter-birth times be Rayleigh random variables, i.e. for any $i \in \mathbb{N}$ and $w \geq 0$, let $X_w(i) \sim \text{Rayleigh}(\sqrt{2/\pi}/f(i-1, w))$. Recall Y_n from (5.5) (with $w^* = 0$). Then,

$$\mathcal{M}_\lambda(Y_n) = \prod_{\ell=n}^{\infty} \left(1 + \frac{\lambda}{f(\ell, 0)} e^{\lambda^2/(\pi f(\ell, 0)^2)} \left(1 + \text{erf}\left(\frac{\lambda}{\sqrt{\pi} f(\ell, 0)}\right) \right) \right),$$

where $\text{erf}(x) := (2/\sqrt{\pi}) \int_0^x e^{-t^2} dt$ denotes the error function. Since the error function is bounded from above by one, we immediately arrive at the upper bound

$$\mathcal{M}_\lambda(Y_n) \leq \exp\left(2\lambda \sum_{\ell=n}^{\infty} \frac{1}{f(\ell, 0)} e^{\lambda^2/(\pi f(\ell, 0)^2)}\right) \leq \exp\left(2\lambda \sum_{\ell=n}^{\infty} \frac{1}{f(\ell, 0)} \exp\left(\frac{\lambda^2}{\pi} \sum_{\ell=n}^{\infty} \frac{1}{f(\ell, 0)^2}\right)\right).$$

Now, with $\lambda = c\mu_n^{-1}$, we finally arrive at the upper bound $\exp(2 \exp(1/\pi))$, as desired.

We now verify Condition 5 of Assumption 2.2, i.e. Equation (2.8), which is summarised in the following lemma.

Lemma B.1. *Assume the inter-birth time distributions satisfy any of the choices in Remark 3.17. With the same conditions for the fitness type, degree function s , and the vertex-weight distribution, as in Lemmas 6.3 and A.4, Condition 5 of Assumption 2.2 is satisfied.*

We split the proof of Lemma B.1 into the different choices for the inter-birth distributions. We observe that, irrespective of the inter-birth distribution, we assume the mean of $X_W(j)$, conditionally on W , is always $1/f(j-1, W)$ for any $j \in \mathbb{N}$, so that the definition (and asymptotic behaviour) of μ_n remains unchanged and only depends on the fitness function f .

Proof of Lemma B.1, Gamma case. For some sequence $(\ell_n)_{n \in \mathbb{N}}$, we have

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{c\mu_n^{-1}}(\mathcal{P}_n; W)] &= \mathbb{E}\left[\left(\prod_{j=0}^{n-1} \frac{k f(j, W)}{k f(j, W) + c\mu_n^{-1}}\right)^k\right] \\ &\leq \exp\left(-c\mu_n^{-1} \sum_{j=0}^{n-1} \frac{1}{k f(j, \ell_n) + c\mu_n^{-1}}\right) + \mathbb{P}(W \geq \ell_n). \end{aligned}$$

This expression is, up to a constant in the exponential term, equivalent to the upper bound in (6.2). We can thus follow the proof of Lemma 6.3 to obtain the desired result for the gamma case in general. \square

Proof of Lemma B.1, Beta case. Recall $X_w(j)$ from (B.1), with $\alpha \geq 1, \beta \in (0, 1]$. We have that

$$\mathbb{E}[\mathcal{L}_\lambda(X_w(j))] = {}_1F_1\left(\alpha; \alpha + \beta; -\frac{\alpha + \beta}{\alpha} \frac{\lambda}{f(j-1, w)}\right),$$

where ${}_1F_1(a; b; z)$ denotes the confluent hypergeometric function, defined as

$${}_1F_1(a; b; z) := \sum_{k=0}^{\infty} \frac{a^{(k)} z^k}{b^{(k)} k!}, \quad a, b > 0, z \in \mathbb{R},$$

where $a^{(k)} := \Gamma(a+k)/\Gamma(a)$, with Γ the gamma function. We now use the Kummer transform ${}_1F_1(a; b; z) = e^z {}_1F_1(b-a; b; -z)$, to obtain

