



*Citation for published version:*

Mailler, C, Mörters, P & Senkevich, A 2021, 'Competing growth processes with random growth rates and random birth times', *Stochastic Processes and their Applications*, vol. 135, pp. 183-226.  
<https://doi.org/10.1016/j.spa.2021.02.003>

*DOI:*

[10.1016/j.spa.2021.02.003](https://doi.org/10.1016/j.spa.2021.02.003)

*Publication date:*

2021

*Document Version*

Peer reviewed version

[Link to publication](#)

*Publisher Rights*

CC BY-NC-ND

**University of Bath**

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Competing growth processes with random growth rates and random birth times

Cécile Mailler, Peter Mörters and Anna Senkevich

## Abstract

Comparing individual contributions in a strongly interacting system of stochastic growth processes can be a very difficult problem. This is particularly the case when new growth processes are initiated depending on the state of previous ones and the growth rates of the individual processes are themselves random. We propose a novel technique to deal with such problems and show how it can be applied to a broad range of examples where it produces new insight and surprising results. The method relies on two steps: In the first step, which is highly problem dependent, the growth processes are jointly embedded into continuous time so that their evolutions after initiation become approximately independent while we retain some control over the initiation times. Once such an embedding is achieved, the second step is to apply a Poisson limit theorem that enables a comparison of the state of the processes initiated in a critical window and therefore allows an asymptotic description of the extremal process. In this paper we prove a versatile limit theorem of this type and show how this tool can be applied to obtain novel asymptotic results for a variety of interesting stochastic processes. These include (a) the maximal degree in different types of preferential attachment networks with fitnesses like the well-known Bianconi-Barabási tree and a network model of Dereich, (b) the most successful mutant in branching processes evolving by selection and mutation, and (c) the ratio between the largest and second largest cycles in a random permutation with random cycle weights, which can also be interpreted as a disordered version of Pitman’s Chinese restaurant process.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Motivation . . . . .	2
1.2	Our framework and principal technical results . . . . .	4
1.3	Examples of fitness distributions . . . . .	8
<b>2</b>	<b>Examples and applications</b>	<b>9</b>
2.1	Branching processes with selection and mutation . . . . .	9
2.2	Preferential attachment networks with fitness . . . . .	13
2.3	Random permutations with random cycle weights . . . . .	18
<b>3</b>	<b>Local convergence of point processes</b>	<b>21</b>
3.1	Preliminaries on the fitness distribution . . . . .	21
3.2	Convergence of a simpler point process . . . . .	22
3.3	Proof of the local convergence result . . . . .	28
<b>4</b>	<b>Compactification and completion of the proofs</b>	<b>31</b>
4.1	Contribution of young and old families . . . . .	31
4.2	Contribution of unfit families . . . . .	33
4.3	Contribution of old and fit families . . . . .	36
4.4	Proof of Theorem 3 . . . . .	37
4.5	Proof of Corollary 4 . . . . .	37
<b>5</b>	<b>Appendix: an auxiliary lemma</b>	<b>38</b>

# 1 Introduction

## 1.1 Motivation

Suppose a population of immortal individuals evolves as follows: We start with one individual with a fitness sampled from a fixed bounded distribution  $\mu$ . When the population consists of  $n$  individuals, the next individual selects its parent from the  $n$  existing individuals with a probability proportional to their individual fitnesses. With high probability the new individual inherits the fitness from its parent and joins the parent's family, but with small probability  $0 < \beta < 1$  the individual is a mutant and founder of a new family, getting a fitness sampled independently of everything else from the distribution  $\mu$ . Even for such a simple model of a population evolving by selection and mutation the structure at large finite times, i.e. when the system is not in equilibrium, can be hard to analyse. The difficulty is that it takes time until an individual born with high fitness can use its advantage to build a large family. Quantities like the relative size of the largest family when the total population has a given large size depend on these delays and therefore involve a comparison of many different random influences which are typically very hard to control.

In this paper we investigate a broad class of problems loosely similar to the above providing a novel technique to their solution. For the method to work one needs an embedding of the problem into continuous time that makes the growth processes of the individuals approximately independent. Such embeddings have been used as a tool for urn processes since the seminal work of Athreya and Karlin [3] and can be constructed for a wide range of models. In our example the embedding is achieved by equipping every individual with fitness  $f$  with an independent Poisson process of intensity  $f$  initiated at the individual's birthtime. The jump times of the Poisson process correspond to the times when the individual is chosen as a parent. Then, given a population of  $n$  individuals the probability that each individual is the next parent is proportional to its fitness. Each family is equipped with an independent fitness sampled from a distribution  $\mu$  and, starting from its birthtime, grows as an independent Yule process with parameter  $(1 - \beta)f$ , where  $f$  is the fitness of the family and  $\beta$  the mutation probability. The downside of looking at the problem in this time-scale is that the families' birthtimes depend in a complex way on the multitude of independent growth processes and all we can hope for is an asymptotic expansion of the birthtime  $\tau_n$  of the  $n$ th family.

The main step in our technique is to use extreme value theory and the approximate independence of the growth processes in our embedding to provide asymptotic properties of the largest family. As in our example we assume that the growth rates are sampled from an i.i.d. sequence  $F_1, F_2, \dots$  of bounded random variables, while the birth times  $\tau_1, \tau_2, \dots$  may be random and depend in an arbitrarily complex fashion on the growth processes. In the most interesting cases the birth times are themselves arising from an exponentially growing process so that the largest family at time  $t$  arises in competition of the few families born early, which have a longer time to grow, and the many families born late, among which the occurrence of a higher birth rate is more probable. We will give interesting examples below, but first we give a flavour of the problem by a calculation based on the simplest nontrivial scenario.

For this purpose let the birth time of the  $n$ th family be  $\tau_n = \frac{1}{\lambda} \log n$  and its size at time  $t$  be

$$Z_n(t) = \begin{cases} e^{(t-\tau_n)F_n} & \text{if } \tau_n < t, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose  $\mu$  is the law of  $F_n$  on the interval  $(0, 1]$  and  $x > 0$  fixed. If  $1 \ll T(t) \ll t$  then

$$\begin{aligned} \mathbb{P}\left(e^{-(t-T(t))} \max_n Z_n(t) \leq e^x\right) &= \mathbb{P}\left((t - \tau_n)F_n \leq (t - T(t)) + x \quad \forall n: \tau_n \leq t\right) \\ &= \prod_{\tau_n \leq T(t) - x} \mathbb{P}\left(F_n \leq \frac{t - T(t) + x}{t - \tau_n}\right) \end{aligned}$$

using that the probability is one if  $\frac{t-T(t)+x}{t-\tau_n} > 1$ . We continue the calculation with

$$\begin{aligned} &= \exp \left( \sum_{n \leq e^{\lambda(T(t)-x)}} \log \left( 1 - \mu \left( \left( \frac{t-T(t)+x}{t-\tau_n}, 1 \right] \right) \right) \right) \\ &= \exp \left( - (1 + o(1)) \sum_{n \leq e^{\lambda(T(t)-x)}} \mu \left( \left( \frac{t-T(t)+x}{t-\tau_n}, 1 \right] \right) \right). \end{aligned}$$

The task is now to choose  $T(t)$  such that, as  $t \uparrow \infty$ ,

$$\sum_{n \leq e^{\lambda(T(t)-x)}} \mu \left( \left( \frac{t-T(t)+x}{t-\tau_n}, 1 \right] \right) \longrightarrow \phi(x),$$

for some nondegenerate function  $\phi$ . The solution depends on the tail of  $\mu$  at one. Supposing for example that  $\mu((1-x, 1]) \sim x^\alpha$  as  $x \downarrow 0$ , for some index  $\alpha > 0$ , we get

$$\sum_{n \leq e^{\lambda(T(t)-x)}} \mu \left( \left( \frac{t-T(t)+x}{t-\tau_n}, 1 \right] \right) \sim \frac{1}{t^\alpha} \sum_{n \leq e^{\lambda(T(t)-x)}} (T(t) - \tau_n - x)^\alpha.$$

Letting  $T(t) = \frac{\alpha}{\lambda} \log t$  this is equivalent to

$$\frac{1}{t^\alpha} \int_0^{t^\alpha e^{-\lambda x}} \left( -\frac{1}{\lambda} \log \left( \frac{n}{t^\alpha} \right) - x \right)^\alpha dn = e^{-\lambda x} \int_0^\infty \lambda e^{-\lambda u} u^\alpha du = e^{-\lambda x} \lambda^{-\alpha} \Gamma(\alpha + 1),$$

using the substitution  $u = -\frac{1}{\lambda} \log \left( \frac{n}{t^\alpha} \right) - x$ . Hence we have that

$$e^{-t} \left( \frac{(\lambda t)^\alpha}{\Gamma(\alpha+1)} \right)^{\frac{1}{\lambda}} \max_n Z_n(t) \Longrightarrow \Phi_\lambda,$$

where  $\Phi_\lambda$  is the Fréchet distribution with parameter  $\lambda$ .

This result, and further asymptotic results on the birthtime and fitness of the largest family, can be generalised to a framework where

- $\mu$  is a sufficiently smooth distribution in the maximum domain of attraction of either the *Weibull* or the *Gumbel distribution* of extreme value theory,
- the growth processes  $(Z_n(\tau_n + s) : s \geq 0)$  are asymptotically independent random processes with growth rates given as  $\gamma F_n$ , for some  $\gamma > 0$ ,
- the birth times  $\tau_n$  are themselves random and may depend on the growth processes.

Generalising the above calculation to such a setup requires, of course, more sophisticated methods. Our approach is to describe the state of a family at time  $t$  as a point in the space  $(-\infty, \infty) \times (-\infty, \infty) \times (0, \infty)$ , where the first coordinate corresponds to its birth time, the second to its fitness and the third to its size at time  $t$ . If  $\mu$  is in the maximum domain of attraction of the *Weibull distribution*, introducing a  $t$ -dependent scaling of the three coordinates (so that the focus is on a carefully chosen window) and letting  $t \rightarrow \infty$  we obtain a limiting point process, see Theorem 1. In this limiting process the point with the maximal third coordinate identifies the largest family, allowing to read off limit theorems for its size, fitness and birthtime, see Corollary 2. A similar result was proved in [8] in the context of reinforced branching process. Our result extends that of [8] to the more general context of competing growth processes, allowing for a much wider range of applications.

The *main technical results* of the present paper provide corresponding results for the case that  $\mu$  is in the maximum domain of attraction of the *Gumbel distribution*. This case is considerably more difficult than the Weibull case and new ideas are needed. The reason for this is that the window

in which one has to search for the largest family is larger, having unbounded width in the first component. Therefore for a limit theorem the first component requires scaling, and hence the scaling of the second component depends not only on  $t$  but also on  $n$ , the birth rank of the family. Using some additional regularity properties of the fitness distribution  $\mu$  however allows to make the scaling of the third component independent of  $n$ , so that we can still achieve a powerful Poisson limit theorem (Theorem 3) as well as convergence of the scaled family size to a Fréchet distribution and of the standardised birth time to a Gaussian distribution (Corollary 4). Taken together, our results give an essentially complete picture for the behaviour of the largest family for fitness distributions  $\mu$  with bounded support. Fitness distributions with unbounded support lead to superexponentially growing processes, which have more complex behaviour and cannot be treated here.

As application of our main technical result, we obtain results on the extremal behaviour of a variety of models that all fall under our general framework of competing growth processes: Our *main examples* of competing growth processes originate from the study of dynamic network models. In these models new vertices get born at random times and are connected to existing vertices by certain rules. The degree of a vertex grows over time with a growth rate given by the attractiveness, or fitness, of the vertex. We show asymptotic results for the vertex of maximal degree at a large time  $t$  and describe its degree, fitness and birthtime as a function of  $t$ : see Section 2.2.1 for the Bianconi and Barabási network [6] and Section 2.2.2 for a model of Dereich [7]. Applications of our main technical result also include asymptotic results on the largest family in the population process with selection and mutation described above (see Section 2.1), and on the largest tables in a disordered Chinese restaurant process for which we derive a surprising result on the relative sizes of the two largest occupied tables (see Section 2.3). We will explain how to get these results in Section 2.

The paper is structured as follows. In Section 1.2 we give a full definition of our framework and assumptions on the embedded process and state the main results. Section 1.3 gives examples of fitness distributions to which our results apply. Section 2 is devoted to a range of interesting examples of growth processes and describes applications of our general results to these examples. The further sections are devoted to the proofs and their structure will be explained at the end of Section 2.

## 1.2 Our framework and principal technical results

Let  $\mu$  be a probability distribution on the nonnegative real numbers with  $\mathfrak{s} = \text{esssup}(\mu) < \infty$ . To rule out less interesting cases we assume that  $\mu$  has no atom at zero or at  $\mathfrak{s}$ . Without loss of generality we can and will further assume that  $\mathfrak{s} = 1$  and hence that  $\mu$  is supported on the interval  $(0, 1)$ . Let

- $(F_n)_{n \geq 1}$  be i.i.d.  $\mu$ -distributed random variables;
- $(\tau_n)_{n \geq 1}$  be a non-decreasing sequence of positive random variables with  $\tau_1 = 0$ ;
- $Z_n(t) = X_n(F_n(t - \tau_n)) \mathbf{1}_{t \geq \tau_n}$  for a family  $(X_n(t) : t \geq 0)_{n \geq 1}$  of non-decreasing integer-valued processes.

Define  $M(t) := \max\{n \geq 1 : \tau_n \leq t\}$  and  $N(t) := \sum_{n=1}^{M(t)} Z_n(t)$ . We view this as a population of immortal individuals and we refer to  $Z_n(t)$  as the size of the  $n$ th family,  $M(t)$  the number of families in the system and  $N(t)$  the total size of the population respectively, at time  $t$ . From this perspective  $\tau_n$  represents the foundation time of the  $n$ th family. Furthermore, we see  $F_n$  as a fitness parameter of the  $n$ th family, determining the rate at which new offspring are born into it.

In this paper we aim at proving convergence results for the maximal family in the population. For this we require the following assumptions on the growth processes and fitness distribution.

**Assumption (A1)** (Families' foundation times). *There exists  $\lambda > 0$  such that for all  $n \in \mathbb{N}$*

$$\tau_n = \tau_n^* + T + \varepsilon_n,$$

where  $\tau_n^* := \frac{1}{\lambda} \log n$ ,  $T$  is a finite random variable, and  $\varepsilon_n \rightarrow 0$  almost surely as  $n \rightarrow \infty$ .

**Assumption (A2)** (Growth processes). *There exist  $\gamma > 0$  and an i.i.d. sequence of processes  $((Y_n(t) : t \geq 0))_{n \geq 1}$  independent of  $(F_n)_{n \geq 1}$ , such that*

$$\Delta_n(t) := \sup_{u \geq t} e^{-\gamma u} |X_n(u) - Y_n(u)| \quad (\text{defined for all } t \geq 0)$$

satisfies for all  $\varepsilon, \kappa > 0$ ,

$$\sup_{n \in I_\kappa(t)} \mathbb{P}(\Delta_n(t) \geq \varepsilon | (F_m)_{m \in \mathbb{N}}) \rightarrow 0, \quad \text{in probability as } t \rightarrow \infty, \quad (1)$$

where  $I_\kappa(t)$  is a collection of indices specified below in dependence on the fitness distribution  $\mu$ .

**Assumption (A3)** (Growth rate). *There exists a non-negative random variable  $\xi$  such that*

$$\mathbb{E}[\xi^{\frac{\lambda}{\gamma}}] < \infty \quad \text{and}$$

$$e^{-\gamma t} Y_1(t) \rightarrow \xi, \quad \text{almost surely as } t \rightarrow \infty.$$

The distribution of  $\xi$  is absolutely continuous with respect to the Lebesgue measure. By  $\nu$  we denote its density on  $[0, \infty)$ .

**Assumption (A4)** (Concentration of growth). *There exist  $c_0, \eta > 0$  such that, for  $n \in \mathbb{N}$ , we have*

$$\mathbb{P}(\max_{u \geq 0} X_n(u) e^{-\gamma u} \geq x | (F_m)_{m \in \mathbb{N}}) \leq c_0 e^{-\eta x} \quad \text{almost surely, for all } x \geq 0.$$

**Note.** *On the one hand, Assumption (A1) implies that, for all finite times  $t$ , the number of families born before time  $t$  is finite almost surely. On the other hand, Assumption (A4) implies that each family stays finite at all finite times almost surely. Assumptions (A1) and (A4) together thus imply that our competing growth process does not explode in finite time.*

Beyond these four assumptions on the growth processes we need assumptions on the fitness distribution  $\mu$ . We discuss two different possible classes of fitness distributions  $\mu$ . The first class, the main case discussed in this paper, corresponds to  $\mu$  being in the maximum domain of attraction of the Gumbel distribution. We make the following assumptions.

**Assumption (A5)** ( $\mu$  in the maximum domain of attraction of the Gumbel distribution).

*The function  $m : x \mapsto -\log \mu(x, 1]$  defined for all  $x \in [0, 1)$  is twice differentiable on  $[0, 1)$  and satisfies*

$$(A5.1) \quad m'(x) > 0 \text{ and } m''(x) > 0 \text{ for all } x \in [0, 1);$$

$$(A5.2) \quad \lim_{x \uparrow 1} \frac{m''(x)}{(m'(x))^2} = 0;$$

$$(A5.3) \quad \exists \varkappa > 0 \text{ such that } \lim_{x \uparrow 1} \frac{m''(x)m(x)x}{(m'(x))^2} = \varkappa;$$

$$(A5.4) \quad \lim_{x \uparrow 1} \frac{m(x)}{m'(x)} = 0.$$

**Note.** *Assumption (A5) is sufficient for  $\mu$  to be in the maximum domain of attraction of the Gumbel distribution (see [20, Section 1.1]), and contains the most important cases, but it is not formally necessary. We discuss this further in Section 1.3.*

Under Assumption **(A5)**, for all  $t \geq 0$ , we define  $\sigma_t$  as the minimum of 1 (for technical reasons) and the unique solution  $x$  of

$$(\log g)'(\lambda x) = \frac{1}{\lambda(t-x)}, \quad (2)$$

where  $g(x) = m^{-1}(x)$ , see Lemma 8 for a proof of existence and uniqueness of  $\sigma_t$ . We then define the collection of indices in **(A2)** as

$$I_\kappa(t) := \{n : \frac{|\tau_n^* - \sigma_t|}{\sqrt{\sigma_t}} \leq \kappa\}, \quad \text{for all } \kappa > 0, t \geq 0. \quad (3)$$

The other class of distributions  $\mu$  we consider is the maximum domain of attraction of the Weibull distribution class (see [20, Section 1.2]).

**Assumption (B5)** ( $\mu$  in the maximum domain of attraction of the Weibull distribution). *The distribution  $\mu$  has a regularly varying tail in one, meaning that there exists  $\alpha > 0$  and a slowly varying function  $\ell$  such that  $\mu(1 - \varepsilon, 1) = \varepsilon^\alpha \ell(\varepsilon)$  for all  $\varepsilon \in [0, 1]$ .*

For all  $t \geq 0$ , we set

$$\sigma_t := \tau_{n(t)} \quad \text{and} \quad \sigma_t^* := \tau_{n(t)}^*, \quad \text{where} \quad n(t) = \left\lceil \frac{1}{\mu(1-t^{-1}, 1)} \right\rceil, \quad (4)$$

and

$$I_\kappa(t) := \{n : |\tau_n^* - \sigma_t^*| \leq \kappa\}, \quad \text{for all } \kappa > 0, \quad (5)$$

for use in Assumption **(A2)**. Assumption **(B5)** implies that  $n(t) = \lceil t^\alpha / \ell(t^{-1}) \rceil$  and so  $\log n(t) \sim \alpha \log t - \log(\ell(t^{-1}))$ . Using this together with Assumption **(A1)** we can write

$$\tau_{n(t)} = \frac{1}{\lambda} \log n(t) + T + \varepsilon_{n(t)} = \frac{\alpha}{\lambda} \log t - \frac{1}{\lambda} \log(\ell(t^{-1})) + T + o(1),$$

almost surely as  $t \rightarrow \infty$ , by Assumption **(A1)**.

We now state our results, first in the easier case of  $\mu$  satisfying Assumption **(B5)**. For all  $t \geq 0$ , we define the point process

$$\Gamma_t = \sum_{n=1}^{M(t)} \delta(\tau_n - \sigma_t, t(1 - F_n), e^{-\gamma(t-\sigma_t)} Z_n(t)), \quad (6)$$

on  $(-\infty, \infty) \times (0, \infty) \times (0, \infty)$ , where  $\delta(x)$  is the Dirac mass at  $x$ . We look at the limits of  $\Gamma_t$ , strengthening the result considerably by partially compactifying the underlying space.

**Theorem 1** (Poisson limit). *Under assumptions **(A1-4)** and **(B5)** as  $t \rightarrow \infty$  the point process  $(\Gamma_t)_{t \geq 0}$  converges vaguely<sup>1</sup> in distribution on the space  $[-\infty, \infty] \times [0, \infty] \times (0, \infty]$  to the Poisson point process with locally finite intensity measure*

$$d\zeta(s, f, z) = \alpha f^{\alpha-1} \lambda e^{\lambda s} e^{\gamma(s+f)} \nu(z e^{\gamma(s+f)}) ds df dz,$$

where  $\nu$  is as in **(A3)**.

---

<sup>1</sup>We say that a sequence of Radon measures  $(\mu_n)_{n \in \mathbb{N}}$  on a locally compact Polish space  $\mathbb{X}$  converges *vaguely* to  $\mu$  iff  $\int f d\mu_n \rightarrow \int f d\mu$ , as  $n \rightarrow \infty$ , for all continuous functions  $f: \mathbb{X} \rightarrow \mathbb{R}$  with compact support. This makes the space of Radon measures itself a Polish space, and if  $(\mu_n)_{n \geq 0}$  is a sequence of *random* measures in this space we say that it converges vaguely in distribution to a random Radon measure  $\mu$  iff for all continuous bounded functions  $F$  on this space, the sequence  $(\mathbb{E}F(\mu_n))_{n \geq 0}$  converges to  $\mathbb{E}F(\mu)$  as  $n \rightarrow \infty$ . By the Portmanteau theorem the convergence also holds for bounded functions  $F$  that are continuous at *almost every*  $\mu$ .

Observe that the compactification of the intervals in Theorem 1 ensures that the point with the largest  $z$ -component in the Poisson process corresponds asymptotically to the family of maximal size. Theorem 1 therefore implies the following distributional limits (denoted by  $\Rightarrow$ ) for the size, fitness and the foundation time of the largest family. Note that the open bracket in the third coordinate of the domain on which the point process converge is crucial, as the domain of convergence cannot be extended to  $[-\infty, \infty] \times [-\infty, \infty] \times [0, \infty]$ .

**Corollary 2** (Limits of family characteristics).

(i) As  $t \rightarrow \infty$ , we have

$$e^{-\gamma t + \frac{\gamma \alpha}{\lambda} \log t + \gamma T} \max_{n \in \mathbb{N}} Z_n(t) \Rightarrow W,$$

and  $W$  is Fréchet distributed with shape parameter  $\lambda/\gamma$  and scale parameter

$$s = (\Gamma(\alpha + 1) \lambda^{-\alpha} \mathbb{E}[\xi^{\frac{\lambda}{\gamma}}])^{\frac{\gamma}{\lambda}}.$$

(ii) Denoting by  $V(t)$  the fitness of the family of maximal size at time  $t$ , as  $t \rightarrow \infty$ , we have

$$t(1 - V(t)) \Rightarrow V,$$

where  $V$  is Gamma distributed with shape parameter  $\alpha$  and scale parameter  $\lambda$ .

(iii) Denoting by  $S(t)$  the birth time of the family of maximal size at time  $t$ , as  $t \rightarrow \infty$ , we have

$$S(t) - \sigma_t \Rightarrow U,$$

where  $U$  is a real valued random variable.

