CrossMark

The Dynamics of Power laws: Fitness and Aging in Preferential Attachment Trees

Alessandro Garavaglia[1] · **Remco van der Hofstad**[1] ·
Gerhard Woeginger[2]

© The Author(s) 2017. This article is an open access publication

Abstract Continuous-time branching processes describe the evolution of a population whose individuals generate a random number of children according to a birth process. Such branching processes can be used to understand preferential attachment models in which the birth rates are linear functions. We are motivated by citation networks, where power-law citation counts are observed as well as aging in the citation patterns. To model this, we introduce fitness and age-dependence in these birth processes. The multiplicative fitness moderates the rate at which children are born, while the aging is integrable, so that individuals receives a *finite* number of children in their lifetime. We show the existence of a limiting degree distribution for such processes. In the preferential attachment case, where fitness and aging are absent, this limiting degree distribution is known to have power-law tails. We show that the limiting degree distribution has exponential tails for bounded fitnesses in the presence of integrable aging, while the power-law tail is restored when integrable aging is combined with fitness with unbounded support with at most exponential tails. In the absence of integrable aging, such processes are explosive.

Keywords Power law · Branching processes · Citation networks · Preferential attachment

✉ Alessandro Garavaglia
   a.garavaglia@tue.nl

   Remco van der Hofstad
   rhofstad@win.tue.nl

   Gerhard Woeginger
   woeginger@cs.rwth-aachen.de

[1] Department of Mathematics and Computer Science, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

[2] Department of Computer Science, RWTH Aachen, Ahornstrasse 55, 52074 Aachen, Germany

1 Introduction

Preferential attachment models (PAMs) aim to describe dynamical networks. As for many real-world networks, PAMs present power-law degree distributions that arise directly from the dynamics, and are not artificially imposed as, for instance, in configuration models or inhomogeneous random graphs.

PAMs were first proposed by Albert and Barabási [1], who defined a random graph model where, at every discrete time step, a new vertex is added with one or more edges, that are attached to existing vertices with probability proportional to the degrees, i.e.,

$$\mathbb{P}\left(\text{vertex } (n+1) \text{ is attached to vertex } i \mid \text{graph at time } n\right) \propto D_i(n),$$

where $D_i(n)$ denotes the degree of a vertex $i \in \{1, \ldots, n\} = [n]$ at time n. In general, the dependence of the attachment probabilities on the degree can be through a *preferential attachment function* of the degree, also called *preferential attachment weights*. Such models are called PAMs with *general weight function*. According to the asymptotics of the weight function $w(\cdot)$, the limiting degree distribution of the graph can behave rather differently. There is an enormous body of literature showing that PAMs present power-law decay in the limiting degree distribution precisely when the weight function is affine, i.e., it is a constant plus a linear function. See e.g., [23, Chap. 8] and the references therein. In addition, these models show the so-called *old-get-richer* effect, meaning that the vertices of highest degrees are the vertices present early in the network formation. An extension of this model is called preferential attachment models with a *random number of edges* [8], where new vertices are added to the graph with a different number of edges according to a fixed distribution, and again power-law degree sequences arise. A generalization that also gives younger vertices the chance to have high degrees is given by PAMs with *fitness* as studied in [9,10]. Borgs et al. [6] present a complete description of the limiting degree distribution of such models, with different regimes according to the distribution of the fitness, using *generalized Polyá's urns*. An interesting variant of a multi-type PAM is investigated in [20], where the author consider PAMs where fitnesses are not i.i.d. across the vertices, but they are sampled according to distributions depending on the fitnesses of the ancestors.

This work is motivated by *citation networks*, where vertices denote papers and the directed edges correspond to citations. For such networks, other models using preferential attachment schemes and adaptations of them have been proposed mainly in the physics literature. Aging effects, i.e., considering the *age of a vertex* in its likelihood to obtain children, have been extensively considered as the starting point to investigate their dynamics [7,11,12,25,26]. Here the idea is that old papers are less likely to be cited than new papers. Such aging has been observed in many citation network datasets and makes PAMs with weight functions depending only on the degree ill-suited for them. As mentioned above, such models could more aptly be called *old-get-richer* models, i.e., in general *old* vertices have the highest degrees. In citation networks, instead, papers with many citations appear all the time. Barabási, Wang and Song [24] investigate a model that incorporates these effects. On the basis of empirical data, they suggest a model where the aging function follows a lognormal distribution with paper-dependent parameters, and the preferential attachment function is the identity. In [24], the fitness function is estimated rather than the more classical approach where it is taken to be i.i.d. Hazoglou et al. in [13] propose a similar dynamics for citation evolution , but only considering the presence of aging and cumulative advantage without fitness.

Tree models, arising when new vertices are added with only one edge, have been analyzed in [2,3,21,22] and lead to continuous-time branching processes (CTBP). The degree distributions in tree models show identical qualitative behavior as for the non-tree setting,

while their analysis is much simpler. Motivated by this and the wish to understand the qualitative behavior of PAMs with general aging and fitness, the starting point of our model is the CTBP or tree setting. Such processes have been intensively studied, due to their applications in other fields, such as biology. Detailed and rigorous analysis of CTBPs can be found in [2–5,14,17,22]. A CTBP consists of individuals, whose children are born according to certain birth processes, these processes being i.i.d. across the individuals in the population. The birth processes $(V_t)_{t\geq 0}$ are defined in term of point or jump processes on $\mathbb{N}$ [14,17], where the birth times of children are the jump times of the process, and the number of children of an individual at time $t \in \mathbb{R}^+$ is given by V_t.

In the literature, the CTBPs are used as a technical tool to study PAMs [3,20,22]. Indeed, the CTBP at the nth birth time follows the same law as the PAM consisting of n vertices. In [3,22], the authors prove an embedding theorem between branching processes and preferential attachment trees, and give a description of the degree distribution in terms of the asymptotic behavior of the weight function $w(\cdot)$. In particular, a power-law degree distribution is present in the case of (asymptotically) linear weight functions [21]. In the sub-linear case, instead, the degree distribution is *stretched-exponential*, while in the super-linear case it collapses, in the sense that one of the first vertices will receive all the incoming new edges after a certain step [18]. Due to the apparent exponential growth of the number of nodes in citation networks, we view the continuous-time process as the real network, which deviates from the usual perspective. Because of its motivating role in this paper, let us now discuss the empirical properties of citation networks in detail.

1.1 Citation Networks Data

Let us now discuss the empirical properties of citation networks in more detail. We analyze the Web Of Science database, focusing on three different fields of science: *Probability and Statistics* (PS), *Electrical Engineering* (EE) and *Biotechnology and Applied Microbiology* (BT). We first point out some characteristics of citation networks that we wish to replicate in our models.

Real-world citation networks possess five main characteristics:

(1) In Fig. 1, we see that the number of scientific publications grows exponentially in time. While this is quite prominent in the data, it is unclear how this exponential growth arises.

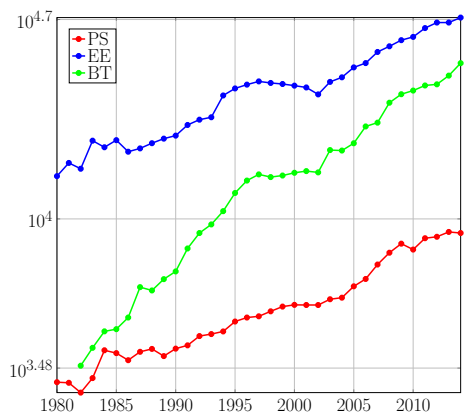

Fig. 1 Number of publication per year (logarithmic Y axis)

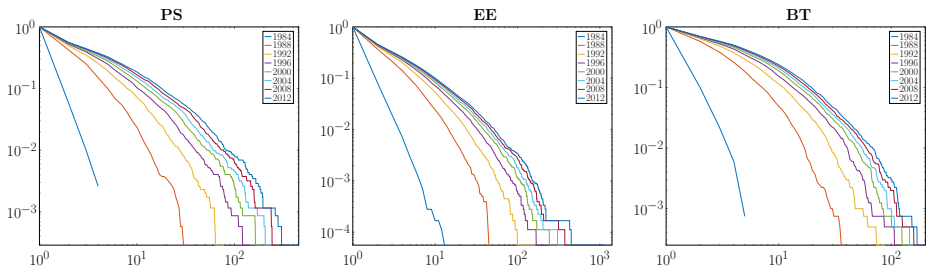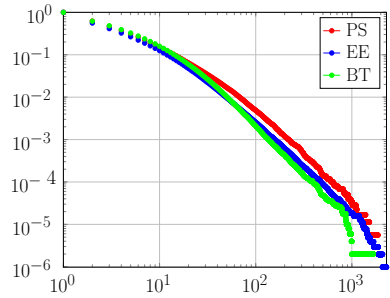**Fig. 2** Loglog plot for the in-degree distribution tail in citation networks

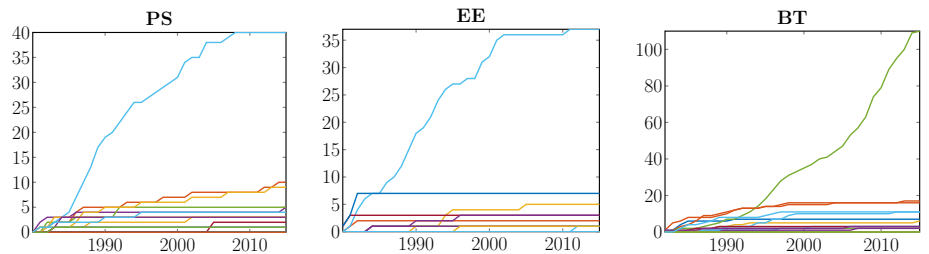**Fig. 3** Degree distribution for papers from 1984 over time

Fig. 4 Time evolution for the number of citations of samples of 20 randomly chosen papers from 1980 for PS and EE, and from 1982 for BT

This could either be due to the fact that the number of journals that are listed in Web Of Science grows over time, or that journals contain more and more papers.

(2) In Fig. 2, we notice that these datasets have empirical power-law citation distributions. Thus, most papers attract few citations, but the amount of variability in the number of citations is rather substantial. We are also interested in the dynamics of the citation distribution of the papers published in a given year, as time proceeds. This can be observed in Fig. 3. We see a *dynamical power law*, meaning that at any time the degree distribution is close to a power law, but the exponent changes over time (and in fact decreases, which corresponds to heavier tails). When time grows quite large, the power law approaches a fixed value.

(3) In Fig. 4, we see that the majority of papers stop receiving citations after some time, while few others keep being cited for longer times. This inhomogeneity in the evolution of node degrees is not present in classical PAMs, where the degree of *every* fixed vertex grows as a positive power of the graph size. Figure 4 shows that the number of citations of papers published in the same year can be rather different, and the majority of papers
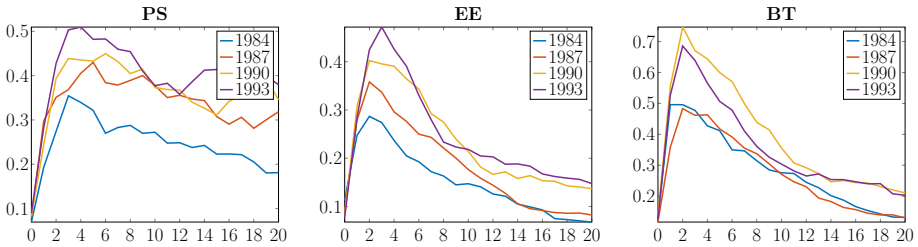
Fig. 5 Average degree increment over a 20-years time window for papers published in different years. PS presents an aging effect different from EE and BT, showing that papers in PS receive citations longer than papers in EE and BT
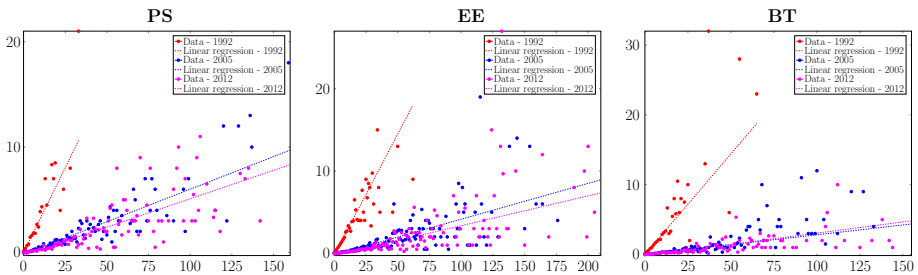

Fig. 6 Linear dependence between past and future number of citations for papers from 1988

actually stop receiving citations quite soon. In particular, after a first increase, the average increment of citations decreases over time (see Fig. 5). We observe a difference in this aging effect between the PS dataset and the other two datasets, due to the fact that in PS, scientists tend to cite older papers than in EE or BT. Nevertheless the average increment of citations received by papers in different years tends to decrease over time for all three datasets.

(4) Figure 6 shows the linear dependence between the past number of citations of a paper and the future ones. Each plot represents the average number of citations received by papers published in 1984 in the years 1993, 2006 and 2013 according to the initial number of citations in the same year. At least for low values of the starting number of citations, we see that the average number of citations received during a year grows linearly. This suggests that the attractiveness of a paper depends on the past number of citations through an affine function.

(5) A last characteristic that we observe is the lognormal distribution of the age of cited papers. In Fig. 7, we plot the distribution of cited papers, looking at references made by papers in different years. We have used a 20 years time window in order to compare different citing years. Notice that this lognormal distribution seems to be very similar within different years, and the shape is similar over different fields.

Let us now explain how we translate the above empirical characteristics into our model. First, CTBPs grow exponentially over time, as observed in citation networks. Secondly, the aging present in citation networks, as seen both in Figs. 4 and 5, suggests that citation rates become smaller for large times, in such a way that typical papers stop receiving citations at some (random) point in time. The hardest characteristic to explain is the power-law degree sequence. For this, we note that citations of papers are influenced by many *external factors*
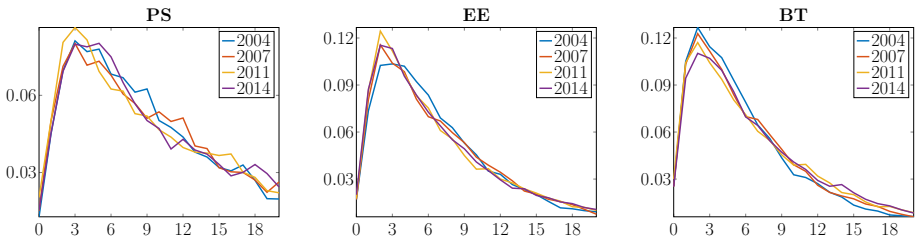
Fig. 7 Distribution of the age of cited papers for different citing years

that affect the attractiveness of papers (the journal, the authors, the topic,…). Since this cannot be quantified explicitly, we introduce another source of randomness in our birth processes that we call *fitness*. This appears in the form of multiplicative factors of the attractiveness of a paper, and for lack of better knowledge, we take these factors to be i.i.d. across papers, as often assumed in the literature. These assumptions are similar in spirit as the ones by Barabási et al. [24], which were also motivated by citation data, and we formalize and extend their results considerably. In particular, we give the precise conditions under which power-law citation counts are observed in this model.

Our main goal is to define CTBPs with both aging as well as random fitness that keep having a power-law decay in the in-degree distribution. Before discussing our model in detail in Sect. 2, we present the heuristic ideas behind it as well as the main results of this paper.

1.2 Our Main Contribution

The crucial point of this work is to show that it is possible to obtain power-law degree distributions in preferential attachment trees where the birth process *is not just depending on an asymptotically linear weight sequence*, in the presence of *integrable aging* and *fitness*. Let us now briefly explain how these two effects change the behavior of the degree distribution.

Integrable Aging and Affine Preferential Attachment Without Fitness In the presence of aging but without fitness, we show that the aging effect substantially slows down the birth process. In the case of affine weights, aging destroys the power-law of the stationary regime, generating a limiting distribution that consists of a power law with exponential truncation. We prove this under reasonable conditions on the underlying aging function (see Lemma 5.1).

Integrable Aging and Super-Linear Preferential Attachment Without Fitness Since the aging destroys the power-law of the affine PA case, it is natural to ask whether the combination of integrable aging and *super-linear* weights restores the power-law limiting degree distribution. Theorem 2.3 states that this is not the case, as super-linear weights imply explosiveness of the branching process, which is clearly unrealistic in the setting of citation networks (here, we call a weight sequence $k \mapsto f_k$ *super-linear* when $\sum_{k \geq 1} 1/f_k < \infty$). This result is quite general, because it holds for *any* integrable aging function. Due to this, it is impossible to obtain power-laws from super-linear preferential attachment weights. This suggests that (apart from slowly-varying functions), affine preferential attachment weights have the strongest possible growth, while maintaining exponential (and thus, in particular, non-explosive) growth.

Integrable Aging and Affine Preferential Attachment with Unbounded Fitness In the case of aging and fitness, the asymptotic behavior of the limiting degree distribution is rather involved. We estimate the asymptotic decay of the limiting degree distribution with affine

weights in Proposition 5.5. With the example fitness classes analyzed in Sect. 5.3, we prove that power-law tails are possible in the setting of aging and fitness, at least when the fitness has roughly exponential tail. So far, PAMs with fitness required the support of the fitness distribution to be *bounded*. The addition of aging allows the support of the fitness distribution to be unbounded, a feature that seems reasonable to us in the context of citation networks. Indeed, the relative attractivity of one paper compared to another one can be enormous, which is inconsistent with a bounded fitness distribution. While we do not know precisely what the necessary and sufficient conditions are on the aging and the fitness distribution to assure a power-law degree distribution, our results suggests that affine PA weights with integrable aging and fitnesses with at most an exponential tail in general do so, a feature that was not observed before.

Dynamical Power Laws In the case of fitness with exponential tails, we further observe that the number of citations of a paper of age t has a power-law distribution with an exponent that depends on t. We call this a *dynamical power law*, and it is a possible explanation of the dynamical power laws observed in citation data (see Fig. 3).

Universality An interesting and highly relevant observation in this paper is that the limiting degree distribution of preferential attachment trees with aging and fitness shows a high amount of *universality*. Indeed, for integrable aging functions, the dependence on the precise choice of the aging function seems to be minor, except for the total integral of the aging function. Further, the dependence on fitness is quite robust as well.

2 Our Model and Main Results

In this paper we introduce the effect of aging and fitness in CTBP populations, giving rise to directed trees. Our model is motivated by the study of *citation networks*, which can be seen as directed graphs. Trees are the simplest case in which we can see the effects of aging and fitness. Previous work has shown that PAMs can be obtained from PA trees by collapsing, and their general degree structure can be quite well understood from those in trees. For example, PAMs with fixed out-degree $m \geq 2$ can be defined through a collapsing procedure, where a vertex in the multigraph is formed by $m \in \mathbb{N}$ vertices in the tree (see [23, Sect. 8.2]). In this case, the limiting degree distribution of the PAM preserve the structure of the tree case ([23, Sect. 8.4], [5, Sect. 5.7]). This explains the relevance of the tree case results for the study of the effect of aging and fitness in PAMs. It could be highly interesting to prove this rigorously.