$$\begin{aligned} \mathbb{E}[\mathcal{L}_\lambda(X_w(j))] &= \exp\left(-\frac{\alpha+\beta}{\alpha} \frac{\lambda}{f(j-1, w)}\right) {}_1F_1\left(\beta; \alpha+\beta; \frac{\alpha+\beta}{\alpha} \frac{\lambda}{f(j-1, w)}\right) \\ &= \exp\left(-\frac{\alpha+\beta}{\alpha} \frac{\lambda}{f(j-1, w)}\right) \sum_{k=0}^{\infty} \frac{\Gamma(\alpha+\beta)\Gamma(\beta+k)}{\Gamma(\beta)\Gamma(\alpha+\beta+k)} \frac{1}{k!} \left(\frac{\alpha+\beta}{\alpha} \frac{\lambda}{f(j-1, w)}\right)^k. \end{aligned}$$

We then use that, since $\beta \leq 1$ and thus $[\beta] = 1$,

$$\frac{\Gamma(\alpha+\beta)\Gamma(\beta+k)}{\Gamma(\beta)\Gamma(\alpha+\beta+k)} = \prod_{\ell=0}^{k-1} \frac{\beta+\ell}{\alpha+\beta+\ell} \leq \prod_{\ell=0}^{k-1} \frac{1+\ell}{[\alpha]+1+\ell} = \prod_{\ell=1}^{[\alpha]} \frac{\ell}{\ell+k} = \frac{k!([\alpha])!}{(k+[\alpha])!}.$$

This yields the upper bound

$$\mathbb{E}[\mathcal{L}_\lambda(X_w(j))] \leq \exp\left(-\frac{\alpha+\beta}{\alpha} \frac{\lambda}{f(j-1, w)}\right) \sum_{k=0}^{\infty} \frac{([\alpha])!}{(k+[\alpha])!} \left(\frac{\alpha+\beta}{\alpha} \frac{\lambda}{f(j-1, w)}\right)^k. \tag{B.2}$$

We then observe that, for any $z > 0$ we have

$$\frac{1}{z}(e^z - 1) \leq e^{z/2+z^2}.$$

Indeed, this is easy to check for $z \geq 1$, whereas for $z < 1$ we use the inequalities $\log x \leq x - 1$ and $e^z - z - 1 \leq \frac{z^2}{2} + z^3$ (the latter using $\sum_{j=3}^{\infty} \frac{1}{j!} \leq 1$). Thus, for $z > 0$ and $\alpha \geq 1$,

$$\sum_{k=0}^{\infty} \frac{1}{(k+[\alpha])!} z^k \leq \frac{1}{([\alpha])!} \sum_{k=0}^{\infty} \frac{1}{(k+1)!} z^k = \frac{1}{([\alpha])!} \frac{1}{z}(e^z - 1) \leq \frac{1}{([\alpha])!} e^{z/2+z^2}. \tag{B.3}$$

Combining (B.2) and (B.3), we thus arrive at

$$\mathbb{E}[\mathcal{L}_\lambda(X_w(j))] \leq \exp\left(-\frac{1}{2} \frac{\alpha+\beta}{\alpha} \frac{\lambda}{f(j-1, w)} + \left(\frac{\alpha+\beta}{\alpha} \frac{\lambda}{f(j-1, w)}\right)^2\right).$$

We then take some sequence $(J_n)_{n \in \mathbb{N}}$ with $J_n \leq n$ to obtain

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{c\mu_n^{-1}}(\mathcal{P}_n; W)] &\leq \mathbb{E}[\mathcal{L}_{c\mu_n^{-1}}(\mathcal{P}_n - \mathcal{P}_{J_n}; W)] \\ &\leq \exp\left(-c\mu_n^{-1} \frac{\alpha+\beta}{2\alpha} \sum_{j=J_n}^{n-1} \frac{1}{f(j, k_n)} + \left(\frac{\alpha+\beta}{\alpha}\right)^2 c^2 \mu_n^{-2} \sum_{j=J_n}^{n-1} \frac{1}{f(j, k_n)^2}\right) + \mathbb{P}(W \geq k_n) \\ &= \exp\left(-c\mu_n^{-1} \frac{\alpha+\beta}{2\alpha} \sum_{j=J_n}^{n-1} \frac{1}{f(j, k_n)} \left(1 - \frac{2(\alpha+\beta)}{\alpha} \frac{c}{\mu_n f(j, k_n)}\right)\right) + \mathbb{P}(W \geq k_n). \end{aligned}$$