**Note.** By Theorem 1, we know that, with high probability as  $t \rightarrow \infty$ , the family that realises  $\max_{n \geq 1} Z_n(t)$  is unique, and thus the definitions of  $S(t)$  and  $V(t)$  are not ambiguous.

The proofs of Theorem 1 and Corollary 2 are carried out in the third author's PhD thesis [21] and can be found at the online repository linked in the bibliography to item [21]. They are not repeated here to limit the length of this paper. The proofs use ideas analogous to those in the present paper, but the execution of these ideas is much simpler. A similar result is contained in [8] in the less general context of reinforced branching processes (see Section 2.1 for details about these processes) using methods that can neither be generalised to the broader class of models considered here, nor at all to the Gumbel case.

To now state *our main technical result* we look at fitness distributions satisfying Assumption **(A5)**. For all  $t \geq 0$ , we define

$$\Gamma_t = \sum_{n=1}^{M(t)} \delta\left(\frac{\tau_n - \sigma_t}{\sqrt{\sigma_t}}, \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}\right), e^{-\gamma g(\lambda\sigma_t)(t - \sigma_t) - a_1 g(\lambda\sigma_t) \log \sigma_t + \gamma T} Z_n(t), \quad (7)$$

where  $\delta(x)$  is the Dirac mass at  $x$ , and  $a_1 := \frac{\gamma}{2\lambda}$ . Note that, by definition (see Equation (2)),  $\sigma_t \geq 1$  almost surely, and thus  $\log(n\sqrt{\sigma_t}) \geq 0$  lies in the domain of definition of  $g$  for all  $t \geq 0$  and  $n \geq 1$ , implying that  $\Gamma_t$  is well defined for all  $t \geq 0$ .

**Theorem 3** (Poisson limit). *Under Assumptions **(A1-5)**, as  $t \rightarrow \infty$ , the point process  $(\Gamma_t)_{t \geq 0}$  converges vaguely in distribution on the space  $[-\infty, \infty] \times [-\infty, \infty] \times (0, \infty]$  to the Poisson point process with locally finite intensity measure*

$$d\zeta(s, f, z) = \lambda e^{-f} e^{s^2 a_2 - f a_3} \nu(z e^{s^2 a_2 - f a_3}) ds df dz,$$

where  $a_2 := \frac{\gamma}{2}\varkappa$ ,  $a_3 := \frac{\gamma}{\lambda}$  and  $\nu$  is as in (A.3).



**Note.** The existence of a density for the random variable  $\xi$  is assumed in **(A3)** for convenience. For example, Theorems 1 and 3 continue to hold if  $\nu = \delta_1$  as in our motivating example.

The technical difference between Theorems 1 and 3 is that in the latter the first (birthtime) coordinate needs to be scaled. As a result the scaling of the second (fitness) component depends on the birth rank  $n$  of the family as well as on the observation time  $t$ . Therefore we cannot derive a general scaling limit for the fitness of the largest family as in Corollary 2. Results for the birth time and size of this family, however, are still possible.

**Corollary 4** (Limits of family characteristics).

(i) As  $t \rightarrow \infty$ , we have

$$e^{-\gamma g(\lambda\sigma_t)(t-\sigma_t) - a_1 g(\lambda\sigma_t) \log \sigma_t + \gamma T} \max_{n \in \mathbb{N}} Z_n(t) \Rightarrow W,$$

where  $W$  is Fréchet distributed with shape parameter  $\lambda/\gamma$  and scale parameter

$$s = \left( \sqrt{\frac{2\pi\lambda}{\lambda}} \mathbb{E} \left[ \xi^{\frac{\lambda}{\gamma}} \right] \right)^{\frac{\gamma}{\lambda}}.$$

(ii) Denoting by  $S(t)$  the birth time of the family of maximal size at time  $t$ , as  $t \rightarrow \infty$ , we have

$$\frac{S(t) - \sigma_t}{\sqrt{\sigma_t}} \Rightarrow U,$$

where  $U$  is normally-distributed with mean 0 and variance  $\frac{1}{\lambda\lambda}$ .

**Note.** Observe that irrespective of whether  $\mu$  is in the maximum domain of attraction of the Weibull or Gumbel distribution, the size of the largest family scaled by a deterministic function of time and the random factor  $e^{\gamma T}$  converges to a Fréchet distribution.

The intuition behind Theorem 3 (see Sections 3 and 4 for the proof, and Figure 1 for a visual aid) is that the only families that have a chance to be the largest at (large) time  $t$  are the ones born at time  $\sigma_t \pm \mathcal{O}(\sqrt{\sigma_t})$  and whose fitness is of leading order  $g(\lambda\sigma_t)$ . This fixes a “moving window” in which we have to look for the representation of the largest family by its birth time and fitness. We prove in Section 3 that the point process  $\Gamma_t$  restricted to the moving window converges to a Poisson point process, and, in Section 4, that the probability that a family outside the window is largest converges to zero.

### 1.3 Examples of fitness distributions

The five following functions  $m(x) = -\log \mu(x, 1)$ , defined for all  $x \in [0, 1)$ , satisfy Assumption **(A5)**:

(1)  $m(x) = (1 - x)^{-\varrho} - 1$ , where  $\varrho > 0$ ;

(2)  $m(x) = e^{\frac{1}{1-x}} - e$ ;

(3)  $m(x) = \frac{x}{1-x}$ ;

(4)  $m(x) = e^{\frac{1}{\sqrt{1-x}}} - e$ ;

(5)  $m(x) = \tan\left(\frac{\pi x}{2}\right)$ .

Assumptions **(A5.1)** and **(A5.2)** imply that the fitness distribution  $\mu$  lies in the maximum domain of attraction of the Gumbel distribution, see [11, ch. 3.3.3]. Although most of the natural examples satisfy Assumptions **(A5.3)** and **(A5.4)**, some probability distributions in the maximum domain of attraction of the Gumbel distribution do not fall into our framework, for example

(6)  $m(x) = \log\left(\frac{e}{1-x}\right) \log \log\left(\frac{e}{1-x}\right)$ ,

see [21, 20] for details.

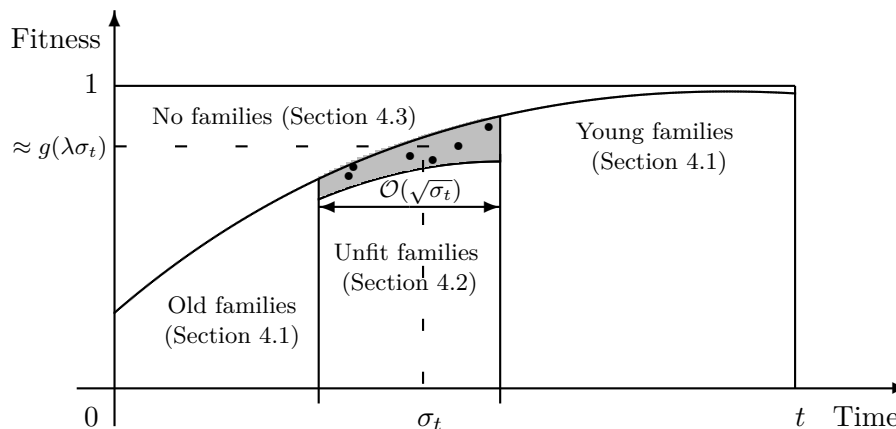


Figure 1: A graphical representation of the proof of Theorem 3. Families are represented by their birth time and fitness. The largest family at large time  $t$  is most likely born at time  $\sigma_t \pm \mathcal{O}(\sqrt{\sigma_t})$ , and has fitness of order  $g(\lambda\sigma_t)$ . Loosely speaking, families that are too old are not fit enough to be large enough, families that are too young have not had sufficient time to grow, and families with a small fitness grow too slowly to compete. Only families in the shaded window appear in the limiting Poisson point process and compete to be the largest.

## 2 Examples and applications

In this section we present a selection of examples covered by our main results. We emphasise that our framework goes well beyond the setup of reinforced branching processes treated in [8] and also that we pick only a small number of representative results out of a wealth of consequences that we can draw from Theorem 3 and Corollary 4. Throughout this section  $\mu$  satisfies our standing assumptions.

### 2.1 Branching processes with selection and mutation

#### 2.1.1 A simple selection and mutation model

Our first example is the model of a population evolving by selection and mutation mentioned at the beginning of Section 1, which we embed into continuous time as follows. We start with one individual with genetic fitness sampled from  $\mu$ . Individuals never die and give birth at a rate given by their fitness to an independent random number of offspring. Note that variations in individual fitness lead to a selection effect: an individual born at time  $t$  selects its parent from the population alive at time  $t$  with a probability proportional to their fitness. At birth each individual independently either inherits the parent's fitness or, with probability  $0 < \beta < 1$ , is a mutant getting a fitness sampled from  $\mu$  independently of everything else. Similar to the deterministic Kingman's model [16, 9] at mutation all genetic information from a particle's ancestry is lost. For a discussion of the relevance of these models in the theory of evolution see [14].

In our framework the non-decreasing sequence of birth times  $\tau_1, \tau_2, \dots$  of mutants constitute the foundation times of new families, their fitnesses are  $F_1, F_2, \dots$  and  $Z_n(t)$  is the number of non-mutant offspring of the  $n$ th mutant at time  $t$ . If  $(p_k)_{k \geq 1}$  is the distribution of offspring numbers at a birth event denote by  $\mathbf{m} = \sum_{k \geq 1} k p_k$  the mean offspring number and assume that  $(p_k)_{k \geq 1}$  has finite support. We assume that mutations have a reasonable chance to produce fit individuals, as expressed in the

Malthusian condition

$$\beta \int_0^1 \frac{d\mu(x)}{1-x} > 1.$$

Under this condition there is a unique solution  $\lambda > (1 - \beta)\mathbf{m}$  of the equation

$$\beta \mathbf{m} \int_0^1 \frac{x}{\lambda - (1 - \beta)\mathbf{m}x} d\mu(x) = 1.$$

In Section 2.1.2, we show that **(A1-4)** are satisfied with  $\gamma = (1 - \beta)\mathbf{m}$ . If  $p_1 = 1$  this is a reinforced branching process as studied in [8]. The generalisation to arbitrary offspring distribution is not difficult (see Section 2.1.2 for details). As an example of the limit theorems implied by our main result, we look at the birth time  $S(t)$  of the largest family at time  $t$  in the case of Gnedenko's distribution (Example (3) in Section 1.3)

$$\mu(x, 1) = e^{-\frac{x}{1-x}}, \quad \text{for } x \in [0, 1),$$

see [12, Exemple 2]. We find a leading order term for  $S(t)$  of

$$\sigma_t = \frac{1}{\lambda} (\sqrt{\lambda t + 1} - 1)$$

and  $\varkappa = 2$ . Corollary 4 therefore gives a central limit theorem of the form

$$\frac{S(t) - \sqrt{t/\lambda}}{\sqrt[4]{t/\lambda}} \rightarrow \mathcal{N}(0, (2\lambda)^{-1}) \text{ in distribution as } t \rightarrow \infty.$$

### 2.1.2 General reinforced branching processes

We now give a general construction for the reinforced branching process where at a birth event, for all  $i, j \in \{0, 1, \dots\}$ , with probability  $p_{ij}$  we create  $i$  new offspring of the same family and  $j$  new families. We assume  $p_{00} = 0$  and denote the first and second marginal by  $(p_i^{(1)})$  and  $(p_j^{(2)})$ , with positive and finite means  $m^{(1)}$  and  $m^{(2)}$ , respectively. Hence, as individuals are immortal, the branching process is supercritical. We also assume that the first marginal is bounded, that is, it has finite support. We can construct the model on an explicit probability space. Let

- $F$  be a  $\mu$ -distributed random variable,
- independently of  $F$  construct a continuous time jump process  $Y = (Y(t) : t \geq 0)$  as follows
  - start at time 0 in state  $Y(0) = 1$ ,
  - if  $Y$  is in state  $k \in \mathbb{N}$  the next jump event follows at rate  $k$ ,
  - let  $0 < t_1 < t_2 < t_3 < \dots$  be the increasing sequence of times at which jump events happen,
  - at jump time  $t_n$  sample a pair  $(J_n, L_n) \in \mathbb{N}_0 \times \mathbb{N}_0$  from  $(p_{ij})$  and increase  $(Y(t) : t \geq 0)$  by  $J_n$  (which may be zero), i.e. set  $Y(t) = Y(t_{n-1}) + J_n$  for all  $t \in [t_n, t_{n+1})$ .
- given the above let  $\Pi = (\Pi(t) : t \geq 0)$  be the jump process which has a jump of height  $L_n$  (which may be zero) at time  $t_n$ .

We let  $((F_n, Y_n, \Pi_n))_{n \geq 1}$  be a sequence of i.i.d. copies of  $(F, Y, \Pi)$ . The process  $(Y_n(t) : t \geq 0)$  describes the creation of new family members, and the process  $(\Pi_n(t) : t \geq 0)$  the creation of new families descending from the  $n$ th family (in a standardised time-scale). To construct our original objects on this probability space we let  $\tau_1 = 0$  and  $Z_1(t) = Y_1(F_1 t)$  and, for  $n \geq 2$  and  $\tau_1, \dots, \tau_{n-1}$  already constructed, iteratively define (recall that, for all  $t \geq 0$  and  $m \geq 1$ , we denote by  $\Delta \Pi_m(t) = \Pi_m(t) - \Pi_m(t-)$ )

$$\tau_n = \inf\{t > \tau_{n-1} : \exists m \in \{1, \dots, n-1\} \text{ with } \Delta \Pi_m(F_m(t - \tau_m)) > 0\},$$

and if  $\Delta\Pi_m(F_m(t - \tau_m)) = k \geq 2$  also set  $\tau_{n+k-1} = \dots = \tau_{n+1} = \tau_n$ . For  $j = 0, \dots, k-1$  let

$$Z_{n+j}(t) = \begin{cases} Y_{n+j}(F_{n+j}(t - \tau_{n+j})), & \text{if } t \geq \tau_{n+j} \\ 0, & \text{otherwise.} \end{cases}$$

We let  $M(t) = \max\{n : \tau_n \leq t\}$  and  $N(t) = \sum_{n=1}^{M(t)} Z_n(t)$ . Now  $(Y_n(F_n(t - \tau_n)) : t \geq \tau_n)$  gives the sizes of the  $n$ th family, and  $(\Pi_n(F_n(t - \tau_n)) : t \geq \tau_n)$  the times of creation of the new families which descend directly from the  $n$ th family. This construction defines a reinforced branching process in a slightly more general way than in [8].

We now check that reinforced branching processes with, for some  $\eta' > 0$ ,

$$\sum_{i,j \geq 0} e^{\eta'(i+j)} p_{ij} < \infty \quad (8)$$

satisfy Assumptions **(A1-4)**. The process  $(M(t) : t > 0)$  is a general branching process, also known as a Crump-Mode-Jagers process, with the laws of offspring times given by the random point process  $(\Pi^*(t) : t > 0)$  given by  $\Pi^*(t) = \Pi(Ft)$ . Assuming that there exists  $\lambda > 0$ , called *Malthusian parameter*, such that

$$\int_0^\infty e^{-\lambda s} \mathbb{E}\Pi^*(ds) = 1, \quad (9)$$

we can apply a strong law of large numbers by Nerman (see [18]) which shows that under an  $x \log x$  condition on  $\Pi^*$  there exists a positive, finite random variable  $W$ , such that

$$\lim_{t \rightarrow \infty} e^{-\lambda t} M(t) = W \quad \text{almost surely.}$$

This gives us that  $\log M(t) = \log W + \lambda t + o(1)$  almost surely as  $t \uparrow \infty$ . Hence  $\tau_n \uparrow \infty$  as  $n \rightarrow \infty$  since  $M(t) < \infty$  for all  $t \geq 0$ , otherwise  $W$  would be infinite with non-zero probability. Plugging  $t = \tau_n$  yields that  $\tau_n = \frac{1}{\lambda} \log n + T + \varepsilon_n$  for  $T = -\frac{1}{\lambda} \log W$  and a sequence  $(\varepsilon_n)$  converging to 0 almost surely.

Note that  $(Y(t) : t > 0)$  is a supercritical continuous-time Galton-Watson process with  $\mathbb{E}Y(t) = e^{m^{(1)}t}$ , see also Lemma 5 below. Given  $(Y(s) : s < t)$  the jump rate of  $\Pi$  at time  $t$  is  $Y(t-)$  and when it jumps, its increment is distributed as  $p^{(2)}$ , and thus we get that

$$\mathbb{E}\Pi^*(ds) = \int_0^1 m^{(2)} f e^{m^{(1)}fs} d\mu(f) ds.$$

Therefore, the Malthusian condition (9) reads as

$$1 = \int_0^\infty e^{-\lambda s} \mathbb{E}\Pi^*(ds) = m^{(2)} \int \frac{f}{\lambda - fm^{(1)}} \mu(df),$$

which has a solution  $\lambda > m^{(1)}$  if and only if

$$m^{(2)} \int_0^1 \frac{f}{1-f} \mu(df) > m^{(1)}.$$

The  $x \log x$  condition states that for the random variable  $X = \int_0^\infty e^{-\lambda s} \Pi^*(ds)$  we have  $\mathbb{E}X \log^+ X < \infty$ . It is straightforward to check that under our assumption on the moments of  $(p_{ij})$  we even have  $\mathbb{E}X^2 < \infty$  so that this condition and hence **(A1)** holds. Indeed, we have

$$\begin{aligned} \mathbb{E}X^2 &= \int_0^\infty \int_0^\infty e^{-\lambda(s+u)} \mathbb{E}[\Pi^*(ds)\Pi^*(du)] \\ &= \int_0^1 d\mu(f) \left( \int_0^\infty \int_0^\infty (m^{(2)})^2 f^2 e^{-(\lambda - m^{(1)}f)(s+u)} ds du + \int_0^\infty \mathbb{E}[\zeta^2] f e^{-(2\lambda - m^{(1)}f)s} ds \right), \end{aligned}$$

where  $\zeta$  has the law  $p^{(2)}$ . We thus get

$$\mathbb{E}X^2 = \int_0^1 \frac{f^2(m^{(2)})^2}{(\lambda - m^{(1)}f)^2} + \frac{f\mathbb{E}\zeta^2}{2\lambda - m^{(1)}f} d\mu(f) \leq \frac{(m^{(2)})^2}{\lambda - m^{(1)}} \int_0^1 \frac{f}{\lambda - m^{(1)}f} d\mu(f) + \mathbb{E}\zeta^2 \int_0^1 \frac{f}{\lambda - m^{(1)}f} d\mu(f),$$

where, for the first integral, we have used the fact that if  $\lambda > m^{(1)}$ , then  $\frac{f}{\lambda - m^{(1)}f}$  is bounded by  $1/(\lambda - m^{(1)})$ . Since  $\lambda$  is the Malthusian parameter, we get

$$\mathbb{E}X^2 \leq \frac{(m^{(2)})^2}{\lambda - m^{(1)}} + \mathbb{E}\zeta^2 < \infty,$$

because  $\mathbb{E}\zeta^2 < \infty$ , by Equation (8); this implies that the  $x \log x$  condition is indeed satisfied.

We let  $Y_n = X_n$  so that  $\Delta_n(t) = 0$  for all  $t \geq 0$ , so the convergence in Assumption **(A2)** is trivially satisfied. The process  $(Y_n(t) : t \geq 0)$  is a continuous-time Galton-Watson process with offspring distribution  $(p_i^{(1)})$ , where the immortal individual itself is not counted as offspring, and hence Assumptions **(A3-4)** follow from Lemma 5 below, parts (c),(d) and (e), respectively.

**Lemma 5** (Galton-Watson process  $(Y(t) : t \geq 0)$  with bounded offspring distribution  $(p_i^{(1)})$ ).  
*With  $\gamma = m^{(1)}$ , we have*

(a)  $\mathbb{E}[Y(t)] = e^{\gamma t}$ .

(b)  $(e^{-\gamma t}Y(t))_{t \geq 0}$  is a uniformly integrable martingale.

(c) The almost sure limit of  $\lim_{t \rightarrow \infty} e^{-\gamma t}Y(t)$  is an absolutely continuous random variable  $\xi$ .

(d) There exists  $\eta > 0$  such that  $\mathbb{E} \exp\{\eta\xi\} < \infty$ .

(e) There exists  $c_0 > 0$  such that  $\mathbb{P}(\max_{t \geq 0} e^{-\gamma t}Y(t) \geq x) \leq c_0 e^{-\eta x}$ , for all  $x \geq 0$ .

*Proof.* (a), (b), (c) are standard. See Athreya and Ney [4] for (a) and Theorem III.7.2 therein for (c), and Asmussen and Hering [2, Theorem 2.1] for (b). Note that the latter is stated for the discrete-time Galton-Watson process, but this implies uniform integrability also for the continuous-time process. Denote the martingale limit in (c) by  $\xi = e^{-\gamma \infty}Y(\infty)$ . (d) follows from [17, Corollary 2.2]. By Jensen's inequality,  $(\exp\{\eta e^{-\gamma t}Y(t)\} : t \in [0, \infty])$  is a sub-martingale. Doob's weak maximal inequality, see [10, Page 443], gives  $\mathbb{P}(\max_{t \geq 0} e^{-\gamma t}Y(t) \geq x) = \mathbb{P}(\max_{t \geq 0} \exp\{\eta e^{-\gamma t}Y(t)\} \geq e^{\eta x}) \leq \mathbb{E}[\exp(\eta\xi)] e^{-\eta x}$ .  $\square$

The selection and mutation model of Section 2.1.1 is a particular case of this more general framework of reinforced branching processes. Recall that at every reproduction event, there is a random number of offspring distributed as  $(p_k)_{k \geq 1}$ , and each of them is a mutant with probability  $\beta$ , independently from the rest. Therefore, for all  $i, j \geq 0$ ,

$$p_{ij} = p_{i+j} \binom{i+j}{i} (1-\beta)^i \beta^j,$$

so that  $m^{(1)} = (1-\beta)\mathbf{m}$  and  $m^{(2)} = \beta\mathbf{m}$ , where  $\mathbf{m} = \sum_{k \geq 1} k p_k$ . Moreover, if  $(p_k)_{k \geq 1}$  has finite support, so has the first marginal  $(p_k^{(1)})_{k \geq 0}$ .

## 2.2 Preferential attachment networks with fitness

### 2.2.1 Preferential attachment tree of Bianconi and Barabási

This model is a random tree where at each step a new vertex is added and connected to an existing vertex with a probability depending on the fitness of the vertices. The model was introduced by Bianconi and Barabási in [6]. We start with two vertices connected by an edge, and endowed with fitnesses sampled independently from  $\mu$ . At every step  $n \geq 3$  a new vertex arrives, gets a fitness sampled from  $\mu$  independently of everything else, and connects to one existing vertex chosen randomly from the  $n - 1$  existing vertices with a probability proportional to the product of their fitness and their degree.