2.1 Our CTBP Model

CTBPs represent a population made of individuals producing children independently from each other, according to i.i.d. copies of a birth process on $\mathbb{N}$. We present the general theory of CTBPs in Sect. 3, where we define such processes in detail and we refer to general results that are used throughout the paper. In general, considering a birth process $(V_t)_{t \geq 0}$ on $\mathbb{N}$, every individual in the population has an i.i.d. copy of the process $(V_t)_{t \geq 0}$, and the number of children of individual x at time t is given by the value of the process V_t^x. We consider birth processes defined by a sequence of weights $(f_k)_{k \in \mathbb{N}}$ describing the birth rates. Here, the time between the kth and the $(k + 1)$st jump is exponentially distributed with parameter f_k. The behavior of the whole population is determined by this sequence.

The fundamental theorem for the CTBPs that we study is Theorem 3.10 quoted in Sect. 3. It states that, under some hypotheses on the birth process $(V_t)_{t \geq 0}$, the population grows exponentially in time, which nicely fits the exponential growth of scientific publications as indicated in Fig. 1. Further, using a so-called *random vertex characteristic* as introduced in [14], a complete class of properties of the population can be described, such as the fraction of individuals having k children, as we investigate in this paper. The two main properties are stated in Definitions 3.8 and 3.9, and are called *supercritical* and *Malthusian* properties. These properties require that there exists a positive value α^* such that

$$\mathbb{E}\left[V_{T_{\alpha^*}}\right] = 1, \qquad \text{and} \qquad -\frac{d}{d\alpha}\mathbb{E}\left[V_{T_\alpha}\right]\Big|_{\alpha=\alpha^*} < \infty,$$

where T_α denotes an exponentially distributed random variable with rate α independent of the process $(V_t)_{t \geq 0}$. The unique value α^* that satisfies both conditions is called the *Malthusian parameter*, and it describes the exponential growth rate of the population size. The aim is to investigate the ratio

$$\frac{\text{number of individuals with } k \text{ children at time } t}{\text{size total population at time } t}.$$

According to Theorem 3.10, this ratio converges almost surely to a deterministic limiting value p_k. The sequence $(p_k)_{k \in \mathbb{N}}$, which we refer to as the limiting degree distribution of the CTBP (see Definition 3.12), is given by

$$p_k = \mathbb{E}\left[\mathbb{P}\left(V_u = k\right)_{u=T_{\alpha^*}}\right].$$

The starting idea of our model of citation networks is that, given the history of the process up to time t,

the rate of an individual of age t and k children to generate a new child is $Y f_k g(t)$, (2.1)

where f_k is a non-decreasing PA function of the degree, g is an integrable function of time, and Y is a positive random variable called fitness. Therefore, the likelihood to generate children increases by having many children and/or a high fitness, while it is reduced by age.

Recalling Fig. 6, we assume that the PA function f is affine, so $f_k = ak + b$. In terms of a PA scheme, this implies

$$\mathbb{P}\left(\text{a paper cites another with past } k \text{ citations} \mid \text{past}\right) \approx \frac{n(k)(ak+b)}{A},$$

where $n(k)$ denotes the number of papers with k past citations, and A is the normalization factor. Such behavior has already been observed by Redner [19] and Barabási et al. [15]).

We assume throughout the paper that the aging function g is integrable. In fact, we start by the fact that the age of cited papers is lognormally distributed (recall Fig. 7). By normalizing such a distribution by the average increment in the number of citations of papers in the selected time window, we identify a universal function $g(t)$. Such function can be approximated by a lognormal shape of the form

$$g(t) \approx c_1 e^{-c_2(\log(t+1)-c_3)^2},$$

for c_1, c_2 and c_3 field-dependent parameters. In particular, from the procedure used to define $g(t)$, we observe that

$$g(t) \approx \frac{\text{number of references to year } t}{\text{number of papers of age } t} \frac{\text{total number of papers considered}}{\text{total number of references considered}},$$

which means in terms of PA mechanisms that

$$\mathbb{P}\left(\text{a paper cites another of age } t \mid \text{past}\right) \approx \frac{n(t)g(t)}{B},$$

where B is the normalization factor, while this time $n(t)$ is the number of paper of age t. This suggests that the citing probability depends on age through a lognormal aging function $g(t)$, which is integrable. This is one of the main assumptions in our model, as we discuss in Sect. 1.2.

It is known from the literature [2,21,22] that CTBPs show power-law limiting degree distributions when the infinitesimal rates of jump depend only on a sequence $(f_k)_{k \in \mathbb{N}}$ that is asymptotically *linear*. Our main aim is to investigate whether power-laws can also arise in branching processes that include aging and fitness. The results are organized as follows. In Sect. 2.2, we discuss the results for CTBPs with aging in the absence of fitness. In Sect. 2.3, we present the results with aging and fitness. In Sect. 2.4, we specialize to fitness with distributions with exponential tails, where we show that the limiting degree distribution is a power law with a *dynamic* power-law exponent.

2.2 Results with Aging Without Fitness

In this section, we focus on aging in PA trees in the absence of fitness. The aging process can then be viewed as a time-changed stationary birth process (see Definition 3.13). A stationary birth process is a stochastic process $(V_t)_{t \geq 0}$ such that, for h small enough,

$$\mathbb{P}\left(V_{t+h} = k + 1 \mid V_t = k\right) = f_k h + o(h).$$

In general, we assume that $k \mapsto f_k$ is increasing. The *affine case* arises when $f_k = ak + b$ with $a, b > 0$. By our observations in Fig. 6, as well as related works [15,19], the affine case is a reasonable approximation for the attachment rates in citation networks.

For a stationary birth process $(V_t)_{t \geq 0}$, under the assumption that it is supercritical and Malthusian, the limiting degree distribution $(p_k)_{k \in \mathbb{N}}$ of the corresponding branching process is given by

$$p_k = \frac{\alpha^*}{\alpha^* + f_k} \prod_{i=0}^{k-1} \frac{f_i}{\alpha^* + f_i}. \tag{2.2}$$

For a more detailed description, we refer to Sect. 3.2. Branching processes defined by stationary processes (with no aging effect) have a so-called *old-get-richer* effect. As this is not what we observe in citation networks (recall Fig. 4), we want to introduce *aging* in the reproduction process of individuals. The aging process arises by adding age-dependence in the infinitesimal transition probabilities:

Definition 2.1 (*Aging birth processes*) Consider a non-decreasing PA sequence $(f_k)_{k \in \mathbb{N}}$ of positive real numbers and an aging function $g \colon \mathbb{R}^+ \to \mathbb{R}^+$. We call a stochastic process $(N_t)_{t \geq 0}$ an *aging birth process* (without fitness) when

(1) $N_0 = 0$, and $N_t \in \mathbb{N}$ for all $t \in \mathbb{N}$;
(2) $N_t \leq N_s$ for every $t \leq s$;
(3) for fixed $k \in \mathbb{N}$ and $t \geq 0$, as $h \to 0$,

$$\mathbb{P}\left(N_{t+h} = k + 1 \mid N_t = k\right) = f_k g(t) h + o(h).$$

Aging processes are time-rescaled versions of the corresponding stationary process defined by the same sequence $(f_k)_{k \in \mathbb{N}}$. In particular, for any $t \geq 0$, N_t has the same distribution as

$V_{G(t)}$, where $G(t) = \int_0^t g(s)ds$. In general, we assume that the aging function is *integrable*, which means that $G(\infty) := \int_0^\infty g(s)ds < \infty$. This implies that the number of children of a single individual in its entire lifetime has distribution $V_{G(\infty)}$, which is finite in expectation. In terms of citation networks, this assumption is reasonable since we do not expect papers to receive an infinite number of citations ever (recall Fig. 5). Instead, for the stationary process $(V_t)_{t\geq 0}$ in Definition 3.13, we have that $\mathbb{P}$-a.s. $V_t \to \infty$, so that also the aging process diverges $\mathbb{P}$-a.s. when $G(\infty) = \infty$.

For aging processes, the main result is the following theorem, proven in Sect. 4. In its statement, we rely on the Laplace transform of a function. For a precise definition of this notion, we refer to Sect. 3:

Theorem 2.2 (Limiting distribution for aging branching processes) *Consider an integrable aging function g and a PA sequence $(f_k)_{k\in\mathbb{N}}$. Denote the corresponding aging birth process by $(N_t)_{t\geq 0}$. Then, assuming that $(N_t)_{t\geq 0}$ is supercritical and Malthusian, the limiting degree distribution of the branching process $\mathbf{N}$ defined by the birth process $(N_t)_{t\geq 0}$ is given by*

$$p_k = \frac{\alpha^*}{\alpha^* + f_k \hat{\mathcal{L}}^g(k, \alpha^*)} \prod_{i=0}^{k-1} \frac{f_i \hat{\mathcal{L}}^g(i, \alpha^*)}{\alpha^* + f_i \hat{\mathcal{L}}^g(i, \alpha^*)}, \tag{2.3}$$

where α^* is the Malthusian parameter of $\mathbf{N}$. Here, the sequence of coefficients $(\hat{\mathcal{L}}^g(k, \alpha^*))_{k\in\mathbb{N}}$ appearing in (2.3) is given by

$$\hat{\mathcal{L}}^g(k, \alpha^*) = \frac{\mathcal{L}(\mathbb{P}(N_. = k) g(\cdot))(\alpha^*)}{\mathcal{L}(\mathbb{P}(N_. = k))(\alpha^*)}, \tag{2.4}$$

where, for $h \colon \mathbb{R}^+ \to \mathbb{R}$, $\mathcal{L}(h(\cdot))(\alpha)$ denotes the Laplace transform of h.

Further, considering a fixed individual in the branching population, the total number of children in its entire lifetime is distributed as $V_{G(\infty)}$, where $G(\infty)$ is the L^1-norm of g.

The limiting degree distribution maintains a product structure as in the stationary case (see (2.2) for comparison). Unfortunately, the analytic expression for the probability distribution $(p_k)_{k\in\mathbb{N}}$ in (2.3) given by the previous theorem is not explicit. In the stationary case, the form reduces to the simple expression in (2.2).

In general, the asymptotics of the coefficients $(\hat{\mathcal{L}}^g(k, \alpha^*))_{k\in\mathbb{N}}$ is unclear, since it depends both on the aging function g as well as the PA weight sequence $(f_k)_{k\in\mathbb{N}}$ itself in an intricate
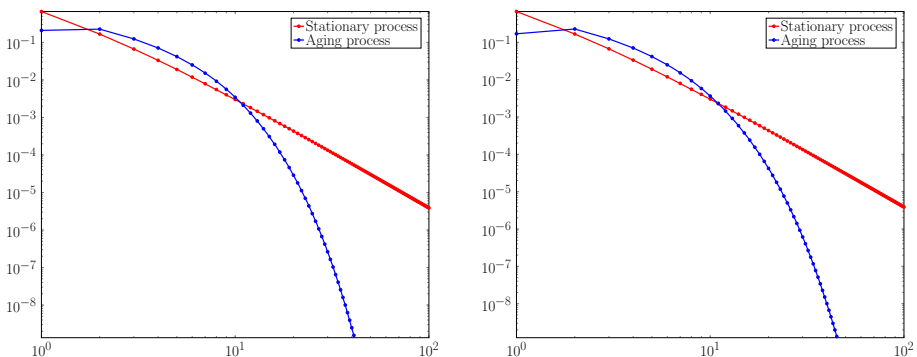

Fig. 8 Examples of stationary and aging limit degree distributions

way. In particular, we have no explicit expression for the ratio in (2.4), except in special cases. In this type of birth process, the cumulative advantage given by $(f_k)_{k \in \mathbb{N}}$ and the aging effect given by g cannot be separated from each other.

Numerical examples in Fig. 8 show how aging destroys the power-law degree distribution. In each of the two plots, the limiting degree distribution of a stationary process with affine PA weights gives a power-law degree distribution, while the process with two different integrable aging functions does not. In the examples we have used $g(t) = e^{-\lambda t}$ and $g(t) = (1 + t)^{-\lambda}$ for some $\lambda > 1$, and we observe the insensitivity of the limiting degree distribution with respect to g. The distribution given by (2.3) can be seen as the limiting degree distribution of a CTBP defined by preferential attachment weight $(f_k \hat{\mathcal{L}}^g(k, \alpha^*))_{k \in \mathbb{N}}$. This suggests that $f_k \hat{\mathcal{L}}^g(k, \alpha^*)$ is not asymptotically linear in k.

In Sect. A.2, we investigate the two examples in Fig. 8, showing that the limiting degree distribution has exponential tails, a fact that we know in general just as an upper bound (see Lemma 5.3).

In order to apply the general CTBP result in Theorem 3.10 below, we need to prove that an aging process $(N_t)_{t \geq 0}$ is supercritical and Malthusian. We show in Sect. 4 that, for an integrable aging function g, the corresponding process is supercritical if and only if

$$\lim_{t \to \infty} \mathbb{E}\left[V_{G(t)}\right] = \mathbb{E}\left[V_{G(\infty)}\right] > 1. \tag{2.5}$$

Condition (2.5) heuristically suggests that the process $(N_t)_{t \geq 0}$ has a Malthusian parameter if and only if the expected number of children in the entire lifetime of a fixed individual is larger than one, which seems quite reasonable. In particular, such a result follows from the fact that if g is integrable, then the Laplace transform is always finite for every $\alpha > 0$. In other words, since $N_{T_{\alpha^*}}$ has the same distribution as $V_{G(T_{\alpha^*})}$, $\mathbb{E}[N_{T_{\alpha^*}}]$ is always bounded by $\mathbb{E}[V_{G(\infty)}]$. This implies that $G(\infty)$ cannot be too small, as otherwise the Malthusian parameter would not exist, and the CTBP would die out $\mathbb{P}$-a.s.

The aging effect obviously slows down the birth process, and makes the limiting degree distribution have exponential tails for affine preferential attachment weights. One may wonder whether the power-law degree distribution could be restored when $(f_k)_{k \in \mathbb{N}}$ grows super-linearly instead. Here, we say that a sequence of weights $(f_k)_{k \in \mathbb{N}}$ grows super-linearly when $\sum_{k \geq 1} 1/f_k < \infty$ (see Definition 3.16). In the super-linear case, however, the branching process is *explosive*, i.e., for every individual the probability of generating an infinite number of children in finite time is 1. In this situation, the Malthusian parameter does not exist, since the Laplace transform of the process is always infinite. One could ask whether, by using an integrable aging function, this explosive behavior is destroyed. The answer to this question is given by the following theorem:

Theorem 2.3 (Explosive aging branching processes for super-linear attachment weights) *Consider a stationary process $(V_t)_{t \geq 0}$ defined by super-linear PA weights $(f_k)_{k \in \mathbb{N}}$. For any aging function g, the corresponding non-stationary process $(N_t)_{t \geq 0}$ is explosive.*

The proof of Theorem 2.3 is rather simple, and is given in Sect. 4.2. We investigate the case of affine PA weights $f_k = ak + b$ in more detail in Sect. 5.1. Under a hypothesis on the regularity of the integrable aging function, in Proposition 5.2, we give the asymptotic behavior of the corresponding limiting degree distribution. In particular, as $k \to \infty$,

$$p_k = C_1 \frac{\Gamma(k + b/a)}{\Gamma(k + 1)} e^{-C_2 k} \mathcal{G}(k, g)(1 + o(1)),$$

for some positive constants C_1, C_2. The term $\mathcal{G}(k, g)$ is a function of k, the aging function g and its derivative. The precise behavior of such term depends crucially on the aging function.

Apart from this, we notice that aging generates an exponential term in the distribution, which explains the two examples in Fig. 8. In Sect. A.2, we prove that the two limiting degree distributions in Fig. 8 indeed have exponential tails.

2.3 Results with Aging and Fitness

The analysis of birth processes becomes harder when we also consider fitness. First of all, we define the birth process with aging and fitness as follows:

Definition 2.4 (*Aging birth process with fitness*) Consider a birth process $(V_t)_{t \geq 0}$. Let $g \colon \mathbb{R}^+ \to \mathbb{R}^+$ be an aging function, and Y a positive random variable. The process $M_t := V_{YG(t)}$ is called a birth process with *aging and fitness*.

Definition 2.4 implies that the infinitesimal jump rates of the process $(M_t)_{t \geq 0}$ are as in (2.1), so that the birth probabilities of an individual depend on the PA weights, the age of the individual and on its fitness. Assuming that the process $(M_t)_{t \geq 0}$ is supercritical and Malthusian, we can prove the following theorem:

Theorem 2.5 (Limiting degree distribution for aging and fitness) *Consider a process* $(M_t)_{t \geq 0}$ *with integrable aging function* g, *fitnesses that are i.i.d. across the population, and assume that it is supercritical and Malthusian with Malthusian parameter* α^*. *Then, the limiting degree distribution for the corresponding branching process is given by*

$$p_k = \mathbb{E}\left[\frac{\alpha^*}{\alpha^* + f_k Y \hat{\mathcal{L}}(k, \alpha^*, Y)} \prod_{i=0}^{k-1} \frac{f_i Y \hat{\mathcal{L}}(i, \alpha^*, Y)}{\alpha^* + f_i Y \hat{\mathcal{L}}(i, \alpha^*, Y)} \right].$$

For a fixed individual, the distribution $(q_k)_{k \in \mathbb{N}}$ *of the number of children it generates over its entire lifetime is given by*

$$q_k = \mathbb{P}\left(V_{YG(\infty)} = k\right).$$

Similarly to Theorem 2.2, the sequence $(\hat{\mathcal{L}}(k, \alpha^*, Y))_{k \in \mathbb{N}}$ is given by

$$\hat{\mathcal{L}}(k, \alpha^*, Y) = \left(\frac{\mathcal{L}(\mathbb{P}\left(V_{uG(\cdot)} = k\right) g(\cdot))(\alpha^*)}{\mathcal{L}(\mathbb{P}\left(V_{uG(\cdot)} = k\right))(\alpha^*)} \right)_{u=Y},$$

where again $\mathcal{L}(h(\cdot))(\alpha)$ denotes the Laplace transform of a function h. Notice that in this case, with the presence of the fitness Y, this sequence is no longer deterministic but random instead. We still have the product structure for $(p_k)_{k \in \mathbb{N}}$ as in the stationary case, but now we have to average over the fitness distribution.

We point out that Theorem 2.2 is a particular case of Theorem 2.5, when we consider $Y \equiv 1$. We state the two results as separate theorems to improve the logic of the presentation. We prove Theorem 2.5 in Sect. 4.1. In Sect. 4.2 we show how Theorem 2.2 can be obtained from Theorem 2.5, and in particular how Condition (2.5) is obtained from the analogous Condition (2.6) stated below for general fitness distributions.

With affine PA weights, in Proposition 5.5, we can identify the asymptotics of the limiting degree distribution we obtain. This is proved by similar techniques as in the case of aging only, even though the result cannot be expressed so easily. In particular, we prove

$$p_k = \frac{\Gamma(k + b/a)}{\Gamma(b/a)\Gamma(k+1)} \frac{2\pi}{\sqrt{\det(k H_k(t_k, s_k))}} e^{-k\Psi_k(t_k, s_k)} \mathbb{P}\left(\mathcal{N}_1 \geq -t_k, \mathcal{N}_2 \geq -s_k\right)(1 + o(1)),$$

where the function $\Psi_k(t, s)$ depends on the aging function, the density μ of the fitness and k. The point (t_k, s_k) is the absolute minimum of $\Psi_k(t, s)$, $H_k(t, s)$ is the Hessian matrix of $\Psi_k(t, s)$, and $(\mathcal{N}_1, \mathcal{N}_2)$ is a bivariate normal vector with covariance matrix related to $H_k(t, s)$. We do not know the necessary and sufficient conditions for the existence of such a minimum (t_k, s_k). However, in Sect. 5.3, we consider two examples where we can apply this result, and we show that it is possible to obtain power-laws for them.

