



**HAL**  
open science

# The genealogy of an exactly solvable Ornstein-Uhlenbeck type branching process with selection

Aser Cortines, Bastien Mallein

► **To cite this version:**

Aser Cortines, Bastien Mallein. The genealogy of an exactly solvable Ornstein-Uhlenbeck type branching process with selection. 2018. hal-01790945v1

**HAL Id: hal-01790945**

**<https://hal.science/hal-01790945v1>**

Preprint submitted on 14 May 2018 (v1), last revised 18 May 2019 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The genealogy of an exactly solvable Ornstein-Uhlenbeck type branching process with selection

Aser Cortines\*  
Universität Zürich

Bastien Mallein†  
LAGA, Université Paris 13

April 20, 2018

## Abstract

We study the genealogy of a solvable population model with  $N$  particles on the real line which evolves according to a discrete-time branching process with selection. At each time step, every particle gives birth to children around  $a$  times its current position, where  $a > 0$  is a parameter of the model. Then, the  $N$  rightmost new-born children are selected to form the next generation. We show that the genealogical trees of the process converge to those of a Beta coalescent as  $N \rightarrow \infty$ . The process we consider can be seen as a toy-model version of a continuous-time branching process with selection, in which particles move according to independent Ornstein-Uhlenbeck processes. The parameter  $a$  is akin to the pulling strength of the Ornstein-Uhlenbeck motion.

## 1 Introduction

A *branching-selection particle system* is a Markov process of particles on the real line that evolves through the repeated application of the two following steps:

**Branching step:** each particle currently alive in the process independently gives birth to children according to a point process whose law depends on the position of the particle.

**Selection step:** some of the newborn children are selected to form the next generation and reproduce at the next *branching step*, while the other particles are “discarded” from the process.

From a biological perspective, such models can be thought of as toy-models for the competition between individuals in a population evolving in an environment with limited resources. In this sense, the positions of particles (also seen as individuals) may be interpreted as their *fitness*: individuals with large fitness have more propensity to reproduce and transfer their genetic advantage to their offspring.

Branching-selection particle systems are also of physical relevance. They are related to noisy reaction-diffusion phenomena and to the F-KPP equation [5, 6], hence they are often used to describe the evolution of disordered systems having two homogeneous steady states. The prototypical example of such systems is the so-called  $N$ -branching random walk: in this process, particles make independently children around their current position, and at each step only the  $N$  rightmost children are kept alive. Based on numerical simulations and the analysis of solvable models (see [5, 6]), it has been conjectured that many branching random walks with similar selection procedure satisfy universal properties. The cloud of particles travels at a deterministic speed  $v_N$  that should satisfy

$$v_N - v_\infty = \frac{-\chi}{(\log N + 3 \log \log N + o(\log \log N))^2}. \quad (1.1)$$

Moreover, the genealogical trees of such models should converge to those of a Bolthausen-Sznitman coalescent.

---

\*aser.cortinespeixoto@math.uzh.ch

†mallein@math.univ-paris13.fr

Some of these conjectures have been verified for individual models. Bérard and Gouéré [1] proved that the corrections to the speed  $v - v_N$  of many  $N$ -branching random walks are of the order  $(\log N)^{-2}$ , in accordance with the conjectures from [6]. These results have been extended to branching random walks with different integrability conditions in [2, 10, 15], and to other related models in [7, 14, 17]. Maillard [13] proved that the barycentre of a  $N$ -branching Brownian motion converges, after proper scaling, to a Lévy process. This result indicates the validity of the second order corrections in (1.1). Other results on the hydrodynamic limit of the shape of the front were obtained in [11, 12].

As for the genealogical structure of such models, the authors in [3] show that the genealogy of a branching Brownian motion with quasi-critical absorption converges toward the Bolthausen-Sznitman coalescent. To the best of our knowledge this is the only example lying on the  $N$ -branching random walk universality class for which the conjectures from [6] about the genealogy have been verified. Nevertheless results on related models [6, 8, 9] indicate the robustness of this conjecture. We studied in [9] the so-called *exponential model* from [6], which shares many common features with  $N$ -branching random walks, even though it does not belong to the same universality class. We showed that when the selection procedure favours the rightmost individuals, the genealogy of the process converges toward the Bolthausen-Sznitman coalescent. This result lead to the natural question: under which conditions do different genealogical behaviours arise?