The preferential attachment tree of Bianconi and Barabási can be embedded in continuous time and then represents a reinforced branching process as in [8], its continuous-time embedding is the reinforced branching process (see Section 2.1.2) with  $p_{11} = 1$  so that  $m^{(1)} = m^{(2)} = 1$ , see [8] for details. Here families correspond to vertices and the family size is the vertex degree. At every birth event a new vertex of degree one (equivalently a new family) is created and by establishing an edge to an existing vertex the degree of this vertex is increased by one (equivalently one existing family is getting a new member). At time  $\tau_n$  the  $n$ th vertex is introduced and, for  $m > n$ , the degree of this vertex when the  $m$ th vertex is introduced is  $Z_n(\tau_m)$ . In this embedding  $\tau_n$  is the birthtime of the  $n$ th vertex,  $F_n$  its fitness and  $Z_n(t)$  its degree at time  $t$ . We showed in Section 2.1.2 that under the Malthusian condition

$$\int_0^1 \frac{\mu(dx)}{1-x} > 2$$

the process satisfies Assumptions **(A1-4)** with  $\gamma = 1$  and  $\lambda > 1$  the unique solution of the equation

$$\int_0^1 \frac{x}{\lambda - x} d\mu(x) = 1.$$

We now give an application of our result for the network with fitness distribution

$$\mu(x, 1) = e^{1-(1-x)^{-\varrho}}, \quad \text{for } x \in [0, 1),$$

where  $0 < \varrho < 1$ , see Example (1) in Section 1.3. We estimate  $\sigma_t$ , as defined in Equation (2). Using that  $g(x) = m^{-1}(x) = 1 - (x+1)^{-\frac{1}{\varrho}}$ , we have that, for all  $t \geq 0$  large enough,  $x = \lambda\sigma_t$  is the unique solution of

$$(\log g)'(x) = \frac{1}{\lambda t + 1 - (x+1)},$$

which we can rewrite as  $\lambda t + 1 = \varrho(x+1)^{\frac{\varrho+1}{\varrho}} + (1-\varrho)(x+1)$ . From this we get

$$\sigma_t = x_0 t^{\frac{\varrho}{\varrho+1}} + \mathcal{O}\left(t^{\frac{\varrho-1}{\varrho+1}}\right) \quad \text{as } t \rightarrow \infty, \quad (10)$$

where  $x_0 = \lambda^{-\frac{1}{\varrho+1}} \varrho^{-\frac{\varrho}{\varrho+1}}$ . By definition of  $\varkappa$  in Assumption **(A5.3)** we get

$$\varkappa = \lim_{x \uparrow 1} \frac{m''(x)m(x)x}{(m'(x))^2} = \lim_{x \uparrow 1} \frac{(\varrho+1)x(1-(1-x)^\varrho)}{\varrho} = \frac{\varrho+1}{\varrho}.$$

As an example we apply Corollary 4(i). Denoting by  $a_4 := \varrho^{-\frac{\varrho}{\varrho+1}} + \varrho^{\frac{1}{\varrho+1}}$  and  $a_5 := \frac{\varrho}{2(\varrho+1)}$ , we get the following distributional limit for the size of the largest family. Asymptotically as  $t \rightarrow \infty$ ,

$$e^{-\left(t - \lambda^{-\frac{1}{\varrho+1}} a_4 t^{\frac{\varrho}{\varrho+1}} + \frac{1}{\lambda}\right) - \frac{1}{\lambda} a_5 \log t + T} \max_{n \in \mathbb{N}} Z_n(t) \Rightarrow W,$$

where  $W$  is a Fréchet distributed random variable with shape parameter  $\lambda$  and scale parameter  $s$  given by  $s^\lambda = \sqrt{\frac{2\pi e}{e+1}} \Gamma(\lambda + 1)$ . To get a result, which is independent of the continuous time embedding we look at the time  $\tau_n$  when the  $(n + 1)$ st vertex is introduced. The largest degree at this instance satisfies

$$cn^{\frac{1}{\lambda}} e^{\frac{1}{\lambda} a_4 (\log n)^{\frac{e}{e+1}} - \frac{1}{\lambda} a_5 \log \log n} \leq \max_{m \leq n} Z_m(\tau_n) \leq Cn^{\frac{1}{\lambda}} e^{\frac{1}{\lambda} a_4 (\log n)^{\frac{e}{e+1}} - \frac{1}{\lambda} a_5 \log \log n},$$

where the constants  $c, C$  are positive, finite random variables.

### 2.2.2 Preferential attachment network of Dereich

Dereich in [7] defined an alternative preferential attachment model with fitness that can be studied without a Malthusian condition. In the model a new vertex is connected to each existing vertex independently by a random number of edges, defining a multigraph.

Start with one vertex labelled one, with fitness  $F_1$  drawn from  $\mu$  and no edges. Denote the graph by  $\mathcal{G}_1$ . Given  $\mathcal{G}_m$  with vertex set  $\{1, \dots, m\}$  we build  $\mathcal{G}_{m+1}$  by introducing the vertex labelled  $m + 1$ , giving it fitness  $F_{m+1}$  drawn from  $\mu$  and connecting it independently to each vertex  $n \in \{1, \dots, m\}$  by a random number  $E_{n,m+1}$  of directed edges (from vertex  $m + 1$  to  $n$ ), which is Poisson distributed with rate

$$r_{n,m} := \beta F_n \frac{1 + \text{indegree of } n \text{ in } \mathcal{G}_m}{m},$$

where  $0 < \beta < 1$  is a fixed parameter.

This model can be embedded into continuous space by letting  $\tau_n = \frac{1}{\lambda} \sum_{i=1}^{n-1} \frac{1}{i}$ , for  $\lambda > 0$ , be the time when the  $n$ th vertex is introduced and defining  $Z_n(\tau_m)$ ,  $m \geq n$  to be the indegree of vertex  $n$  prior to the establishment of vertex  $m + 1$ , or in other words the number of edges pointing from vertices  $n + 1, \dots, m$  to vertex  $n$ . Note that the indegree process  $(Z_n(t) : t \geq \tau_n)$  has  $Z_n(\tau_n) = 0$  and it is actually the process  $(1 + Z_n(t) : t \geq \tau_n)$  that corresponds to the family sizes in our general framework. This is taken into account when we check below that this model satisfies assumptions **(A1-4)** without any Malthusian condition for  $\gamma = \lambda\beta$ . But we first show what kind of information can be obtained by applying our main results to this model.

We look at the fitness  $V(t)$  of the vertex  $m \in \{1, \dots, n - 1\}$  with largest degree at the time  $t = \frac{1}{\lambda} \log n + C + o(1)$  when the  $n$ th vertex is introduced (where  $C$  denotes the Euler-Mascheroni constant) again in the case Gnedenko's distribution (Example (3) in Section 1.3). Recall that in this case  $g(x) = \frac{x}{1+x}$  and  $\lambda\sigma_t = \sqrt{\lambda t + 1} - 1$ . We denote by  $S(t)$  the time of creation of this vertex; by Corollary 4, we have  $S(t) = \sigma_t + (W + o(1))\sqrt{\sigma_t/\lambda}$  in distribution when  $t \uparrow \infty$ , where  $W$  is a centred Gaussian random variable of variance  $1/2$ . Theorem 3 gives that, in distribution when  $t \uparrow \infty$ ,

$$\begin{aligned} V(t) &= g\left(\lambda\sigma_t + \sqrt{\lambda\sigma_t}(W + o(1))\right) + \mathcal{O}\left(g'\left(\lambda\sigma_t + \sqrt{\lambda\sigma_t}(W + o(1))\right)\right) \\ &= 1 - \frac{1}{\lambda\sigma_t} + \frac{W + o(1)}{\sqrt{\lambda\sigma_t}} = 1 - \frac{1}{\sqrt{\lambda t}} + \frac{W + o(1)}{(\lambda t)^{1/4}}, \end{aligned}$$

so that there is asymptotic normality for the fitness of the vertex of maximal degree. This is in contrast to the result in Corollary 2(ii) for the case of  $\mu$  in the maximum domain of attraction of the Weibull distribution, where  $t(1 - V(t))$  converges to a Gamma distribution.

The rest of this section is devoted to the proof of **(A1-4)** for the model of Dereich. Assumption **(A1)** is straightforward for the deterministic choice

$$\tau_n = \frac{1}{\lambda} \sum_{i=1}^{n-1} \frac{1}{i} = \frac{1}{\lambda} \log n + \frac{C}{\lambda} + o(1),$$

where  $C$  is the Euler-Mascheroni constant. To show that Assumption **(A2)** is satisfied we introduce a coupling of the indegree processes  $(Z_n(t): t \geq 0)$  to independent Yule processes. For all  $n \geq 1$ ,  $u \geq 0$ , we set

$$X_n(u) = Z_n\left(\frac{u}{F_n} + \tau_n\right).$$

**Proposition 6.** *There exists a coupling of the processes  $(X_n(u): u \geq 0)$  and a sequence  $(Y_n(u): u \geq 0)$  of independent Yule processes with parameter  $\gamma = \beta\lambda$  such that (see Equation (3) for the definition of  $I_\kappa(t)$ ), as  $t \rightarrow \infty$ , in probability,*

$$\sup_{n \in I_\kappa(t)} \mathbb{P}\left(\sup_{u \geq t} e^{-\gamma u} |1 + X_n(u) - Y_n(u)| \geq \varepsilon | (F_i)_i \right) \rightarrow 0.$$

To prove this start with a sequence  $(Y_n(u): u \geq 0)$  of independent Yule processes with parameter  $\gamma$ . For  $m \geq n + 1$  we take

$$J_n(m) = Y_n(F_n(\tau_m - \tau_n)) - Y_n(F_n(\tau_{m-1} - \tau_n)).$$

We need the following lemma.

**Lemma 7.** *Given  $n$  there is a coupling of  $J_n(m)$  and random variables  $P_n(m)$ ,  $m \geq n + 1$ , such that*

- *conditionally on  $F_n = f \in (0, 1)$  the random variable  $P_n(n + 1)$  is Poisson distributed with parameter  $\beta f \frac{1}{n+1}$ , and*
- *for  $m \geq n + 2$ , conditionally on  $F_n = f \in (0, 1)$  and  $\sum_{\ell=n+1}^{m-1} P_n(\ell) = k \in \{0, 1, \dots\}$ , the random variable  $P_n(m)$  is Poisson distributed with parameter  $\beta f \frac{1+k}{m}$ ,*

*and  $\sup_{n \in I_\kappa(t)} \mathbb{P}(J_n(m) \neq P_n(m) \text{ for some } m \geq n + 1) \rightarrow 0$ , as  $t \rightarrow \infty$ .*

*Proof.* We abbreviate  $Y_n^*(t) = Y_n(F_n(t - \tau_n))$  and note that, conditionally on  $F_n$ ,  $(Y_n^*(t): t \geq \tau_n)$  is a continuous time Galton-Watson process starting with one individual at time  $\tau_n$  and individuals performing binary branching at rate  $\gamma F_n$ . The coupling is now performed in two steps.

- (a) For  $m \geq n + 1$ , we let  $\mathcal{E}_{n,m}$  be the event that all of the individuals alive at time  $\tau_{m-1}$  have at most one descendant in the interval  $[\tau_{m-1}, \tau_m)$ . This means that an individual existing at time  $\tau_{m-1}$  can only give birth to at most one individual, which in turn does not reproduce before  $\tau_m$ . Denoting

$$\mathcal{E}_n(t) = \bigcap_{\substack{m \geq n+1 \\ \tau_{m+1} < t}} \mathcal{E}_{n,m},$$

we show that

$$\sup_{n \in I_\kappa(t)} \mathbb{P}(\mathcal{E}_n^c(t)) \rightarrow 0, \quad \text{as } t \rightarrow \infty.$$

- (b) For all  $m \geq n + 1$ , there are random variables  $J_n^*(m)$ , which, conditionally on  $F_n = f$  and  $\sum_{\ell=n+1}^{m-1} J_n^*(\ell) = k$ , are binomially distributed with parameters  $1 + k$  and  $\beta f/m$ , such that  $J_n^*(m) = J_n(m)$  on  $\mathcal{E}_{n,n+1} \cap \dots \cap \mathcal{E}_{n,m}$ . We can couple  $J_n^*(m)$  to random variables  $P_n(m)$ , which given  $F_n = f \in (0, 1)$  and  $\sum_{\ell=n+1}^{m-1} P_n(\ell) = k \in \{0, 1, \dots\}$  are Poisson distributed with parameter  $\beta f \frac{1+k}{m}$  such that

$$\sup_{n \in I_\kappa(t)} \mathbb{P}(J_n^*(m) \neq P_n(m) \text{ for some } m \geq n + 1) \rightarrow 0, \quad \text{as } t \rightarrow \infty.$$

It is clear that the lemma follows from claims (a) and (b).



We now prove (a). Fix  $n \in I_\kappa(t)$  and let  $m > n$ . Denote by  $\eta = \gamma F_n$  and by  $W_\theta$  an independent random variable, exponentially distributed with parameter  $\theta$ . Recall that in a Yule process of rate  $\eta$  each particle gives birth to one offspring after an exponentially distributed waiting time with rate  $\eta$ , independently of everything else. Thus the conditional probability that a fixed particle has at least one offspring in the interval  $[\tau_m, \tau_{m+1})$  is equal to

$$\mathbb{P}(W_\eta \leq \tau_{m+1} - \tau_m | (F_i)_i) = 1 - e^{-\frac{\eta}{\lambda m}} \leq \frac{\beta F_n}{m}.$$

Furthermore, the probability of a given particle having at least 2 descendants in  $[\tau_m, \tau_{m+1})$  is equal to

$$\begin{aligned} \mathbb{P}(W_\eta + W_{2\eta} \leq \tau_{m+1} - \tau_m | (F_i)_i) &= 1 + e^{-\frac{2\eta}{\lambda m}} - 2e^{-\frac{\eta}{\lambda m}} \\ &= (1 - e^{-\frac{\beta F_n}{m}})^2 \leq \frac{(\beta F_n)^2}{m^2}, \end{aligned} \quad (11)$$

where  $W_{2\eta}$  is the minimum of two independent exponentially distributed waiting times with rate  $\eta$ . Using the law of total probability we can express the probability that at least one particle of  $(Y_n^*(t) : t \geq 0)$  at time  $\tau_m$  has at least 2 descendants in the interval  $[\tau_m, \tau_{m+1})$ ,

$$\mathbb{P}(\mathcal{E}_{n,m+1}^c | (F_i)_i) = \sum_{k=1}^{\infty} \mathbb{P}(\mathcal{E}_{n,m+1}^c | Y_n^*(\tau_m) = k, (F_i)_i) \mathbb{P}(Y_n^*(\tau_m) = k) \leq \frac{(\beta F_n)^2}{m^2} \mathbb{E}[Y_n^*(\tau_m) | (F_i)_i]. \quad (12)$$

By Lemma 5(a), we have

$$\mathbb{E}[Y_n^*(\tau_m) | (F_n)_n] = \mathbb{E}[Y_n(F_n(\tau_m - \tau_n)) | (F_i)_i] = e^{\beta \lambda F_n (\tau_m - \tau_n)} \sim \left(\frac{m}{n}\right)^{\beta F_n}, \quad (13)$$

where we have used the fact that  $\tau_m - \tau_n = \frac{1}{\lambda} \log\left(\frac{m}{n}\right) + o(1)$  almost surely for all  $m \geq n$  and  $n$  large (see Lemma 5(a)). We now look at  $n$  such that  $n \in I_\kappa(t)$  or, equivalently,

$$e^{\lambda(\sigma_t - \kappa\sqrt{\sigma_t})} \leq n \leq e^{\lambda(\sigma_t + \kappa\sqrt{\sigma_t})}.$$

Putting this together with Equations (12) and (13) we get

$$\begin{aligned} \mathbb{P}(\mathcal{E}_n^c(t) | (F_i)_i) &\leq \sum_{m=n}^{\infty} \mathbb{P}(\mathcal{E}_{n,m+1}^c | (F_i)_i) \leq \text{const.} \sum_{m \geq e^{\lambda(\sigma_t - \kappa\sqrt{\sigma_t})}} \frac{(\beta F_n)^2}{m^2} \left(\frac{m}{n}\right)^{\beta F_n} \\ &\leq \text{const.} \frac{(\beta F_n)^2}{e^{\beta \lambda F_n (\sigma_t - \kappa\sqrt{\sigma_t})}} \sum_{m \geq e^{\lambda(\sigma_t - \kappa\sqrt{\sigma_t})}} m^{\beta F_n - 2} \\ &\leq \text{const.} \frac{(\beta F_n)^2}{e^{\beta \lambda F_n (\sigma_t - \kappa\sqrt{\sigma_t})}} \int_{e^{\lambda(\sigma_t - \kappa\sqrt{\sigma_t})}}^{\infty} x^{\beta F_n - 2} dx \leq \text{const.} \frac{(\beta F_n)^2}{1 - \beta F_n} e^{-\lambda(\sigma_t - \kappa\sqrt{\sigma_t})}, \end{aligned}$$

which goes to zero, as  $t \rightarrow \infty$ . This completes the proof of (a).

To show (b) fix  $n \in I_\kappa(t)$  and let  $m \geq n + 1$ . Note that the existence of  $J_n^*(m)$  binomially distributed with parameters  $k + 1$  and  $\beta f/m$  such that  $J_n^*(m) = J_n(m)$  on  $\mathcal{E}_{n,n+1} \cap \dots \cap \mathcal{E}_{n,m}$  is easy because on this event there are  $k + 1$  individuals alive at time  $\tau_{m-1}$  and each independently produces offspring with probability  $\beta f/m$ .

Moreover, if for all  $n \geq 1$  we let

$$A_n := \sup_{s \geq \tau_n} \frac{Y_n^*(s)}{e^{\gamma F_n s}},$$

then the  $(A_n)_{n \geq 1}$  are i.i.d. random variables that satisfy Lemma 5(e) and, almost surely for all  $m \geq n$ ,  $Y_n^*(\tau_m) \leq A_n e^{\gamma F_n (\tau_m - \tau_n)}$ . Applying Theorem 9 of [13, ch.4.12], we get that the conditional total variation distance between  $J_n^*(m)$  and  $P_n(m)$  conditionally on  $(F_i)_i$  and  $Y_n^*(\tau_{m-1})$  satisfies

$$d_{\text{TV}}(J_n^*(m), P_n(m)) \leq 2 \sum_{i=1}^{Y_n^*(\tau_{m-1})} \left(\frac{\beta F_n}{m}\right)^2 = \frac{2(\beta F_n)^2 Y_n^*(\tau_{m-1})}{m^2} \leq \frac{A_n (\beta F_n)^2}{m^{2 - \beta F_n} \eta^{\beta F_n}},$$

almost surely. This implies that (see, e.g. [13, Exercise 4.12.5]) there exists a coupling of  $J_n^*(m)$  and  $P_n(m)$ ,  $m \geq n + 1$ , such that

$$\mathbb{P}(J_n^*(m) \neq P_n(m) \text{ for some } m \geq n + 1 | (F_i)_i) \leq A_n \sum_{m=n}^{\infty} \frac{(\beta F_n)^2}{m^{2-\beta F_n} n^{\beta F_n}} \leq \frac{A_n}{n},$$

using again that  $\beta < 1$ . This implies that

$$\sup_{n \in I_\kappa(t)} \mathbb{P}(J_n^*(m) \neq P_n(m) \text{ for some } m \geq n + 1 | (F_i)_i) \leq \sup_{n \in I_\kappa(t)} \left\{ \frac{A_n}{n} \right\} \leq \frac{\sup_{n \in I_\kappa(t)} A_n}{\inf(I_\kappa(t))},$$

where  $\inf(I_\kappa(t))$  is the smallest element of the set  $I_\kappa(t)$ , i.e.  $\lceil \exp(\lambda(\sigma_t - \kappa\sqrt{\sigma_t})) \rceil$  (see Equation (3)). Note that, in distribution,  $\sup_{i \in I_\kappa(t)} A_i = \sup_{1 \leq i \leq |I_\kappa(t)|} A_i$ . Moreover, by Lemma 5(e), there exists  $c_0, \eta > 0$  such that for all  $\varepsilon > 0$ , for all  $n \geq 1$ ,

$$\mathbb{P}\left(\max_{1 \leq i \leq n} A_i \geq (1 + \varepsilon) \log n / \eta\right) \leq 1 - (1 - c_0 n^{-1-\varepsilon})^n \sim c_0 n^{-\varepsilon} \rightarrow 0 \quad \text{as } n \uparrow \infty.$$

By definition of  $I_\kappa(t)$ , we have  $|I_\kappa(t)| \sim 2\kappa\sqrt{\sigma_t} \uparrow \infty$ , and thus

$$\lim_{t \uparrow \infty} \mathbb{P}\left(\sup_{i \in I_\kappa(t)} A_i \geq (1 + \varepsilon) \log |I_\kappa(t)|\right) \rightarrow 0.$$

Since we also have that  $\inf(I_\kappa(t)) = \sigma_t - \kappa\sqrt{\sigma_t}$ , we get that, in probability

$$\sup_{n \in I_\kappa(t)} \mathbb{P}(J_n^*(m) \neq P_n(m) \text{ for some } m \geq n + 1 | (F_i)_i) \rightarrow 0 \text{ when } t \uparrow \infty,$$

which concludes the proof.  $\square$

To complete the proof of Proposition 6 we define

$$X_n^*(F_n(t - \tau_n)) = \sum_{k=n+1}^m P_n(k), \quad \text{for all } m \geq n + 1 \text{ and } \tau_m \leq t < \tau_{m+1},$$

and note that, for all  $n \geq 1$ , the process  $(X_n^*(t) : t \geq \tau_n)$  has the same distribution as  $(X_n(t) : t \geq \tau_n)$ . Moreover,

$$\begin{aligned} \mathbb{P}(1 + X_n(F_n(\tau_m - \tau_n)) = Y_n(F_n(\tau_m - \tau_n)) \text{ for all } m \geq n + 1) \\ = 1 - \mathbb{P}(J_n(m) \neq P_n(m) \text{ for some } m \geq n + 1) \end{aligned}$$

because  $X_n(F_n(\tau_m - \tau_n)) = \sum_{k=n+1}^m P_n(k)$  and  $Y_n(F_n(\tau_m - \tau_n)) = 1 + \sum_{k=n+1}^m J_n(k)$ . Suppose now that  $\tau_m \leq t < \tau_{m+1}$  and  $1 + X_n(F_n(\tau_m - \tau_n)) = Y_n(F_n(\tau_m - \tau_n))$ . Then, almost surely, as  $m \uparrow \infty$ ,

$$\begin{aligned} |1 + X_n(F_n(t - \tau_n)) - Y_n(F_n(t - \tau_n))| &= |Y_n(F_n(\tau_m - \tau_n)) - Y_n(F_n(t - \tau_n))| \\ &= (\xi_n + o(1))e^{\gamma F_n(t - \tau_n)} - \xi_n e^{\gamma F_n(\tau_m - \tau_n)} \\ &\leq (\xi_n + o(1))e^{\gamma F_n(t - \tau_n)}(1 - e^{-\beta F_n/m}), \end{aligned}$$

and hence

$$\begin{aligned} \sup_{n \in I_\kappa(t)} \mathbb{P}\left(\sup_{u \geq t} e^{-\gamma u} |1 + X_n(u) - Y_n(u)| \geq \varepsilon | (F_i)_i \right) \\ \leq \sup_{n \in I_\kappa(t)} \mathbb{P}\left(\xi_n(1 - e^{-\beta F_n/n}) \geq \varepsilon/2 | (F_i)_i\right) + \sup_{n \in I_\kappa(t)} \mathbb{P}(J_n(m) \neq P_n(m) \text{ for some } m \geq n + 1 | (F_i)_i) \\ \rightarrow 0 \text{ in probability as } t \uparrow \infty. \end{aligned}$$

This completes the proof of Proposition 6 and hence of Assumption **(A2)**. Further, from Lemma 5(c) we see that Assumption **(A3)** holds.