In the case of aging and fitness, the supercriticality condition in (2.5) is replaced by the analogous condition that

$$\mathbb{E}\left[V_{YG(t)}\right] < \infty \quad \text{for every } t \geq 0 \quad \text{and} \quad \lim_{t \to \infty} \mathbb{E}\left[V_{YG(t)}\right] > 1. \tag{2.6}$$

Borgs et al. [6] and Dereich [9, 10] prove results on stationary CTBPs with fitness. In these works, the authors investigate models with affine dependence on the degree and bounded fitness distributions. This is necessary since unbounded distributions with affine weights are explosive and thus *do not have Malthusian parameter*. We refer to Sect. 3.3 for a more precise discussion of the conditions on fitness distributions.

In the case of integrable aging and fitness, it is possible to consider affine PA weights, even with unbounded fitness distributions, as exemplified by (2.6). In particular, for $f_k = ak + b$,

$$\mathbb{E}[V_t] = \frac{b}{a}\left(e^{at} - 1\right).$$

As a consequence, Condition (2.6) can be written as

$$\forall t \geq 0 \quad \mathbb{E}\left[e^{aYG(t)}\right] < \infty \quad \text{and} \quad \lim_{t \to \infty} \mathbb{E}\left[e^{aYG(t)}\right] > 1 + \frac{a}{b}. \tag{2.7}$$

The expected value $\mathbb{E}\left[e^{aYG(t)}\right]$ is the moment generating function of Y evaluated in $aG(t)$. In particular, a necessary condition to have a Malthusian parameter is that the moment generating function is finite on the interval $[0, aG(\infty))$. As a consequence, denoting $\mathbb{E}[e^{sY}]$ by $\varphi_Y(s)$, we have effectively moved from the condition of having bounded distributions to the condition

$$\varphi_Y(x) < +\infty \quad \text{on} \quad [0, aG(\infty)), \quad \text{and} \quad \lim_{x \to aG(\infty)} \varphi_Y(x) > \frac{a+b}{a}. \tag{2.8}$$

Condition (2.8) is weaker than assuming a bounded distribution for the fitness Y, which means we can consider a larger class of distributions for the aging and fitness birth processes. Particularly for citation networks, it seems reasonable to have unbounded fitnesses, as the relative popularity of papers varies substantially.

2.4 Dynamical Power-Laws for Exponential Fitness and Integrable Aging

In Sect. 5.3 we introduce three different classes of fitness distributions, for which we give the asymptotics for the limiting degree distribution of the corresponding CTBP.

The first class is called *heavy-tailed*. Recalling (2.8), any distribution Y in this class satisfies, for any $t > 0$,

$$\varphi_Y(t) = \mathbb{E}\left[e^{tY}\right] = +\infty. \tag{2.9}$$

These distributions have a tail that is thicker than exponential. For instance, power-law distributions belong to this first class. Similarly to unbounded distributions in the stationary regime, such distributions generate *explosive* birth processes, independent of the choice of the integrable aging functions.

The second class is called *sub-exponential*. The density μ of a distribution Y in this class satisfies

$$\forall\,\beta > 0, \qquad \lim_{s \to +\infty} \mu(s)\mathrm{e}^{\beta s} = 0. \tag{2.10}$$

An example of this class is the density $\mu(s) = C\mathrm{e}^{-\theta s^{1+\varepsilon}}$, for some $\varepsilon, C, \theta > 0$. For such density, we show in Proposition 5.7 that the corresponding limiting degree distribution has a thinner tail than a power-law.

The third class is called *general-exponential*. The density μ of a distribution Y in this class is of the form

$$\mu(s) = Ch(s)\mathrm{e}^{-\theta s}, \tag{2.11}$$

where $h(s)$ is a twice differentiable function such that $h'(s)/h(s) \to 0$ and $h''(s)/h(s) \to 0$ as $s \to \infty$, and C is a normalization constant. For instance, exponential and Gamma distributions belong to this class. From (2.8), we know that in order to obtain a non-explosive process, it is necessary to consider the exponential rate $\theta > aG(\infty)$. We will see that the limiting degree distribution obeys a power law as $\theta > aG(\infty)$ with tails becoming thinner when θ increases.

For a distribution in the general exponential class, as proven in Proposition 5.6, the limiting degree distribution of the corresponding CTBP has a power-law term, with slowly-varying corrections given by the aging function g and the function h. We do not state Propositions 5.6 and 5.7 here, as these need notation and results from Sect. 5.1. For this reason, we only state the result for the special case of purely exponential fitness distribution:

Corollary 2.6 (Exponential fitness distribution) *Let the fitness distribution Y be exponentially distributed with parameter θ, and let g be an integrable aging function. Assume that the corresponding birth process $(M_t)_{t \geq 0}$ is supercritical and Malthusian. Then, the limiting degree distribution $(p_k)_{k \in \mathbb{N}}$ of the corresponding CTBP $\boldsymbol{M}$ is*

$$p_k = \mathbb{E}\left[\frac{\theta}{\theta + f_k G(T_{\alpha^*})} \prod_{i=0}^{k-1} \frac{f_i G(T_{\alpha^*})}{\theta + f_i G(T_{\alpha^*})} \right].$$

The distribution $(q_k)_{k \in \mathbb{N}}$ of the number of children of a fixed individual in its entire lifetime is given by

$$q_k = \frac{\theta}{\theta + G(\infty)f_k} \prod_{i=0}^{k-1} \frac{G(\infty)f_i}{\theta + G(\infty)f_i}.$$

Using exponential fitness makes the computation of the Laplace transform and the limiting degree distribution easier. We refer to Sect. 5.4 for the precise proof. In particular, the sequence defined in Corollary 2.6 is very similar to the limiting degree distribution of a stationary process with a bounded fitness. Let $(\xi_t^Y)_{t \geq 0}$ be a birth process with PA weights $(f_k)_{k \in \mathbb{N}}$ and fitness Y with bounded support. As proved in [10, Corollary 2.8], and as we show in Sect. 3.3, the limiting degree distribution of the corresponding branching process, assuming that $(\xi_t^Y)_{t \geq 0}$ is supercritical and Malthusian, has the form

$$p_k = \mathbb{E}\left[\frac{\alpha^*}{\alpha^* + Yf_k} \prod_{i=0}^{k-1} \frac{Yf_i}{\alpha^* + Yf_i} \right] = \mathbb{P}\left(\xi_{T_{\alpha^*}}^Y = k \right).$$

We notice the similarities with the limiting degree sequence given by Corollary 2.6. When g is integrable, the random variable $G(T_{\alpha^*})$ has bounded support. In particular, we can rewrite the sequence of the Corollary 2.6 as

$$p_k = \mathbb{P}\left(\xi_{T_\theta}^{G(T_{\alpha^*})} = k\right).$$

As a consequence, the limiting degree distribution of the process $(M_t)_{t\geq 0}$ equals that of a stationary process with fitness $G(T_{\alpha^*})$ and Malthusian parameter θ.

In the case where Y has exponential distribution and the PA weights are affine, we can also investigate the occurrence of *dynamical power laws*. In fact, with $(M_t)_{t\geq 0}$ such a process, the exponential distribution Y leads to

$$P_k[M](t) = \mathbb{P}\left(M_t = k\right) = \frac{\theta}{\theta + f_k G(t)} \prod_{i=0}^{k-1} \frac{f_i G(t)}{\theta + f_i G(t)}$$

$$= \frac{\theta}{aG(t)} \frac{\Gamma((b+\theta)/(aG(t)))}{\Gamma(aG(t))} \frac{\Gamma(k+b/(aG(t)))}{\Gamma(k + b/(aG(t)) + 1 + \theta/(aG(t)))}.$$
(2.12)

Here, M_t describes the number of children of an individual of age t. In other words, $(\mathbb{P}(M_t = k))_{k\in\mathbb{N}}$ is a distribution such that, as $k \to \infty$,

$$P_k[M](t) = \mathbb{P}\left(M_t = k\right) = k^{-(1+\theta/aG(t))}(1 + o(1)).$$

This means that for every time $t \geq 0$, the random variable M_t has a power-law distribution with exponent $\tau(t) = 1 + \theta/aG(t) > 2$. In particular, for every $t \geq 0$, M_t has finite expectation. We call this behavior where power laws occur that vary with the age of the individuals a *dynamical power law*. This occurs not only in the case of pure exponential fitness, but in general for every distribution as in (2.11), as shown in Proposition 5.6 below.

Further, we see that when $t \to \infty$, the dynamical power-law exponent coincides with the power-law exponent of the entire population. Indeed, the limiting degree distribution equals

$$p_k = \mathbb{E}\left[\theta/(aG(T_{\alpha^*})))\frac{\Gamma(\theta/(aG(T_{\alpha^*})) + b/(aG(T_{\alpha^*})))}{\Gamma(b/(aG(T_{\alpha^*})))} \frac{\Gamma(k + b/(aG(T_{\alpha^*})))}{\Gamma(k + b/(aG(T_{\alpha^*})) + 1 + \theta/(aG(T_{\alpha^*})))}\right].$$
(2.13)

In Fig. 9, we show a numerical example of the dynamical power-law for a process with exponential fitness distribution and affine weights. When time increases, the power-law expo-
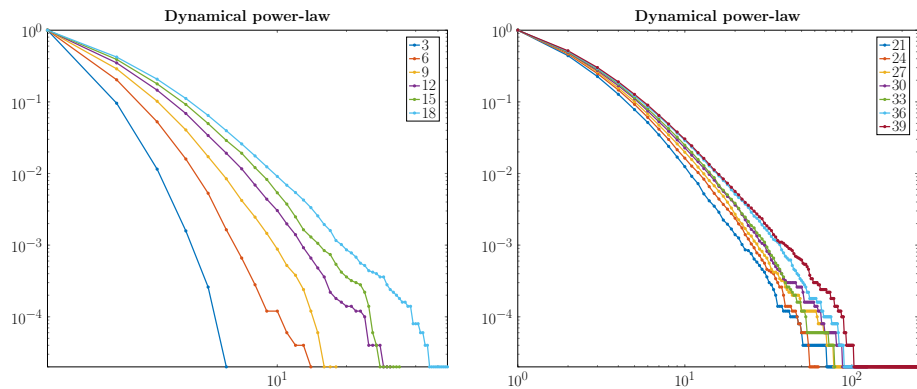

Fig. 9 Degree distribution for simulated processes with fitness, aging and affine weights. We considered $a = 1$, $b = 3.1567$, exponentially distributed fitness with parameter $\theta = 2.3866$, and normalized lognormal aging function. Simulations made at different times show the change of the power-law exponent
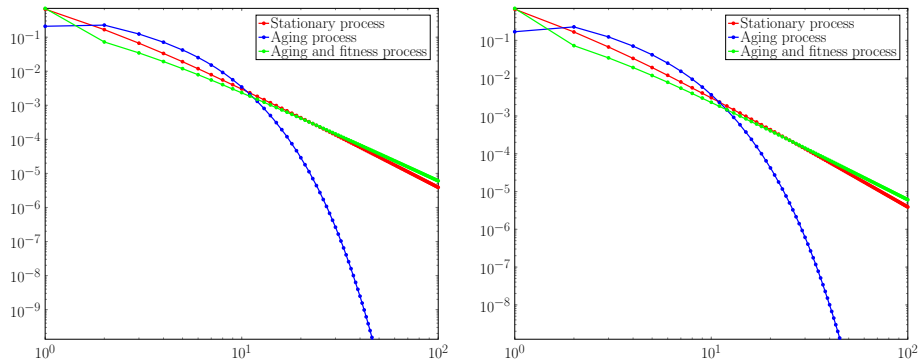
Fig. 10 Example of limiting degree distribution for branching processes

nent monotonically decreases to the limiting exponent $\tau \equiv \tau(\infty) > 2$, which means that the limiting distribution still has finite first moment. Note the similarity to the case of citation networks in Fig. 3.

When $t \to \infty$, the power-law exponent converges, and also M_t converges in distribution to a limiting random distribution M_∞ given by

$$q_k = \mathbb{P}\left(M_\infty = k\right) = \frac{\theta}{aG(\infty)} \frac{\Gamma((b+\theta)/(aG(\infty)))}{\Gamma(b/(aG(\infty)))} \frac{\Gamma(k+b/(aG(\infty)))}{\Gamma(k+b/(aG(\infty))+1+\theta/(aG(\infty)))}. \tag{2.14}$$

M_∞ has a power-law distribution, where the power-law exponent is

$$\tau = \lim_{t \to \infty} \tau(t) = 1 + \theta/(aG(\infty)) > 2.$$

In particular, since $\tau > 2$, a fixed individual has finite expected number of children also in its entire lifetime, unlike the stationary case with affine weights. In terms of citation networks, this type of processes predicts that papers do not receive an infinite number of citations after they are published (recall Fig. 5).

Figure 8 shows the effect of aging on the stationary process with affine weights, where the power-law is lost due to the aging effect. Thus, aging *slows down* the stationary process, and it is not possible to create the amount of high-degree vertices that are present in power-law distributions. Fitness can *speed up* the aging process to gain high-degree vertices, so that the power-law distribution is restored. This is shown in Fig. 10, where aging is combined with exponential fitness for the same aging functions as in Fig. 8.

In the stationary case, it is not possible to use unbounded distributions for the fitness to obtain a Malthusian process if the PA weights $(f_k)_{k \in \mathbb{N}}$ are affine. In fact, using unbounded distributions, the expected number of children at exponential time T_α is not finite *for any* $\alpha > 0$, i.e., the branching process is *explosive*. The aging effect allows us to relax the condition on the fitness, and the restriction to bounded distributions is relaxed to a condition on its moment generating function.

2.5 Conclusion and Open Problems

Beyond the Tree Setting In this paper, we only consider the *tree setting*, which is clearly unrealistic for citation networks. However, the analysis of PAMs has shown that the qualitative features of the degree distribution for PAMs are identical to those in the tree setting. Proving

this remains an open problem that we hope to address hereafter. Should this indeed be the case, then we could summarize our findings in the following simple way: The power-law tail distribution of PAMs is destroyed by integrable aging, and cannot be restored either by super-linear weights or by adding bounded fitnesses. However, it *is* restored by *unbounded* fitnesses with at most an exponential tail. Part of these results are example based, while we have general results proving that the limiting degree distribution exists.

Structure of the Paper The present paper is organized as follows. In Sect. 3, we quote general results on CTBPs, in particular Theorem 3.10 that we use throughout our proofs. In Sect. 3.2, we describe known properties of the stationary regime. In Sect. 3.3, we briefly discuss the Malthusian parameter, focusing on conditions on fitness distributions to obtain supercritical processes. In Sect. 4, we prove Theorems 2.3 and 2.5, and we show how Theorem 2.2 is a particular case of Theorem 2.5. In Sect. 5 we specialize to the case of affine PA function, giving precise asymptotics.

3 General Theory of Continuous-Time Branching Processes

3.1 General Set-Up of the Model

In this section we present the general theory of continuous-time branching processes (CTBPs). In such models, individuals produce children according to i.i.d. copies of the same birth process. We now define birth processes in terms of point processes:

Definition 3.1 (*Point process*) A *point process* ξ is a random variable from a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ to the space of integer-valued measures on $\mathbb{R}^+$.

A point process ξ is defined by a sequence of positive real-valued random variables $(T_k)_{k \in \mathbb{N}}$. With abuse of notation, we can denote the density of the point process ξ by

$$\xi(dt) = \sum_{k \in \mathbb{N}} \delta_{T_k}(dt),$$

where $\delta_x(dt)$ is the delta measure in x, and the random measure ξ evaluated on $[0, t]$ as

$$\xi(t) = \xi([0, t]) = \sum_{k \in \mathbb{N}} \mathbb{1}_{[0,t]}(T_k).$$

We suppose throughout the paper that $T_k < T_{k+1}$ with probability 1 for every $k \in \mathbb{N}$.

Remark 3.2 Equivalently, considering a sequence $(T_k)_{k \in \mathbb{N}}$ (where $T_0 = 0$) of positive real-valued random variables, such that $T_k < T_{k+1}$ with probability 1, we can define

$$\xi(t) = \xi([0, t]) = k \quad \text{when} \quad t \in [T_k, T_{k+1}).$$

We will often define a point process from the jump-times sequence of an integer-valued process $(V_t)_{t \geq 0}$. For instance, consider $(V_t)_{t \geq 0}$ as a Poisson process, and denote $T_k = \inf\{t > 0 : V_t \geq k\}$. Then we can use the sequence $(T_k)_{k \in \mathbb{N}}$ to define a point process ξ. The point process defined from the jump times of a process $(V_t)_{t \geq 0}$ will be denoted by ξ_V.

We now introduce some notation before giving the definition of CTBP. We denote the set of individuals in the population using Ulam-Harris notation for trees. The set of individuals

is

$$\mathcal{N} = \bigcup_{n \in \mathbb{N}} \mathbb{N}^n.$$

For $x \in \mathbb{N}^n$ and $k \in \mathbb{N}$ we denote the k-th child of x by $xk \in \mathbb{N}^{n+1}$. This construction is well known, and has been used in other works on branching processes (see [14,17,22] for more details).

We now are ready to define our branching process:

Definition 3.3 (*Continuous-time branching process*) Given a point process ξ, we define the CTBP associated to ξ as the pair of a probability space

$$(\Omega, \mathcal{A}, \mathbb{P}) = \prod_{x \in \mathcal{N}} (\Omega_x, \mathcal{A}_x, \mathbb{P}_x),$$

and an infinite set $(\xi^x)_{x \in \mathcal{N}}$ of i.i.d. copies of the process ξ. We will denote the branching process by $\boldsymbol{\xi}$.

Remark 3.4 (Point processes and their jump times) Throughout the paper, we will define point processes in terms of jump times of processes $(V_t)_{t \geq 0}$. In order to keep the notation light, we will denote branching processes defined by point processes given by jump times of the process V_t by $\boldsymbol{V}$. To make it more clear, by $\boldsymbol{V}$ we denote a probability space as in Definition 3.3 and an infinite set of measures $(\xi_V^x)_{x \in \mathbb{N}}$, where ξ_V is the point process defined by the process V.

According to Definition 3.3, a branching process is a pair of a probability space and a sequence of random measures. It is possible though to define an *evolution* of the branching population. At time $t = 0$, our population consists only of the root, denoted by $\varnothing$. Every time t an individual x gives birth to its k-th child, i.e., $\xi^x(t) = k + 1$, assuming that $\xi^x(t-) = k$, we start the process ξ^{xk}. Formally:

Definition 3.5 (*Population birth times*) We define the sequence of birth times for the process $\boldsymbol{\xi}$ as $\tau_\varnothing^\xi = 0$, and for $x \in \mathcal{N}$,

$$\tau_{xk}^\xi = \tau_x^\xi + \inf \left\{ s \geq 0 : \xi^x(s) \geq k \right\}.$$

In this way we have defined the set of individuals, their birth times and the processes according to which they reproduce. We still need a way to count how many individuals are alive at a certain time t.