It is well known that the genealogy of neutral population models, such as Wright-Fisher and Moran models, converge to the Kingman coalescent [16]. On the other hand, if the rightmost particles are favoured by the selection procedure, then the Bolthausen-Sznitman coalescent is expected. Hence we look for coalescent processes arising in branching-selection particle systems, which interpolate between these two behaviours. Such a family appears in the context of Galton-Watson trees [20], where the so-called Beta coalescent forms a 1-parameter family interpolating between the Kingman and the Bolthausen-Sznitman coalescent.

In this paper, we consider a variant of the exponential model from [6], in which particles are subjected to a pulling force attracting them to zero. Let  $N \in \mathbb{N}$  denote the size of the population and  $a \in \mathbb{R}_+$  be a positive parameter governing the intensity of the attractive force. The process is defined as follows: it starts with  $N$  particles scattered on the on the real line. At each discrete time  $n$ , every particle gives birth to children whose position are determined by independent Poisson point processes. More precisely, the offspring of an individual located at  $x \in \mathbb{R}$  are positioned according to a Poisson point process with intensity  $e^{ax-y}dy$  on  $\mathbb{R}$ . We then select the  $N$  rightmost newborn individuals to form the next generation of the process. We call this process the  $(N, a)$ -*exponential model*.

In the coming section, we will show that the  $(N, a)$ -exponential model is well-defined for all  $n \in \mathbb{N}$ , and that it is a reversible Markov process. Hence we can construct a version of it for all  $n \in \mathbb{Z}$ . Thanks to this bi-infinite construction, we define the *ancestral partition process*  $\Pi^N$  of the process as follows: for every  $n \in \mathbb{N}$ , let  $\Pi_n^N$  be the partition of  $\{1, \dots, N\}$  such that  $i$  and  $j$  belong to the same block if and only if the  $i$ th and the  $j$ th rightmost particles at time 0 share a common ancestor at time  $-n$ . Our main result concerns the asymptotic behaviour of  $\Pi^N$  as  $N \rightarrow \infty$ .

**Theorem 1.1.** *As  $N \rightarrow \infty$ , we have:*

- a) *If  $0 < a < 1/2$ , then  $(\Pi_{[tN]}^N, t \geq 0)$  converges in law to the Kingman coalescent.*
- b) *If  $a = 1/2$ , then  $(\Pi_{[tN/\log N]}^N, t \geq 0)$  converges in law to the Kingman coalescent.*
- c) *If  $1/2 < a < 1$ , then  $(\Pi_{[tN^{(1-a)/a}]}^N, t \geq 0)$  converges in law to the Beta( $2 - a^{-1}, a^{-1}$ )-coalescent.*
- d) *If  $a = 1$ , then  $(\Pi_{[t \log N]}^N, t \geq 0)$  converges in law to the Bolthausen-Sznitman coalescent.*
- e) *If  $a > 1$ , then  $\Pi^N$  converges in law toward a discrete coalescent on  $\mathbb{N}$ .*

We refer to [4, Example 3, p72] for the precise definitions of the limiting processes, and Section 2.2 of the same book for the topology in which the above convergence takes place.

In view of the above result,  $a = 1$  marks a phases transition in the coalescent behaviour. We can also observe this transition in the dynamical behaviour of the cloud of particles. As long as

$a < 1$ , the cloud of particles remains within finite distance from 0. When  $a = 1$ , we proved in [9] that it drifts toward  $\infty$  at positive speed. Finally, one can easily show that when  $a > 1$  the cloud moves away from 0 at exponential rate. For  $a < 1$ , we have the following more precise estimates on the extremal positions of the cloud of particles.

**Proposition 1.2.** *Given  $a \in (0, 1)$ , we denote by  $M_n$  and  $m_n$  the largest and smallest positions at time  $n$  in the  $(N, a)$ -exponential model. Then, if the initial conditions are bounded in  $L^1$  uniformly in  $N$ , we have*

$$\lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} (\mathbf{E}(M_n) - \log N) = \gamma - \frac{\log(1-a)}{1-a} \quad \text{and} \quad \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbf{E}(m_n) = -\frac{\log(1-a)}{1-a},$$

where  $\gamma$  is the Euler-Mascheroni constant.