Finally, to prove Assumption **(A4)** we fix the fitnesses  $(F_n)_{n \geq 1}$  and work conditionally on this sequence of random variables. Note that, by definition, the jump of  $(X_n(t) : t \geq \tau_n)$  at time  $t = F_n(\tau_m - \tau_n)$  given  $X_n(F_n(\tau_{m-1} - \tau_n)) = k$  is Poisson distributed with parameter  $\beta F_n \frac{1+k}{m}$ . Hence the processes  $(M_m^{(n)} : m \geq n)$  given by

$$M_m^{(n)} := (1 + X_n(F_n(\tau_m - \tau_n))) \prod_{\ell=n+1}^m (1 + \frac{\beta F_n}{\ell})^{-1}$$

are martingales, i.e. for all  $n \geq 1$ , for all  $m \geq n + 1$ ,  $\mathbb{E}[M_{m+1}^{(n)} | (F_i)_i, M_m^{(n)}] = M_m^{(n)}$ . The scaling factor satisfies

$$f_m^{-1} := \prod_{\ell=n+1}^m (1 + \frac{\beta F_n}{\ell}) = e^{\gamma(F_n(\tau_m - \tau_n))} (1 + o(1)).$$

Hence almost sure limits  $M_\infty^{(n)} = \lim_{m \rightarrow \infty} M_m^{(n)}$  exist and Doob's submartingale inequality yields

$$\begin{aligned} \mathbb{P}(\max_{u \geq 0} X_n(u) e^{-\gamma u} \geq x) &\leq \mathbb{P}(\max_{m \geq n+1} M_m^{(n)} \geq x/2(1 + o(1)) | (F_i)_i) \\ &\leq \mathbb{E}[\max_{m \geq n+1} e^{2\varpi M_m^{(n)}} | (F_i)_i] e^{-\varpi x} \leq \mathbb{E}[e^{2\varpi M_\infty^{(n)}} | (F_i)_i] e^{-\varpi x}, \end{aligned}$$

for all  $\varpi > 0$ . It remains to show that there exists  $\varpi > 0$  such that  $\mathbb{E}[e^{\varpi M_\infty^{(n)}}] < \infty$  or, using Fatou's lemma, that  $\mathbb{E}[\exp(\varpi M_m^{(n)})]$  remains bounded. Using the generating function for Poisson variables we get

$$\mathbb{E}[e^{\varpi M_m^{(n)}} | (F_i)_i, X_n(F_n(\tau_{m-1} - \tau_n)) = k] = \exp((1+k)(\eta f_m + \beta F_n \frac{1}{m}(e^{\varpi f_m} - 1))).$$

Hence, using that  $e^{\varpi f_m} - 1 \leq \varpi f_m + C\varpi^2 f_m^2$  for some constant  $C > 0$ , we get

$$\mathbb{E}[e^{\varpi M_m^{(n)}} | (F_i)_i] \leq \mathbb{E}[e^{(\varpi + C\varpi^2 \frac{1}{m} f_m) M_{m-1}^{(n)}}],$$

and iterating this we get an upper bound of  $e^{a_{m-n}}$  for the recursion  $a_0 = \varpi$  and

$$a_{i+1} = a_i + C a_i^2 \frac{1}{m-i} f_{m-i}, \quad \text{for } i \geq 0.$$

As  $f_\ell \sim (n/\ell)^{\beta F_n}$  almost surely when  $\ell \geq n + 1 \rightarrow \infty$ , there exists an almost-surely finite  $(F_i)_i$ -measurable random variable  $A$  such that

$$\prod_{\ell=n+1}^m (1 + C \frac{1}{\ell} f_\ell) \leq A \quad \text{for all } m \geq n \text{ and } n.$$

Hence  $(a_{m-n} : m \geq n)$  is bounded by one if  $0 < \varpi < 1/A$ . This completes the proof of **(A4)**.

### 2.3 Random permutations with random cycle weights

Let  $\theta \geq 0$  be a fixed parameter and suppose we are given a permutation  $\sigma$  of the indices  $\{1, \dots, n\}$  and, for each of the  $k$  cycles of the permutation, a weight  $W_j$ ,  $j = 1, \dots, k$ . Denote the length of the cycles by  $Z_1, \dots, Z_k$ . We create a permutation  $\sigma'$  of the indices  $\{1, \dots, n+1\}$  from this as follows

- either pick one of the indices  $m \in \{1, \dots, n\}$  from the  $j$ th cycle with probability  $\frac{W_j}{n+\theta}$  and insert the new index into its cycle so that we have  $\sigma'(m) = n+1$ ,  $\sigma'(n+1) = \sigma(m)$  and  $\sigma'(i) = \sigma(i)$  for all  $i \neq m, n+1$ ;

- with the remaining probability  $1 - \frac{\sum_{j=1}^k Z_j W_j}{n+\theta}$  the new index  $n+1$  is mapped onto itself, creating a new cycle of length one. This cycle is given a weight  $W_{k+1}$  sampled, independently of everything else, from  $\mu$ .

The resulting process  $(\sigma_n)$  can be seen as a *disordered chinese restaurant process*. The idea is that the cycles correspond to tables and new customers either join a table with a probability proportional to both the weight and the number of seats on the table, or sit at a new table. In the original chinese restaurant process customers chose to sit on a table with a probability proportional to the number of seats and the probability of introducing a new table is  $\frac{\theta}{n+\theta}$ , see [1, p. 92]. This corresponds to all weights being equal to one in our scenario. We briefly mention that this model differs from the model of Betz, Ueltschi and Velenik on random permutations with cycle weights, as in their case the weight of a cycle is not random and instead depends on the size of the cycle, see [5].

Let us show that this model falls into our framework of competing growth processes and satisfies Assumptions **(A1-4)**. Key is again an embedding of the process in continuous time such that  $T_n$  is the time when the  $n$ th customer enters the restaurant. We let  $T_1 = 0$  and define  $T_{n+1}$ ,  $n \in \mathbb{N}$ , inductively as follows. At time  $T_n$  we start  $n+1$  independent exponential clocks, one clock of parameter one for each of the  $n$  customers seated in the restaurant and one additional clock of parameter  $\theta$  for the creation of additional tables. We let  $T_{n+1}$  be the time when the first of these clocks rings.

- If it is the clock corresponding to customer  $m$  sitting at table  $j$  we toss a coin with success probability  $W_j$ .
  - If there is a success the  $(n+1)$ st customer joins this table, resp. in the language of random permutations the element  $n+1$  is inserted in this cycle between elements  $m$  and  $\sigma_n(m)$ ,
  - if there is no success the  $(n+1)$ st customer sits at a new table which, if it is the  $(k+1)$ st occupied table, gets weight  $W_{k+1}$ .
- If it is the clock for the creation of additional tables, the  $(n+1)$ st customer also sits at a new table which, if it is the  $(k+1)$ st occupied table, gets weight  $W_{k+1}$ .

Suppose  $W_1, W_2, \dots$  are given. We note that, as required, the overall probability that a new table is created at time  $T_{n+1}$  is

$$\frac{\sum_{j=1}^k Z_j(T_n)(1 - W_j) + \theta}{n + \theta} = 1 - \frac{\sum_{j=1}^k Z_j(T_n)W_j}{n + \theta},$$

where  $Z_j(T_n)$  is the number of occupants at the  $j$ th table at time  $T_n$ , and the probability that the  $(n+1)$ st customer joins the  $j$ th table is  $Z_j(T_n)W_j/(n+\theta)$ . Looking at the  $j$ th table, we let  $\tau_j$  be the time when it is first occupied. If at time  $t$  this table is occupied by  $m$  customers the rate at which new customers join this table is  $mW_j$ , independent of the occupancy of other tables. The processes  $(Z_j(t+\tau_j) : t \geq 0)$  are therefore independent Yule processes with rate  $W_j$ . Hence Assumptions **(A2-4)** are satisfied for  $\gamma = 1$  and where  $X_n(u) = Y_n(u)$ ,  $u \geq 0$ , are given by  $Z_n(t) = X_n(W_n(t - \tau_n))$ .

Finally, to check Assumption **(A1)** we note that the process of introduction of new tables is a general branching process with immigration. The immigration process corresponds to the creation of the additional tables, which is a homogeneous Poisson process with rate  $\theta$ . The point process of creation of tables by unsuccessful coin tossing is a Cox process  $(\Pi(t) : t \geq 0)$ , i.e. a Poisson process with random intensity. Its intensity is given by  $(1 - W)Y(t) dt$  where  $W$  has distribution  $\mu$  and given  $W$  the process  $(Y(t) : t \geq 0)$  is a Yule process with parameter  $W$ . The relevant results for general branching processes can be found in [18] with the case of branching processes with immigration treated in [19]. The crucial assumption is the existence of a Malthusian parameter  $\alpha \geq 0$  such that

$$1 = \int e^{-\alpha t} \mathbb{E}\Pi(dt) = \int \int_0^\infty (1 - w)e^{-\alpha t} e^{wt} dt \mu(dw) = \int \frac{1 - w}{\alpha - w} \mu(dw),$$

which is always satisfied for  $\alpha = 1$ . As above it is a routine exercise to check the  $x \log x$  condition on  $\int e^{-t} \Pi(dt)$ . We obtain from [18, Theorem 5.4] for general branching processes without immigration (our case  $\theta = 0$ ) and modifications described in [19, Theorem 4.2] for the general case (stated there only for convergence in  $L^1$ ) that there exists a positive random variable  $M_\theta$  such that the total number  $M(t)$  of tables which have been occupied by time  $t$  satisfies

$$e^{-t} M(t) \longrightarrow M_\theta \quad \text{almost surely,}$$

from which we infer that  $\tau_n = \log n - \log M_\theta + o(1)$ , which is Assumption **(A1)** with  $\lambda = 1$ .

We now give an example of a result that follows from our main technical result. We look at the ratio  $R(t)$  of the size of the largest and second largest cycle in the permutation at time  $t$ . We have

$$\mathbf{1}_{R(t) \geq x} = \int \mathbf{1}_{\Gamma_t([-\infty, \infty] \times [-\infty, \infty] \times (z, \infty))=0} \mathbf{1}_{\Gamma_t([-\infty, \infty] \times [-\infty, \infty] \times (z/x, z))=0} d\Gamma_t(s, f, z).$$

If  $\mu$  satisfies Assumption **(A5)**, then, by Theorem 3, we hence have, for  $x > 1$ , with  $N$  a Poisson point process with intensity measure  $\zeta$ ,

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{P}(R(t) \geq x) &= \mathbb{E} \int \mathbf{1}_{N([-\infty, \infty] \times [-\infty, \infty] \times (z, \infty))=0} \mathbf{1}_{N([-\infty, \infty] \times [-\infty, \infty] \times (z/x, z))=0} N(ds df dz) \\ &= \int \exp(-\zeta((-\infty, \infty) \times (-\infty, \infty) \times (z/x, \infty))) \zeta(ds df dz). \end{aligned}$$

Using that  $\nu(x) = e^{-x}$  and  $a_3 = 1$  in the first equality (similar as in (35) below) and the change of variable  $v = f - \log y$  in the second, we get that

$$\begin{aligned} \zeta((-\infty, \infty) \times (-\infty, \infty) \times (z/x, \infty)) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ds df e^{s^2 a_2 - 2f} \int_{z/x}^{\infty} e^{-y e^{s^2 a_2 - f}} dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ds dv e^{s^2 a_2 - 2v} e^{-e^{s^2 a_2 - v}} \int_{z/x}^{\infty} y^{-2} dy = a_5 \frac{x}{z}, \end{aligned}$$

where  $a_5$  is a positive constant. Hence, substituting  $f$  by  $f + \log x$  in the final step,

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{P}(R(t) \geq x) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ds df \int_0^{\infty} dz e^{-f} e^{s^2 a_2 - f} e^{-z(e^{s^2 a_2 - f}) - a_5 \frac{x}{z}} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ds df \int_0^{\infty} dw e^{-f} e^{-w - a_5 \frac{1}{w} e^{s^2 a_2 - f + \log x}} = \frac{1}{x}. \end{aligned}$$

Similarly, if  $\mu$  satisfies the assumptions (B.5), we have  $\zeta((-\infty, \infty) \times (0, \infty) \times (z/x, \infty)) = a_6 \frac{x}{z}$ , and hence by Theorem 1,

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{P}(R(t) \geq x) &= \int_{-\infty}^{\infty} ds \int_0^{\infty} df \int_0^{\infty} dz \alpha f^{\alpha-1} e^{2s+f} e^{-ze^{s+f} - a_6 \frac{x}{z}} \\ &= \int_{-\infty}^{\infty} ds \int_0^{\infty} df \int_0^{\infty} dz x \alpha f^{\alpha-1} e^{2s+f} e^{-ze^{s+f+\log x} - a_6 \frac{1}{z}} = \frac{1}{x}. \end{aligned}$$

substituting  $s$  by  $s - \log x$  in the final step. Note that this is in contrast to the case without disorder where the cycles have macroscopic size and the distribution of the asymptotic ratio is given by the ratio of the two largest elements in the Poisson-Dirichlet distribution.

The remainder of the paper is devoted to the proofs of Theorem 3 and Corollary 4 and is structured as follows. In Section 3 we look at the Poisson limit theorem given in Theorem 3, but first in a space without compactifications. After some preparations we prove in Section 3.2 a basic form of the limit theorem, see Proposition 10. This is derived from an approximation which corresponds to a classical Poisson convergence result for extremes in the first two components and an independent

third component. In Section 3.3 a further approximation turns the basic form into the original form of the Poisson limit theorem, the crucial difference being that the scaling of the third component becomes independent of the birth rank  $n$  of the family. Section 4 is devoted to the compactification of the space, effectively showing that the points suppressed by the scalings do not provide the largest families. These points are either born too late (Section 4.1) or not fit enough (Section 4.2). In Section 4.3 we show that there are no points outside our scaling window that are competitive in age and fitness. The proof of Theorem 3 is completed in Section 4.4 and the proof of Corollary 4, which crucially uses the compactification, in Section 4.5.

### 3 Local convergence of point processes

In this section we prove a convergence result for the point processes  $(\Gamma_t)$  in a space without compactification. The strengthening of the results by compactification will follow in the next section. We begin by noting some preliminary results on the fitness distribution.

#### 3.1 Preliminaries on the fitness distribution

First of all we show the existence and uniqueness of  $\sigma_t$  as defined in Equation (2):

**Lemma 8.** *For all  $t$  large enough, there exists a unique  $\hat{\sigma}_t \in [0, t)$ , such that*

$$(\log g)'(\lambda \hat{\sigma}_t) = \frac{1}{\lambda(t - \hat{\sigma}_t)}.$$

Furthermore, as  $t \uparrow \infty$ , we have  $\hat{\sigma}_t \rightarrow \infty$  (and thus, for all  $t$  large enough  $\sigma_t = \hat{\sigma}_t$ ) and  $\frac{\hat{\sigma}_t}{t} \rightarrow 0$ .

*Proof.* For all  $t \geq 0$ , for all  $x \in [0, \lambda t)$ , we set

$$F(x) := (\log g)'(x) - \frac{1}{\lambda t - x},$$

so  $F$  is continuous on  $(0, \lambda t)$ , since, by Assumption **(A5)**,  $m$ , and thus  $g$  are continuous and non-zero on, respectively  $(0, 1)$  and  $(0, \infty)$ . Since  $g = m^{-1} : [0, \infty) \rightarrow [0, 1)$  and  $g(0) = 0$ , we have

$$\lim_{x \downarrow 0} (\log g)'(x) = \lim_{x \downarrow 0} \frac{g'(x)}{g(x)} = \lim_{x \downarrow 0} \frac{1}{xm'(x)} = \infty,$$

because  $m'(0) < \infty$ , since, by Assumption **(A5)**,  $m$  is differentiable on  $[0, 1)$ . Therefore we get

$$\lim_{x \downarrow 0} F(x) = \infty, \quad \text{and} \quad \lim_{x \uparrow \lambda t} F(x) = -\infty.$$

Hence by continuity of  $F$ , there exists  $x \in (0, \lambda t)$  such that  $F(x) = 0$ . Furthermore such  $x$  is unique because  $F$  is a decreasing function: indeed, for all  $x \in (0, \lambda t)$

$$F'(x) = \frac{g''(x)}{g(x)} - \left( \frac{g'(x)}{g(x)} \right)^2 - \left( \frac{1}{\lambda t - x} \right)^2 < 0 \quad \text{for all } x \in (0, \lambda t),$$

since  $g''(x) = -\frac{m''(g(x))}{(m'(g(x)))^3} < 0$  by Assumption **(A5.1)**. Setting  $\hat{\sigma}_t = \frac{1}{\lambda}x$  proves existence and uniqueness as required, moreover  $\hat{\sigma}_t$  is increasing in  $t$ .

It remains to show that  $\hat{\sigma}_t \rightarrow \infty$  as  $t \rightarrow \infty$ . If  $\hat{\sigma}_t$  was bounded, we had  $\frac{1}{\lambda t - \lambda \hat{\sigma}_t} \rightarrow 0$  as  $t \rightarrow \infty$ . This implies that  $(\log g)'(\lambda \hat{\sigma}_t) = g'(\lambda \hat{\sigma}_t)/g(\lambda \hat{\sigma}_t) \rightarrow 0$  and hence  $g'(\lambda \hat{\sigma}_t) \rightarrow 0$ . This implies  $\frac{1}{m'(g(\lambda \hat{\sigma}_t))} \rightarrow 0$ , i.e.  $m'(g(\lambda \hat{\sigma}_t)) \rightarrow \infty$ . From Assumption **(A5.4)**, we know that  $m'(x) \uparrow \infty$  as  $x \uparrow 1$  and therefore  $g(\lambda \hat{\sigma}_t) \rightarrow 1$  and hence  $\hat{\sigma}_t \rightarrow \infty$  as required.

Finally we show that  $\frac{\hat{\sigma}_t}{t} \rightarrow 0$ . By definition of  $\hat{\sigma}_t$ , we have  $t = \hat{\sigma}_t + \frac{g(\lambda\hat{\sigma}_t)}{\lambda g'(\lambda\hat{\sigma}_t)}$ , so we can write

$$\lim_{t \rightarrow \infty} \frac{\hat{\sigma}_t}{t} = \lim_{t \rightarrow \infty} \frac{\hat{\sigma}_t}{\hat{\sigma}_t + \frac{g(\lambda\hat{\sigma}_t)}{\lambda g'(\lambda\hat{\sigma}_t)}} = \lim_{t \rightarrow \infty} \frac{1}{1 + \frac{g(\lambda\hat{\sigma}_t)}{\lambda \hat{\sigma}_t g'(\lambda\hat{\sigma}_t)}}.$$

As  $\hat{\sigma}_t \rightarrow \infty$  as have  $g(\lambda\hat{\sigma}_t) \rightarrow 1$  as  $t \rightarrow \infty$  we get

$$\lim_{t \rightarrow \infty} \frac{\hat{\sigma}_t}{t} = \lim_{x \uparrow 1} \frac{1}{1 + \frac{x}{\frac{m(x)}{m'(x)}}} = 0,$$

since  $\lim_{x \uparrow 1} \frac{m(x)}{m'(x)} = 0$  by Assumption **(A5.4)**. □

From Lemma 8, it follows that  $\lambda t \sim \frac{g(\lambda\sigma_t)}{g'(\lambda\sigma_t)}$  as  $t \rightarrow \infty$ . Since  $\lim_{t \uparrow \infty} g(\lambda\sigma_t) = 1$ , we get that

$$g'(\lambda\sigma_t) \sim \frac{1}{\lambda t} \quad \text{when } t \uparrow \infty. \quad (14)$$

**Lemma 9.** *We have*

$$\lim_{t \rightarrow \infty} \sigma_t t g''(\lambda\sigma_t) = -\varkappa \lambda^{-2}, \quad (15)$$

where  $\varkappa$  is defined in Assumption **(A5.3)**, and

$$\lim_{t \rightarrow \infty} \sigma_t g'(\lambda\sigma_t) = 0. \quad (16)$$

*Proof.* Recall that, by Lemma 8, for all  $t$  large enough,  $\sigma_t = m(g(\lambda\sigma_t))/\lambda$ ,  $t \sim \frac{g(\lambda\sigma_t)}{\lambda g'(\lambda\sigma_t)}$  as  $t \uparrow \infty$ , and

$$g''(x) = -\frac{m''(g(x))}{(m'(g(x)))^3} = -\frac{m''(g(x))g'(x)}{(m'(g(x)))^2} \quad \text{for all } x \in [0, \infty),$$

since  $m' \circ g = 1/g'$ . Substituting these into (15) and substituting  $x = g(\lambda\sigma_t)$ , we get

$$\lim_{t \rightarrow \infty} \sigma_t t g''(\lambda\sigma_t) = \lim_{t \rightarrow \infty} -\frac{m(g(\lambda\sigma_t))g(\lambda\sigma_t)m''(g(\lambda\sigma_t))}{(\lambda m'(g(\lambda\sigma_t)))^2} = \lim_{x \uparrow 1} -\frac{m''(x)m(x)x}{(\lambda m'(x))^2} = -\varkappa \lambda^{-2},$$

by Assumption **(A5.3)**. Similarly, using  $g'(\lambda\sigma_t) = \frac{1}{m'(g(\lambda\sigma_t))}$ , we have

$$\lim_{t \rightarrow \infty} \sigma_t g'(\lambda\sigma_t) = \lim_{t \rightarrow \infty} \frac{m(g(\lambda\sigma_t))}{\lambda m'(g(\lambda\sigma_t))} = \lim_{x \uparrow 1} \frac{m(x)}{\lambda m'(x)} = 0,$$

by Assumption **(A5.4)**. □

### 3.2 Convergence of a simpler point process

In this section we prove the following proposition, which gives a more basic form of the Poisson limit in a space without compactification.

**Proposition 10.** *As  $t \rightarrow \infty$  we have vague convergence in distribution of the point process*

$$\Psi_t = \sum_{n=1}^{M(t)} \delta\left(\frac{\tau_n - \sigma_t}{\sqrt{\sigma_t}}, \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}\right), e^{-\gamma F_n(t-\tau_n)} Z_n(t)$$

to the Poisson point process with intensity

$$\zeta^*(ds, df, dz) = \lambda e^{-f} \nu(z) ds df dz,$$

on  $(-\infty, \infty) \times (-\infty, \infty) \times [0, \infty]$ .

We prove Proposition 10 in two steps:

(1) In Lemma 12 we approximate  $\Psi_t$  by the point process

$$\Psi_t^* = \sum_{n \in \mathbb{N}} \delta \left( \frac{\frac{1}{\lambda} \log n - \sigma_t}{\sqrt{\sigma_t}}, \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}, \xi_n \right),$$

where we have replaced the rescaled family sizes by

$$\xi_n := \lim_{t \rightarrow \infty} e^{-\gamma t} Y_n(t),$$

and the birth times  $\tau_n$  by the approximate birth times  $\frac{1}{\lambda} \log n$ , using **(A1-3)**.

(2) In Lemma 11 we prove that  $\Psi_t^*$  converges to the Poisson point process with intensity  $\zeta^*$ .

**Lemma 11.** *As  $t \rightarrow \infty$  the process  $(\Psi_t^*)_{t \geq 0}$  converges vaguely in distribution on  $(-\infty, \infty) \times (-\infty, \infty] \times [0, \infty]$  to the Poisson point process with intensity  $\zeta^*$ .*

*Proof.* We apply Kallenberg's theorem, see [20, Proposition 3.22]. Since  $\zeta^*$  is diffuse it suffices to show that, for every precompact relatively open box  $B \subset (-\infty, \infty) \times (-\infty, \infty] \times [0, \infty]$ , we have

- (a)  $\mathbb{P}(\Psi_t^*(B) = 0) \rightarrow \exp(-\zeta^*(B))$ , as  $t \uparrow \infty$ , and
- (b)  $\mathbb{E}[\Psi_t^*(B)] \rightarrow \zeta^*(B)$ , as  $t \uparrow \infty$ .