Definition 3.6 (*Random characteristic*) A *random characteristic* is a real-valued process $\Phi \colon \Omega \times \mathbb{R} \to \mathbb{R}$ such that $\Phi(\omega, s) = 0$ for any $s < 0$, and $\Phi(\omega, s) = \Phi(s)$ is a deterministic bounded function for every $s \geq 0$ that only depends on ω through the birth time of the individual, as well as the birth process of its children.

An important example of a random characteristic is obtained by the function $\mathbb{1}_{\mathbb{R}^+}(s)$, which measures whether the individual has been born at time s. Another example is $\mathbb{1}_{\mathbb{R}^+}(s)\mathbb{1}_{\{k\}}(\xi)$, which measures whether the individual has been born or not at time s and whether it has k children presently.

For each individual $x \in \mathcal{N}$, $\Phi_x(\omega, s)$ denotes the value of Φ evaluated on the progeny of x, regarding x as ancestor, when the age of x is s. In other words, $\Phi_x(\omega, s)$ is the evaluation of Φ on the tree rooted at x, ignoring the rest of the population. If we do not specify the individual x, then we assume that $\Phi = \Phi_\varnothing$. We use random characteristics to describe the properties of the branching population.

Definition 3.7 (*Evaluated branching processes*) Consider a random characteristic Φ as in Definition 3.6. We define the evaluated branching processes with respect to Φ at time $t \in \mathbb{R}^+$ as

$$\xi_t^\Phi = \sum_{x \in \mathcal{N}} \Phi_x(t - \tau_x^\xi).$$

The meaning of the evaluated branching process is clear when we consider the random characteristic $\Phi(t) = \mathbb{1}_{\mathbb{R}^+}(t)$, for which

$$\xi_t^{\mathbb{1}_{\mathbb{R}^+}} = \sum_{x \in \mathcal{N}} (\mathbb{1}_{\mathbb{R}^+})_x(t - \tau_x^\xi),$$

which is the number of $x \in \mathcal{N}$ such that $t - \tau_x^\xi \geq 0$, i.e., the total number of individuals already born up to time t. Another characteristic that we consider in this paper is, for $k \in \mathbb{N}$, $\Phi_k(t) = \mathbb{1}_{\{k\}}(\xi_t)$, for which

$$\xi_t^{\Phi_k} = \sum_{x \in \mathcal{N}} \mathbb{1}_{\{k\}}\left(\xi_{t-\tau_x^\xi}^x\right)$$

is the number of individuals with k children at time t.

As known from the literature, the properties of the branching process are determined by the behavior of the point process ξ. First of all, we need to introduce some notation. Consider a function $f : \mathbb{R}^+ \to \mathbb{R}$. We denote the Laplace transform of f by

$$\mathcal{L}(f(\cdot))(\alpha) = \int_0^\infty \mathrm{e}^{-\alpha t} f(t)dt.$$

With a slight abuse of notation, if μ is a positive measure on $\mathbb{R}^+$, then we denote

$$\mathcal{L}(\mu(d\cdot))(\alpha) = \int_0^\infty \mathrm{e}^{-\alpha t} \mu(dt).$$

We use the Laplace transform to analyze the point process ξ:

Definition 3.8 (*Supercritical property*) Consider a point process ξ on $\mathbb{R}^+$. We say ξ is *supercritical* when there exists $\alpha^* > 0$ such that

$$\mathcal{L}(\mathbb{E}\xi(d\cdot))(\alpha^*) = \int_0^\infty \mathrm{e}^{-\alpha^* t} \mathbb{E}\xi(dt) = \sum_{k \in \mathbb{N}} \mathbb{E}\left[\int_0^\infty \mathrm{e}^{-\alpha^* t} \delta_{T_k}(dt)\right] = \sum_{k \in \mathbb{N}} \mathbb{E}\left[\mathrm{e}^{-\alpha^* T_k}\right] = 1.$$

We call α^* the *Malthusian parameter* of the process ξ.

We point out that $\mathbb{E}\xi(d\cdot)$ is an abuse of notation to denote the density of the *averaged* measure $\mathbb{E}[\xi([0, t])]$. A second fundamental property for the analysis of branching processes is the following:

Definition 3.9 (*Malthusian property*) Consider a supercritical point process ξ, with Malthusian parameter α^*. The process ξ is *Malthusian* when

$$-\frac{d}{d\alpha}\left(\mathcal{L}(\mathbb{E}\xi(dt))\right)(\alpha)\bigg|_{\alpha^*} = \int_0^\infty t\mathrm{e}^{-\alpha^* t} \mathbb{E}\xi(d\cdot) = \sum_{k \in \mathbb{N}} \mathbb{E}\left[T_k \mathrm{e}^{-\alpha^* T_k}\right] < \infty.$$

We denote

$$\tilde{\alpha} = \inf\{\alpha > 0 \;:\; \mathcal{L}\left(\mathbb{E}\xi(d\cdot)\right)(\alpha) < \infty\}, \tag{3.1}$$

and we will also assume that the process satisfies the condition

$$\lim_{\alpha \searrow \tilde{\alpha}} \mathcal{L}\left(\mathbb{E}\xi(d\cdot)\right)(\alpha) > 1. \tag{3.2}$$

Integrating by parts, it is possible to show that, for a point process ξ,

$$\mathcal{L}\left(\mathbb{E}\xi(d\cdot)\right)(\alpha) = \mathbb{E}\left[V_{T_\alpha}\right],$$

where T_α is an exponentially distributed random variable independent of the process $(V_t)_{t\geq 0}$. Heuristically, the Laplace transform of a point process ξ_V is the expected number of children born at exponentially distributed time T_α. In this case the Malthusian parameter is the exponential rate α^* such that at time T_{α^*} exactly one children has been born.

These two conditions are required to prove the main result on branching processes that we rely upon:

Theorem 3.10 (Population exponential growth) *Consider the point process ξ, and the corresponding branching process $\boldsymbol{\xi}$. Assume that ξ is supercritical and Malthusian with parameter α^*, and suppose that there exists $\bar{\alpha} < \alpha^*$ such that*

$$\int_0^\infty \mathrm{e}^{-\bar{\alpha}t}\mathbb{E}\xi(dt) < \infty.$$

Then

(1) *there exists a random variable Θ such that as $t \to \infty$,*

$$\mathrm{e}^{-\alpha^* t}\boldsymbol{\xi}_t^{\mathbb{1}_{\mathbb{R}^+}} \xrightarrow{\mathbb{P}-as} \Theta; \tag{3.3}$$

(2) *for any two random characteristics Φ and Ψ,*

$$\frac{\boldsymbol{\xi}_t^\Phi}{\boldsymbol{\xi}_t^\Psi} \xrightarrow{\mathbb{P}-as} \frac{\mathcal{L}(\mathbb{E}[\Phi(\cdot)])(\alpha^*)}{\mathcal{L}(\mathbb{E}[\Psi(\cdot)])(\alpha^*)}. \tag{3.4}$$

This result is stated in [22, Theorem A], which is a weaker version of [17, Theorem 6.3]. Formula (3.3) implies that, $\mathbb{P}$-a.s., the population size grows exponentially with time. It is relevant though to give a description of the distribution of the random variable Θ:

Theorem 3.11 (Positivity of Θ) *Under the hypothesis of Theorem 3.10, if*

$$\mathbb{E}\left[\mathcal{L}(\xi(d\cdot))(\alpha^*)\log^+\left(\mathcal{L}(\xi(d\cdot))(\alpha^*)\right)\right] < \infty, \tag{3.5}$$

then, on the event $\{\boldsymbol{\xi}_t^{\mathbb{1}_{\mathbb{R}^+}} \to \infty\}$, i.e., on the event that the branching population keeps growing in time, the random variable Θ in (3.3) is positive with probability 1, and $\mathbb{E}[\Theta] = 1$. Otherwise, $\Theta = 0$ with probability 1. Condition (3.5) is called the (xlogx) *condition.*

This result is proven in [14, Theorem 5.3], and it is the CTBPs equivalent of the Kesten-Stigum theorem for Galton-Watson processes ([16, Theorem 1.1]).

Formula (3.4) says that the ratio between the evaluation of the branching process with two different characteristics converges $\mathbb{P}$-a.s. to a constant that depends only on the two characteristics involved. In particular, if we consider, for $k \in \mathbb{N}$,

$$\Phi(t) = \mathbb{1}_{\{k\}}(\xi_t), \quad \text{and} \quad \Psi(t) = \mathbb{1}_{\mathbb{R}^+}(t),$$

then Theorem 3.10 gives

$$\frac{\xi_t^{\Phi}}{\xi_t^{\mathbb{1}_{\mathbb{R}^+}}} \xrightarrow{\mathbb{P}-as} \alpha^* \mathcal{L}(\mathbb{P}(\xi(\cdot) = k))(\alpha^*), \tag{3.6}$$

since $\mathcal{L}(\mathbb{E}[\mathbb{1}_{\mathbb{R}^+}(\cdot)])(\alpha^*) = 1/\alpha^*$. The ratio in the previous formula is the fraction of individuals with k children in the whole population:

Definition 3.12 (*Limiting degree distribution for CTBP*) The sequence $(p_k)_{k \in \mathbb{N}}$, where

$$p_k = \alpha^* \mathcal{L}(\mathbb{P}(\xi(\cdot) = k))(\alpha^*) = \alpha^* \int_0^\infty e^{-\alpha^* t} \mathbb{P}(\xi(t) = k) \, dt$$

is the *limiting degree distribution* for the branching process $\boldsymbol{\xi}$.

The aim of the following sections will be to study when point processes satisfy the conditions of Theorem 3.10, in order to analyze the limiting degree distribution in Definition 3.12.

3.2 Stationary Birth Processes with No Fitness

In this section we present the theory of birth processes that are stationary and have deterministic rates. This is relevant since the definition of aging processes starts with a stationary process. In particular, we give description of the affine case, which plays a central role in the present work:

Definition 3.13 (*Stationary non-fitness birth processes*) Consider a non-decreasing sequence $(f_k)_{k \in \mathbb{N}}$ of positive real numbers. A *stationary non-fitness birth process* is a stochastic process $(V_t)_{t \geq 0}$ such that

(1) $V_0 = 0$, and $V_t \in \mathbb{N}$ for all $t \in \mathbb{R}^+$;
(2) $V_t \leq V_s$ for every $t \leq s$;
(3) for h small enough,

$$\mathbb{P}(V_{t+h} = k + 1 \mid V_t = k) = f_k h + o(h), \text{ and for } j \geq 2,$$
$$\mathbb{P}(V_{t+h} = k + j \mid V_t = k) = o(h^2). \tag{3.7}$$

We denote the jump times by $(T_k)_{k \in \mathbb{N}}$, i.e.,

$$T_k = \inf\{t \geq 0 : V_t \geq k\}.$$

We denote the point process corresponding to $(V_t)_{t \geq 0}$ by ξ_V. In this case, $(V_t)_{t \geq 0}$ is an inhomogeneous Poisson process, and for every $k \in \mathbb{N}$, $T_{k+1} - T_k$ has exponential law with parameter f_k independent of $(T_{h+1} - T_h)_{h=0}^{k-1}$. It is possible to show the following proposition:

Proposition 3.14 (Probabilities for $(V_t)_{t \geq 0}$) *Consider a stationary non-fitness birth process $(V_t)_{t \geq 0}$. Denote, for every $k \in \mathbb{N}$, $\mathbb{P}(V_t = k) = P_k[V](t)$. Then*

$$P_0[V](t) = \exp(-f_0 t), \tag{3.8}$$

and, for $k \geq 1$,

$$P_k[V](t) = f_{k-1} \exp(-f_k t) \int_0^t \exp(f_k x) P_{k-1}[V](x) dx. \tag{3.9}$$

For a proof, see [4, Chap. 3, Sect. 2]. From the jump times, it is easy to compute the explicit expression for the Laplace transform of ξ_V as

$$\mathcal{L}(\mathbb{E}\xi_V(d\cdot))(\alpha) = \sum_{k\in\mathbb{N}} \mathbb{E}\left[\int_0^\infty e^{-\alpha t}\delta_{T_k}(dt)\right] = \sum_{k\in\mathbb{N}} \mathbb{E}\left[e^{-\alpha T_k}\right] = \sum_{k\in\mathbb{N}}\prod_{i=0}^{k-1}\frac{f_i}{\alpha+f_i},$$

since every T_k can be seen as sum of independent exponential random variables with parameters given by the sequence $(f_k)_{k\in\mathbb{N}}$. Assuming now that ξ_V is supercritical and Malthusian with parameter α^*, we have the explicit expression for the limit distribution $(p_k)_{k\in\mathbb{N}}$, given by (2.2).

An analysis of the behavior of the limit distribution of branching processes is presented in [2] and [21], where the authors prove that $(p_k)_{k\in\mathbb{N}}$ has a power-law tail only if the sequence of rates $(f_k)_{k\in\mathbb{N}}$ is asymptotically linear with respect to k.

Proposition 3.15 (Characterization of stationary and linear process V) *Consider the sequence $f_k = ak + b$. Then:*

(1) *for every $\alpha \in \mathbb{R}^+$,*

$$\mathcal{L}(\mathbb{E}\xi_V(d\cdot))(\alpha) = \frac{\Gamma(\alpha^*/a + b/a)}{\Gamma(b/a)}\sum_{k\in\mathbb{N}}\frac{\Gamma(k+b/a)}{\Gamma(k+b/a+\alpha/a)} = \frac{b}{\alpha-a}.$$

(2) *The Malthusian parameter is $\alpha^* = a + b$, and $\tilde{\alpha} = a$, where $\tilde{\alpha}$ is defined as in (3.1).*
(3) *The derivative of the Laplace transform is*

$$-\frac{b}{(\alpha-a)^2},$$

   which is finite whenever $\alpha > a$;
(4) *The process $(V_t)_{t\geq 0}$ satisfies the (xlogx) condition (3.5).*

Proof The proof can be found in [22, Theorem 2], or [2, Theorem 2.6]. ☐

For affine PA weights $(f_k)_{k\in\mathbb{N}} = (ak+b)_{k\in\mathbb{N}}$, the Malthusian parameter α^* exists. Since $\alpha^* = a + b$, the limiting degree distribution of the branching process V is given by

$$p_k = (1+b/a)\frac{\Gamma(1+2b/a)}{\Gamma(b/a)}\frac{\Gamma(k+b/a)}{\Gamma(k+b/a+2+b/a)}. \tag{3.10}$$

Notice that p_k has a power-law decay with exponent $\tau = 2 + \frac{b}{a}$. Branching processes of this type are related to PAM, also called the Barabási-Albert model [1]. This model shows the so-called *old-get-richer* effect. Clearly this is not true for real-world citation networks. In Fig. 5, we notice that, on average, the increment of the citation received by old papers is smaller than the increment of younger papers. Rephrasing it, old papers tend to be cited less and less over time.

3.3 The Malthusian Parameter

The existence of the Malthusian parameter is a necessary condition to have a branching process growing at exponential rate. In particular, the Malthusian parameter does not exist in two cases: when the process is subcritical and grows slower than exponential, or when it is explosive. In the first case, the branching population might either die out or grow indefinitely with positive probability, but slower than at exponential rate. In the second case, the population

size explodes in finite time with probability one. In both cases, the behavior of the branching population is different from what we observe in citation networks (Fig. 1). For this reason, we focus on supercritical processes, i.e., on the case where the Malthusian parameter exists.

Denote by $(V_t)_{t \geq 0}$ a stationary birth process defined by PA weights $(f_k)_{k \in \mathbb{N}}$. In general, we assume $f_k \to \infty$. Denote the sequence of jump times by $(T_k)_{k \in \mathbb{N}}$. As we quote in Sect. 3.2, the Laplace transform of a birth process $(V_t)_{t \geq 0}$ is given by

$$\mathcal{L}(\mathbb{E}V(d\cdot))(\alpha) = \mathbb{E}\left[\sum_{k \in \mathbb{N}} e^{-\alpha T_k}\right] = \mathbb{E}\left[V_{T_\alpha}\right] = \sum_{k \in \mathbb{N}} \prod_{i=0}^{k-1} \frac{f_i}{\alpha + f_i}.$$

Such expression comes from the fact that, in stationary regime, T_k is the sum of k independent exponential random variables. We can write

$$\sum_{k \in \mathbb{N}} \exp\left(-\sum_{i=0}^{k-1} \log\left(1 + \frac{\alpha}{f_i}\right)\right) = \sum_{k \in \mathbb{N}} \exp\left(-\alpha \sum_{i=0}^{k-1} \frac{1}{f_i}(1 + o(1))\right).$$

The behavior of the Laplace transform depends on the asymptotic behavior of the PA weights. We define now the terminology we use:

Definition 3.16 (*Superlinear PA weights*) Consider a PA weight sequence $(f_k)_{k \in \mathbb{N}}$. We say that the PA weights are *superlinear* if $\sum_{i=0}^{\infty} 1/f_i < \infty$.

As a general example, consider $f_k = ak^q + b$, where $q > 0$. In this case, the sequence is affine when $q = 1$, superlinear when $q > 1$ and sublinear when $q < 1$.

When the weights are superlinear, since $C = \sum_{i=0}^{\infty} 1/f_i < \infty$, we have

$$\sum_{k \in \mathbb{N}} \exp\left(-\alpha \sum_{i=0}^{k-1} \frac{1}{f_i}(1 + o(1))\right) \geq \sum_{k \in \mathbb{N}} \exp(-\alpha C) = +\infty. \tag{3.11}$$

This holds for every $\alpha > 0$. As a consequence, the Laplace transform $\mathcal{L}(\mathbb{E}V(d\cdot))(\alpha)$ is always infinite, and there exist no Malthusian parameter. In particular, if we denote by $T_\infty = \lim_{k \to \infty} T_k$, then $T_\infty < \infty$ a.s. This means that the birth process $(V_t)_{t \geq 0}$ explodes in a finite time.

When the weights are at most linear, the bound in (3.11) does not hold anymore. In fact, consider as example affine weights $f_k = ak + b$. We have that $\sum_{i=0}^{k-1} \frac{1}{f_i} = (1/a) \log k (1 + o(1))$. As a consequence, the Laplace transform can be written as

$$\sum_{k \in \mathbb{N}} \exp\left(-\frac{\alpha}{a} \log k (1 + o(1))\right) = \sum_{k \in \mathbb{N}} k^{-\frac{\alpha}{a}}(1 + o(1)). \tag{3.12}$$

In this case, the Laplace transform is finite for $\alpha > a$. For the sublinear case, for which $\sum_{i=0}^{k-1} 1/f_i = Ck^{(1-q)}(1 + o(1))$, we obtain

$$\sum_{k \in \mathbb{N}} \exp\left(-C\alpha k^{1-q}\right).$$

This sum is finite for any $\alpha > 0$.

We can now introduce fitness in the stationary process:

Remark 3.17 Consider the process $(V_t)_{t \geq 0}$ defined by the sequence of PA weights $(f_k)_{k \in \mathbb{N}}$ as in Sect. 3.2. For $u \in \mathbb{R}^+$ we denote by $(V_t^u)_{t \geq 0}$ the process defined by the sequence $(uf_k)_{k \in \mathbb{N}}$. It is easy to show that

$$\mathcal{L}(\mathbb{E}\xi_{V^u}(d\cdot))(\alpha) = \mathcal{L}(\mathbb{E}\xi_V(d\cdot))(\alpha/u).$$

The behavior of the degree sequence of $(V_t^u)_{t \geq 0}$ is the same of the process V_t.