We now choose J_n for the different cases of the degree function s (according to Assumption 3.13) such that for all $j \geq J_n$, we have $\mu_n f(j, k_n) \geq 4c(\alpha+\beta)/\alpha$. This yields the upper bound

$$\begin{aligned} &\exp\left(-c \frac{\alpha+\beta}{4\alpha} \mu_n^{-1} \sum_{j=J_n}^{n-1} \frac{1}{f(j, k_n)}\right) + \mathbb{P}(W \geq k_n) \\ &\leq \exp\left(-c \frac{\alpha+\beta}{4\alpha} \mu_n^{-1} \sum_{j=J_n}^{n-1} \frac{1}{f(j, k_n) + \mu_n^{-1}}\right) + \mathbb{P}(W \geq k_n). \end{aligned}$$

With this upper bound at hand, we also claim that we can use the proof of Lemma 6.3 to obtain the desired upper bounds, despite the fact that the sum in the exponential term starts from J_n rather than from 0 (which is the case in Lemma 6.3).

For all additive cases we have $J_n = 0$, so that we can directly use the proof Lemma 6.3. This is due to the fact that $\mu_n f(j, k_n) \geq \mu_n f(0, k_n)$ diverges with n , so that it is at least $4(\alpha + \beta)/\alpha$ for all n sufficiently large. Indeed, for all additive weights types, for some small $\varepsilon > 0$ and some constant $C > 0$,

$$\begin{aligned} \text{Super-linear:} & \quad \mu_n = (C + o(1))ns(n)^{-1}, & k_n & := n^{(1-\varepsilon)p}. \\ \text{Log-stretched:} & \quad \mu_n = (C + o(1))(\log n)^{1-\beta} \exp(-(\log n)^\beta), & k_n & := n \exp(-(1-\varepsilon)\bar{c}(\log n)^\tau). \\ \text{Poly-log:} & \quad \mu_n = (C + o(1))(\log n)^{-(\beta-1)}, & k_n & := \exp((\log n)^{(1+\varepsilon)/\tau}). \end{aligned}$$

One can directly verify that $\mu_n f(0, k_n) = \Theta(\mu_n h(k_n))$ diverges with n , where h is some regularly-varying function with exponent 1.

For the three mixed weights cases (depending on the degree function s), we can set J_n as follows, where K is a sufficiently large constant:

$$\begin{aligned} \text{Super-linear:} & \quad J_n := Kn^{\varepsilon(p-1)/p}. \\ \text{Log-stretched:} & \quad J_n := KI_n, \text{ with } I_n \text{ as in (6.13)}. \\ \text{Poly-log:} & \quad J_n := KI_n, \text{ with } I_n \text{ as in (A.19)}. \end{aligned}$$

Using that for all $j \geq J_n$,

$$\mu_n f(j, k_n) \geq \mu_n f(J_n, k_n) \geq \mu_n g(k_n) s(J_n),$$

it is readily verified in all three cases that $\mu_n f(j, k_n) \geq 4(\alpha + \beta)/\alpha$ for all $j \geq J_n$ and all sufficiently large n . It also holds that the bounds used in the proof of Lemma 6.3 still hold when using this choice of J_n in the mixed weights cases, so that the conclusions from Lemma 6.3 are valid here, too. \square

Proof of Lemma B.1, Rayleigh case. We have

$$\begin{aligned} \mathbb{E} [\mathcal{L}_{c\mu_n^{-1}}(\mathcal{P}_n; W)] & \leq \prod_{j=0}^{n-1} \left(1 - \frac{c\mu_n^{-1}}{f(j, k_n)} e^{\mu_n^{-2}/(\pi f(j, k_n)^2)} \left(1 - \operatorname{erf} \left(\frac{c\mu_n^{-1}}{\sqrt{\pi} f(j, k_n)} \right) \right) \right) \quad (\text{B.4}) \\ & \quad + \mathbb{P}(W \geq k_n). \end{aligned}$$