It further suffices to consider nonempty boxes  $B$  of the form  $(s_0, s_1) \times (f_0, f_1) \times (z_0, z_1)$ , where  $s_0, s_1 \in (-\infty, \infty)$ ,  $f_0, f_1 \in (-\infty, \infty]$ ,  $z_0, z_1 \in [0, \infty]$ , and  $s_0 < s_1$ ,  $f_0 < f_1$ ,  $z_0 < z_1$ . Note that

$$\zeta^*(B) = \lambda(s_1 - s_0)(e^{-f_0} - e^{-f_1}) \int_{z_0}^{z_1} \nu(x) dx.$$

(a) Let

$$r(a) := \exp(\lambda(a\sqrt{\sigma_t} + \sigma_t)), \quad \text{for all } a \in \mathbb{R},$$

and consider

$$\hat{\Psi}_t^* = \sum_{n \in \mathbb{N}} \delta \left( \frac{\frac{1}{\lambda} \log n - \sigma_t}{\sqrt{\sigma_t}}, \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))} \right),$$

$$\hat{\zeta}^*(ds, df) = \lambda e^{-f} ds df.$$

So that for  $\hat{B} = (s_0, s_1) \times (f_0, f_1)$ , we get  $\hat{\zeta}^*(\hat{B}) = \lambda(s_1 - s_0)(e^{-f_0} - e^{-f_1})$ . Denote, for all  $a > 0$ ,  $x \in \mathbb{R}$ ,

$$\hat{f}_a(x) = g(\log(a\sqrt{\sigma_t})) + xg'(\log(a\sqrt{\sigma_t})).$$

Then we have

$$\begin{aligned} \mathbb{P}(\hat{\Psi}_t^*(\hat{B}) = 0) &= \prod_{n=r(s_0)}^{r(s_1)} \mathbb{P} \left( \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))} \notin (f_0, f_1) \right) \\ &= \prod_{n=r(s_0)}^{r(s_1)} \left[ 1 - \mu(\hat{f}_n(f_0), \hat{f}_n(f_1)) \right] \\ &\sim \prod_{n=r(s_0)}^{r(s_1)} \exp \left\{ -\mu(\hat{f}_n(f_0), 1) + \mu(\hat{f}_n(f_1), 1) \right\}. \end{aligned}$$



Recalling that  $\mu(x, 1) = e^{-m(x)}$ , we get the following: when  $t \rightarrow \infty$ ,

$$\begin{aligned} \mathbb{P}(\hat{\Psi}_t^*(\hat{B}) = 0) &\sim \exp \left\{ \sum_{n=r(s_0)}^{r(s_1)} -e^{-m(\hat{f}_n(f_0))} + e^{-m(\hat{f}_n(f_1))} \right\} \\ &\sim \exp \left\{ - \int_{r(s_0)}^{r(s_1)} e^{-m(\hat{f}_x(f_0))} dx + \int_{r(s_0)}^{r(s_1)} e^{-m(\hat{f}_x(f_1))} dx \right\}. \end{aligned}$$

We now evaluate the integrals in the exponent. For  $i = 0, 1$  we have

$$\begin{aligned} \int_{r(s_0)}^{r(s_1)} e^{-m(\hat{f}_x(f_i))} dx &= \int_{r(s_0)}^{r(s_1)} \exp \left\{ -m \left( g(\log(x\sqrt{\sigma_t})) + f_i g'(\log(x\sqrt{\sigma_t})) \right) \right\} dx \\ &= \int_{s_0}^{s_1} \lambda \sqrt{\sigma_t} r(y) \exp \left\{ -m \left( g(\log(r(y)\sqrt{\sigma_t})) + f_i g'(\log(r(y)\sqrt{\sigma_t})) \right) \right\} dy, \end{aligned}$$

by a change of variables, with  $x = e^{\lambda(y\sqrt{\sigma_t} + \sigma_t)} = r(y)$ . By the mean value theorem, for each  $i \in \{0, 1\}$ , there exists a constant  $c_3 \in [g(\log(r(y)\sqrt{\sigma_t})), g(\log(r(y)\sqrt{\sigma_t})) + f_i g'(\log(r(y)\sqrt{\sigma_t}))]$ , such that

$$\begin{aligned} &m \left( g(\log(r(y)\sqrt{\sigma_t})) + f_i g'(\log(r(y)\sqrt{\sigma_t})) \right) \\ &= m \left( g(\log(r(y)\sqrt{\sigma_t})) \right) + f_i g'(\log(r(y)\sqrt{\sigma_t})) m' \left( g(\log(r(y)\sqrt{\sigma_t})) \right) \\ &\quad + \frac{1}{2} \left( f_i g'(\log(r(y)\sqrt{\sigma_t})) \right)^2 m''(c_3). \end{aligned}$$

Recall that, for  $x \in \mathbb{R}$ , we have  $m(g(x)) = x$  and  $g'(x) = \frac{1}{m'(g(x))}$ , so the integral simplifies to

$$\begin{aligned} \int_{r(s_0)}^{r(s_1)} e^{-m(\hat{f}_x(f_i))} dx &= \int_{s_0}^{s_1} \lambda r(y) \sqrt{\sigma_t} e^{-\log(r(y)\sqrt{\sigma_t}) - f_i - \frac{1}{2} (f_i g'(\log(r(y)\sqrt{\sigma_t})))^2 m''(c_3)} dy \\ &= \lambda \int_{s_0}^{s_1} \exp \left\{ -f_i - \frac{f_i^2}{2} g'(\log(r(y)\sqrt{\sigma_t}))^2 m''(c_3) \right\} dy. \end{aligned}$$

Recall that  $c_3 \in [g(\log(r(y)\sqrt{\sigma_t})), g(\log(r(y)\sqrt{\sigma_t})) + f_i g'(\log(r(y)\sqrt{\sigma_t}))]$ . By Assumption **(A5.4)** and since  $\lim_{x \uparrow 1} m(x) = \infty$ , we have  $\lim_{x \uparrow \infty} g'(x) = 0$ , and thus  $c_3 \sim g(\log(r(y)\sqrt{\sigma_t}))$  when  $t \uparrow \infty$ . By Assumption **(A5.2)**, we thus get

$$g'(\log(r(y)\sqrt{\sigma_t}))^2 m''(c_3) = \frac{m''(c_3)}{m'(g(\log(r(y)\sqrt{\sigma_t})))^2} \rightarrow 0,$$

as  $t \rightarrow \infty$ . By the dominated convergence theorem, as  $t \rightarrow \infty$ , we get

$$\int_{r(s_0)}^{r(s_1)} e^{-m(\hat{f}_x(f_i))} dx = \lambda(s_1 - s_0)e^{-f_i} + o(1).$$

Therefore, as  $t \rightarrow \infty$  we get

$$\begin{aligned} \mathbb{P}(\hat{\Psi}_t^*(\hat{B}) = 0) &\sim \exp \left\{ -\lambda(s_1 - s_0)e^{-f_0} + \lambda(s_1 - s_0)e^{-f_1} + o(1) \right\} \\ &\rightarrow \exp \left\{ -\lambda(s_1 - s_0)(e^{-f_0} - e^{-f_1}) \right\} = \exp \left\{ -\hat{\zeta}^*(\hat{B}) \right\}. \end{aligned}$$

(b) To calculate the limit of  $\mathbb{E}[\hat{\Psi}_t^*(\hat{B})]$  we apply similar asymptotic estimates as in part (a), and get that, when  $t \uparrow \infty$

$$\begin{aligned}\mathbb{E}[\hat{\Psi}_t^*(\hat{B})] &= \sum_{r(a_0) \leq n \leq r(a_1)} \mu(\hat{f}_n(f_0), \hat{f}_n(f_1)) \\ &\sim \int_{r(s_0)}^{r(s_1)} \mu(\hat{f}_x(f_0), \hat{f}_x(f_1)) dx \\ &\sim \lambda(s_1 - s_0)(e^{-f_0} - e^{-f_1}) = \hat{\zeta}^*(\hat{B}).\end{aligned}\tag{17}$$

Using Kallenberg's theorem, we thus get that, in distribution when  $t \rightarrow \infty$ ,  $\hat{\Psi}_t^*$  converge vaguely on  $(-\infty, \infty) \times (-\infty, \infty]$  to the Poisson point process of intensity  $\hat{\zeta}^*$ . By our assumptions,  $(F_n, \xi_n)_{n \geq 1}$  is a sequence of i.i.d. random variables with each  $F_n$  being independent of  $\xi_n$ . Together with the fact that  $\mathbb{P}(\xi_n \in (z_0, z_1)) = \int_{z_0}^{z_1} \nu(x) dx$ , this completes the proof.  $\square$

**Lemma 12.** *For all Lipschitz continuous, compactly supported functions  $f : (-\infty, \infty) \times (-\infty, \infty] \times [0, \infty] \rightarrow \mathbb{R}$ ,*

$$\left| \int f d\Psi_t^* - \int f d\Psi_t \right| \rightarrow 0 \text{ in probability, as } t \uparrow \infty.$$

**Note.** *By density of the set of Lipschitz-continuous compactly-supported functions in the set of continuous compactly supported functions for the topology of the  $L^\infty$  norm, Lemma 12 implies that, for all continuous compactly supported functions  $f : (-\infty, \infty) \times (-\infty, \infty] \times [0, \infty] \rightarrow \mathbb{R}$ ,*

$$\left| \int f d\Psi_t^* - \int f d\Psi_t \right| \rightarrow 0 \text{ in probability, as } t \uparrow \infty.$$

*Proof.* Let  $f$  be a Lipschitz continuous function supported on  $K = [-a, a] \times [-b, \infty] \times [0, \infty]$  for  $1 \leq a, b < \infty$ . We have, for all  $t \geq 0$ ,

$$\begin{aligned}&\left| \int f d\Psi_t^* - \int f d\Psi_t \right| \\ &\leq \sum_{n=1}^{M(t)} \left| f\left(\frac{\tau_n - \sigma_t}{\sqrt{\sigma_t}}, \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}, e^{-\gamma F_n(t - \tau_n)} Z_n(t)\right) - f\left(\frac{\frac{1}{\lambda} \log n - \sigma_t}{\sqrt{\sigma_t}}, \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}, \xi_n\right) \right| \\ &\leq c_L \sum_{n \in \hat{I}(t)} \left( \left| \frac{\tau_n - \tau_n^*}{\sqrt{\sigma_t}} \right| + \left| e^{-\gamma F_n(t - \tau_n)} Z_n(t) - \xi_n \right| \right),\end{aligned}\tag{18}$$

where  $c_L$  is the Lipschitz constant of the function  $f$ ,  $\xi_n$  are i.i.d. copies of  $\xi$  (defined in Assumption **(A3)**),  $\tau_n^* = \frac{1}{\lambda} \log n$ , and  $\hat{I}(t)$  is the random set of indices  $n \in \mathbb{N}$  such that

$$(a) \quad \left| \frac{\tau_n - \sigma_t}{\sqrt{\sigma_t}} \right| \leq a \text{ and } \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))} \geq -b \text{ or}$$

$$(b) \quad \left| \frac{\tau_n^* - \sigma_t}{\sqrt{\sigma_t}} \right| \leq a \text{ and } \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))} \geq -b.$$

The last inequality of Equation (18) comes from the fact that, by definition of  $\hat{I}(t)$ , all summands associated to integers  $n \notin \hat{I}(t)$  are zero because the support of  $f$  is included in  $[-a, a] \times [-b, \infty] \times [0, \infty]$ . By Lemma 8, there exists  $t_0$  such that, for all  $t \geq t_0$ ,  $\sigma_t \leq \frac{t}{3}$  and  $\sqrt{\sigma_t} \leq \sigma_t$ . For  $\varepsilon \in (0, 1/2)$  we denote by  $\Upsilon_\varepsilon(t)$  the event that

$$|\tau_n - \tau_n^*| \leq \varepsilon \sqrt{\sigma_t} \quad \text{for all } n \in \mathbb{N}.$$

Assumption **(A1)** together with Lemma 8 implies that  $\mathbb{P}(\Upsilon_\varepsilon(t)) \rightarrow 1$ , as  $t \rightarrow \infty$  for all  $\varepsilon \in (0, 1/2)$ . Set

$$\bar{I}(t) := \left\{ n \in \mathbb{N} : \frac{|\tau_n^* - \sigma_t|}{\sqrt{\sigma_t}} \leq 2a, \text{ and } \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))} \geq -b \right\}.$$

We have that  $\hat{I}(t) \subset \bar{I}(t)$  on  $\Upsilon_\varepsilon(t)$ . Indeed, if (a) and  $\Upsilon_\varepsilon(t)$  hold then

$$\frac{|\tau_n^* - \sigma_t|}{\sqrt{\sigma_t}} \leq \frac{|\tau_n^* - \tau_n|}{\sqrt{\sigma_t}} + \frac{|\tau_n - \sigma_t|}{\sqrt{\sigma_t}} \leq \varepsilon + a \leq 2a,$$

and similarly if (b) hold. We now consider the sum on the right hand side of Equation (18), but taken over all  $n \in \bar{I}(t)$ . First note that, for  $n \in \bar{I}(t)$  on  $\Upsilon_\varepsilon(t)$ , we have

$$\tau_n \leq 2a\sqrt{\sigma_t} + \sigma_t \leq 2a\sigma_t + \sigma_t = \sigma_t(2a + 1) \leq \frac{t}{2}, \quad (19)$$

for all  $t \geq t_0$ . Since  $(\log g)'(\log(n\sqrt{\sigma_t})) \rightarrow 0$  as  $t \rightarrow \infty$ , and  $g(\log(n\sqrt{\sigma_t})) \rightarrow 1$ , we have

$$\begin{aligned} F_n &\geq g(\log(n\sqrt{\sigma_t})) - bg'(\log(n\sqrt{\sigma_t})) \\ &= g(\log(n\sqrt{\sigma_t})) \left( 1 - b(\log g)'(\log(n\sqrt{\sigma_t})) \right) \rightarrow 1, \end{aligned} \quad (20)$$

as  $t \rightarrow \infty$ . Recall  $\Delta_n(t)$  from Assumption **(A2)**,  $\xi_n = \lim_{t \rightarrow \infty} e^{-\gamma t} Y_n(t)$ , and define

$$R_n(t) := \sup_{w > t} |e^{-\gamma w} Y_n(w) - \xi_n|.$$

By Assumption **(A3)** we have  $R_n(t) \rightarrow 0$  in probability and, for all  $t$  large enough, we have

$$\begin{aligned} \left| e^{-\gamma F_n(t - \tau_n)} Z_n(t) - \xi_n \right| &\leq \left| e^{-\gamma F_n(t - \tau_n)} Z_n(t) - e^{-\gamma F_n(t - \tau_n)} Y_n(F_n(t - \tau_n)) \right| \\ &\quad + \left| e^{-\gamma F_n(t - \tau_n)} Y_n(F_n(t - \tau_n)) - \xi_n \right| \\ &\leq \Delta_n(F_n(t - \tau_n)) + R_n(F_n(t - \tau_n)) \\ &\leq \Delta_n\left(\frac{t}{2}\right) + R_n\left(\frac{t}{2}\right), \end{aligned}$$

where we have used Equations (19) and (20). Hence we get that, for sufficiently large  $t$ , on  $\Upsilon_\varepsilon(t)$ ,

$$\begin{aligned} \left| \int f d\Psi_t - \int f d\Psi_t^* \right| &\leq c_L \sum_{n \in \bar{I}(t)} \left( \frac{|\tau_n - \tau_n^*|}{\sqrt{\sigma_t}} + \left| e^{-\gamma F_n(t - \tau_n)} Z_n(t) - \xi_n \right| \right) \\ &\leq c_L \sum_{n \in \bar{I}(t)} \left( \frac{\sup_n |\tau_n - \tau_n^*|}{\sqrt{\sigma_t}} + \Delta_n\left(\frac{t}{2}\right) + R_n\left(\frac{t}{2}\right) \right) \\ &\leq c_L \frac{|\bar{I}(t)| \sup_n |\tau_n - \tau_n^*|}{\sqrt{\sigma_t}} + c_L \sum_{n \in \bar{I}(t)} \Delta_n\left(\frac{t}{2}\right) + c_L \sum_{n \in \bar{I}(t)} R_n\left(\frac{t}{2}\right). \end{aligned}$$

By assumption, the random processes  $(R_n)_{n \geq 1}$  are independent of  $(F_n)_{n \geq 1}$  and thus also of the random set  $\bar{I}(t)$ . Recall that, by Lemma 11,  $|\bar{I}(t)|$  converges in distribution to a Poisson distribution and hence

$$\lim_{t \rightarrow \infty} \sum_{n \in \bar{I}(t)} R_n\left(\frac{t}{2}\right) = 0, \quad \text{in probability.}$$

To prove that  $\sum_{n \in \bar{I}(t)} \Delta_n\left(\frac{t}{2}\right) \rightarrow 0$  in probability as  $t \rightarrow \infty$  we use Assumption (A.2). For all  $t \geq 0$ ,

we have

$$\begin{aligned}
\mathbb{P}\left(\sum_{n \in \bar{I}(t)} \Delta_n\left(\frac{t}{2}\right) \geq \varepsilon\right) &= \mathbb{E}\left[\mathbb{P}\left(\sum_{n \in \bar{I}(t)} \Delta_n\left(\frac{t}{2}\right) \geq \varepsilon \mid (F_n)\right)\right] \\
&\leq \sum_{k=0}^{\infty} \mathbb{E}\left[\mathbb{P}\left(\exists n \in \bar{I}(t) : \Delta_n\left(\frac{t}{2}\right) \geq \frac{\varepsilon}{k} \mid (F_n)\right) \mathbf{1}_{\{|\bar{I}(t)|=k\}}\right] \\
&\leq \sum_{k=0}^{\infty} \mathbb{E}\left[\sum_{n \in \bar{I}(t)} \mathbb{P}\left(\Delta_n\left(\frac{t}{2}\right) \geq \frac{\varepsilon}{k} \mid (F_n)\right) \mathbf{1}_{\{|\bar{I}(t)|=k\}}\right] \\
&\leq \sum_{k=0}^{\infty} \mathbb{E}\left[k \max_{n \in \bar{I}(t)} \mathbb{P}\left(\Delta_n\left(\frac{t}{2}\right) \geq \frac{\varepsilon}{k} \mid (F_n)\right) \mathbf{1}_{\{|\bar{I}(t)|=k\}}\right].
\end{aligned}$$

Now, given  $\delta > 0$  pick  $K \in \mathbb{N}$  such that, for sufficiently large  $t$ ,

$$\sum_{k=K+1}^{\infty} \mathbb{E}\left[k \max_{n \in \bar{I}(t)} \mathbb{P}\left(\Delta_n\left(\frac{t}{2}\right) \geq \frac{\varepsilon}{k} \mid (F_n)\right) \mathbf{1}_{\{|\bar{I}(t)|=k\}}\right] \leq \mathbb{E}\left[|\bar{I}(t)| \mathbf{1}_{|\bar{I}(t)| > K}\right]. \quad (21)$$

Note that, by definition,  $|\bar{I}(t)| = \Psi_t^*(B)$ , with  $B = [-2a, 2a] \times [-b, \infty] \times [0, \infty]$ . In (17), we showed that  $\mathbb{E}[|\bar{I}(t)|] = \zeta^*(B)$ . Similarly, we have

$$\mathbb{E}[|\bar{I}(t)|^2] = \mathbb{E}\left[\left(\sum_{r(-2a) \leq n \leq r(2a)} \mathbf{1}_{\hat{F}_n \geq -b}\right)^2\right],$$

where for all  $n \geq 1$ , we have set  $\hat{F}_n = [F_n - g(\log(n\sqrt{\sigma_t}))]/g'(\log(n\sqrt{\sigma_t}))$ . Recall that the  $(F_n)_{n \geq 1}$  are i.i.d. We thus get

$$\begin{aligned}
\mathbb{E}[|\bar{I}(t)|^2] &= \mathbb{E}\left[\sum_{r(-2a) \leq n \leq r(2a)} \mathbf{1}_{\hat{F}_n \geq -b}\right] + \mathbb{E}\left[\sum_{r(-2a) \leq n \neq m \leq r(2a)} \mathbf{1}_{\hat{F}_n \geq -b} \mathbf{1}_{\hat{F}_m \geq -b}\right] \\
&\leq \mathbb{E}|\bar{I}(t)| + \sum_{r(-2a) \leq n, m \leq r(2a)} \mathbb{P}(\hat{F}_n \geq -b) \mathbb{P}(\hat{F}_m \geq -b) \\
&= \mathbb{E}|\bar{I}(t)| + \mathbb{E}[|\bar{I}(t)|^2] \rightarrow \zeta^*(B) + \zeta^*(B)^2,
\end{aligned}$$

as  $t \uparrow \infty$ . Hence the family  $(|\bar{I}(t)|)_{t \geq 1}$  is  $L^2$ -bounded and therefore uniformly integrable. Using this in (21), we get that, for all  $\delta > 0$ , there exists  $K$  sufficiently large such that, for all  $t$  large enough,

$$\sum_{k=K+1}^{\infty} \mathbb{E}\left[k \max_{n \in \bar{I}(t)} \mathbb{P}\left(\Delta_n\left(\frac{t}{2}\right) \geq \frac{\varepsilon}{k} \mid (F_n)\right) \mathbf{1}_{\{|\bar{I}(t)|=k\}}\right] \leq \frac{\delta}{2}.$$

Since, by definition,  $\bar{I}(t) \subseteq I_\kappa(t)$  for  $\kappa = 2a$ , we get

$$\sum_{k=0}^K \mathbb{E}\left[k \max_{n \in \bar{I}(t)} \mathbb{P}\left(\Delta_n\left(\frac{t}{2}\right) \geq \frac{\varepsilon}{k} \mid (F_n)\right) \mathbf{1}_{\{|\bar{I}(t)|=k\}}\right] \leq K(K+1) \mathbb{E}\left[\max_{n \in I_\kappa(t)} \mathbb{P}\left(\Delta_n\left(\frac{t}{2}\right) \geq \frac{\varepsilon}{K} \mid (F_n)\right)\right],$$

which converges to zero by (A.2) and dominated convergence. This shows that  $\sum_{n \in \bar{I}(t)} \Delta_n\left(\frac{t}{2}\right) \rightarrow 0$  in probability. Summarising, we get that, in probability when  $t \uparrow \infty$ ,

$$\left|\int f d\Psi_t - \int f d\Psi_t^*\right| \leq c_L |\bar{I}(t)| \frac{\sup_n |\tau_n - \tau_n^*|}{\sqrt{\sigma_t}} + o(1),$$

which converges to zero in probability, as  $t \uparrow \infty$ .  $\square$

*Proof of Proposition 10.* Let  $f: (-\infty, \infty) \times (-\infty, \infty] \times [0, \infty] \rightarrow \mathbb{R}$  be Lipschitz continuous and compactly supported. Combining Lemmas 11 and 12, together with Slutsky's theorem (see for example [13, ch.7.2]) we get  $\int f d\Psi_t \Rightarrow \int f d\text{PPP}(\zeta^*)$  as  $t \rightarrow \infty$ , where  $\text{PPP}(\zeta^*)$  denotes the Poisson point process with intensity  $\zeta^*$ .  $\square$

### 3.3 Proof of the local convergence result

**Proposition 13.** *Asymptotically when  $t \rightarrow \infty$ , the point process*

$$\Gamma_t = \sum_{n=1}^{M(t)} \delta\left(\frac{\tau_n - \sigma_t}{\sqrt{\sigma_t}}, \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}, e^{-\gamma g(\lambda\sigma_t)(t-\sigma_t) - a_1 g(\lambda\sigma_t) \log \sigma_t + \gamma T} Z_n(t)\right),$$

*converges vaguely in distribution on  $(-\infty, \infty) \times (-\infty, \infty) \times [0, \infty]$  to the Poisson point process with intensity*

$$\zeta(ds, df, dz) = \lambda e^{-f} e^{s^2 a_2 - f a_3} \nu(z e^{s^2 a_2 - f a_3}) ds df dz.$$

*Proof of Proposition 13.* Consider the continuous function

$$\phi: (s, f, z) \rightarrow (s, f, e^{-s^2 a_2 + f a_3} z),$$

so that  $\zeta \circ \phi^{-1} = \zeta^*$ . We argue that  $\Psi_t \circ \phi^{-1}$  is asymptotically equivalent to  $\Gamma_t$ , i.e. for all Lipschitz continuous, compactly supported functions  $f: (-\infty, \infty) \times (-\infty, \infty) \times [0, \infty] \rightarrow \mathbb{R}$ ,

$$\left| \int f d\Psi_t \circ \phi^{-1} - \int f d\Gamma_t \right| \rightarrow 0 \quad \text{in probability, as } t \uparrow \infty.$$