Remark 3.17 shows a sort of monotonicity of the Laplace transform with respect to the sequence $(f_k)_{k \in \mathbb{N}}$. This is very useful to describe the Laplace transform of a birth process with fitness, which we define now:

Definition 3.18 (*Stationary fitness birth processes*) Consider a birth process $(V_t)_{t \geq 0}$ defined by a sequence of weights $(f_k)_{k \in \mathbb{N}}$. Let Y be a positive random variable. We call stationary fitness birth processes the process $(V_t^Y)_{t \geq 0}$, defined by the random sequence of weights $(Y f_k)_{k \in \mathbb{N}}$, i.e., conditionally on Y,

$$\mathbb{P}\left(V_{t+h}^Y = k + 1 \mid V_t^Y = k, Y\right) = Y f_k h + o(h).$$

By Definition 3.18, it is obvious that the properties of the process $(V_t^Y)_{t \geq 0}$ are related to the properties of $(V_t)_{t \geq 0}$. Since we consider a random fitness Y independent of the process $(V_t)_{t \geq 0}$, from Remark 3.17 it follows that

$$\mathcal{L}(\mathbb{E}V^Y(d\cdot))(\alpha) = \mathbb{E}\left[\mathcal{L}(\mathbb{E}\xi_{V^u}(d\cdot))(\alpha)_{u=Y}\right] = \mathbb{E}\left[\sum_{k \in \mathbb{N}} \prod_{i=0}^{k-1} \frac{Y f_i}{\alpha + Y f_i}\right]. \tag{3.13}$$

For affine weights the fitness distribution needs to be bounded, as discussed in Sect. 2.4. In this section we give a qualitative explanation of this fact. Consider the sum in the expectation in the right hand term of (3.13). We can rewrite the sum as

$$\sum_{k \in \mathbb{N}} \prod_{i=0}^{k-1} \frac{Y f_i}{\alpha + Y f_i} = \sum_{k \in \mathbb{N}} \exp\left(-\sum_{i=0}^{k-1} \log\left(1 + \frac{\alpha}{Y f_i}\right)\right) = \sum_{k \in \mathbb{N}} \exp\left(-\frac{\alpha}{Y} \sum_{i=0}^{k-1} \frac{1}{f_i}(1 + o(1))\right). \tag{3.14}$$

The behavior depends sensitively on the asymptotic behavior of the PA weights. In particular, a necessary condition for the existence of the Malthusian parameter is that the sum in (3.13) is finite on an interval of the type $(\tilde{\alpha}, +\infty)$. In other words, since the Laplace transform is a decreasing function (when finite), we need to prove the existence of a minimum value $\tilde{\alpha}$ such that it is finite for every $\alpha > \tilde{\alpha}$. Using (3.14) in (3.13), we just need to find a value α such that the right hand side of (3.14) equals 1.

In the case of affine weights $f_k = ak + b$, we have $\sum_{i=0}^{k-1} \frac{1}{f_i} = C \log k(1 + o(1))$, for a constant C. As a consequence, (3.14) is equal to

$$\mathbb{E}\left[\sum_{k \in \mathbb{N}} \exp\left(-C\frac{\alpha}{Y}\log k\right)\right] = \mathbb{E}\left[\sum_{k \in \mathbb{N}} k^{-C\alpha/Y}\right]. \tag{3.15}$$

The sum inside the last expectation is finite only on the event $\{Y < C\alpha\}$. If Y has an unbounded distribution, then for every value of $\alpha > 0$ we have that $\{Y \geq C\alpha\}$ is an event of positive probability. As a consequence, for every $\alpha > 0$, the Laplace transform of the birth process $(V_y^Y)_{t \geq 0}$ is infinite, which means there exists no Malthusian parameter.

This is why a bounded fitness distribution is necessary to have a Malthusian parameter using affine PA weights. The situation is different in the case of sublinear weights. For example, consider $f_k = (1 + k)^q$, where $q \in (0, 1)$. Then, the difference to affine weights is that now $\sum_{i=0}^{k-1} 1/f_i = Ck^{1-q}(1 + o(1))$. Using this in (3.14), we obtain

$$\mathbb{E}\left[\sum_{k \in \mathbb{N}} \exp\left(-C\frac{\alpha}{Y}k^{(1-q)}\right)\right].$$

In this case, since both α and Y are always positive, the last sum is finite with probability 1, and the expectation might be finite under appropriate moment assumptions on Y.

Assume now that the fitness Y satisfies the necessary conditions, so that the process $(V_t^Y)_{t \geq 0}$ is supercritical and Malthusian with parameter α^*. We can evaluate the limiting degree distribution. Conditioning on Y, the Laplace transform of $\mathbb{E}\xi_{V^Y}(dx)$ is

$$\sum_{k \in \mathbb{N}} \prod_{i=0}^{k-1} \frac{Y f_i}{\alpha + Y f_i},$$

so, as a consequence, the limiting degree distribution of the branching processes is

$$p_k = \mathbb{E}\left[\frac{\alpha^*}{\alpha^* + Y f_k} \prod_{i=0}^{k-1} \frac{Y f_i}{\alpha^* + Y f_i}\right]. \tag{3.16}$$

It is possible to see that the right-hand side of (3.16) is similar to the distribution of the simpler case with no fitness given by (2.2). We still have a product structure for the limit distribution, but in the fitness case it has to be averaged over the fitness distribution. This result is similar to [10, Theorem 2.7, Corollary 2.8].

Considering affine weights $f_k = ak + b$, we can rewrite (3.16) as

$$p_k = \mathbb{E}\left[\frac{\Gamma((\alpha^* + b)/(aY))}{\Gamma(b/(aY))} \frac{\Gamma(k + b/(aY))}{\Gamma(k + b/(aY) + 1 + \alpha^*/(aY))}\right].$$

Asymptotically in k, the argument of the expectation in the previous expression is random with a power-law exponent $\tau(Y) = 1 + \alpha^*/(aY)$. For example, in this case averaging over the fitness distribution, it is possible to obtain power-laws with logarithmic corrections (see eg [5, Corollary 32]).

4 Existence of Limiting Distributions

In this section, we give the proof of Theorems 2.2, 2.3 and 2.5, proving that the branching processes defined in Sect. 2 do have a limiting degree distribution. As mentioned, we start by proving Theorem 2.5, and then explain how Theorem 2.2 follows as special case.

Before proving the result, we do need some remarks on the processes we consider. Birth process with aging alone and aging with fitness are defined respectively in Definition 2.1 and 2.4. Consider then a process with aging and fitness $(M_t)_{t \geq 0}$ as in Definition 2.4. Let $(T_k)_{k \in \mathbb{N}}$ denote the sequence of birth times, i.e.,

$$T_k = \inf\{t \geq 0 : M_t \geq k\}.$$

It is an immediate consequence of the definition that, for every $k \in \mathbb{N}$,

$$\mathbb{P}(T_k \leq t) = \mathbb{P}(\bar{T}_k \leq Y G(t)), \tag{4.1}$$

where $(\bar{T}_k)_{k \in \mathbb{N}}$ is the sequence of birth times of a stationary birth process $(V_t)_{t \geq 0}$ defined bu the same PA function f. Consider then the sequence of functions $(P_k[V](t))_{k \in \mathbb{N}}$ associated with the stationary process $(V_t)_{t \geq 0}$ defined by the same sequence of weights $(f_k)_{k \in \mathbb{N}}$ (see Proposition 3.14). As a consequence, for every $k \in \mathbb{N}$, $\mathbb{P}(M_t = k) = \mathbb{E}[P_k[V](Y G(t))]$, and the same holds for an aging process just considering $Y \equiv 1$. Formula (4.1) implies that the aging process is the stationary process with a deterministic time-change given by $G(t)$. A

process with aging and fitness is the stationary process with a random time-change given by $YG(t)$.

Assume now that g is integrable, i.e. $\lim_{t\to\infty} G(t) = G(\infty) < \infty$. Using (4.1) we can describe the limiting degree distribution $(q_k)_{k\in\mathbb{N}}$ of a fixed individual in the branching population, i.e., the distribution N_∞ (or M_∞) of the total number of children an individual will generate in its entire lifetime. In fact, for every $k \in \mathbb{N}$,

$$\lim_{t\to\infty} \mathbb{P}\left(N_t = k\right) = \lim_{t\to\infty} P[V](G(t)) = \mathbb{P}\left(V_{G(\infty)} = k\right), \tag{4.2}$$

which means that N_∞ has the same distribution as $V_{G(\infty)}$. With fitness,

$$\lim_{t\to\infty} \mathbb{P}\left(M_t = k\right) = \lim_{t\to\infty} \mathbb{E}[P[V](YG(t))] = \mathbb{P}\left(V_{YG(\infty)} = k\right).$$

For example, in the case of aging only, this is rather different from the stationary case, where the number of children of a fixed individual diverges as the individual gets old (see e.g [2, Theorem 2.6]).

4.1 Proof of Theorem 2.5

Birth processes with continuous aging effect and fitness are defined in Definition 2.4. We now identify conditions on the fitness distribution to have a Malthusian parameter:

Lemma 4.1 (Condition (2.6)) *Consider a stationary process $(V_t)_{t\geq0}$, an integrable aging function g and a random fitness Y. Assume that $\mathbb{E}[V_t] < \infty$ for every $t \geq 0$. Then the process $(V_{YG(t)})_{t\geq0}$ is supercritical if and only if Condition (2.6) holds, i.e.,*

$$\mathbb{E}\left[V_{YG(t)}\right] < \infty \ \ \text{for every } t \geq 0 \quad \text{and} \quad \lim_{t\to\infty} \mathbb{E}\left[V_{YG(t)}\right] > 1.$$

Proof For the if part, we need to prove that

$$\lim_{\alpha\to0^+} \mathbb{E}\left[V_{YG(T_{\alpha^*})}\right] > 1 \quad \text{and} \quad \lim_{\alpha\to\infty} \mathbb{E}\left[V_{YG(T_{\alpha^*})}\right] = 0.$$

As before, $(\bar{T}_k)_{k\in\mathbb{N}}$ are the jump times of the process $(V_{G(t)})_{t\geq0}$. Then

$$\mathbb{E}\left[V_{YG(T_{\alpha^*})}\right] = \sum_{k\in\mathbb{N}} \mathbb{E}\left[e^{-\alpha\bar{T}_k/Y}\right].$$

When $\alpha \to 0$, we have $\mathbb{E}\left[e^{-\alpha\bar{T}_k}\right] \to \mathbb{P}\left(\bar{T}_k/Y < \infty\right)$. Now,

$$\sum_{k\in\mathbb{N}} \mathbb{P}\left(\bar{T}_k < \infty\right) = \lim_{t\to\infty} \sum_{k\in\mathbb{N}} \mathbb{P}\left(\bar{T}_k/Y \leq t\right) = \lim_{t\to\infty} \mathbb{E}\left[V_{YG(t)}\right] > 1.$$

For $\alpha \to \infty$,

$$\int_0^\infty \alpha e^{-\alpha t} \mathbb{E}\left[V_{YG(t)}\right] dt = \int_0^\infty e^{-u} \mathbb{E}\left[V_{YG(u/\alpha)}\right] du.$$

When $\alpha \to \infty$ we have $\mathbb{E}\left[V_{YG(u/\alpha)}\right] \to 0$. Then, fix $\alpha_0 > 0$ such that $\mathbb{E}\left[V_{YG(u/\alpha)}\right] < 1$ for every $\alpha > \alpha_0$. As a consequence, $e^{-u}\mathbb{E}\left[V_{YG(u/\alpha)}\right] du \leq e^{-u}$ for any $\alpha > \alpha_0$. By dominated convergence,

$$\lim_{\alpha\to\infty} \int_0^\infty \alpha e^{-\alpha t} \mathbb{E}\left[V_{YG(t)}\right] dt = 0.$$

Now suppose Condition (2.6) does not hold. This means that $\mathbb{E}[V_{YG(t_0)}] = +\infty$ for some $t_0 \in [0, G(\infty))$ or $\lim_{t \to \infty} \mathbb{E}[V_{YG(t)}] \leq 1$.

If the first condition holds, then there exists $t_0 \in (0, aG(\infty))$ such that $\mathbb{E}\left[V_{YG(t)}\right] = +\infty$ for every $t \geq t_0$ (recall that $\mathbb{E}\left[V_{YG(t)}\right]$ in an increasing function in t). As a consequence, for every $\alpha > 0$, we have $\mathbb{E}\left[V_{YG(T_\alpha)}\right] = +\infty$, which means that the process is explosive.

If the second condition holds, then for every $\alpha > 0$ the Laplace transform of the process is strictly less than 1, which means there exists no Malthusian parameter. □

Lemma 4.1 gives a weaker condition on the distribution Y than requiring it to be bounded. Now, we want to investigate the degree distribution of the branching process, assuming that the process $(M_t)_{t \geq 0}$ is supercritical and Malthusian. Denote the Malthusian parameter by α^*. The above allows us to complete the proof of Theorem 2.5:

Proof of Theorem 2.5 We start from

$$p_k = \mathbb{E}\left[P_k[V](YG(T_{\alpha^*}))\right]. \tag{4.3}$$

Conditioning on Y and integrating by parts in the integral given by the expectation in (4.3), gives

$$-f_k Y \int_0^\infty e^{-\alpha^* t} P_k[V](YG(t))g(t)dt + f_{k-1} Y \int_0^\infty e^{-\alpha^* t} P_{k-1}[V](YG(t))g(t)dt.$$

Now, we define

$$\hat{\mathcal{L}}(k, \alpha^*, Y) = \left(\frac{\mathcal{L}(\mathbb{P}\left(V_{uG(\cdot)} = k\right) g(\cdot))(\alpha^*)}{\mathcal{L}(\mathbb{P}\left(V_{uG(\cdot)} = k\right))(\alpha^*)}\right)_{u=Y} \tag{4.4}$$

Notice that the sequence $(\hat{\mathcal{L}}(k, \alpha^*, Y))_{k \in \mathbb{N}}$ is a sequence of random variables. Multiplying both sides of the equation by α^*, on the right hand side we have

$$-f_k Y \hat{\mathcal{L}}(k, \alpha^*, Y)\mathbb{E}[P_k[V](uG(T_{\alpha^*}))]_{u=Y} + f_{k-1} Y \hat{\mathcal{L}}(k-1, \alpha^*, Y)\mathbb{E}\left[P_{k-1}[V](uG(T_{\alpha^*}))\right]_{u=Y},$$

while on the left hand side we have

$$\alpha^* \mathbb{E}[P_k[V](uG(T_{\alpha^*}))]_{u=Y}.$$

As a consequence,

$$\mathbb{E}[P_k[V](uG(T_{\alpha^*}))]_{u=Y} = \frac{f_{k-1} Y \hat{\mathcal{L}}(k-1, \alpha^*, Y)}{\alpha^* + f_k Y \hat{\mathcal{L}}(k, \alpha^*, Y)} \mathbb{E}\left[P_{k-1}[V](uG(T_{\alpha^*}))\right]_{u=Y}. \tag{4.5}$$

We start from p_0, that is given by

$$\mathbb{E}[P_0[V](uG(T_{\alpha^*}))]_{u=Y} = \frac{\alpha^*}{\alpha^* + f_0 Y \hat{\mathcal{L}}(0, \alpha^*, Y)}.$$

Recursively using (4.5), gives

$$\mathbb{E}[P_k[V](uG(T_{\alpha^*}))]_{u=Y} = \frac{\alpha^*}{\alpha^* + f_k Y \hat{\mathcal{L}}(k, \alpha^*, Y)} \prod_{i=0}^{k-1} \frac{f_i Y \hat{\mathcal{L}}(i, \alpha^*, Y)}{\alpha^* + f_i Y \hat{\mathcal{L}}(i, \alpha^*, Y)}.$$

Taking expectation on both sides gives

$$p_k = \mathbb{E}\left[\frac{\alpha^*}{\alpha^* + f_k Y \hat{\mathcal{L}}(k, \alpha^*, Y)} \prod_{i=0}^{k-1} \frac{f_i Y \hat{\mathcal{L}}(i, \alpha^*, Y)}{\alpha^* + f_i Y \hat{\mathcal{L}}(i, \alpha^*, Y)}\right].$$

□

Now the sequence $(\hat{\mathcal{L}}(k, \alpha^*, Y))_{k \in \mathbb{N}}$ creates a relation among the sequence of weights, the aging function and the fitness distribution, so that these three ingredients are deeply related.

4.2 Proof of Theorems 2.2 and 2.3

As mentioned, Theorem 2.2 follows immediately by considering $Y \equiv 1$. The proof in fact is the same, since we can express the probabilities $\mathbb{P}(N_t = k)$ as function of the stationary process $(V_t)_{t \geq 0}$ defined by the same PA function f.

Condition (2.5) immediately follows from Condition (2.6). In fact, considering $Y \equiv 1$, Condition (2.6) becomes

$$\mathbb{E}\left[V_{G(t)}\right] < \infty \quad \text{for every } t \geq 0 \quad \text{and} \quad \lim_{t \to \infty} \mathbb{E}\left[V_{G(t)}\right] > 1. \tag{4.6}$$

The first inequality in general true for the type of stationary process we consider (for instance with affine f). The second inequality is exaclty Condition (2.5).

The expression of the sequence $(\hat{\mathcal{L}}^g(k, \alpha^*))_{k \in \mathbb{N}}$ is simplier than the general case given in (4.4). In fact, in (4.4), the sequence $(\hat{\mathcal{L}}(k, \alpha^*, Y))_{k \in \mathbb{N}}$ is actually a sequence of random variables. In the case of aging alone,

$$\hat{\mathcal{L}}^g(k, \alpha^*) = \frac{\mathcal{L}(\mathbb{P}\left(V_{G(\cdot)} = k\right) g(\cdot))(\alpha^*)}{\mathcal{L}(\mathbb{P}\left(V_{G(\cdot)} = k\right))(\alpha^*)},$$

which is a deterministic sequence.

Remark 4.2 Notice that $\hat{\mathcal{L}}^g(k, \alpha^*) = 1$ when $g(t) \equiv 1$, so that $G(t) = t$ for every $t \in \mathbb{R}^+$ and there is no aging, and we retrieve the stationary process $(V_t)_{t \geq 0}$.

Unfortunately, the explicit expression of the coefficients $(\hat{\mathcal{L}}^g(k, \alpha^*))_{k \in \mathbb{N}}$ is not easy to find, even though they are deterministic.