We now use that, for any $x > 0$,

$$\begin{aligned} e^{x^2}(1 - \operatorname{erf}(x)) & = \frac{2}{\sqrt{\pi}} e^{x^2} \int_x^\infty e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-s(s+2x)} ds \\ & \geq \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-(s+2x)^2} ds = 1 - \operatorname{erf}(2x), \end{aligned}$$

where we use a variable substitution $s = t - x$. Also using that $1 - x \leq e^{-x}$ for all $x \in \mathbb{R}$, we can bound the first term on the right-hand side of (B.4) from above from

$$\exp \left(-c\mu_n^{-1} \sum_{j=0}^{n-1} \frac{1}{f(j, k_n)} \left(1 - \operatorname{erf} \left(\frac{2c\mu_n^{-1}}{\sqrt{\pi} f(j, k_n)} \right) \right) \right).$$

Since the error function increases and $\lim_{x \rightarrow \infty} \operatorname{erf}(x) = 1$, we can now repeat the same argument as in the **Beta** case: we create an upper bound by starting the sum from some index J_n , such that we can bound

$$1 - \operatorname{erf}\left(\frac{2c\mu_n^{-1}}{\sqrt{\pi}f(j, k_n)}\right) \geq \delta, \quad \text{for all } j \geq J_n,$$

for some small constant $\delta > 0$. The choice of J_n can be the same as in the **Beta** case, and the result thus follows through here, as well. \square

B.2 Conditions for an infinite path, Assumption 2.6.

We need to verify Conditions (2.9) and (2.10). Here, we choose $\nu_n^w = d\mu_n^w$ for some arbitrary constant $d \in (0, 1)$, as in the proof of Theorem 3.4. We recall from the proof of Lemma 6.4 that the lower bound in (6.3) uses Markov's inequality only. Hence, since we assume that the mean of the inter-birth time $X_w(j)$ equals $1/f(j-1, w)$, irrespective of its distribution, it follows that the proofs of Lemmas 6.4 and A.5 immediately follow for the other choices of inter-birth distributions in Remark 3.17.

It thus remains to verify (2.10). We observe that the proof of (2.10) for exponentially distributed inter-birth times, as in the proof of Theorem 3.4, holds more generally, as long as for any $i \in \mathbb{N}$ and $w \geq 0$, and for some $K > 0$,

$$\operatorname{Var}(X_w(i)) \leq K \mathbb{E}[X_w(i)]^2 = \frac{K}{f(i-1, w)^2}.$$

This is readily checked for all the cases in Remark 3.17, and is related to Remark 3.6.

C Proof of Lemma 6.6

Proof of Lemma 6.6. The aim of the proof is to provide an upper and lower bound for $\mathbb{P}(r(W) \geq x)$ that are asymptotically equivalent (i.e. the same up to a $(1 + o(1))$ multiplicative factor). First, we define, for a function $r : [0, \infty) \rightarrow (0, \infty)$, the *generalised inverse* as

$$r^\leftarrow(x) := \inf\{y \geq 0 : r(y) \geq x\}, \quad x \geq 0.$$

We start by proving an upper bound. Suppose $r(W) \geq x$. Then $W \in \{y \geq 0 : r(y) \geq x\}$, so that by the definition of the generalised inverse, it follows that $W \geq r^\leftarrow(x)$. As a result

$$\mathbb{P}(r(W) \geq x) \leq \mathbb{P}(W \geq r^\leftarrow(x)), \tag{C.1}$$

which concludes the upper bound. For a lower bound, suppose $r(W) < x$. Then,

$$\{y \geq 0 : r(y) \geq x\} \subseteq \{y \geq 0 : r(y) \geq r(W)\}.$$

Hence, the infimum of the left-hand side is larger than the infimum of the right-hand side, so that $r^\leftarrow(x) \geq r^\leftarrow(r(W))$. As a result,

$$\mathbb{P}(r(W) \geq x) \geq \mathbb{P}(r^\leftarrow(r(W)) > r^\leftarrow(x))$$

We use [13, Proposition 1.5.12] to obtain that $r^\leftarrow(r(x)) = (1 + o(1))x$ as $x \rightarrow \infty$. Hence, since $r^\leftarrow(r(W)) > r^\leftarrow(x)$ implies $W \geq r^\leftarrow(x)$, we can write $r^\leftarrow(r(W)) = (1 + t(W))W$, for some function t such that $t(x) \rightarrow 0$ as $x \rightarrow \infty$. We thus obtain

$$\mathbb{P}(r^\leftarrow(r(W)) > r^\leftarrow(x)) = \mathbb{P}((1 + t(W))W > r^\leftarrow(x)).$$