To prove this let  $f$  be a Lipschitz continuous function with Lipschitz constant  $c_L$ , supported on  $K = [-a, a] \times [-b, b] \times [0, \infty]$  for  $1 \leq a, b < \infty$  and abbreviate

$$s_n = \frac{\tau_n - \sigma_t}{\sqrt{\sigma_t}} \quad \text{and} \quad f_n = \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}, \quad \text{for } n \geq 1.$$

For all  $t \geq 0$ , we have

$$\begin{aligned} & \left| \int f d\Psi_t \circ \phi^{-1} - \int f d\Gamma_t \right| \\ & \leq \sum_{n=1}^{M(t)} \left| f\left(\frac{\tau_n - \sigma_t}{\sqrt{\sigma_t}}, \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}, e^{-a_2 s_n^2 + a_3 f_n} e^{-\gamma F_n(t-\tau_n)} Z_n(t)\right) \right. \\ & \quad \left. - f\left(\frac{\tau_n - \sigma_t}{\sqrt{\sigma_t}}, \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}, e^{-\gamma g(\lambda\sigma_t)(t-\sigma_t) - a_1 g(\lambda\sigma_t) \log \sigma_t + \gamma T} Z_n(t)\right) \right| \\ & \leq c_L \sum_{n \in \tilde{I}(t)} \left| e^{-\gamma F_n(t-\tau_n) - a_2 s_n^2 + a_3 f_n} Z_n(t) - e^{-\gamma g(\lambda\sigma_t)(t-\sigma_t) - a_1 g(\lambda\sigma_t) \log \sigma_t + \gamma T} Z_n(t) \right|, \quad (22) \end{aligned}$$

where  $\tilde{I}(t)$  is the random set of indices  $n \in \mathbb{N}$  such that  $|s_n| \leq a$  and  $|f_n| \leq b$  (this definition implies that all summands associated to integers  $n \notin \tilde{I}(t)$  are zero because the support of  $f$  is contained in  $K$ ). We now show that the exponents of (22) are asymptotically equivalent, namely

$$-\gamma F_n(t-\tau_n) - a_2 s_n^2 + a_3 f_n = -\gamma g(\lambda\sigma_t)(t-\sigma_t) - a_1 g(\lambda\sigma_t) \log \sigma_t + \gamma T + o(1), \quad (23)$$

almost surely when  $t \uparrow \infty$ , where the  $o(1)$ -term is uniform in  $n$ . Indeed, by definition of  $s_n$  and using Assumption **(A1)**, we get

$$\log n = \lambda(\sigma_t + s_n \sqrt{\sigma_t} - T_n), \quad \text{for } n \geq 1,$$

where we set  $T_n = T + \varepsilon_n$ . Therefore, we have

$$\begin{aligned} F_n &= g(\log(n\sqrt{\sigma_t})) + f_n g'(\log(n\sqrt{\sigma_t})) \\ &= g\left(\lambda(\sigma_t + s_n \sqrt{\sigma_t} - T_n) + \frac{1}{2} \log \sigma_t\right) + f_n g'\left(\lambda(\sigma_t + s_n \sqrt{\sigma_t} - T_n) + \frac{1}{2} \log \sigma_t\right). \end{aligned}$$

Let  $x_n := \lambda s_n \sqrt{\sigma_t} + \frac{1}{2} \log \sigma_t - \lambda T_n$ , so that

$$F_n(t - \tau_n) = \left( g(\lambda \sigma_t + x_n) + f_n g'(\lambda \sigma_t + x_n) \right) (t - \sigma_t - s_n \sqrt{\sigma_t}).$$

By the mean value theorem, there exist  $c_1, c_2 \in [\lambda \sigma_t, \lambda \sigma_t + x_n]$ , such that

$$g(\lambda \sigma_t + x_n) = g(\lambda \sigma_t) + x_n g'(\lambda \sigma_t) + \frac{1}{2} x_n^2 g''(c_1), \quad \text{and} \quad (24)$$

$$g'(\lambda \sigma_t + x_n) = g'(\lambda \sigma_t) + x_n g''(c_2). \quad (25)$$

Hence, for  $n \in \tilde{I}(t)$  we can rewrite

$$\begin{aligned} F_n(t - \tau_n) &= \left( g(\lambda \sigma_t) + x_n g'(\lambda \sigma_t) + \frac{1}{2} x_n^2 g''(c_1) + f_n g'(\lambda \sigma_t) + x_n f_n g''(c_2) \right) (t - \sigma_t - s_n \sqrt{\sigma_t}) \\ &= g(\lambda \sigma_t)(t - \sigma_t) - g(\lambda \sigma_t) s_n \sqrt{\sigma_t} + \lambda s_n \sqrt{\sigma_t} g'(\lambda \sigma_t)(t - \sigma_t) - \lambda s_n^2 \sigma_t g'(\lambda \sigma_t) \\ &\quad + \left( \frac{1}{2} \log \sigma_t - \lambda T_n \right) g'(\lambda \sigma_t)(t - \sigma_t - s_n \sqrt{\sigma_t}) + \frac{1}{2} x_n^2 g''(c_1)(t - \sigma_t - s_n \sqrt{\sigma_t}) \\ &\quad + f_n g'(\lambda \sigma_t)(t - \sigma_t) - f_n g'(\lambda \sigma_t) s_n \sqrt{\sigma_t} + f_n x_n g''(c_2)(t - \sigma_t - s_n \sqrt{\sigma_t}). \end{aligned}$$

Recall that by definition  $g'(\lambda \sigma_t)(t - \sigma_t) = \frac{g(\lambda \sigma_t)}{\lambda}$ , and  $g(\lambda \sigma_t) = 1 + o(1)$  when  $t \rightarrow \infty$ . We get

$$f_n g'(\lambda \sigma_t)(t - \sigma_t) = \frac{f_n}{\lambda} + o(1) \quad \text{almost surely when } t \uparrow \infty.$$

By definition  $g(\lambda \sigma_t) \uparrow 1$  as  $t \uparrow \infty$  and by Lemma 8, we have  $\sigma_t = o(t)$  and  $g'(\lambda \sigma_t) \sim \frac{1}{\lambda t}$  (see Equations (2) and (14)). Furthermore, for  $n \in \tilde{I}(t)$ , Assumption **(A1)** implies  $T_n = T + \varepsilon_n \rightarrow T$ , as  $t \rightarrow \infty$ . Combining these with the fact that for all  $n \in \tilde{I}(t)$ ,  $|s_n| \leq a$  and  $|f_n| \leq b$ , we can show that for all  $n \in \tilde{I}(t)$ , almost surely as  $t \rightarrow \infty$ , the following terms go to zero:

$$\begin{aligned} |\lambda s_n^2 \sigma_t g'(\lambda \sigma_t)| &\leq \frac{a^2 \sigma_t}{t - \sigma_t} = \mathcal{O}\left(\frac{\sigma_t}{t}\right) = o(1), \\ \left| \left( \frac{1}{2} \log \sigma_t - \lambda T_n \right) g'(\lambda \sigma_t) s_n \sqrt{\sigma_t} \right| &\leq \left| \frac{1}{2} \log \sigma_t - \lambda T_n \right| \frac{a \sqrt{\sigma_t}}{\lambda(t - \sigma_t)} \sim \left( \frac{1}{2} \log \sigma_t - \lambda T \right) \frac{a \sqrt{\sigma_t}}{\lambda(t - \sigma_t)} = o(1), \\ |f_n g'(\lambda \sigma_t) s_n \sqrt{\sigma_t}| &\leq \frac{ab \sqrt{\sigma_t}}{\lambda(t - \sigma_t)} = \mathcal{O}\left(\frac{\sqrt{\sigma_t}}{t}\right) = o(1). \end{aligned}$$

Therefore, almost surely as  $t \uparrow \infty$ ,

$$\begin{aligned} F_n(t - \tau_n) &= g(\lambda \sigma_t)(t - \sigma_t) + \frac{g(\lambda \sigma_t)}{2\lambda} \log \sigma_t - g(\lambda \sigma_t) T_n + \frac{f_n}{\lambda} + \frac{1}{2} x_n^2 g''(c_1)(t - \sigma_t - s_n \sqrt{\sigma_t}) \\ &\quad + f_n x_n g''(c_2)(t - \sigma_t - s_n \sqrt{\sigma_t}) + o(1). \end{aligned} \quad (26)$$

We can write  $g(\lambda \sigma_t) = 1 + o(1)$ , as  $t \uparrow \infty$ , and by Assumption **(A1)**,  $T_n = T + o(1)$  uniformly in  $n \in \tilde{I}(t)$  where the  $o(1)$ -term converges to zero almost surely as  $t \rightarrow \infty$ . Therefore we get

$$g(\lambda \sigma_n) T_n = T + o(1) \quad \text{as } t \rightarrow \infty. \quad (27)$$

To simplify the last two terms in Equation (26), we recall that Lemma 9 implies  $g''(c_i) \sim \frac{-\varkappa}{\lambda^2 \sigma_t t}$  for  $i = 1, 2$ . Combing this with the fact that  $\sigma_t \rightarrow \infty$  as  $t \rightarrow \infty$  (by Lemma 8), we get for  $n \in \tilde{I}(t)$ ,

$$\begin{aligned} |f_n x_n g''(c_2)(t - \sigma_t - s_n \sqrt{\sigma_t})| &= |f_n (\lambda s_n \sqrt{\sigma_t} + \frac{1}{2} \log \sigma_t - \lambda T_n) g''(c_2)(t - \sigma_t - s_n \sqrt{\sigma_t})| \\ &\leq |b(\lambda a \sqrt{\sigma_t} + \frac{1}{2} \log \sigma_t - \lambda T_n) g''(c_2)(t - \sigma_t + a \sqrt{\sigma_t})| \\ &= \left| b(\lambda a \sqrt{\sigma_t} + \frac{1}{2} \log \sigma_t - \lambda T + o(1)) \frac{\varkappa(t - \sigma_t + a \sqrt{\sigma_t})}{\lambda^2 \sigma_t t} + o(1) \right| \\ &= \mathcal{O}\left(\frac{1}{\sqrt{\sigma_t}}\right) = o(1), \quad \text{almost surely as } t \rightarrow \infty. \end{aligned} \quad (28)$$

Consider the penultimate term of Equation (26). By definition of  $x_n$  we can rewrite it as follows,

$$\begin{aligned} \frac{1}{2}x_n^2 g''(c_1)(t - \sigma_t - s_n \sqrt{\sigma_t}) &= \frac{1}{2} \left( \lambda s_n \sqrt{\sigma_t} + \frac{1}{2} \log \sigma_t - \lambda T_n \right)^2 g''(c_1)(t - \sigma_t - s_n \sqrt{\sigma_t}) \\ &= \frac{1}{2} \lambda^2 s_n^2 \sigma_t g''(c_1)(t - \sigma_t) - \frac{1}{2} \lambda^2 s_n^3 \sigma_t^{3/2} g''(c_1) \\ &\quad + \frac{1}{2} \left( 2\lambda s_n \sqrt{\sigma_t} \left( \frac{1}{2} \log \sigma_t - \lambda T_n \right) + \left( \frac{1}{2} \log \sigma_t - \lambda T_n \right)^2 \right) g''(c_1)(t - \sigma_t - s_n \sqrt{\sigma_t}). \end{aligned} \quad (29)$$

The first summand is the largest term and, by Lemma 9, we get that, almost surely as  $t \rightarrow \infty$ ,

$$\frac{1}{2} \lambda^2 s_n^2 \sigma_t g''(c_1)(t - \sigma_t) = -\frac{\lambda^2 s_n^2 \sigma_t \varkappa(t - \sigma_t)}{2\lambda^2 \sigma_t t} + o(1) = -\frac{1}{2} s_n^2 \varkappa + o(1).$$

Using Lemmas 8, 9 and Assumption **(A1)**, we show that the second and third summands in Equation (29) go to zero almost surely as  $t \rightarrow \infty$ . Indeed,

$$\begin{aligned} &\left| -\frac{1}{2} \lambda^2 s_n^3 \sigma_t^{3/2} g''(c_1) + \frac{1}{2} \left( 2\lambda s_n \sqrt{\sigma_t} \left( \frac{1}{2} \log \sigma_t - \lambda T_n \right) + \left( \frac{1}{2} \log \sigma_t - \lambda T_n \right)^2 \right) g''(c_1)(t - \sigma_t - s_n \sqrt{\sigma_t}) \right| \\ &\leq \left| -\frac{1}{2} \lambda^2 a^3 \sigma_t^{3/2} g''(c_1) + \frac{1}{2} \left( 2\lambda a \sqrt{\sigma_t} \left( \frac{1}{2} \log \sigma_t - \lambda T_n \right) + \left( \frac{1}{2} \log \sigma_t - \lambda T_n \right)^2 \right) g''(c_1)(t - \sigma_t + a \sqrt{\sigma_t}) \right| \\ &= \left| \frac{1}{2} a^3 \sqrt{\sigma_t} \frac{\varkappa}{t} - \frac{1}{2} \left( 2\lambda a \sqrt{\sigma_t} \left( \frac{1}{2} \log \sigma_t - \lambda T + o(1) \right) + \left( \frac{1}{2} \log \sigma_t - \lambda T + o(1) \right)^2 \right) \frac{\varkappa(t - \sigma_t + a \sqrt{\sigma_t})}{\lambda^2 \sigma_t t} + o(1) \right| \\ &= \mathcal{O} \left( \frac{t \sqrt{\sigma_t} \log \sigma_t}{\sigma_t t} \right) = o(1), \end{aligned}$$

almost surely when  $t \rightarrow \infty$ . Therefore, for all  $n \in \tilde{I}(t)$ , we have

$$\frac{1}{2} x_n^2 g''(c_1)(t - \sigma_t - s_n \sqrt{\sigma_t}) = -\frac{1}{2} s_n^2 \varkappa + o(1), \quad \text{almost surely as } t \rightarrow \infty. \quad (30)$$

Combining (27), (28) and (30), Equation (26) becomes

$$F_n(t - \tau_n) = g(\lambda \sigma_t)(t - \sigma_t) + \frac{g(\lambda \sigma_t)}{2\lambda} \log \sigma_t - T - \frac{1}{2} s_n^2 \varkappa + \frac{1}{\lambda} f_n + o(1),$$

and thus

$$-\gamma F_n(t - \tau_n) = -\gamma g(\lambda \sigma_t)(t - \sigma_t) - a_1 g(\lambda \sigma_t) \log \sigma_t + \gamma T + a_2 s_n^2 - a_3 f_n + o(1),$$

almost surely as  $t \rightarrow \infty$ , where  $a_1 = \gamma/2\lambda$ ,  $a_2 = \gamma\varkappa/2$  and  $a_3 = \gamma/\lambda$ . Rearranging we get Equation (23).

Substituting Equation (23) into (22) we get

$$\left| \int f d\Psi_t \circ \phi^{-1} - \int f d\Gamma_t \right| \leq c_L \sum_{n \in \tilde{I}(t)} Z_n(t) e^{-\gamma F_n(t - \tau_n)} e^{-a_2 s_n^2 + a_3 f_n} |1 - e^{o(1)}|.$$

Since the almost-sure  $o(1)$ -term is uniform in  $n \in \tilde{I}(t)$ , we get

$$\left| \int f d\Psi_t \circ \phi^{-1} - \int f d\Gamma_t \right| = o \left( \sum_{n \in \tilde{I}(t)} Z_n(t) e^{-\gamma F_n(t - \tau_n)} e^{-a_2 s_n^2 + a_3 f_n} \right),$$

almost surely as  $t \rightarrow \infty$ . Furthermore, by definitions of  $\Psi_t \circ \phi^{-1}$  and  $\tilde{I}(t)$ ,

$$\begin{aligned} \sum_{n \in \tilde{I}(t)} Z_n(t) e^{-\gamma F_n(t - \tau_n)} e^{-a_2 s_n^2 + a_3 f_n} &= \int_0^\infty \mathbf{1}_{|s| \leq a} \mathbf{1}_{|f| \leq b} z \, d\Psi_t \circ \phi^{-1}(s, f, z) \\ &\rightarrow \int_0^\infty \mathbf{1}_{|s| \leq a} \mathbf{1}_{|f| \leq b} z \, d\text{PPP}(\zeta^* \circ \phi^{-1}) \\ &= \int_0^\infty \mathbf{1}_{|s| \leq a} \mathbf{1}_{|f| \leq b} z \, d\text{PPP}(\zeta), \end{aligned}$$

in distribution as  $t \rightarrow \infty$ , by Proposition 10, since the function  $(s, f, z) \mapsto \mathbf{1}_{|s| \leq a} \mathbf{1}_{|f| \leq b} z$  has compact support in  $(-\infty, \infty) \times (-\infty, \infty) \times [0, \infty]$ . Recalling the definition of  $\zeta$ , and substituting  $w = ze^{s^2 a_2 - f a_3}$ , we get

$$\begin{aligned} \mathbb{E} \left[ \int_0^\infty z \mathbf{1}_{|s| \leq a} \mathbf{1}_{|f| \leq b} d\text{PPP}(\zeta) \right] &= \int_{-a}^a \int_{-b}^b \int_0^\infty \lambda e^{-f} e^{s^2 a_2 - f a_3} z \nu(z e^{s^2 a_2 - f a_3}) ds df dz \\ &= \int_{-a}^a \lambda e^{-s^2 a_2} ds \int_{-b}^b e^{(a_3 - 1)f} df \int_0^\infty w \nu(w) dw \\ &= \lambda \sqrt{\frac{\pi}{a_2}} \operatorname{erf}(a \sqrt{a_2}) \frac{1}{a_3 - 1} \left( e^{(a_3 - 1)b} - e^{-(a_3 - 1)b} \right) \int_0^\infty w \nu(w) dw =: C_1, \end{aligned}$$

where  $\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^x e^{-t^2} dt$ . Note that  $C_1 < \infty$  since  $\int_0^\infty w \nu(w) dw < \infty$  by Assumption **(A3)**. This implies that  $\sum_{n \in \tilde{I}(t)} Z_n(t) e^{-\gamma F_n(t - \tau_n)} e^{-a_2 s_n^2 + a_3 f_n}$  converges in distribution to an almost surely finite random variable, and thus

$$\left| \int f d\Psi_t \circ \phi^{-1} - \int f d\Gamma_t \right| \rightarrow 0, \quad \text{in distribution, and thus in probability as } t \rightarrow \infty,$$

which means that the point process  $\Gamma_t$  is asymptotically equivalent to  $\Psi_t \circ \phi^{-1}$ . Note that, by a change of variable,  $\int f d\Psi_t \circ \phi^{-1} = \int f \circ \phi d\Psi_t$ . For all functions  $f$  continuous and compactly supported, since  $\phi$  is continuous, the function  $f \circ \phi$  is also continuous and compactly supported, implying that

$$\int f d\Psi_t \circ \phi^{-1} = \int f \circ \phi d\Psi_t \rightarrow \int f \circ \phi d\text{PPP}(\zeta^*) = \int f d\text{PPP}(\phi(\zeta^*)),$$

where we have used Proposition 10. One can check that  $\zeta$  is the image of  $\zeta^*$  by  $\phi$ . This implies that  $\Psi_t \circ \phi^{-1}$  converges vaguely in distribution on  $(-\infty, \infty) \times (-\infty, \infty) \times [0, \infty]$  to  $\text{PPP}(\zeta)$ , implying that  $\Gamma_t$  does too, since the two point processes are asymptotically equivalent.  $\square$

## 4 Compactification and completion of the proofs

To deduce Theorem 3 from Proposition 13, one has to control the contribution of the point process near the closed boundaries of  $[-\infty, \infty] \times [-\infty, \infty] \times (0, \infty]$ . We prove that the families that are born outside of the main window, namely the ones that are unfit or born late, are too small to contribute in the limit. We first consider families which are born either early or late. We then show the negligibility of families lying under the main window by looking at families with small fitness (see Figure 1).

### 4.1 Contribution of young and old families

**Lemma 14** (Contribution of young and old families). *For every  $\eta > 0$  and  $\varepsilon > 0$  there exists  $v > 1$  such that, for all sufficiently large  $t$ , we have*

$$\mathbb{P} \left( \max_{n \in E_t(v)} e^{-\gamma g(\lambda \sigma_t)(t - \sigma_t) - a_1 g(\lambda \sigma_t) \log \sigma_t + \gamma T} Z_n(t) \geq \varepsilon \right) \leq \eta,$$

where  $E_t(v) := [0, n_t(-v)] \cup [n_t(v), \infty]$ ,  $n_t(\pm v) := \exp \{ \lambda(\sigma_t \pm v \sqrt{\sigma_t}) \}$ .