Theorem 2.3, which states that even if g is integrable, the aging does not affect the explosive behavior of a birth process with superlinear weights, is a direct consequence of (4.2):

Proof of Theorem 2.3 Consider a birth process $(V_t)_{t \geq 0}$, defined by a sequence of superlinear weights $(f_k)_{k \in \mathbb{N}}$ (in the sense of Definition 3.16), and an integrable aging function g. Then, for every $t > 0$,

$$\mathbb{P}\left(N_t = \infty\right) = \mathbb{P}\left(V_{G(t)} = \infty\right) > 0.$$

Since this holds for every $t > 0$, the process $(N_t)_{t \geq 0}$ is explosive. As a consequence, for any $\alpha > 0$, $\mathbb{E}\left[N_{T_\alpha}\right] = \infty$, which means that there exists no Malthusian parameter. $\qquad \square$

5 Affine Weights and Adapted Laplace Method

5.1 Aging and No Fitness

In this section, we consider affine PA weights, i.e., we consider $f_k = ak + b$. The main aim is to identify the asymptotic behavior of the limiting degree distribution of the branching process with aging. Consider a stationary process $(V_t)_{t \geq 0}$, where $f_k = ak + b$. Then, for any $t \geq 0$, it is possible to show by induction and the recursions in (3.8) and (3.9) that

$$P_k[V](t) = \mathbb{P}\left(V_t = k\right) = \frac{1}{\Gamma(b/a)} \frac{\Gamma(k + b/a)}{\Gamma(k + 1)} e^{-bt} \left(1 - e^{-at}\right)^k. \tag{5.1}$$

We omit the proof of (5.1). As a consequence, since the corresponding aging process is $(V_{G(t)})_{t \geq 0}$, the limiting degree distribution is given by

$$p_k = \mathbb{P}(V_t = k) = \frac{\Gamma(k + b/a)}{\Gamma(b/a)\Gamma(k+1)} \int_0^\infty \alpha^* e^{-\alpha^* t} e^{-bG(t)} \left(1 - e^{-aG(t)}\right)^k dt. \quad (5.2)$$

We can obtain an immediate upper bound for p_k, in fact

$$p_k = \frac{\Gamma(k + b/a)}{\Gamma(b/a)\Gamma(k+1)} \int_0^\infty \alpha^* e^{-\alpha^* t} e^{-bG(t)} \left(1 - e^{-aG(t)}\right)^k dt$$

$$\leq \frac{\Gamma(k + b/a)}{\Gamma(b/a)\Gamma(k+1)} (1 - e^{-aG(\infty)})^k,$$

which implies that the distribution $(p_k)_{k \in \mathbb{N}}$ has at most an exponential tail. A more precise analysis is hard. Instead we will give an asymptotic approximation, by adapting the Laplace method for integrals to our case.

The Laplace method states that, for a function f that is twice differentiable and with a unique absolute minimum $x_0 \in (a, b)$, as $k \to \infty$,

$$\int_a^b e^{-k\Psi(x)} dx = \sqrt{\frac{2\pi}{k\Psi''(x_0)}} e^{-k\Psi(x_0)}(1 + o(1)). \quad (5.3)$$

In this situation, the interval $[a, b]$ can be infinite. The idea behind this result is that, when $k \gg 1$, the major contribution to the integral comes from a neighborhood of x_0 where $e^{-k\Psi(x)}$ is maximized. In the integral in (5.2), we do not have this situation, since we do not have an integral of the type (5.3). Defining

$$\Psi_k(t) := \frac{\alpha^*}{k} t + \frac{b}{k} G(t) - \log\left(1 - e^{-aG(t)}\right), \quad (5.4)$$

we can rewrite the integral in (5.2) as

$$I(k) := \int_0^\infty \alpha^* e^{-k\Psi_k(t)} dt. \quad (5.5)$$

The derivative of the function $\Psi_k(t)$ is

$$\Psi_k'(t) = \frac{\alpha^*}{k} + \frac{b}{k} g(t) - \frac{ag(t)e^{-aG(t)}}{1 - e^{-aG(t)}}. \quad (5.6)$$

In particular, if there exists a minimum t_k, then it depends on k. In this framework, we cannot directly apply the Laplace method. We now show that we can apply a result similar to (5.3) even to our case:

Lemma 5.1 (Adapted Laplace method 1) *Consider $\alpha, a, b > 0$. Let the integrable aging function g be such that*

(1) for every $t \geq 0$, $0 < g(t) \leq A < \infty$;
(2) g is differentiable on $\mathbb{R}^+$, and g' is finite almost everywhere;
(3) there exists a positive constant $B < \infty$ such that $g(t)$ is decreasing for $t \geq B$;
(4) assume that the solution t_k of $\Psi_k'(t) = 0$, for $\Psi_k'(t)$ as in (5.6), is unique, and $g'(t_k) < 0$.

Then, for $\sigma_k^2 = (k\Psi_k''(t_k))^{-1}$, there exists a constant C such that, as $k \to \infty$,

$$I(k) = C\sqrt{2\pi\sigma_k^2} e^{-k\Psi_k(t_k)} \left(\frac{1}{2} + \mathbb{P}\left(\mathcal{N}(0, \sigma_k^2) \geq t_k\right)\right)(1 + o(1)),$$

where $\mathcal{N}(0, \sigma_k^2)$ denotes a normal distribution with zero mean and variance σ_k^2.

Since Lemma 5.1 is an adapted version of the classical Laplace method, we move the proof to Appendix B. We can use the result of Lemma 5.1 to prove:

Proposition 5.2 (Asymptotics—affine weights, aging, no fitness) *Consider the affine PA weights $f_k = ak + b$, an integrable aging function g, and denote the limiting degree distribution of the corresponding branching process by $(p_k)_{k\in\mathbb{N}}$. Then, under the hypotheses of Lemma 5.1, there exists a constant $C > 0$ such that, as $k \to \infty$,*

$$p_k = \frac{\Gamma(k+b/a)}{\Gamma(k+1)} \left(Cg(t_k) - \frac{g'(t_k)}{g(t_k)}\right)^{1/2} e^{-\alpha^* t_k}(1 - e^{-aG(\infty)})^k D_k(g)(1 + o(1)), \quad (5.7)$$

where

$$D_k(g) = \frac{1}{2} + \frac{1}{2\sqrt{\pi}} \int_{-C_k(g)}^{C_k(g)} e^{-\frac{u^2}{2}} du,$$

and $C_k(g) = t_k \left(Cg(t_k) - \frac{g'(t_k)}{g(t_k)}\right)^{1/2}$.

5.2 Aging and Fitness Case

In this section, we investigate the asymptotic behavior of the limiting degree distribution of a CTBP, in the case of affine PA weights. The method we use is analogous to that in Sect. 5.1.

We assume that the fitness Y is absolutely continuous with respect to the Lebesgue measure, and we denote its density function by μ. The limiting degree distribution of this type of branching process is given by

$$p_k = \mathbb{P}\left(V_{YG(T_{\alpha^*})} = k\right) = \frac{\Gamma(k+b/a)}{\Gamma(b/a)\Gamma(k+1)} \int_{\mathbb{R}^+\times\mathbb{R}^+} \alpha^* e^{-\alpha^* t} \mu(s) e^{-bsG(t)} \left(1 - e^{-asG(t)}\right)^k ds dt. \quad (5.8)$$

We immediately see that the degree distribution has exponential tails when the fitness distribution is bounded:

Lemma 5.3 (Exponential tails for integrable aging and bounded fitnesses) *When there exists γ such that $\mu([0, \gamma]) = 1$, i.e., the fitness has a bounded support, then*

$$p_k \leq \frac{\Gamma(k+b/a)}{\Gamma(b/a)\Gamma(k+1)} \left(1 - e^{-a\gamma G(\infty)}\right)^k. \quad (5.9)$$

In particular, p_k has exponential tails.

Proof Obvious. □

Like in the situation with only aging, the explicit solution of the integral in (5.8) may be hard to find. We again have to adapt the Laplace method to estimate the asymptotic behavior of the integral. We write

$$I(k) := \int_{\mathbb{R}^+\times\mathbb{R}^+} e^{-k\Psi_k(t,s)} ds dt, \quad (5.10)$$

where

$$\Psi_k(t,s) := \frac{\alpha^*}{k} t + \frac{b}{k} sG(t) - \frac{1}{k} \log\mu(s) - \log(1 - e^{-saG(t)}). \quad (5.11)$$

As before, we want to minimize the function Ψ_k. We state here the lemma:

Lemma 5.4 (Adapted Laplace method 2) *Let* $\Psi_k(t,s)$ *as in* (5.11). *Assume that*

(1) *g satisfies the assumptions of Lemma 5.1;*
(2) *μ is twice differentiable on $\mathbb{R}^+$;*
(3) *there exists a constant $B' > 0$ such that, for every $s \geq B'$, μ is monotonically decreasing;*
(4) *(t_k, s_k) is the unique point where both partial derivatives are zero;*
(5) *(t_k, s_k) is the absolute minimum for $\Psi_k(t,s)$;*
(6) *the hessian matrix $H_k(t_k, s_k)$ of $\Psi_k(t,s)$ evaluated in (t_k, s_k) is positive definite.*

Then,

$$I(k) = e^{-k\Psi_k(t_k, s_k)} \frac{2\pi}{\sqrt{\det(kH_k(t_k, s_k))}} \mathbb{P}\left(\mathcal{N}_1(k) \geq -t_k, \mathcal{N}_2(k) \geq -s_k\right)(1 + o(1)),$$

where $(\mathcal{N}_1(k), \mathcal{N}_2(k)) := \mathcal{N}(\mathbf{0}, (kH_k(t_k, s_k))^{-1})$ is a bivariate normal distributed vector and $\mathbf{0} = (0,0)$.

The proof of Lemma 5.4 can be found in Appendix B.1. Using Lemma 5.4 we can describe the limiting degree distribution $(p_k)_{k \in \mathbb{N}}$:

Proposition 5.5 (Asymptotics—affine weights, aging, fitness) *Consider affine PA weights $f_k = ak + b$, an integrable aging function g and a fitness distribution density μ. Assume that the corresponding branching process is supercritical and Malthusian. Under the hypotheses of Lemma 5.4, the limiting degree distribution $(p_k)_{k \in \mathbb{N}}$ of the corresponding* CTBP *satisfies*

$$p_k = \frac{k^{b/a-1}}{\Gamma(b/a)} \frac{2\pi}{\sqrt{\det(kH_k(t_k, s_k))}} e^{-k\Psi_k(t_k, s_k)} \mathbb{P}\left(\mathcal{N}_1 \geq -t_k, \mathcal{N}_2 \geq -s_k\right)(1 + o(1)).$$

5.3 Three Classes of Fitness Distributions

Proposition 5.5 in Sect. 5.2 gives the asymptotic behavior of the limiting degree distribution of a CTBP with integrable aging and fitness. Lemma 5.4 requires conditions under which the function $\Psi_k(t,s)$ as in (5.11) has a unique minimum point denoted by (t_k, s_k). In this section we consider the three different classes of fitness distributions that we have introduced in Sect. 2.4.

For the heavy-tailed class, i.e., for distributions with tail thicker than exponential, there is nothing to prove. In fact, (2.7) immediately implies that such distributions are explosive.

For the other two cases, we apply Proposition 5.5, giving the precise asymptotic behavior of the limiting degree distributions of the corresponding CTBPs. Propositions 5.6 and 5.7 contain the results respectively on the general-exponential and sub-exponential classes. The proof of these propositions are moved to Appendix C.

Proposition 5.6 *Consider a general exponential fitness distribution as in* (2.11). *Let $(M_t)_{t \geq 0}$ be the corresponding birth process. Denote the unique minimum point of $\Psi_k(t,s)$ as in* (5.11) *by (t_k, s_k). Then*

(1) *for every $t \geq 0$, M_t has a dynamical power law with exponent $\tau(t) = 1 + \frac{\theta}{aG(t)}$;*
(2) *the asymptotic behavior of the limiting degree distribution $(p_k)_{k \in \mathbb{N}}$ is given by*

$$p_k = e^{-\alpha^* t_k} h(s_k) \left(\tilde{C} - \alpha^* \frac{g'(t_k)}{g(t_k)}\right)^{-1/2} k^{-(1+\theta/(aG(\infty)))}(1 + o(1)),$$

where the power law term has exponent $\tau = 1 + \theta/aG(\infty)$;

(3) *the distribution $(q_k)_{k \in \mathbb{N}}$ of the total number of children of a fixed individual has a power law behavior with exponent $\tau = 1 + \theta / aG(\infty)$.*

By (2.8) it is necessary to consider the exponential rate $\theta > aG(\infty)$ to obtain a non-explosive process. In particular, this implies that, for every $t \geq 0$, $\tau(t)$, as well as τ, are strictly larger than 2. As a consequence, the three distributions $(P_k[M](t))_{k \in \mathbb{N}}$, $(p_k)_{k \in \mathbb{N}}$ and $(q_k)_{k \in \mathbb{N}}$ have finite first moment. Increasing the value of θ leads to power-law distributions with exponent larger than 3, so with finite variance.

A second observation is that, independently of the aging function g, the point s_k is of order $\log k$. In particular, this has two consequences. First the correction to the power law given by $h(s_k)$ is a power of $\log k$. Since $h'(s)/h(s) \to 0$ as $s \to \infty$. Second the power-law term $k^{-(1+\theta/(aG(\infty)))}$ arises from $\mu(s_k)$. This means that the exponential term in the fitness distribution μ not only is necessary to obtain a non-explosive process, but also generates the power law.

The third observation is that the behavior of the three distributions $(P_k[M](t))_{k \in \mathbb{N}}$, $(p_k)_{k \in \mathbb{N}}$ and $(q_k)_{k \in \mathbb{N}}$ depends on the integrability of the aging function, *but does only marginally depends on its precise shape*. The contribution of the aging function g to the exponent of the power law in fact is given only by the value $G(\infty)$. The other terms that depend directly on the shape of g are $e^{-\alpha^* t_k}$ and the ratio $g'(t_k)/g(t_k)$. The ratio g'/g does not contribute for any function g whose decay is in between power law and exponential. The term $e^{-\alpha^* t_k}$ depends on the behavior of t_k, that can be seen as roughly $g^{-1}(1/\log k)$. For any function between power law and exponential, $e^{-\alpha^* t_k}$ is asymptotic to a power of $\log k$.

The last observation is that every distribution in the general exponential class shows a *dynamical power law* as for the pure exponential distribution, as shown in Sect. 5.4. The pure exponential distribution is a special case where we consider $h(s) \equiv 1$. Interesting is the fact that τ actually does not depend on the choice of $h(s)$, but only on the exponential rate $\theta > aG(\infty)$. In particular, Proposition 5.6 proves that the limiting degree distribution of the two examples in Fig. 9 have power-law decay.

We move to the class of sub-exponential fitness. We show that the power law is lost due to the absence of a pure exponential term. We prove the result using densities of the form

$$\mu(s) = Ce^{-s^{1+\varepsilon}}, \tag{5.12}$$

for $\varepsilon > 0$ and C the normalization constant. The result is the following:

Proposition 5.7 *Consider a sub-exponential fitness distribution as in* (5.12). *Let $(M_t)_{t \geq 0}$ be the corresponding birth process. Denote the minimum point of $\Psi_k(t, s)$ as in* (5.11) *by (t_k, s_k). Then*

(1) *for every $t \geq 0$, M_t satisfies*

$$\mathbb{P}(M_t = k) = k^{-1}(\log k)^{-\varepsilon/2} e^{-\frac{\theta}{(aG(t))^{1+\varepsilon}}(\log k)^{1+\varepsilon}} (1 + o(1));$$

(2) *the limiting degree distribution $(p_k)_{k \in \mathbb{N}}$ of the* CTBP *has asymptotic behavior given by*

$$p_k = e^{-\alpha^* t_k} k^{-1} \left(C_1 - s_k^\varepsilon \frac{g'(t_k)}{g(t_k)} \right) e^{-\frac{\theta}{(aG(\infty))^{1+\varepsilon}}(\log k)^{1+\varepsilon}} (1 + o(1));$$

(3) *the distribution $(q_k)_{k \in \mathbb{N}}$ of the total number of children of a fixed individual satisfies*

$$q_k = k^{-1}(\log k)^{-\varepsilon/2} e^{-\frac{\theta}{(aG(\infty))^{1+\varepsilon}}(\log k)^{1+\varepsilon}} (1 + o(1)).$$

In Proposition 5.7 the distributions $(P_k[M](t))_{k\in\mathbb{N}}$, $(p_k)_{k\in\mathbb{N}}$ and $(q_k)_{k\in\mathbb{N}}$ decay faster than a power law. This is due to the fact that a sub-exponential tail for the fitness distribution does not allow the presence of sufficiently many individuals in the branching population whose fitness value is sufficiently high to restore the power law.

In this case, we have that s_k is roughly $c_1 \log k - c_2 \log \log k$. Hence, as first approximation, s_k is still of logarithmic order. The power-law term is lost because there is no pure exponential term in the distribution μ. In fact, in this case $\mu(s_k)$ generates the dominant term $e^{-\theta(\log k)^{1+\varepsilon}}$.

5.4 The Case of Exponentially Distributed Fitness: Proof of Corollary 2.6

The case when the fitness Y is exponentially distributed turns out to be simpler. In this section, denote the fitness by T_θ, where θ is the parameter of the exponential distribution. First of all, we investigate the Laplace transform of the process. In fact, we can write

$$\mathbb{E}\left[M_{T_\alpha}\right] = \int_0^\infty \theta e^{-\theta s} \mathbb{E}\left[V_{sG(T_\alpha)}\right] ds,$$

which is the Laplace transform of the stationary process $(V_{sG(T_\alpha)})_{s\geq 0}$ with bounded fitness $G(T_\alpha)$ in θ. As a consequence,

$$\mathbb{E}\left[M_{T_\alpha}\right] = \sum_{k\in\mathbb{N}} \mathbb{E}\left[\prod_{i=0}^{k-1} \frac{f_i G(T_\alpha)}{\theta + f_i G(T_\alpha)}\right].$$

Suppose that there exists a Malthusian parameter α^*. This means that, for fixed $(f_k)_{k\in\mathbb{N}}$, g and θ, α^* is the unique value such that $\mathbb{E}\left[M_{T_{\alpha^*}}\right] = 1$. As a consequence, if we fix $(f_k)_{k\in\mathbb{N}}$, g and α^*, θ is the unique value such that

$$\sum_{k\in\mathbb{N}} \mathbb{E}\left[\prod_{i=0}^{k-1} \frac{f_i G(T_\alpha)}{\theta + f_i G(T_\alpha)}\right] = 1.$$

Therefore θ is the Malthusian parameter of the process $(V_{sG(T_\alpha)})_{s\geq 0}$. We are now ready to prove Corollary 2.6:

Proof of Corollary 2.6 We can write $\mathbb{P}(M_t = k) = \mathbb{P}\left(V_{T_\theta G(t)} = k\right)$, which means that we have to evaluate the Laplace transform of $\mathbb{P}\left(V_{sG(t)} = k\right)$ in θ. Using (5.1) the first part follows immediately by simple calculations. For the second part, we just need to take the limit as $t \to \infty$. For the sequence $(p_k)_{k\in\mathbb{N}}$, the result is immediate since $p_k = \mathbb{E}[P_k[M](T_{\alpha^*})]$.   □

The case of affine PA weights $f_k = ak + b$ is particularly nice. As already mentioned in Sect. 2, the process $(M_t)_{t\geq 0}$ has a power-law distribution at every $t \in \mathbb{R}^+$ and (2.12) follows immediately.

Further, (2.13) and (2.14) follow directly.   □

A. Limiting Distribution with Aging Effect, No Fitness

In this section, we analyze the limiting degree distribution $(p_k)_{k \in \mathbb{N}}$ of CTBPs with aging but no fitness. In Sect. A.1 we prove the adapted Laplace method for the general asymptotic behavior of p_k. In Sect. A.2 we consider some examples of aging function g, giving the asymptotics for the corresponding distributions.