Now, for any $\varepsilon > 0$ we can take x sufficiently large so that, since $W \geq r^\leftarrow(x)$, $|t(W)| < \varepsilon$. We thus obtain the lower bound,

$$\mathbb{P}(W > r^\leftarrow(x)/(1 - \varepsilon)).$$

As the tail distribution of W is regularly varying with exponent $-(\zeta - 1)$, we obtain

$$\mathbb{P}(W > r^\leftarrow(x)/(1 - \varepsilon)) = ((1 - \varepsilon)^{\zeta-1} + o(1))\mathbb{P}(W > r^\leftarrow(x)).$$

As ε is arbitrary, combining this with the upper bound in (C.1), we obtain

$$\mathbb{P}(r(W) \geq x) = (1 + o(1))\mathbb{P}(W \geq r^\leftarrow(x)).$$

We now use that, by [13, Theorem 1.5.12], the function r^\leftarrow is regularly varying with exponent $1/\rho$. This implies, by Lemma 6.1 and the assumption on the tail distribution of W , that the right-hand side is regularly varying with exponent $-(\zeta - 1)/\rho$ as desired. \square

References

- [1] O. Amini, L. Devroye, S. Griffiths, and N. Olver. On explosions in heavy-tailed branching random walks. *Ann. Probab.*, 41(3B):1864–1899, 2013.
- [2] K. B. Athreya. Preferential attachment random graphs with general weight function. *Internet Math.*, 4(4):401–418, 2007.
- [3] K. B. Athreya and S. Karlin. Limit theorems for the split times of branching processes. *J. Math. Mech.*, 17:257–277, 1967.
- [4] K. B. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39:1801–1817, 1968.
- [5] K. B. Athreya and P. E. Ney. *Branching processes*, volume Band 196 of *Die Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York-Heidelberg, 1972.
- [6] F. Ball, M. González, R. Martínez, and M. Slavtchova-Bojkova. Stochastic monotonicity and continuity properties of functions defined on Crump-Mode-Jagers branching processes, with application to vaccination in epidemic modelling. *Bernoulli*, 20(4):2076–2101, 2014.
- [7] F. Ball, M. González, R. Martínez, and M. Slavtchova-Bojkova. Total progeny of Crump-Mode-Jagers branching processes: an application to vaccination in epidemic modelling. In *Branching processes and their applications*, volume 219 of *Lect. Notes Stat.*, pages 257–267. Springer, [Cham], 2016.
- [8] S. Banerjee and S. Bhamidi. Persistence of hubs in growing random networks. *Probab. Theory Related Fields*, 180(3-4):891–953, 2021.