*Proof.* Let  $\eta, \varepsilon > 0$ . For all  $n \geq 1$ , we define

$$A_n := \max_{u \geq \tau_n} Z_n(u) e^{-\gamma F_n(u - \tau_n)}.$$

If there exists  $t \geq \tau_n$  such that

$$Z_n(t) \geq \varepsilon e^{\gamma g(\lambda \sigma_t)(t - \sigma_t) + a_1 g(\lambda \sigma_t) \log \sigma_t - \gamma T}, \quad (31)$$



then we get,

$$A_n \geq Z_n(t)e^{-\gamma F_n(t-\tau_n)} \geq \varepsilon e^{\gamma g(\lambda\sigma_t)(t-\sigma_t)+a_1g(\lambda\sigma_t)\log\sigma_t-\gamma T-\gamma F_n(t-\tau_n)}. \quad (32)$$

By Assumption **(A1)**, we have  $\tau_n = \frac{1}{\lambda} \log n + T_n$ , where  $T_n = T + \varepsilon_n$ ; therefore (32) is equivalent to

$$A_n \geq c_{n,t}e^{-\gamma(1-F_n)T+\gamma F_n\varepsilon_n},$$

where we have set

$$c_{n,t} := \varepsilon \exp(\gamma g(\lambda\sigma_t) - \gamma F_n)t + (\gamma F_n - \gamma g(\lambda\sigma_t))\sigma_t + a_1g(\lambda\sigma_t)\log\sigma_t - \gamma F_n(\sigma_t - \frac{1}{\lambda} \log n).$$

Hence,

$$\mathbb{P}\left(\max_{n \in E_t(v)} Z_n(t) \geq \varepsilon e^{\gamma g(\lambda\sigma_t)(t-\sigma_t)+a_1g(\lambda\sigma_t)\log\sigma_t-\gamma T}\right) \leq \mathbb{P}\left(\bigcup_{n \in E_t(v)} \{A_n \geq c_{n,t}e^{-\gamma(1-F_n)T+\gamma F_n\varepsilon_n}\}\right).$$

Moreover, for any  $y > 0$ , we have

$$\begin{aligned} \mathbb{P}\left(\bigcup_{n \in E_t(v)} \{A_n \geq c_{n,t}e^{-\gamma(1-F_n)T+\gamma F_n\varepsilon_n}\}\right) \\ \leq \sum_{n \in E_t(v)} \mathbb{P}(A_n \geq c_{n,t}e^{-\gamma y}) + \mathbb{P}(|T| \geq y) + \mathbb{P}\left(\sup_{n \in E_t(v)} |\varepsilon_n| \geq y\right). \end{aligned}$$

Since  $\varepsilon_n \rightarrow 0$  almost surely and  $|T|$  is finite, we can fix  $y > 0$  large enough, such that  $\mathbb{P}(|T| \geq y) \leq \frac{\eta}{3}$  and  $\mathbb{P}(\sup_{n \in E_t(v)} |\varepsilon_n| \geq y) \leq \frac{\eta}{3}$ . Consider

$$S := \sum_{n \in E_t(v)} \mathbb{P}(A_n \geq c_{n,t}e^{-\gamma y}) = \sum_{n \in E_t(v)} \mathbb{E}[\mathbb{P}(A_n \geq c_{n,t}e^{-\gamma y} | (F_m)_{m \in \mathbb{N}})].$$

By Assumption **(A4)**,  $\mathbb{P}(A_n \geq u | (F_m)_{m \in \mathbb{N}}) \leq c_0 e^{-\eta u}$ , so we get

$$\begin{aligned} S &\leq c_0 \sum_{n \in E_t(v)} \mathbb{E}\left[\exp\left\{-\eta \varepsilon e^{(\gamma g(\lambda\sigma_t)-\gamma F_n)t+(\gamma F_n-\gamma g(\lambda\sigma_t))\sigma_t+a_1g(\lambda\sigma_t)\log\sigma_t-\gamma y-\gamma F_n(\sigma_t-\frac{1}{\lambda}\log n)}\right\}\right] \\ &\leq c_0 \int_{E_t(v)} \mathbb{E}\left[\exp\left\{-\eta \varepsilon e^{(\gamma g(\lambda\sigma_t)-\gamma F)t+(\gamma F-\gamma g(\lambda\sigma_t))\sigma_t+a_1g(\lambda\sigma_t)\log\sigma_t-\gamma y-\gamma F(\sigma_t-\frac{1}{\lambda}\log x)}\right\}\right] dx, \end{aligned}$$

where  $F$  is a random variable of law  $\mu$ . Let  $x = \exp\{\lambda(\sigma_t + w\sqrt{\sigma_t})\}$ , therefore we can write

$$\begin{aligned} S &\leq c_0 \int_{|w| \geq v} \lambda \sqrt{\sigma_t} e^{\lambda(\sigma_t + w\sqrt{\sigma_t})} \mathbb{E}\left[\exp\left\{-\eta \varepsilon e^{(\gamma g(\lambda\sigma_t)-\gamma F)t+(\gamma F-\gamma g(\lambda\sigma_t))\sigma_t+a_1g(\lambda\sigma_t)\log\sigma_t-\gamma y+\gamma F w\sqrt{\sigma_t}}\right\}\right] dw \\ &\leq c_0 \int_{|w| \geq v} \lambda \sqrt{\sigma_t} e^{\lambda(\sigma_t + w\sqrt{\sigma_t})} \\ &\quad \times \int_0^1 \mathbb{P}\left(\exp\left\{-\eta \varepsilon e^{(\gamma g(\lambda\sigma_t)-\gamma F)t+(\gamma F-\gamma g(\lambda\sigma_t))\sigma_t+a_1g(\lambda\sigma_t)\log\sigma_t-\gamma y+\gamma F w\sqrt{\sigma_t}}\right\} \geq x\right) dx dw \\ &=: c_0 \int_{|w| \geq v} \lambda \sqrt{\sigma_t} e^{\lambda(\sigma_t + w\sqrt{\sigma_t})} \int_0^1 P(x) dx dw. \end{aligned}$$

Letting  $\tilde{x}_0 = 1 + w\sigma_t^{-1/2}$  and substituting into  $\mu(x, 1) = \exp\{-m(x)\}$ , we get

$$\begin{aligned} P(x) &= \mathbb{P}\left(\exp\left\{-\eta \varepsilon e^{\gamma g(\lambda\sigma_t)t-\gamma g(\lambda\sigma_t)\sigma_t+a_1g(\lambda\sigma_t)\log\sigma_t-\gamma y+\gamma F(-t+\tilde{x}_0\sigma_t)}\right\} \geq x\right) \\ &= \mathbb{P}\left(F \geq (\gamma t - \gamma \tilde{x}_0\sigma_t)^{-1} \left(\gamma g(\lambda\sigma_t)t - \gamma g(\lambda\sigma_t)\sigma_t + a_1g(\lambda\sigma_t)\log\sigma_t - \gamma y - \log\left(-\frac{1}{\eta\varepsilon}\log x\right)\right)\right) \\ &= \exp\left\{-m\left((\gamma t - \gamma \tilde{x}_0\sigma_t)^{-1} \left(\gamma g(\lambda\sigma_t)t - \gamma g(\lambda\sigma_t)\sigma_t + a_1g(\lambda\sigma_t)\log\sigma_t - \gamma y - \log\left(-\frac{1}{\eta\varepsilon}\log x\right)\right)\right)\right\} \\ &= \exp\left\{-m\left(\left(1 - \tilde{x}_0\frac{\sigma_t}{t}\right)^{-1} \left(g(\lambda\sigma_t) - \frac{g(\lambda\sigma_t)}{t}\sigma_t + \frac{a_1g(\lambda\sigma_t)}{\gamma t}\log\sigma_t - \frac{y}{t} - \frac{1}{\gamma t}\log\left(-\frac{1}{\eta\varepsilon}\log x\right)\right)\right)\right\}. \end{aligned}$$

We can approximate  $P(x)$  by

$$\begin{aligned} P(x) &= \exp \left\{ -m \left( \left( 1 + \tilde{x}_0 \frac{\sigma_t}{t} \right) \left( g(\lambda \sigma_t) - \frac{\gamma g(\lambda \sigma_t)}{\gamma t} \sigma_t + \frac{a_1 g(\lambda \sigma_t)}{\gamma t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log \left( -\frac{1}{\eta \varepsilon} \log x \right) \right) + \mathcal{O} \left( \frac{\sigma_t}{t} \right)^2 \right\} \\ &= \exp \left\{ -m \left( g(\lambda \sigma_t) + \frac{w \sqrt{\sigma_t}}{t} g(\lambda \sigma_t) + \frac{a_1 g(\lambda \sigma_t)}{\gamma t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log \left( -\frac{1}{\eta \varepsilon} \log x \right) + \mathcal{O} \left( \frac{\sigma_t}{t} \right)^2 \right) \right\}. \end{aligned}$$

Lemma 17 (Equation (36)) implies

$$\begin{aligned} P(x) &= \exp \left\{ -m(g(\lambda \sigma_t)) - m'(g(\lambda \sigma_t)) \left( \frac{w \sqrt{\sigma_t}}{t} g(\lambda \sigma_t) + \frac{a_1 g(\lambda \sigma_t)}{\gamma t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log \left( -\frac{1}{\eta \varepsilon} \log x \right) \right) \right. \\ &\quad \left. - \frac{1}{2} m''(c_1) \left( \frac{w \sqrt{\sigma_t}}{t} g(\lambda \sigma_t) + \frac{a_1 g(\lambda \sigma_t)}{\gamma t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log \left( -\frac{1}{\eta \varepsilon} \log x \right) \right)^2 \right\}. \end{aligned}$$

Recall that  $m(g(\lambda \sigma_t)) = \lambda \sigma_t$  and  $m'(g(\lambda \sigma_t)) = \frac{\lambda(t - \sigma_t)}{g(\lambda \sigma_t)}$ . Using Assumption **(A5.3)**, one can show that  $m''(g(\lambda \sigma_t)) \sim \frac{\lambda \varkappa t^2}{\sigma_t (g(\lambda \sigma_t))^3}$  as  $t$  goes to infinity. Therefore we get

$$\begin{aligned} P(x) &= \exp \left\{ - \left( \lambda \sigma_t + \frac{\lambda(t - \sigma_t)}{g(\lambda \sigma_t)} \left( \frac{w \sqrt{\sigma_t}}{t} g(\lambda \sigma_t) + \frac{g(\lambda \sigma_t)}{2\lambda t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log \left( -\frac{1}{\eta \varepsilon} \log x \right) \right) \right. \right. \\ &\quad \left. \left. + \frac{\lambda \varkappa t^2}{2\sigma_t (g(\lambda \sigma_t))^3} \left( \frac{w \sqrt{\sigma_t}}{t} g(\lambda \sigma_t) + \frac{g(\lambda \sigma_t)}{2\lambda t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log \left( -\frac{1}{\eta \varepsilon} \log x \right) \right)^2 \right) + o(1) \right\} \\ &= \sigma_t^{-1/2} \exp \left\{ -\lambda \sigma_t - \lambda w \sqrt{\sigma_t} + \frac{\lambda y}{g(\lambda \sigma_t)} + \frac{\lambda}{\gamma g(\lambda \sigma_t)} \log \left( -\frac{1}{\eta \varepsilon} \log x \right) - \frac{\lambda \varkappa w^2}{2g(\lambda \sigma_t)} + o(1) \right\}. \end{aligned}$$

Hence we get

$$\begin{aligned} S &\leq c_0 \int_{|w| \geq v} \lambda \sqrt{\sigma_t} e^{\lambda(\sigma_t + w \sqrt{\sigma_t})} \sigma_t^{-\frac{1}{2}} e^{-\lambda(\sigma_t + w \sqrt{\sigma_t}) + \frac{\lambda y}{g(\lambda \sigma_t)} - \frac{\lambda \varkappa w^2}{2g(\lambda \sigma_t)} + o(1)} \int_0^1 e^{\frac{\lambda}{\gamma g(\lambda \sigma_t)} \log \left( -\frac{1}{\eta \varepsilon} \log x \right)} dx dw \\ &\leq c_0 \int_{|w| \geq v} \lambda e^{\frac{\lambda y}{g(\lambda \sigma_t)} - \frac{\lambda \varkappa w^2}{2g(\lambda \sigma_t)} + o(1)} \Gamma \left( \frac{\lambda}{\gamma g(\lambda \sigma_t)} + 1 \right) dw = \mathcal{O} \left( \int_{|w| \geq v} \exp \left\{ \frac{\lambda y}{g(\lambda \sigma_t)} - \frac{\lambda \varkappa w^2}{2g(\lambda \sigma_t)} \right\} dw \right), \end{aligned}$$

which goes to 0 as  $v$  goes to infinity, uniformly for all  $t \geq 1$ .  $\square$

## 4.2 Contribution of unfit families

**Lemma 15** (Negligibility of families with small fitnesses). *For every  $\eta > 0$  and  $\varepsilon > 0$ , there exists  $\kappa > 0$  such that for all sufficiently large  $t$ , we have*

$$\mathbb{P} \left( \max_{n \leq M(t)} \mathbf{1} \left\{ \frac{F_n - g(\log(n \sqrt{\sigma_t}))}{g'(\log(n \sqrt{\sigma_t}))} \leq -\kappa \right\} e^{-\gamma g(\lambda \sigma_t)(t - \sigma_t) - a_1 g(\lambda \sigma_t) \log \sigma_t + \gamma T} Z_n(t) \geq \varepsilon \right) \leq \eta.$$

*Proof.* Let  $\varepsilon, \eta > 0$  and  $\kappa > 0$ . We analyse the event that there exists a family with fitness at most

$$f_n(\kappa) := g(\log(n \sqrt{\sigma_t})) - \kappa g'(\log(n \sqrt{\sigma_t}))$$

and size at least  $\varepsilon \exp\{\gamma g(\lambda \sigma_t)(t - \sigma_t) + a_1 g(\lambda \sigma_t) \log \sigma_t - \gamma T\}$ . Similarly to the proof of Lemma 14 we define, for all  $n \geq 1$ ,

$$A_n := \max_{u \geq \tau_n} Z_n(u) e^{-\gamma F_n(u - \tau_n)},$$

and as before we define

$$c_{n,t} := \varepsilon \exp \left\{ (\gamma g(\lambda \sigma_t) - \gamma F_n) t + (\gamma F_n - \gamma g(\lambda \sigma_t)) \sigma_t + a_1 g(\lambda \sigma_t) \log \sigma_t - \gamma F_n \left( \sigma_t - \frac{1}{\lambda} \log n \right) \right\}.$$

It can be shown that

$$\begin{aligned} \mathbb{P}\left(\max_{n \leq M(t)} \mathbf{1}_{F_n \leq f_n(\kappa)} Z_n(t) \geq \varepsilon e^{\gamma g(\lambda \sigma_t)(t - \sigma_t) + a_1 g(\lambda \sigma_t) \log \sigma_t}\right) \\ \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n \mathbf{1}_{F_n \leq f_n(\kappa)} \geq c_{n,t} e^{-\gamma y}) + \mathbb{P}(|T| \geq y) + \mathbb{P}\left(\sup_{n \in \mathbb{N}} |\varepsilon_n| \geq y\right), \end{aligned}$$

where  $y > 0$  is large enough, so that  $\mathbb{P}(|T| \geq y) \leq \frac{\eta}{3}$  and  $\mathbb{P}(\sup_{n \in \mathbb{N}} |\varepsilon_n| \geq y) \leq \frac{\eta}{3}$ . Set

$$S := \sum_{n=1}^{\infty} \mathbb{P}(A_n \mathbf{1}_{F_n \leq f_n(\kappa)} \geq c_{n,t} e^{-\gamma y}) = \sum_{n=1}^{\infty} \mathbb{E}[\mathbf{1}_{F_n \leq f_n(\kappa)} \mathbb{P}(A_n \geq c_{n,t} e^{-\gamma y} \mid (F_m)_{m \in \mathbb{N}})].$$

By Assumption **(A4)**,  $\mathbb{P}(A_n \geq u \mid (F_m)_{m \in \mathbb{N}}) \leq c_0 e^{-\eta u}$ , which implies

$$\begin{aligned} S &\leq c_0 \sum_{n=1}^{\infty} \mathbb{E}\left[\mathbf{1}_{F_n \leq f_n(\kappa)} \exp\left\{-\eta \varepsilon e^{(\gamma g(\lambda \sigma_t) - \gamma F_n)t + (\gamma F_n - \gamma g(\lambda \sigma_t))\sigma_t + a_1 g(\lambda \sigma_t) \log \sigma_t - \gamma y - \gamma F_n(\sigma_t - \frac{1}{\lambda} \log n)}\right\}\right] \\ &\leq c_0 \int_0^{\infty} \mathbb{E}\left[\mathbf{1}_{F \leq f_x(\kappa)} \exp\left\{-\eta \varepsilon e^{(\gamma g(\lambda \sigma_t) - \gamma F)t + (\gamma F - \gamma g(\lambda \sigma_t))\sigma_t + a_1 g(\lambda \sigma_t) \log \sigma_t - \gamma y - \gamma F(\sigma_t - \frac{1}{\lambda} \log x)}\right\}\right] dx, \end{aligned}$$

where  $F$  is a random variable of law  $\mu$ . We change variables by setting  $x = \exp\{\lambda(\sigma_t + w\sqrt{\sigma_t})\}$  and set  $\hat{f}_w(\kappa) := f_{\exp\{\lambda(\sigma_t + w\sqrt{\sigma_t})\}}(\kappa)$ . This yields

$$\begin{aligned} S &\leq c_0 \int_{-\infty}^{\infty} \lambda \sqrt{\sigma_t} e^{\lambda(\sigma_t + w\sqrt{\sigma_t})} \\ &\quad \times \mathbb{E}\left[\mathbf{1}_{F \leq \hat{f}_w(\kappa)} \exp\left\{-\eta \varepsilon e^{(\gamma g(\lambda \sigma_t) - \gamma F)t + (\gamma F - \gamma g(\lambda \sigma_t))\sigma_t + a_1 g(\lambda \sigma_t) \log \sigma_t - \gamma y + \gamma F w \sqrt{\sigma_t}}\right\}\right] dw. \end{aligned}$$

Denoting by  $\tilde{x}_0 := 1 + w\sigma_t^{-1/2}$  and

$$E := \mathbb{E}\left[\mathbf{1}_{F \leq \hat{f}_w(\kappa)} \exp\left\{-\eta \varepsilon e^{(\gamma g(\lambda \sigma_t) - \gamma F)t + (\gamma F - \gamma g(\lambda \sigma_t))\sigma_t + a_1 g(\lambda \sigma_t) \log \sigma_t - \gamma y + \gamma F w \sqrt{\sigma_t}}\right\}\right],$$

we get

$$\begin{aligned} E &= \int_0^1 \mathbb{P}\left(F \leq \hat{f}_w(\kappa); \exp\left\{-\eta \varepsilon e^{\gamma g(\lambda \sigma_t)t - \gamma g(\lambda \sigma_t)\sigma_t + a_1 g(\lambda \sigma_t) \log \sigma_t - \gamma y + \gamma F(-t + \tilde{x}_0 \sigma_t)}\right\} \geq x\right) dx \\ &= \int_0^1 \mathbb{P}\left(\frac{\gamma g(\lambda \sigma_t)t - \gamma g(\lambda \sigma_t)\sigma_t + a_1 g(\lambda \sigma_t) \log \sigma_t - \gamma y - \log\left(\frac{1}{\eta \varepsilon} \log\left(\frac{1}{x}\right)\right)}{\gamma(t - \tilde{x}_0 \sigma_t)} \leq F \leq \hat{f}_w(\kappa)\right) dx. \end{aligned}$$

This integral is zero if the lower bound in the probability is larger than the upper bound. Note that

$$\begin{aligned} f_x &:= \frac{\gamma g(\lambda \sigma_t)t - \gamma g(\lambda \sigma_t)\sigma_t + a_1 g(\lambda \sigma_t) \log \sigma_t - \gamma y - \log\left(\frac{1}{\eta \varepsilon} \log\left(\frac{1}{x}\right)\right)}{\gamma(t - \tilde{x}_0 \sigma_t)} \\ &= \left(g(\lambda \sigma_t) - \frac{g(\lambda \sigma_t)}{t} \sigma_t + \frac{a_1 g(\lambda \sigma_t)}{\gamma t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log\left(-\frac{1}{\eta \varepsilon} \log x\right)\right) \left(1 + \frac{\tilde{x}_0}{t} \sigma_t + \mathcal{O}\left(\frac{\sigma_t}{t}\right)^2\right) \\ &= g(\lambda \sigma_t) - (g(\lambda \sigma_t) - g(\lambda \sigma_t) \tilde{x}_0) \frac{\sigma_t}{t} + \frac{a_1 g(\lambda \sigma_t)}{\gamma t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log\left(-\frac{1}{\eta \varepsilon} \log x\right) + o\left(\frac{1}{t}\right). \end{aligned}$$

We have  $E = \int_0^1 (\mu(f_x, 1) - \mu(\hat{f}_w(\kappa), 1)) dx$ . By Lemma 17 (Equation (36)),

$$\begin{aligned} \mu(f_x, 1) &= \exp\left\{-m\left(g(\lambda \sigma_t) - (g(\lambda \sigma_t) - g(\lambda \sigma_t) \tilde{x}_0) \frac{\sigma_t}{t} + \frac{a_1 g(\lambda \sigma_t)}{\gamma t} \log \sigma_t - \frac{y}{t} \right. \right. \\ &\quad \left. \left. - \frac{1}{\gamma t} \log\left(-\frac{1}{\eta \varepsilon} \log x\right) + o\left(\frac{1}{t}\right)\right)\right\} \\ &= \exp\left\{-m\left(g(\lambda \sigma_t) + g(\lambda \sigma_t) \frac{w\sqrt{\sigma_t}}{t} + \frac{g(\lambda \sigma_t)}{2\lambda t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log\left(-\frac{1}{\eta \varepsilon} \log x\right) + o\left(\frac{1}{t}\right)\right)\right\} \\ &= \exp\left\{-m\left(g(\lambda \sigma_t)\right) - m'\left(g(\lambda \sigma_t)\right)\left(g(\lambda \sigma_t) \frac{w\sqrt{\sigma_t}}{t} + \frac{g(\lambda \sigma_t)}{2\lambda t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log\left(-\frac{1}{\eta \varepsilon} \log x\right)\right) \right. \\ &\quad \left. - \frac{1}{2} m''(c_3) \left(g(\lambda \sigma_t) \frac{w\sqrt{\sigma_t}}{t} + \frac{g(\lambda \sigma_t)}{2\lambda t} \log \sigma_t - \frac{y}{t} - \frac{1}{\gamma t} \log\left(-\frac{1}{\eta \varepsilon} \log x\right)\right)^2\right\} \\ &= \exp\left\{-\lambda \sigma_t - \lambda w \sqrt{\sigma_t} - \log \sqrt{\sigma_t} + \frac{\lambda y}{g(\lambda \sigma_t)} + \frac{\lambda}{\gamma g(\lambda \sigma_t)} \log\left(-\frac{1}{\eta \varepsilon} \log x\right) - \frac{\lambda \kappa w^2}{2g(\lambda \sigma_t)} + o(1)\right\}, \end{aligned}$$

since  $m''(g(\lambda\sigma_t)) \sim \frac{\lambda\kappa t^2}{\sigma_t(g(\lambda\sigma_t))^3}$ , by Assumption **(A5.3)**. Using Lemma 17, (Equation (37)), and the fact that  $m'(g(x))g'(x) = 1$  for all  $x > 0$ , we get

$$\begin{aligned}
\mu(\hat{f}_w(\kappa), 1) &= \exp \left\{ -m \left( g(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t}) - \kappa g'(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t}) \right) \right\} \\
&= \exp \left\{ -m \left( g(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t}) \right) + m' \left( g(\lambda\tilde{x}_0\sqrt{\sigma_t} + \log \sqrt{\sigma_t}) \right) \kappa g'(\lambda\tilde{x}_0\sigma_t + \log \sigma_t) \right. \\
&\quad \left. - \frac{1}{2} m''(c_4) \left( \kappa g'(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t}) \right)^2 \right\} \\
&= \exp \left\{ -\lambda\tilde{x}_0\sigma_t - \log \sqrt{\sigma_t} + \kappa - \frac{\lambda\kappa t^2}{2\sigma_t(g(\lambda\sigma_t))^3} \left( \kappa g'(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t}) \right)^2 \right\} \\
&= \exp \left\{ -\lambda\tilde{x}_0\sigma_t - \log \sqrt{\sigma_t} + \kappa + o(1) \right\},
\end{aligned}$$

as  $t \rightarrow \infty$ . This last equality holds in view of Lemma 17, (Equation (38)), since

$$\begin{aligned}
\frac{\lambda\kappa t^2}{2(g(\lambda\sigma_t))^3\sigma_t} \left( \kappa g'(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t}) \right)^2 &= \frac{\lambda\kappa t^2}{2(g(\lambda\sigma_t))^3\sigma_t} \kappa^2 \left( g'(\lambda\sigma_t) + g''(c_2)(\lambda w\sigma_t + \log \sqrt{\sigma_t}) \right)^2 \\
&= \frac{\lambda\kappa t^2}{2(g(\lambda\sigma_t))^3\sigma_t} \kappa^2 \left( \frac{g(\lambda\sigma_t)}{\lambda(t - \sigma_t)} - \frac{\kappa}{\sigma_t t} (\lambda w\sqrt{\sigma_t} + \log \sqrt{\sigma_t}) \right)^2 \\
&= \mathcal{O}(\sigma_t^{-1}) = o(1).
\end{aligned}$$

For  $E > 0$ , we need  $\mu(f_x, 1) > \mu(\hat{f}_w(\kappa), 1)$ , which holds if and only if

$$x \leq \exp \left\{ -\frac{\varepsilon}{2} \exp \left\{ \frac{\gamma}{\lambda} g(\lambda\sigma_t) \left( \kappa + \frac{\lambda\kappa w^2}{2g(\lambda\sigma_t)} - \frac{\lambda y}{g(\lambda\sigma_t)} \right) \right\} \right\} =: f_1.$$

Since  $g(\lambda\sigma_t) \rightarrow 1$  as  $t \rightarrow \infty$ , we get

$$f_1 = \exp \left\{ -\eta\varepsilon \exp \left\{ \frac{\gamma}{\lambda} \left( \kappa + \frac{\lambda\kappa w^2}{2} \right) - \lambda y + o(1) \right\} \right\}.$$