A.1 Proofs of Lemma 5.1 and Proposition 5.2

Proof of Lemma 5.1 First of all, we show that t_k is actually a minimum. In fact,

$$\lim_{t \to 0} \frac{d}{dt} \Psi_k(t) = -\infty, \quad \text{and} \quad \lim_{t \to \infty} \frac{d}{dt} \Psi_k(t) = \frac{\alpha}{k} > 0.$$

As a consequence, t_k is a minimum. Then,

$$\lim_{k \to \infty} g(t_k) \left(\frac{\alpha^*}{k} \frac{1 - e^{-aG(\infty)}}{ae^{-aG(\infty)}} \right)^{-1} = \lim_{k \to \infty} g(t_k) \frac{ak}{\alpha^*(e^{aG(\infty)} - 1)} = 1. \quad (A.1)$$

In particular, $g(t_k)$ is of order $1/k$. Then, since t_k is the actual minimum, and g is monotonically decreasing for $t \geq B$,

$$\Psi_k''(t_k) = \frac{b}{k} g'(t_k) + g(t_k)^2 \frac{a^2 e^{-aG(t_k)}(2 - e^{-aG(t_k)})}{(1 - e^{-aG(t_k)})^2} - g'(t_k) \frac{ae^{-aG(t_k)}}{1 - e^{-aG(t_k)}} > 0. \quad (A.2)$$

We use the fact that we are evaluating the second derivative in the point t_k where the first derivative is zero. This means

$$g(t_k) \frac{ae^{-aG(t_k)})}{1 - e^{-aG(t_k)}} = \frac{\alpha}{k} + \frac{b}{k} g(t_k).$$

We use this in (A.2) to obtain

$$
\begin{aligned}
k \Psi_k''(t_k) &= bg'(t_k) + g(t_k) \frac{a(2 - e^{-aG(t_k)})}{1 - e^{-aG(t_k)}} (\alpha + bg(t_k)) - \frac{g'(t_k)}{g(t_k)} (\alpha + bg(t_k)) \\
&= g(t_k) \frac{a(2 - e^{-aG(t_k)})}{1 - e^{-aG(t_k)}} (\alpha + bg(t_k)) - \alpha \frac{g'(t_k)}{g(t_k)}. \quad (A.3)
\end{aligned}
$$

Now, we use Taylor expansion around t_k of $\Psi_k(t)$ in the integral in (5.5). Since we use the expansion around t_k, which is the minimum of $\Psi_k(t)$, the first derivative of Ψ_k is zero. As a consequence, we have

$$I(k) = \int_0^\infty e^{-k \left(\Psi_k(t_k) + \frac{1}{2} \Psi_k''(t_k)(t - t_k)^2 + o((t - t_k)^2) \right)} dt.$$

First of all, notice that the contribution of the terms with $|t - t_k| \gg 1$ is negligible. In fact, we have

$$e^{-k \Psi_k(t)} \leq e^{-\alpha^* t} (1 - e^{-aG(\infty)})^k, \quad (A.4)$$

which means that such terms are exponentially small, so we can ignore them. Now we make a change of variable $u = t - t_k$. Then

$$I(k) = \int_{-t_k}^\infty e^{-k \left(\Psi_k(t_k) + \frac{1}{2} \Psi_k''(t_k) u^2 + o(u^2) \right)} du.$$

In particular, since the term $e^{-k\Psi_k(t_k)}$ does not depend on u, we can write

$$I(k) = e^{-k\Psi_k(t_k)} \int_{-t_k}^{\infty} e^{-k\left(\frac{1}{2}\Psi_k''(t_k)u^2 + o(u^2)\right)} du.$$

We use the notation $k\Psi_k(t_k) = \frac{1}{\sigma_k^2}$, which means we can rewrite the integral as

$$e^{-k\Psi_k(t_k)} \sqrt{2\pi\sigma_k^2} \int_{-\infty}^{t_k} \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{u^2}{2\sigma_k^2}} dt = e^{-k\Psi_k(t_k)} \sqrt{2\pi\sigma_k^2} \, \mathbb{P}\left(\mathcal{N}(0, \sigma_k^2) \leq t_k\right).$$

Since the distribution $\mathcal{N}(0, \sigma_k^2)$ is symmetric with respect to 0, for every $k \in \mathbb{N}$,

$$\mathbb{P}\left(\mathcal{N}(0, \sigma_k^2) \leq t_k\right) = \frac{1}{2}\left[1 + \frac{1}{\sqrt{\pi}} \int_{-t_k/\sigma_k}^{t_k/\sigma_k} e^{-\frac{u^2}{2}} du\right]. \tag{A.5}$$

The behavior of the above integral depends on the ratio t_k/σ_k, which is bounded between 0 and 1. As a consequence, the term $\mathbb{P}\left(\mathcal{N}(0, \sigma_k^2) \leq t_k\right)$ is bounded between $1/2$ and 1. □

Using Lemma 5.1, we can prove Proposition 5.2:

Proof of Proposition 5.2 Recall that $\sigma_k^2 = (k\Psi(t_k)'')^{-1}$. Using (A.3), the fact that g is bounded almost everywhere, and $g'(t_k) < 0$, we can write

$$k\Psi(t_k)'' = \alpha \left(\frac{a(2 - e^{-aG(\infty)})}{1 - e^{-aG(\infty)}} g(t_k) - \frac{g'(t_k)}{g(t_k)}\right)(1 + o(1)). \tag{A.6}$$

Notice that in (A.6) the terms $g(t_k) - \frac{g'(t_k)}{g(t_k)}$ are always strictly positive, since $g(t)$ is decreasing and $t_k \to \infty$ as $k \to \infty$. As a consequence, we can replace the term $\sqrt{2\pi/\sigma_k^2}$ by $\left(Cg(t_k) - \frac{g'(t_k)}{g(t_k)}\right)^{1/2}$, for $C = \frac{a(2-e^{-aG(\infty)})}{1-e^{-aG(\infty)}}$. We also have that

$$e^{-k\Psi_k(t_k)} = \exp\left[-\alpha^* t_k - bG(t_k) + k\log\left(1 - e^{-aG(t_k)}\right)\right]$$
$$= e^{-\alpha^* t_k}(1 - e^{-aG(\infty)})^k(1 + o(1)),$$

since $G(t_k)$ converges to $G(\infty)$. For the term in (A.5), it is easy to show that it is asymptotic to $D_k(g)$. This completes the proof. □

A.2 Examples of Aging Functions

In this section, we analyze two examples of aging functions, in order to give examples of the limiting degree distribution of the branching process. We consider affine weights $f_k = ak+b$, and three different aging functions:

$$g(t) = e^{-\lambda t}, \qquad g(t) = (1 + t)^{-\lambda}, \qquad \text{and} \qquad g(t) = \lambda_1 e^{-\lambda_2(\log(t+1) - \lambda_3)^2}.$$

We assume that in every case the aging function g is integrable, so we consider $\lambda > 0$ for the exponential case, $\lambda > 1$ for the power-law case and $\lambda_1, \lambda_2, \lambda_3 > 0$ for the lognormal case. We assume that g satisfies Condition (2.5) in order to have a supercritical process.

We now apply (5.7) to these three examples, giving their asymptotics. In general, we approximate t_k with the solution of, for $c_1 = \frac{ae^{-aG(\infty)}}{1-e^{-aG(\infty)}}$,

$$\frac{\alpha^*}{k} + \frac{b}{k}g(t) - c_1 g(t) = 0. \tag{A.7}$$

We start considering the exponential case $g(t) = e^{-\lambda t}$. In this case, from (A.7) we obtain that, ignoring constants,

$$t_k = \log k(1 + o(1)). \tag{A.8}$$

As we expected, $t_k \to \infty$. We now use (A.6), which gives a bound on σ_k^2 in (A.1) in terms of g and its derivatives. As a consequence,

$$\left(g(t_k) - \frac{g'(t_k)}{g(t_k)}\right)^{-1/2} = \left(e^{-\lambda t_k} + \lambda\right)^{1/2} \sim \lambda^{1/2}(1 + o(1)).$$

Looking at $e^{-k\Psi_k(t_k)}$, it is easy to compute that, with t_k as in (A.8),

$$\exp\left[-\alpha^* \log k + bG(t_k) + k\log(1 - e^{-aG(t_k)})\right] = k^{-\alpha^*}(1 + o(1)).$$

Since $t_k/\sigma_k \to \infty$, then $\mathbb{P}\left(\mathcal{N}(0, \sigma_k^2) \le t_k\right) \to 1$, so that

$$p_k = \frac{\Gamma(k + b/a)}{\Gamma(b/a)} \frac{1}{\Gamma(k+1)} C_1 k^{-\alpha^*} e^{-C_2 k}(1 + o(1)),$$

which means that p_k has an exponential tail with power-law corrections.

We now apply the same result to the power-law aging function, so $g(t) = (1+t)^{-\lambda}$, and $G(t) = \frac{1}{\lambda-1}(1+t)^{1-\lambda}$. In this case

$$(1 + t_k) = \left(\frac{\alpha^*}{c_1 k}\right)^{-1/\lambda}(1 + o(1)).$$

We use again (A.1), so

$$\left(g(t_k) - \frac{g'(t_k)}{g(t_k)}\right) = \left(\frac{\alpha^*}{c_1 k} + \lambda\left(\frac{c_1 k}{\alpha^*}\right)^{1/\lambda}\right)^{1/2} \sim k^{\alpha^*/2\lambda}(1 + o(1)).$$

In conclusion,

$$p_k = \frac{\Gamma(k + b/a)}{\Gamma(b/a)} \frac{1}{\Gamma(k+1)} k^{\alpha^*/2\lambda} e^{-\alpha^*\left(\frac{\alpha^*}{c_1 k}\right)^{-1/\lambda} - C_2 k}(1 + o(1)),$$

which means that also in this case we have a power-law with exponential truncation.

In the case of the lognormal aging function, (A.7) implies that

$$[\log(t_k + 1) - \lambda_3]^2 \approx +\frac{1}{\lambda_2}\log\left(\frac{c_1}{\alpha^*}k\right).$$

By (A.1) we can say that

$$\left(g(t_k) - \frac{g'(t_k)}{g(t_k)}\right) = \left(\lambda_1 \log\left(\frac{c_1}{\alpha^*}k\right) + 2\lambda_2\frac{\log(t_k+1)}{t_k+1}\right)(1 + o(1)) = \lambda_1 \log\left(\frac{c_1}{\alpha^*}k\right)(1 + o(1)).$$

We conclude then, for some constant $C_3 > 0$,

$$p_k = \frac{\Gamma(k + b/a)}{\Gamma(b/a)} \frac{1}{\Gamma(k+1)} \left(\lambda_1 \log\left(\frac{c_1}{\alpha^*}k\right)\right)^{1/2} e^{-\alpha^* e^{(\log(\frac{c_1}{\alpha^*}k))^{1/2}}} e^{-C_3 k}(1 + o(1)).$$

B. Limiting Distribution with Aging and Fitness

In this section, we consider birth processes with aging and fitness. We prove Lemma 5.4, used in the proof of Proposition 5.5. Then we give examples of limiting degree distributions for different aging functions and exponentially distributed fitness.

B.1 Proofs of Lemma 5.4 and Proposition 5.5

Proof of Lemma 5.4 We use again second order Taylor expansion of the function $\Psi_k(t, s)$ centered in (t_k, s_k), where the first order partial derivatives are zero. As a consequence we write

$$\exp\left[-k\Psi_k(t, s)\right] = \exp\left[-k\Psi_k(t_k, s_k) + \frac{1}{2}\boldsymbol{x}^T \left(kH_k(t_k, s_k)\right)\boldsymbol{x} + o(||\boldsymbol{x}||^2)\right],$$

where

$$\boldsymbol{x} = \begin{bmatrix} t - t_k \\ s - s_k \end{bmatrix}, \quad \text{and} \quad H_k(t_k, s_k) = \begin{bmatrix} \dfrac{\partial^2 \Psi_k}{\partial t^2}(t_k, s_k) & \dfrac{\partial^2 \Psi_k}{\partial s \partial t}(t_k, s_k) \\ \dfrac{\partial^2 \Psi_k}{\partial s \partial t}(t_k, s_k) & \dfrac{\partial^2 \Psi_k}{\partial s^2}(t_k, s_k) \end{bmatrix}.$$

As for the proof of Lemma 5.1, we start by showing that we can ignore the terms where $||\boldsymbol{x}||^2 \gg 1$. In fact,

$$e^{-k\Psi_k(t,s)} \leq \exp\left(-\alpha^* t - bsG(t) + k\log(\mu(s))\right).$$

Since μ is a probability density, $\mu(s) < 1$ for $s \gg 1$. As a consequence, $\log(\mu(s)) < 0$, which means that the above bound is exponentially decreasing whenever t and s are very large. As a consequence, we can ignore the contribution given by the terms where $|t - t_k| \gg 1$ and $|s - s_k| \gg 1$.

The term $e^{-k\Psi_k(t_k, s_k)}$ is independent of t and s, so we do not consider it in the integral. Writing $u = t - t_k$ and $v = s - s_k$, we can write

$$\int_{\mathbb{R}^+ \times \mathbb{R}^+} e^{-\frac{1}{2}\boldsymbol{x}^T (kH_k(t_k, s_k))\boldsymbol{x}} \, ds \, dt = \int_{-t_k}^{\infty} \int_{-s_k}^{\infty} e^{-\frac{1}{2}\boldsymbol{y}^T (kH_k(t_k, s_k))\boldsymbol{y}} \, du \, dv,$$

where this time $\boldsymbol{y}^T = [u \ v]$. As a consequence,

$$\int_{-t_k}^{\infty} \int_{-s_k}^{\infty} e^{-\frac{1}{2}\boldsymbol{y}^T (kH_k(t_k, s_k))\boldsymbol{y}} \, du \, dv = \frac{2\pi}{\sqrt{\det(kH_k(t_k, s_k))}} \mathbb{P}\left(\mathcal{N}_1(k) \geq -t_k, \mathcal{N}_2(k) \geq -s_k\right), \tag{B.1}$$

provided that the covariance matrix $(kH_k(t_k, s_k))^{-1}$ is positive definite. □

As a consequence, we can use (B.1) to obtain that, for the corresponding limiting degree distribution of the branching process $(p_k)_{k \in \mathbb{N}}$, as $k \to \infty$,

$$\begin{aligned} p_k &= \frac{\Gamma(k + b/a)}{\Gamma(b/a)} \frac{1}{\Gamma(k + 1)} e^{-k\Psi_k(t_k, s_k)} \frac{2\pi}{\sqrt{\det(kH_k(t_k, s_k))}} \mathbb{P}\left(\mathcal{N}_1(k) \geq -t_k, \mathcal{N}_2(k)\right. \\ &\left. \geq -s_k\right) (1 + o(1)). \end{aligned}$$

This results holds if the point (t_k, s_k) is the absolute minimum of Ψ_k, and the Hessian matrix is positive definite at (t_k, s_k).

B.2 The Hessian Matrix of $\Psi_k(t, s)$

First of all, we need to find a point (t_k, s_k) which is the solution of the system

$$\frac{\partial \Psi_k}{\partial t} = \frac{\alpha^*}{k} + \frac{b}{k} s g(t) - \frac{sag(t)e^{-saG(t)}}{1 - e^{-saG(t)}} = 0, \tag{B.2}$$

$$\frac{\partial \Psi_k}{\partial s} = \frac{b}{k} G(t) - \frac{1}{k} \frac{\mu'(s)}{\mu(s)} - \frac{aG(t)e^{-saG(t)}}{1 - e^{-saG(t)}} = 0. \tag{B.3}$$

Denote the solution by (t_k, s_k). Then

$$\frac{\partial^2 \Psi_k}{\partial t^2} = \frac{b}{k} s g'(t_k) + g(t_k)^2 \frac{s^2 a^2 e^{-asG(t_k)}}{(1 - e^{-asG(t_k)})^2} - g'(t_k) \frac{ase^{-asG(t_k)}}{1 - e^{-asG(t_k)}},$$

$$\frac{\partial^2 \Psi_k}{\partial s^2} = -\frac{1}{k} \frac{\mu''(s)\mu(s) - \mu'(s)^2}{\mu(s)^2} + \frac{a^2 G(t)^2 e^{-saG(t)}}{(1 - e^{-saG(t)})^2},$$

$$\frac{\partial^2 \Psi_k}{\partial s \partial t} = \frac{b}{k} g(t_k) + \left(1 - \frac{1}{as_k}\right)\left(\frac{b}{k} G(t_k) - \frac{1}{k} \frac{\mu'(s_k)}{\mu(s_k)}\right)\left(\frac{\alpha^*}{k} + \frac{b}{k} s_k g(t_k)\right). \tag{B.4}$$

From (B.2) and (B.3) we know

$$\frac{\alpha^*}{k} + \frac{b}{k} s_k g(t_k) = \frac{s_k a g(t_k) e^{-s_k aG(t_k)}}{1 - e^{-s_k aG(t_k)}},$$

$$\frac{b}{k} G(t_k) - \frac{1}{k} \frac{\mu'(s_k)}{\mu(s_k)} = \frac{aG(t_k)e^{-s_k aG(t_k)}}{1 - e^{-s_k aG(t_k)}}. \tag{B.5}$$

Using (B.5) in the expressions for the second derivatives,

$$\frac{\partial^2 \Psi_k}{\partial t^2} = \frac{b}{k} s_k g'(t_k) + \frac{a s_k g(t_k)}{(1 - e^{-a s_k G(t_k)})}\left(\frac{\alpha}{k} + \frac{b}{k} s_k g(t_k)\right) - \frac{g'(t_k)}{g(t_k)}\left(\frac{\alpha}{k} + \frac{b}{k} s_k g(t_k)\right)$$

$$= \frac{a s_k g(t_k)}{(1 - e^{-a s_k G(t_k)})}\left(\frac{\alpha}{k} + \frac{b}{k} s_k g(t_k)\right) - \frac{\alpha}{k} \frac{g'(t_k)}{g(t_k)}, \tag{B.6}$$

$$\frac{\partial^2 \Psi_k}{\partial s^2} = -\frac{1}{k} \frac{\mu''(s_k)\mu(s_k) - \mu'(s_k)^2}{\mu(s_k)^2} + \frac{aG(t_k)}{1 - e^{-a s_k G(t_k)}}\left(\frac{b}{k} G(t_k) - \frac{1}{k} \frac{\mu'(s_k)}{\mu(s_k)}\right)$$

$$= -\frac{1}{k} \frac{\mu''(s_k)}{\mu(s_k)} + \frac{1}{k}\left(\frac{\mu'(s_k)}{\mu(s_k)}\right)^2 - \frac{1}{k} \frac{\mu'(s_k)}{\mu(s_k)} \frac{aG(t_k)}{1 - e^{-a s_k G(t_k)}} + \frac{1}{k} \frac{abG(t_k)^2}{1 - e^{-a s_k G(t_k)}}. \tag{B.7}$$

In conclusion, the matrix $k H_k(t_k, s_k)$ is given by

$$(k H_k(t_k, s_k))_{1,1} = \frac{a s_k g(t_k)}{(1 - e^{-a s_k G(t_k)})}(\alpha + b s_k g(t_k)) - \alpha \frac{g'(t_k)}{g(t_k)};$$

$$(k H_k(t_k, s_k))_{2,2} = -\frac{\mu''(s_k)}{\mu(s_k)} + \left(\frac{\mu'(s_k)}{\mu(s_k)}\right)^2 - \frac{\mu'(s_k)}{\mu(s_k)} \frac{aG(t_k)}{1 - e^{-a s_k G(t_k)}} + \frac{abG(t_k)^2}{1 - e^{-a s_k G(t_k)}};$$

$$(k H_k(t_k, s_k))_{2,1} = bg(t_k) + \left(1 - \frac{1}{as_k}\right)\left(\frac{b}{k} G(t_k) - \frac{1}{k} \frac{\mu'(s_k)}{\mu(s_k)}\right)(\alpha^* + b s_k g(t_k)). \tag{B.8}$$

We point out that, solving (B.3) in terms of s, it follows that

$$s = \frac{1}{aG(t)} \log \left(1 + k \frac{aG(t)}{bG(t) - \frac{\mu'(s)}{\mu(s)}} \right). \tag{B.9}$$

As a consequence,

$$s_k g(t_k) = \alpha^* G(t_k) \frac{\mu(s_k)}{\mu'(s_k)}. \tag{B.10}$$

We use (B.9), (B.10) and the expressions for the elements of the Hessian matrix given in (B.8) for the examples in Sect. B.3. We also use the formulas of this section in the proof of Propositions 5.6 and 5.7 given in Sect. C.