- [9] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. Asymptotic behavior and distributional limits of preferential attachment graphs. *Ann. Probab.*, 42(1):1–40, 2014.
- [10] B. Berndtsson and P. Jagers. Exponential growth of a branching process usually implies stable age distribution. *J. Appl. Probab.*, 16(3):651–656, 1979.
- [11] J. Bertoin and R. Stephenson. Local explosion in self-similar growth-fragmentation processes. *Electron. Commun. Probab.*, 21:Paper No. 66, 12, 2016.
- [12] S. Bhamidi. Universal techniques to analyze preferential attachment trees: global and local analysis, 2007. Preprint available at https://scholar.google.com/citations?view_op=view_citation&hl=en&user=4JsyJK8AAAAJ&citation_for_view=4JsyJK8AAAAJ:IjCSPb-0Ge4C.
- [13] N. H. Bingham, C. M. Goldie, J. L. Teugels, and J. Teugels. *Regular variation*. Number 27. Cambridge university press, 1989.
- [14] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.
- [15] C. Borgs, J. Chayes, C. Daskalakis, and S. Roch. 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*, pages 135–144. ACM, New York, 2007.
- [16] J. Choi and S. Sethuraman. Large deviations for the degree structure in preferential attachment schemes. *Ann. Appl. Probab.*, 23(2):722–763, 2013.
- [17] K. S. Crump and C. J. Mode. A general age-dependent branching process. I, II. *J. Math. Anal. Appl.*, 24:494–508; *ibid.* 25 (1968), 8–17, 1968.
- [18] S. Dereich. Preferential attachment with fitness: unfolding the condensate. *Electron. J. Probab.*, 21:Paper No. 3, 38, 2016.
- [19] S. Dereich, C. Mailler, and P. Mörters. Nonextensive condensation in reinforced branching processes. *Ann. Appl. Probab.*, 27(4):2539–2568, 2017.
- [20] S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: degree evolutions. *Electron. J. Probab.*, 14:no. 43, 1222–1267, 2009.
- [21] S. Dereich and M. Ortgiese. Robust analysis of preferential attachment models with fitness. *Combin. Probab. Comput.*, 23(3):386–411, 2014.
- [22] N. Fountoulakis and T. Iyer. Condensation phenomena in preferential attachment trees with neighbourhood influence. *Electronic Journal of Probability*, 27:1 – 49, 2022.
- [23] P. Galashin. Existence of a persistent hub in the convex preferential attachment model. *arXiv preprint arXiv:1310.7513*, 2013.
- [24] A. Garavaglia, R. S. Hazra, R. van der Hofstad, and R. Ray. Universality of the local limit of preferential attachment models. *arXiv preprint arXiv:2212.05551*, 2022.
- [25] A. Garavaglia, R. van der Hofstad, and N. Litvak. Local weak convergence for PageRank. *Ann. Appl. Probab.*, 30(1):40–79, 2020.

- [26] A. Garavaglia, R. van der Hofstad, and G. Woeginger. The dynamics of power laws: Fitness and aging in preferential attachment trees. *J. Stat. Phys.*, 168:1137–1179, 2017.
- [27] T. Gottfried and S. Grosskinsky. Asymptotics of generalized Pólya urns with non-linear feedback. *arXiv preprint arXiv:2303.01210*, 2023.
- [28] S. Grishechkin. On a relationship between processor-sharing queues and Crump-Mode-Jagers branching processes. *Adv. in Appl. Probab.*, 24(3):653–698, 1992.
- [29] R. van der Hofstad. *Random graphs and complex networks*, volume 43. Cambridge university press, 2016.
- [30] C. Holmgren and S. Janson. Fringe trees, Crump-Mode-Jagers branching processes and m -ary search trees. *Probab. Surv.*, 14:53–154, 2017.
- [31] A. Iksanov, K. Kolesko, and M. Meiners. Gaussian fluctuations and a law of the iterated logarithm for Nerman's martingale in the supercritical general branching process. *Electron. J. Probab.*, 26:Paper No. 160, 22, 2021.
- [32] A. Iksanov, A. Marynych, and B. Rashytov. Stable fluctuations of iterated perturbed random walks in intermediate generations of a general branching process tree. *Lith. Math. J.*, 62(4):447–466, 2022.
- [33] T. Iyer. On explosion in Crump-Mode-Jagers branching processes, with applications to the limiting structure of recursive trees with fitness. *In preparation*.
- [34] T. Iyer. Degree distributions in recursive trees with fitnesses. *Adv. in Appl. Probab.*, 55(2):407–443, 2023.
- [35] P. Jagers. A general stochastic model for population development. *Skand. Aktuarietidskr.*, pages 84–103, 1969.
- [36] P. Jagers. *Branching processes with biological applications*. Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics. Wiley-Interscience [John Wiley & Sons], London-New York-Sydney, 1975.
- [37] P. Jagers. General branching processes as Markov fields. *Stochastic Process. Appl.*, 32(2):183–212, 1989.
- [38] P. Jagers and O. Nerman. The asymptotic composition of supercritical multi-type branching populations. In *Séminaire de Probabilités, XXX*, volume 1626 of *Lecture Notes in Math.*, pages 40–54. Springer, Berlin, 1996.
- [39] P. Jagers, O. Nerman, and Z. Taïb. When did Joe's great...grandfather live? Or: On the time scale of evolution. In *Selected Proceedings of the Sheffield Symposium on Applied Probability (Sheffield, 1989)*, volume 18 of *IMS Lecture Notes Monogr. Ser.*, pages 118–126. Inst. Math. Statist., Hayward, CA, 1991.
- [40] S. Janson. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.*, 9:103–252, 2012.
- [41] S. Janson. Asymptotics of fluctuations in Crump-Mode-Jagers processes: the lattice case. *Adv. in Appl. Probab.*, 50(A):141–171, 2018.