Hence we can rewrite  $E$  as

$$\begin{aligned}
E &= (1 + o(1)) e^{-\lambda\sigma_t - \lambda w\sqrt{\sigma_t} - \log \sqrt{\sigma_t}} \int_0^{f_1} \left( \exp \left\{ \frac{\lambda}{\gamma} \log \left( -\frac{1}{\eta\varepsilon} \log x \right) - \frac{\lambda\kappa w^2}{2} + \lambda y \right\} - e^\kappa \right) dx \\
&= (1 + o(1)) e^{-\lambda\sigma_t - \lambda w\sqrt{\sigma_t} - \log \sqrt{\sigma_t}} \left( e^{\lambda y - \frac{\lambda\kappa w^2}{2}} \left( \frac{1}{\eta\varepsilon} \right)^{\frac{\lambda}{\gamma}} \int_0^{f_1} \left( \log \frac{1}{x} \right)^{\frac{\lambda}{\gamma}} dx - \int_0^{f_1} e^\kappa dx \right) \\
&= (1 + o(1)) e^{-\lambda\sigma_t - \lambda w\sqrt{\sigma_t} - \log \sqrt{\sigma_t}} \left( e^{\lambda y - \frac{\lambda\kappa w^2}{2}} \left( \frac{1}{\eta\varepsilon} \right)^{\frac{\lambda}{\gamma}} \Gamma \left( \frac{\lambda}{\gamma} + 1, \eta\varepsilon e^{\frac{\gamma}{\lambda} \left( \kappa + \frac{\lambda\kappa w^2}{2} \right)} \right) \right. \\
&\quad \left. - \exp \left\{ \kappa - \eta\varepsilon e^{\frac{\gamma}{\lambda} \left( \kappa + \frac{\lambda\kappa w^2}{2} - \lambda y \right)} \right\} \right),
\end{aligned}$$

where  $\Gamma(s, x) = \int_x^\infty z^{s-1} e^{-z} dz$  is the upper incomplete gamma function. So we get

$$S \leq (1 + o(1)) c_0 \int_{-\infty}^\infty \lambda \left( e^{\lambda y - \frac{\lambda\kappa w^2}{2}} \left( \frac{1}{\eta\varepsilon} \right)^{\frac{\lambda}{\gamma}} \Gamma \left( \frac{\lambda}{\gamma} + 1, \eta\varepsilon e^{\frac{\gamma}{\lambda} \left( \kappa + \frac{\lambda\kappa w^2}{2} - \lambda y \right)} \right) - \exp \left\{ \kappa - \eta\varepsilon e^{\frac{\gamma}{\lambda} \left( \kappa + \frac{\lambda\kappa w^2}{2} - \lambda y \right)} \right\} \right) dw.$$

Since  $\frac{\Gamma(s, x)}{x^{s-1}e^{-x}} \rightarrow 1$  as  $x \rightarrow \infty$ , as  $\kappa \rightarrow \infty$  we have

$$\Gamma \left( \frac{\lambda}{\gamma} + 1, \eta\varepsilon e^{\frac{\gamma}{\lambda} \left( \kappa + \frac{\lambda\kappa w^2}{2} - \lambda y \right)} \right) \sim (\eta\varepsilon)^{\frac{\lambda}{\gamma}} \exp \left\{ \kappa + \frac{\lambda\kappa w^2}{2} - \lambda y - \eta\varepsilon e^{\frac{\gamma}{\lambda} \left( \kappa + \frac{\lambda\kappa w^2}{2} - \lambda y \right)} \right\},$$

and so  $S \rightarrow 0$ . □

### 4.3 Contribution of old and fit families

**Lemma 16** (Absence of fit families above the ‘‘window’’). *For every  $\varepsilon > 0$  and  $\nu > 0$ , there exists  $\kappa > 0$  such that for all sufficiently large  $t$ , we have*

$$\mathbb{P}\left(\max_{n \in E_t^c(v)} \left(\frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}\right) \leq \kappa\right) \geq 1 - \varepsilon, \quad (33)$$

where  $E_t^c(v) = [n_t(-v), n_t(v)]$ ,  $n_t(\pm v) := \exp\{\lambda(\sigma_t \pm v\sqrt{\sigma_t})\}$ .

*Proof.* Let  $\varepsilon, \nu > 0$  and  $\kappa > 0$ . We have<sup>2</sup>

$$\begin{aligned} \mathbb{P}\left(\max_{n \in E_t^c(v)} \left(\frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}\right) \leq \kappa\right) &= \prod_{n_t(-v)}^{n_t(v)} \mathbb{P}\left(F_n \leq g(\log(n\sqrt{\sigma_t})) + \kappa g'(\log(n\sqrt{\sigma_t}))\right) \\ &= \prod_{n_t(-v)}^{n_t(v)} \left(1 - \mu\left(g(\log(n\sqrt{\sigma_t})) + \kappa g'(\log(n\sqrt{\sigma_t})), 1\right)\right). \end{aligned}$$

Using the fact that  $e^{-\mu(x,1)} = 1 - \mu(x,1) + o(\mu(x,1))$  when  $x \rightarrow 1$ , we get that when  $t \rightarrow \infty$ ,

$$\mathbb{P}\left(\max_{n \in E_t^c(v)} \left(\frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}\right) \leq \kappa\right) \sim \exp\left\{-\sum_{n_t(-v)}^{n_t(v)} \mu\left(g(\log(n\sqrt{\sigma_t})) + \kappa g'(\log(n\sqrt{\sigma_t})), 1\right)\right\}.$$

Recall that  $\mu(x,1) = e^{-m(x)}$ , which implies that, as  $t \rightarrow \infty$ ,

$$\mathbb{P}\left(\max_{n \in E_t^c(v)} \left(\frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}\right) \leq \kappa\right) \sim \exp\left\{-\int_{n_t(-v)}^{n_t(v)} e^{-m\left(g(\log(x\sqrt{\sigma_t})) + \kappa g'(\log(x\sqrt{\sigma_t}))\right)} dx\right\}.$$

Using the change of variables with  $x = e^{\lambda(\sigma_t + w\sqrt{\sigma_t})} = n_t(w)$ , we get

$$\begin{aligned} \mathbb{P}\left(\max_{n \in E_t^c(v)} \left(\frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}\right) \leq \kappa\right) \\ \sim \exp\left\{-\int_{-v}^v e^{-m\left(g(\log(n_t(w)\sqrt{\sigma_t})) + \kappa g'(\log(n_t(w)\sqrt{\sigma_t}))\right)} \lambda\sqrt{\sigma_t} n_t(w) dw\right\}. \end{aligned}$$

By the same technique as in Lemma 15(a), we get that there exists

$$c_6 \in [g(\log(n_t(w)\sqrt{\sigma_t})), g(\log(n_t(w)\sqrt{\sigma_t})) + \kappa g'(\log(n_t(w)\sqrt{\sigma_t}))]$$

such that

$$\begin{aligned} m\left(g(\log(n_t(w)\sqrt{\sigma_t})) + \kappa g'(\log(n_t(w)\sqrt{\sigma_t}))\right) \\ = m\left(g(\log(n_t(w)\sqrt{\sigma_t}))\right) + \kappa m'\left(g(\log(n_t(w)\sqrt{\sigma_t}))\right) g'(\log(n_t(w)\sqrt{\sigma_t})) \\ + \frac{1}{2} m''(c_6) \left(\kappa g'(\log(n_t(w)\sqrt{\sigma_t}))\right)^2 \\ = \log(n_t(w)\sqrt{\sigma_t}) + \kappa + \mathcal{O}(\sigma_t^{-1}) = \lambda(\sigma_t + w\sqrt{\sigma_t}) + \log(\sqrt{\sigma_t}) + \kappa + \mathcal{O}(\sigma_t^{-1}), \end{aligned}$$

where we have used that  $m(g(x)) = x$  and hence  $m'(g(x))g'(x) = 1$  for all  $x > 0$ . We also used the fact that  $m''(c_6)(\kappa g'(\log(n_t(w)\sqrt{\sigma_t})))^2 \rightarrow 0$  as  $t \rightarrow \infty$ , by Assumption **(A5.2)**. Therefore, the integral becomes

$$\begin{aligned} \int_{-v}^v e^{-m\left(g(\lambda(\sigma_t + w\sqrt{\sigma_t})) + \kappa g'(\lambda(\sigma_t + w\sqrt{\sigma_t}))\right)} \sqrt{\sigma_t} e^{\lambda(\sigma_t + w\sqrt{\sigma_t})} dw \\ = \int_{-v}^v e^{-\lambda(\sigma_t + w\sqrt{\sigma_t}) - \log(\sqrt{\sigma_t}) - \kappa - \mathcal{O}(\sigma_t^{-1})} \sqrt{\sigma_t} e^{\lambda(\sigma_t + w\sqrt{\sigma_t})} dw = \int_{-v}^v e^{-\kappa - \mathcal{O}(\sigma_t^{-1})} dw \sim 2ve^{-\kappa}. \end{aligned}$$

<sup>2</sup>For all  $x, y \in \mathbb{R}$ , we denote by  $\prod_{n=x}^y$  and  $\sum_{n=x}^y$  the product (resp. sum) over integers with  $x \leq n \leq y$ .

Therefore, we get

$$\mathbb{P}\left(\max_{n \in E_t^c(v)} \left(\frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}\right) \leq \kappa\right) \sim \exp\{-2ve^{-\kappa}\} \rightarrow 1, \quad \text{as } \kappa \rightarrow \infty. \quad \square$$

#### 4.4 Proof of Theorem 3

Let  $\eta, \varepsilon > 0$ . By Lemma 15 there exists  $\kappa_1 = \kappa_1(\varepsilon, \eta)$  such that

$$\liminf_{t \rightarrow \infty} \mathbb{P}\left(\Gamma_t([-\infty, \infty] \times [-\infty, -\kappa_1] \times (\varepsilon, \infty]) = 0\right) \geq 1 - \eta.$$

By Lemma 14 there exists  $v = v(\varepsilon, \eta) > 1$  such that

$$\liminf_{t \rightarrow \infty} \mathbb{P}\left(\Gamma_t([-v, \infty] \cup [-\infty, -v] \times [-\infty, \infty] \times (\varepsilon, \infty]) = 0\right) \geq 1 - \eta.$$

By Lemma 16 there exists  $\kappa_2 = \kappa_2(\varepsilon, \eta)$  such that

$$\liminf_{t \rightarrow \infty} \mathbb{P}\left(\Gamma_t([-v, v] \times [\kappa_2, \infty] \times (\varepsilon, \infty]) = 0\right) \geq 1 - \eta.$$

Finally, Proposition 13 gives that  $\Gamma_t$  converges on  $(-v, v) \times (-\kappa_1, \kappa_2) \times (\varepsilon, \infty]$  to the Poisson process with intensity measure  $\zeta$ . Combining these four facts and using that  $\eta > 0$  is arbitrarily small, we get convergence on  $[-\infty, \infty] \times [-\infty, \infty] \times (\varepsilon, \infty]$ . As this holds for all  $\varepsilon > 0$  the proof is complete.

#### 4.5 Proof of Corollary 4

(i) Vague convergence in distribution of  $\Gamma_t$  to  $\text{PPP}(\zeta)$  implies convergence in distribution of  $\Gamma_t(B)$  to  $\text{PPP}(\zeta)(B)$  for compact sets  $B$  with  $\zeta(\partial B) = 0$ , see, e.g., [20, Proposition 3.12]. We fix  $x > 0$  and  $B := [-\infty, \infty] \times [-\infty, \infty] \times [x, \infty]$ . By Theorem 3, we get that, as  $t \uparrow \infty$ ,

$$\sum_{n=1}^{M(t)} \mathbf{1}_B\left(\frac{\tau_n - \sigma_t}{\sqrt{\sigma_t}}, \frac{F_n - g(\log(n\sqrt{\sigma_t}))}{g'(\log(n\sqrt{\sigma_t}))}, e^{-\gamma g(\lambda\sigma_t)(t-\sigma_t) - a_1 g(\lambda\sigma_t) \log \sigma_t + \gamma T} Z_n(t)\right) \Rightarrow \text{Poisson}\left(\int_B d\zeta\right),$$

since  $B$  is a compact set. Hence, as  $t \uparrow \infty$ ,

$$\begin{aligned} & \mathbb{P}\left(e^{-\gamma g(\lambda\sigma_t)(t-\sigma_t) - a_1 g(\lambda\sigma_t) \log \sigma_t + \gamma T} \max_{n \in \{1, \dots, M(t)\}} Z_n(t) \geq x\right) \\ & \rightarrow \mathbb{P}\left(\text{Poisson}\left(\int_B d\zeta\right) \geq 1\right) = 1 - \mathbb{P}\left(\text{Poisson}\left(\int_B d\zeta\right) = 0\right) = 1 - \exp\left(-\int_B d\zeta\right). \end{aligned} \quad (34)$$

Note that

$$\begin{aligned} \int_B d\zeta &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_x^{\infty} \lambda e^{-f} e^{s^2 a_2 - f a_3} \nu(z e^{s^2 a_2 - f a_3}) dz df ds \\ &= \lambda \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{x e^{s^2 a_2 - f a_3}}^{\infty} e^{-f} \nu(w) dw df ds = \lambda \int_{-\infty}^{\infty} \int_0^{\infty} \int_{\frac{1}{a_3}(s^2 a_2 - \log \frac{w}{x})}^{\infty} e^{-f} \nu(w) df dw ds \\ &= \lambda \left(\int_{-\infty}^{\infty} e^{-\frac{a_2}{a_3} s^2} ds\right) \left(\int_0^{\infty} \nu(w) \left(\frac{w}{x}\right)^{\frac{1}{a_3}} dw\right) = \lambda \sqrt{\pi \frac{a_3}{a_2}} \left(\int_0^{\infty} \nu(w) w^{\frac{1}{a_3}} dw\right) x^{-\frac{1}{a_3}}. \end{aligned} \quad (35)$$

Recall that  $a_2 = \gamma^2/2$  and  $a_3 = \frac{\gamma}{\lambda}$ . Thus the right hand side in (34) is  $1 - \exp(-s^\eta x^{-\eta})$ , for  $\eta = \frac{\lambda}{\gamma}$ . In summary, for all  $x > 0$ , we have

$$\mathbb{P}\left(e^{-\gamma g(\lambda\sigma_t)(t-\sigma_t) - a_1 g(\lambda\sigma_t) \log \sigma_t + \gamma T} \max_{n \in \{1, \dots, M(t)\}} Z_n(t) \leq x\right) \rightarrow e^{-\left(\frac{x}{s}\right)^{-\frac{\lambda}{\gamma}}} = \mathbb{P}(W \leq x),$$

where  $W \sim \text{Fréchet}(\frac{\lambda}{\gamma}, s)$ .

(ii) We have

$$\mathbf{1}_{\frac{S(t)-\sigma_t}{\sqrt{\sigma_t}} \geq x} = \int \mathbf{1}_{s \geq x} \mathbf{1}_{\Gamma_t([-\infty, \infty] \times [-\infty, \infty] \times (z, \infty]) = 0} d\Gamma_t(s, f, z),$$

which is an almost everywhere vaguely continuous bounded function of  $\Gamma_t$ . By Theorem 3, we have

$$\lim_{t \rightarrow \infty} \mathbb{P}\left(\frac{S(t)-\sigma_t}{\sqrt{\sigma_t}} \geq x\right) = \int \mathbf{1}_{s \geq x} \mathbb{P}(\text{PPP}(\zeta)([-\infty, \infty] \times [-\infty, \infty] \times [z, \infty])) = 0) d\zeta(s, f, z).$$

Hence, the random variable  $\frac{S(t)-\sigma_t}{\sqrt{\sigma_t}}$  converges to a random variable  $U$  with density

$$\int_{-\infty}^{\infty} \int_0^{\infty} e^{-\zeta([-\infty, \infty] \times [-\infty, \infty] \times [z, \infty])} \zeta(s, df, dz).$$

We recall from above that

$$\zeta([-\infty, \infty] \times [-\infty, \infty] \times [z, \infty]) = \lambda \sqrt{\pi \frac{a_3}{a_2}} \left( \int_0^{\infty} \nu(w) w^{\frac{1}{a_3}} dw \right) z^{-\frac{1}{a_3}} =: c_6 z^{-\frac{1}{a_3}}.$$

We get, substituting  $u = ze^{s^2 a_2 - f a_3}$ ,

$$\begin{aligned} \int_{-\infty}^{\infty} \int_0^{\infty} e^{-\zeta([-\infty, \infty] \times [-\infty, \infty] \times [z, \infty])} d\zeta(s, f, z) \\ = \lambda \int_0^{\infty} \nu(u) \int_{-\infty}^{\infty} \exp\left\{-f - c_6 u^{-\frac{1}{a_3}} e^{s^2 \frac{a_2}{a_3}} e^{-f}\right\} df du ds. \end{aligned}$$

Integrating with respect to  $f$  and simplifying, gives us

$$\int_{-\infty}^{\infty} \int_0^{\infty} e^{-\zeta([-\infty, \infty] \times [-\infty, \infty] \times [z, \infty])} d\zeta(s, f, z) = \frac{\lambda}{c_6} e^{-s^2 \frac{a_2}{a_3}} ds \int_0^{\infty} \nu(u) u^{\frac{1}{a_3}} du = \frac{1}{\sqrt{\frac{2\pi}{\lambda z}}} e^{-s^2 \frac{\lambda z}{2}} ds.$$

## 5 Appendix: an auxiliary lemma

For our proofs we need the following consequence of Taylor's formula.

**Lemma 17.** *For all  $x \in [0, 1]$ , there exists  $c_3 \in [g(\lambda\sigma_t), g(\lambda\sigma_t) + \frac{w\sqrt{\sigma_t}}{t}g(\lambda\sigma_t) + \frac{a_1 g(\lambda\sigma_t)}{\gamma t} \log \sigma_t - \frac{1}{\gamma t} \log(-\frac{1}{\eta^\varepsilon} \log x)]$  such that*

$$\begin{aligned} m\left(g(\lambda\sigma_t) + \frac{w\sqrt{\sigma_t}}{t}g(\lambda\sigma_t) + \frac{a_1 g(\lambda\sigma_t)}{\gamma t} \log \sigma_t - \frac{1}{\gamma t} \log\left(-\frac{1}{\eta^\varepsilon} \log x\right)\right) \\ = m(g(\lambda\sigma_t)) + m'(g(\lambda\sigma_t))\left(\frac{w\sqrt{\sigma_t}}{t}g(\lambda\sigma_t) + \frac{a_1 g(\lambda\sigma_t)}{\gamma t} \log \sigma_t - \frac{1}{\gamma t} \log\left(-\frac{1}{\eta^\varepsilon} \log x\right)\right) \\ + \frac{1}{2}m''(c_3)\left(\frac{w\sqrt{\sigma_t}}{t}g(\lambda\sigma_t) + \frac{a_1 g(\lambda\sigma_t)}{\gamma t} \log \sigma_t - \frac{1}{\gamma t} \log\left(-\frac{1}{\eta^\varepsilon} \log x\right)\right)^2. \end{aligned} \quad (36)$$

For  $\tilde{x}_0 > 0$  and  $\kappa > 0$  there exist  $c_4 \in [g(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t}), g(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t}) - \kappa g'(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t})]$ , such that

$$\begin{aligned} m\left(g(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t}) - \kappa g'(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t})\right) \\ = m\left(g(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t})\right) + m'\left(g(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t})\right)\left(-\kappa g'(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t})\right) \\ + \frac{1}{2}m''(c_4)\left(-\kappa g'(\lambda\tilde{x}_0\sigma_t + \log \sqrt{\sigma_t})\right)^2. \end{aligned} \quad (37)$$

And finally for all  $w \in [-\infty, \infty]$  there exists  $c_5 \in [\lambda\sigma_t, \lambda\sigma_t + \lambda w\sqrt{\sigma_t} + \log \sqrt{\sigma_t}]$  such that

$$g'(\lambda\sigma_t + \lambda w\sqrt{\sigma_t} + \log \sqrt{\sigma_t}) = g'(\lambda\sigma_t) + g''(c_5)(\lambda w\sqrt{\sigma_t} + \log \sqrt{\sigma_t}). \quad (38)$$

*Proof.* This follows from a Taylor expansion. □

**Acknowledgments:** This paper contains the main results of the third author's PhD thesis [21]. The first author is supported by EPSRC Fellowship EP/R022186/1, the second author by EPSRC grant EP/K016075/1, and the third is supported by a scholarship from the EPSRC Centre for Doctoral Training in Statistical Applied Mathematics at Bath (SAMBa), under the project EP/L015684/1. We would like to thank two anonymous referees for their very careful reading of the manuscript, which led to substantial improvements of the paper.

## References

- [1] ALDOUS, D. *Exchangeability and related topics*. Springer Lecture Notes in Mathematics 1117, pp 1-198, 1985.
- [2] ASMUSSEN, S. AND HERING, H. *Branching processes*. Springer, 1983
- [3] ATHREYA, K. B., AND KARLIN, S.. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39 (1968) 1801-1817.
- [4] ATHREYA, K. B., AND NEY, P. E. *Branching Processes*, Springer-Verlag, New York, 1972.
- [5] BETZ, V., UELTSCHI, D., AND VELENIK, Y. Random permutations with cycle weights. *Ann. Appl. Probab.* 21 (2011), 312–331.
- [6] BIANCONI, G., AND BARABÁSI, A.-L. Bose-Einstein condensation in complex networks. *Physical Review Letters* 86 (2001), 5632 – 5635.
- [7] DEREICH, S. Preferential attachment with fitness: Unfolding the condensate. *Electronic Journal of Probability* 21 (2016), Paper 3.
- [8] DEREICH, S., MAILLER, C., AND MÖRTERS, P. Non-extensive condensation in reinforced branching processes. *Annals of Applied Probability* 27 (2017), 2539–2568.
- [9] DEREICH, S., AND MÖRTERS, P. Emergence of condensation in Kingman's model of selection and mutation. *Acta Appl. Math* 127 (2013), 17–26.
- [10] DOOB, J. *Classical potential theory and its probabilistic counterpart*. Springer Verlag, 2001.
- [11] EMBRECHTS, P., KLÜPPELBERG, C., AND MIKOSCH, T. *Modelling Extremal Events for Insurance and Finance*. Springer, 1999.
- [12] GNEDENKO, B. V. Sur la distribution limitée du terme d'une série aléatoire. *Ann. Math.* 44 (1943), 423–453.
- [13] GRIMMETT, G. R., AND STIRZAKER, D. R. *Probability and Random Processes*, second ed. Oxford University Press, Oxford, 2009.



- [14] HODGINS-DAVIS, A., RICE, D. P., AND TOWNSEND, J. P. Gene expression evolves under a house-of-cards model of stabilizing selection. *Mol. Biol. Evol.* (2015), doi: 10.1093/molbev/msv094.
- [15] KENDALL, D.G. Branching processes since 1873. *Journal of the London Mathematical Society*, (1978), 1(1), 385-406.
- [16] KINGMAN, J. A simple model for the balance between selection and mutation. *Journal Appl. Prob.* 15 (1978), 1–12.
- [17] LIU, Q. S. The growth of an entire characteristic function and the tail probabilities of the limit of a tree martingale. In: Chauvin B., Cohen S., Rouault A. (eds) *Trees*. Progress in Probability, vol 40. Birkhäuser Basel (1996), 51–80.
- [18] NERMAN, O. On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 57 (1981), 365–395.
- [19] OLOFSSON, P. General branching processes with immigration. *Journal of Applied Probability* 33 (1996), 940–948.
- [20] RESNICK, S. *Extreme values, regular variation, and point processes*. Springer-Verlag, 1987.
- [21] SENKEVICH, A. *Competing growth processes with applications to networks*. PhD thesis, University of Bath, 2019. Available at <https://researchportal.bath.ac.uk/en/studentTheses/competing-growth-processes-with-applications-to-networks>

Cécile Mailer, Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, UK  
 Peter Mörters, Mathematisches Institut, Universität zu Köln, Weyertal 86-90, 50931 Köln, Germany  
 Anna Senkevich, Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, UK