B.3 Examples of Aging Functions

Here we give examples of limiting degree distributions. We consider the same three examples of aging functions we considered in Sect. A.2, so

$$g(t) = e^{-\lambda t}, \qquad g(t) = (1+t)^{-\lambda}, \qquad \text{and} \qquad g(t) = \lambda_1 e^{-\lambda_2 (\log(t+1) - \lambda_3)^2}.$$

We consider exponentially distributed fitness, so $\mu(s) = \theta e^{-\theta s}$. In order to have a supercritical and Malthusian process, we can rewrite Condition (2.8) for exponentially distributed fitness as $aG(\infty) < \theta < (a+b)G(\infty)$.

In general, we identify the minimum point (t_k, s_k), then use (C.13). For all three examples, replacing $G(t)$ by $G(\infty)$ and using (B.9), it holds that

$$s_k \approx \frac{1}{aG(\infty)} \log \left(k \frac{aG(\infty)}{bG(\infty) + \theta} \right),$$

and $s_k g(t_k) \approx \alpha^* G(\infty) \theta$. For the exponential aging function, using (B.10), it follows that $e^{-\lambda t_k} \approx \log k$. In this case, since $g'(t)/g(t) = -\lambda$, the conclusion is that, ignoring the constants,

$$p_k = k^{-(1+\lambda\theta/a)} (\log k)^{\alpha^*/\lambda} (1 + o(1)).$$

For the inverse-power aging function $t_k \approx (\log k)^{1/\lambda}$, which implies (ignoring again the constants) that

$$p_k = k^{-(1+(\lambda-1)\theta/a)} e^{-\alpha^* (\log k)^{1/\lambda}} (1 + o(1)),$$

where we recall that, for g being integrable, $\lambda > 1$. For the lognormal case,

$$t_k \approx e^{(\log k)^{1/2}},$$

which means that

$$p_k = k^{-(1+\theta/aG(\infty))} e^{-\alpha^* e^{(\log k)^{1/2}}} (1 + o(1)).$$

C. Proof of Propositions 5.6 and 5.7

In the present section, we prove Propositions 5.6 and 5.7. These proofs are applications of Proposition 5.5, and mainly consist of computations. In the proof of the two propositions, we often refer to Appendix B.2 for expressions regarding the Hessian matrix of $\Psi_k(t, s)$ as in (5.11).

C.1 Proof of Proposition 5.6

We start by proving the existence of the dynamical power-law. We already know that

$$\mathbb{P}\left(M_t = k\right) = \frac{\Gamma(k+b/a)}{\Gamma(b/a)\Gamma(k+1)} \int_0^\infty \mu(s) e^{-bsG(t)} \left(1 - e^{-asG(t)}\right)^k ds. \tag{C.1}$$

We write

$$J(k) = \int_0^\infty e^{-k\psi_k(s)} ds, \tag{C.2}$$

where

$$\psi_k(s) = \frac{bG(t)}{k} s - \frac{1}{k} \log(\mu(s)) - \log\left(1 - e^{-asG(t)}\right). \tag{C.3}$$

In order to give asymptotics on $J(k)$ as in (C.2), we can use a Laplace method similar to the one used in the proof of Lemma 5.1, but the analysis is simpler since in this case $\psi_k(s)$ is a function of only one variable. The idea is again to find a minimum point s_k for $\psi_k(s)$, and to use Taylor expansion inside the integral, so

$$\psi_k(s) = \psi_k(s_k) + \frac{1}{2} \psi''_k(s_k)(s - s_k)^2 + o((s - s_k)^2).$$

We can ignore the contribution of the terms where $(s - s_k)^2 \gg 1$, since $e^{-k\psi_k(s)} \le e^{-bsG(t)}$, so that the error is at most exponentially small. As a consequence,

$$J(k) = \sqrt{\frac{\pi}{\psi''_k(s_k)}} e^{-k\psi_k(s_k)} (1 + o(1)). \tag{C.4}$$

The minimum s_k is a solution of

$$\frac{d\psi_k(s)}{ds} = \frac{bG(t)}{k} - \frac{1}{k} \frac{\mu'(s)}{\mu(s)} - \frac{aG(t)e^{-saG(t)}}{1 - e^{-asG(t)}} = 0. \tag{C.5}$$

In particular, s_k satisfies the following equality, which is similar to (B.9):

$$s_k = \frac{1}{aG(t)} \log\left(1 + k \frac{aG(t)}{bG(t) - \mu'(s_k)/\mu(s_k)}\right). \tag{C.6}$$

When $\mu(s) = Ch(s)e^{-\theta s}$,

$$\frac{\mu'(s)}{\mu(s)} = \frac{h'(s)e^{-\theta s} - \theta h(s)e^{-\theta s}}{h(s)e^{-\theta s}} = -\theta\left(1 - \frac{h'(s)}{\theta h(s)}\right) \approx -\theta. \tag{C.7}$$

In particular, this implies

$$s_k = \frac{1}{aG(t)} \log\left(1 + k \frac{aG(t)}{bG(t) + \theta}\right)(1 + o(1)). \tag{C.8}$$

Similarly to the element $(kH_k(t_k, s_k))_{2,2}$ in (B.8),

$$k\frac{d^2\psi_k(s_k)}{ds^2} = -\frac{\mu''(s_k)}{\mu(s_k)} + \left(\frac{\mu'(s_k)}{\mu(s_k)}\right)^2 - \frac{\mu'(s_k)}{\mu(s_k)} \frac{aG(t)}{1 - e^{-as_kG(t)}} + \frac{abG(t)^2}{1 - e^{-as_kG(t)}}. \tag{C.9}$$

For the general exponential class, the ratio

$$\frac{\mu''(s_k)}{\mu(s_k)} = \frac{h''(s)}{h(s)} - 2\theta + \theta^2.$$

As a consequence, $k\frac{d^2\psi_k(s)}{ds^2}$ converges to a positive constant, which means that s_k is an actual minimum. Then $J(k) = c_1 e^{-k\psi_k(s_k)}(1+o(1))$. Using this in (C.1) and ignoring the constants,

$$
\begin{aligned}
\mathbb{P}(M_t = k) &= \frac{\Gamma(k+b/a)}{\Gamma(b/a)\Gamma(k+1)} e^{-s_k bG(t)} \mu(s_k)(1+o(1)) \\
&= k^{-1} k^{-b/a} k^{b/a} h(s_k) k^{-\theta/aG(t)}(1+o(1)) = h(s_k) k^{-(1+\theta/aG(t))}(1+o(1)),
\end{aligned}
\tag{C.10}
$$

which is a power-law distribution with exponent $\tau(t) = 1 + \theta/aG(t)$, and minor corrections given by $h(s_k)$. This holds for every $t \geq 0$. In particular, considering $G(\infty)$ instead of $G(t)$, with the same argument we can also prove that the distribution of the total number of children obeys a power-law tail with exponent $\tau(\infty) = 1 + \theta/aG(\infty)$.

We now prove the result on the limiting distribution $(p_k)_{k\in\mathbb{N}}$ of the CTBP, for which we apply directly Proposition 5.5, using the analysis on the Hessian matrix given in Sect. B.2. First of all, from (B.9) it follows that

$$
s_k = \frac{1}{aG(t_k)} \log\left(1 + k\frac{aG(t_k)}{bG(t_k)+\theta}\right)(1+o(1)),
\tag{C.11}
$$

and by (B.10)

$$
s_k g(t_k) \xrightarrow{k\to\infty} \alpha\frac{G(\infty)}{\theta}.
\tag{C.12}
$$

For the Hessian matrix, using (C.11) and (C.11) in (B.8), for any integrable aging function g we have

$$
(kH_k(t_k,s_k))_{2,2} = C_2 + o(1) > 0, \quad \text{and} \quad (kH_k(t_k,s_k))_{2,1} = o(1),
$$

but $(kH_k(t_k,s_k))_{1,1}$ behaves according to $g'(t_k)/g(t_k)$. If this ratio is bounded, then $(kH_k(t_k,s_k))_{1,1} = C_1 + o(1) > 0$, while $(kH_k(t_k,s_k))_{1,1} \to \infty$ whenever $g'(t_k)/g(t_k)$ diverges. In both cases, (t_k,s_k) is a minimum. In particular, again ignoring the multiplicative constants and using (C.11) and (C.12) in the definition of $\Psi_k(t,s)$, the limiting degree distribution of the CTBP is asymptotic to

$$
k^{-(1+\theta/(aG(t_k)))} h(s_k) e^{-\alpha^* t_k}\left(\tilde{C} - \alpha^*\frac{g'(t_k)}{g(t_k)}\right)^{-1/2},
\tag{C.13}
$$

where the term $\left(\tilde{C} - \alpha^*\frac{g'(t_k)}{g(t_k)}\right)^{-1/2}$, which comes from the determinant of the Hessian matrix, behaves differently according to the aging function. With this, the proof of Proposition 5.6 is complete. □

C.2 Proof of Proposition 5.7

This proof is identical to the proof of Proposition 5.6, but this time we consider a sub-exponential distribution. First, we start looking at the distribution of the birth process at a fixed time $t \geq 0$. We define $\psi_k(s)$ and $J(k)$ as in (C.3) and (C.2). We use again (C.1), so

$$
s_k = \frac{1}{aG(t)} \log\left(1 + k\frac{aG(t)}{bG(t) - \mu'(s_k)/\mu(s_k)}\right).
$$

In this case, we have

$$
\frac{\mu'(s)}{\mu(s)} = -\theta(1+\varepsilon)s^\varepsilon.
\tag{C.14}
$$

Then s_k satisfies

$$s_k = \frac{1}{aG(t)} \log\left(1 + k\frac{aG(t)}{bG(t) + \theta(1+\varepsilon)s_k^\varepsilon}\right). \tag{C.15}$$

By substitution, it is easy to check that s_k is approximatively $c_1 \log k - c_2 \log\log k = \log k(1 - \frac{\log\log k}{\log k})$, for some positive constants c_1 and c_2. This means that as first order approximation, s_k is still of logarithmic order. Then,

$$\frac{\mu''(s)}{\mu(s)} = \theta^2(1+\varepsilon)^2 s^{2\varepsilon} - \theta(1+\varepsilon)\varepsilon s^{\varepsilon-1}. \tag{C.16}$$

Using (C.14) and (C.16), we can write

$$k\frac{d^2\psi_k(s)}{ds^2} = \theta(1+\varepsilon)\varepsilon s_k^{\varepsilon-1} + \theta(1+\varepsilon)s_k^\varepsilon \frac{aG(t)}{1 - e^{-as_kG(t)}} + \frac{abG(t)^2}{1 - e^{-as_kG(t)}}$$

$$= \theta(1+\varepsilon)\varepsilon s_k^{\varepsilon-1} + \theta(1+\varepsilon)\frac{s_k^\varepsilon}{k}(bG(t) + \theta(1+\varepsilon)s_k^\varepsilon)$$

$$+ \frac{bG(t)}{k}(bG(t) + \theta(1+\varepsilon)s_k^\varepsilon). \tag{C.17}$$

The dominant term is $c_1 s_k^\varepsilon$, for some constant c_1. This means $k\frac{d^2\psi_k(s)}{ds^2}$ is of order $(\log k)^\varepsilon$. Now,

$$J(k) = \left(k\frac{d^2\psi_k(s)}{ds^2}\right)^{-1/2} Ce^{-bG(t)s_k - \theta s_k^{1+\varepsilon} + k\log(1 - e^{-aG(t)s_k})}(1 + o(1))$$

$$= (\log k)^{-\varepsilon/2} k^{-b/a} e^{-\theta(\log k)^{1+\varepsilon}}(1 + o(1)). \tag{C.18}$$

As a consequence,

$$\mathbb{P}(M_t = k) = k^{-1}(\log k)^{-\varepsilon/2} e^{-\theta(\log k)^{1+\varepsilon}}(1 + o(1)), \tag{C.19}$$

which is not a power-law distribution. Again using similar arguments, we show that the limiting degree distribution of the CTBP does not show a power-law tail. In this case

$$s_k = \frac{1}{aG(t_k)} \log\left(1 + k\frac{aG(t_k)}{bG(t_k) + \theta(1+\varepsilon)s_k^\varepsilon}\right),$$

and

$$s_k g(t_k) = \frac{\alpha G(t_k)}{\theta(1+\varepsilon)s_k^\varepsilon} = \frac{\alpha G(t_k)}{\log^\varepsilon k}(1 + o(1)) \to 0.$$

The Hessian matrix elements are

$$(kH_k(t_k, s_k))_{1,1} = \frac{a\alpha^2 G(t_k)}{s_k^\varepsilon} - a\alpha\frac{g'(t_k)}{g(t_k)} + o(1),$$

$$(kH_k(t_k, s_k))_{2,2} = \theta(1+\varepsilon)\varepsilon s_k^{\varepsilon-1} + \theta s_k^\varepsilon aG(\infty) + abG(\infty)^2 + o(1),$$

$$(kH_k(t_k, s_k))_{1,2} = o(1). \tag{C.20}$$

This implies that

$$\det(kH_k(t_k, s_k)) = C_1 - s_k^\varepsilon\frac{g'(t_k)}{g(t_k)} + o(1) > 0. \tag{C.21}$$

As a consequence, (t_k, s_k) is an actual minimum. Then using the definition of $\Psi_k(t, s)$,

$$p_k = \mathrm{e}^{-\alpha^* t_k} k^{-1+b/a} k^{-b/a} \mu(s_k) \sim \mathrm{e}^{-\alpha^* t_k} k^{-1} \left(C_1 - s_k^\varepsilon \frac{g'(t_k)}{g(t_k)} \right) \mathrm{e}^{-\frac{\theta}{(aG(\infty))^{1+\varepsilon}} (\log k)^{1+\varepsilon}} (1 + o(1)).$$

(C.22)

This completes the proof. □

References

1. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**(1), 47–97 (2002)
2. Athreya, K.B.: Preferential attachment random graphs with general weight function. Internet Math. **4**(4), 401–418 (2007)
3. Athreya, K.B., Ghosh, A.P., Sethuraman, S.: Growth of preferential attachment random graphs via continuous-time branching processes. Proc. Indian Acad. Sci. Math. Sci. **118**(3), 473–494 (2008)
4. Athreya, K.B., Ney, P.E.: Branching Processes. Dover Publications, NY (2004). Reprint of the 1972 original [Springer, New York; MR0373040], pp. xii+287
5. Bhamidi, S.: Universal techniques to analyze preferential attachment trees: global and Local analysis (preprint). http://www.unc.edu/bhamidi/preferent.pdf (2007)
6. Borgs, C., Chayes, J., Daskalakis, C., Roch, S.: 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, pp. 135–144 (2007)
7. Csáardi, G.: Dynamics of Citation Networks. In: Kollias, S.D., Stafylopatis, A., Duch, W., Oja, E. (eds.) Lecture Notes in Computer Science 4131, pp. 698–709 (2006)
8. Deijfen, M., van den Esker, H., van der Hofstad, R., Hooghiemstra, G.: A preferential attachment model with random initial degrees. Arkiv föor Matematik **47**(1), 41–72 (2008)
9. Dereich, S.: Preferential attachment with fitness: unfolding the condensate. Electron. J. Probab. **21** (2016)
10. Dereich, S., Ortgiese, M.: Robust analysis of preferential attachment models with fitness. Comb. Probab. Comput. **23**(03), 386–411 (2014)
11. Hajra, K.B., Sen, P.: Aging in citation networks. Phys. A: Stat. Mech. Appl. **346**(1–2), 44–48 (2005)
12. Hajra, K.B., Sen, P.: Modelling aging characteristics in citation networks. Phys. A: Stat. Mech. Appl. **368**(2), 575–582 (2006)
13. Hazoglou, M.J., Kulkarni, V., Skiena, S.S., Dill, K.K.: Citation histories of papers: sometimes the rich get richer, sometimes they dont. arXiv:1703.05943v1 (2017)
14. Jagers, P., Nerman, O.: The growth and composition of branching populations. Adv. Appl. Probab. **16**(2), 221–259 (1984)
15. Jeong, H., Néda, Z., Barabási, A.L.: Measuring preferential attachment in evolving networks. Europhys. Lett. **61**(4), 567–572 (2003)
16. Kesten, H., Stigum, B.P.: Limit theorems for decomposable multi-dimensional Galton-Watson processes. J. Math. Anal. Appl. **17**(2), 309–338 (1967)
17. Nerman, O.: On the convergence of supercritical general (C-M-J) branching processes. Probab. Theory Relat. Fields **57**(3), 365–395 (1981)
18. Oliveira, R., Spencer, J.: Connectivity transitions in networks with super-linear preferential attachment. Internet Math. **2**(2), 121–163 (2005)
19. Redner, S.: Citation statistics from 110 years of physical review. Phys. Today **58**(6), 49–54 (2005)
20. Rosengren, S.: A multi-type preferential attachment model. http://arxiv.org/pdf/1704.03256.pdf
21. Rudas, A.: Random tree growth with general weight function. arXiv:math/0410532 (2004)
22. Rudas, A., Tóth, B., Valkó, B.: Random trees and general branching processes. Random Struct. Algorithms **31**(2), 186–202 (2007)
23. van der Hofstad, R.: Random Graphs and Complex Networks, vol. 1. Cambridge University Press, Cambridge (2017)
24. Wang, D., Song, C., Barabási, A.L.: Quantifying long-term scientific impact. Science **342**(6154), 127–132 (2013)
25. Wang, M., Yu, G., Yu, D.: Effect of the age of papers on the preferential attachment in citation networks. Phys. A: Stat. Mech. Appl. **388**(19), 4273–4276 (2009)
26. Wang, M., Yu, G., Yu, D.: Measuring the preferential attachment mechanism in citation networks. Phys. A: Stat. Mech. Appl. **387**(18), 4692–4698 (2008)