- [42] J. Jordan. Preferential attachment graphs with co-existing types of different fitnesses. *J. Appl. Probab.*, 55(4):1211–1227, 2018.
- [43] J. F. C. Kingman. The first birth problem for an age-dependent branching process. *Ann. Probab.*, 3(5):790–801, 10 1975.
- [44] K. Kolesko and E. Sava-Huss. Limit theorems for discrete multitype branching processes counted with a characteristic. *Stochastic Process. Appl.*, 162:49–75, 2023.
- [45] J. Komjáthy. Explosive Crump-Mode-Jagers branching processes. *arXiv preprint arXiv:1602.01657*, 2016.
- [46] P. Krapivsky and D. Krioukov. Scale-free networks as preasymptotic regimes of superlinear preferential attachment. *Physical Review E*, 78(2):026114, 2008.
- [47] J. Kunegis, M. Blattner, and C. Moser. Preferential attachment in online networks: Measurement and explanations. In *Proceedings of the 5th annual ACM web science conference*, pages 205–214, 2013.
- [48] G. Last and M. Penrose. *Lectures on the Poisson process*, volume 7 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, Cambridge, 2018.
- [49] J. Levesque, D. W. Maybury, and R. H. A. D. Shaw. A model of COVID-19 propagation based on a gamma subordinated negative binomial branching process. *J. Theoret. Biol.*, 512:Paper No. 110536, 10, 2021.
- [50] T. Y. Y. Lo. Weak local limit of preferential attachment random trees with additive fitness, 2021.
- [51] B. Lodewijks. Location of high-degree vertices in weighted recursive graphs with bounded random weights. *To appear in the Journal of Applied Probability*, 2021.
- [52] B. Lodewijks and M. Ortgiese. The maximal degree in random recursive graphs with random weights. *arXiv preprint arXiv:2007.05438*, 2020.
- [53] B. Lodewijks and M. Ortgiese. A phase transition for preferential attachment models with additive fitness. *Electron. J. Probab.*, 25:Paper No. 146, 54, 2020.
- [54] T. Mikosch. *Regular variation, subexponentiality and their applications in probability theory*, volume 99. Eindhoven University of Technology Eindhoven, The Netherlands, 1999.
- [55] M. Mitzenmacher, R. Oliveira, and J. Spencer. A scaling result for explosive processes. *Electron. J. Combin.*, 11(1):Research Paper 31, 14, 2004.
- [56] O. Nerman. On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrsch. Verw. Gebiete*, 57(3):365–395, 1981.
- [57] O. Nerman and P. Jagers. The stable double infinite pedigree process of supercritical branching populations. *Z. Wahrsch. Verw. Gebiete*, 65(3):445–460, 1984.
- [58] R. Oliveira and J. Spencer. Connectivity transitions in networks with super-linear preferential attachment. *Internet Math.*, 2(2):121–163, 2005.

- [59] P. Olofsson. The $x \log x$ condition for general branching processes. *J. Appl. Probab.*, 35(3):537–544, 1998.
- [60] T. Pham, P. Sheridan, and H. Shimodaira. Joint estimation of preferential attachment and node fitness in growing complex networks. *Scientific reports*, 6(1):32558, 2016.
- [61] B. Pittel. Note on the heights of random recursive trees and random m -ary search trees. *Random Structures Algorithms*, 5(2):337–347, 1994.
- [62] A. Rudas, B. Tóth, and B. Valkó. Random trees and general branching processes. *Random Structures Algorithms*, 31(2):186–202, 2007.
- [63] D. Sénizergues. Geometry of weighted recursive and affine preferential attachment trees. *Electronic Journal of Probability*, 26:1–56, 2021.
- [64] S. Sethuraman and S. C. Venkataramani. On the growth of a superlinear preferential attachment scheme. In *Probability and analysis in interacting physical systems*, volume 283 of *Springer Proc. Math. Stat.*, pages 243–265. Springer, Cham, 2019.