Proposition 1.2 shows that the cloud of particle in the  $(N, a)$ -exponential model is roughly of size  $\log N$ , which is typical in many branching selection particle systems.

The proofs of both Proposition 1.2 and Theorem 1.1 rely on the observation that the distribution of the children at time  $n+1$  is a Poisson point process with exponential intensity around the position of a unique fictional particle. This construction was introduced in [6], and further developed in [9], from where we borrow the approach.

**Outline of the paper:** In the next section, we prove the two main results of the paper, namely, Theorem 1.1 and Proposition 1.2. In Section 3, we define the *branching Ornstein-Uhlenbeck processes* and discuss its relationship with the  $(N, a)$ -exponential model.

## 2 Proofs of the main results

We start with the mathematical definition of the  $(N, a)$ -exponential model. Let  $(\mathcal{P}_{n,j}, n \in \mathbb{N}, j \in \mathbb{N})$  be an infinite array of i.i.d. copies of a Poisson point processes with intensity measure  $e^{-x} dx$  and  $X_1(0) > X_2(0) > \dots > X_N(0)$  denote the ranked position of the particles at time 0, that we shall assume distinct for the sake of simplicity. At each time  $n$ , the position of the  $i$ th largest individual is denoted by  $X_n(i)$ , and its children at time  $n+1$  are distributed according to the point process  $\mathcal{P}_{n+1,i}$  shifted by  $aX_n(i)$ . Then, we select the  $N$ -rightmost, new-born individuals to form the  $(n+1)$ th generation and denote by  $X_1(n+1) > \dots > X_N(n+1)$  their ranked positions. To keep track of the genealogical structure of the process, we shall label individuals with  $A_{n+1}(k) \in \{1, \dots, N\}$  according to the index of their parents.

In other words, for each  $n \in \mathbb{N}$  and  $k \leq N$ , we have that

- i)  $X_{n+1}(k)$  is the  $k$ th largest atom of the point process  $\sum_{j=1}^N \sum_{p \in \mathcal{P}_{n+1,j}} \delta_{aX_n(j)+p}$ ,
- ii)  $A_{n+1}(k)$  is the (unique) integer  $j \leq N$  such that  $X_{n+1}(k) - aX_n(j) \in \mathcal{P}_{n+1,j}$ .

It can be readily checked using standard properties of Poisson point processes that the above process is well defined for all  $n$ . Moreover, the law of  $(X_{n+1}(k), k \leq N)$  from (i) may be obtained as the  $N$  largest points in a Poisson point process centered at

$$X_n(\text{eq}) := \log \left( \sum_{j=1}^N e^{aX_n(j)} \right). \quad (2.1)$$

Therefore, one may think of  $X_n(\text{eq})$  as the “equivalent” position of the front, that is, a fictional particle that generates the entire front in generation  $n+1$ . In the next lemma we prove the above claim and characterize the (conditional) law of  $A_n(\cdot)$ .

**Lemma 2.1.** *The point processes  $\sum_{j=1}^N \delta_{X_{n+1}(j) - X_n(\text{eq})}$ ,  $n \in \mathbb{N}$  are i.i.d. with common distribution given by the  $N$  rightmost points in a Poisson point process with intensity measure  $e^{-x} dx$ . Moreover, let  $\mathcal{H} := \sigma(X_n(j), j \leq N, n \in \mathbb{N})$  and  $k_1, \dots, k_N \in \mathbb{Z}_+$  such that  $k_1 + \dots + k_N = N$ , then*

$$\mathbf{P}(A_{n+1}(j) = k_j; 1 \leq j \leq N | \mathcal{H}) = \prod_{j=1}^N \theta_n(k_j), \quad \text{where} \quad \theta_n(k) := \frac{e^{aX_n(k)}}{\sum_{i=1}^N e^{aX_n(i)}}.$$

*Proof.* This result is obtained with a reasoning similar to [9, Proposition 1.3 and Lemma 1.6], we will therefore only outline the main parts of the proof and omit the technical details.

First, using the superposition property of Poisson point processes, we obtain that

$$\sum_{j=1}^N \sum_{p \in \mathcal{P}_{n+1,j}} \delta_{aX_n(j)+p} \stackrel{(d)}{=} \sum_{p \in \mathcal{P}} \delta_{X_n(\text{eq})+p},$$

where  $\mathcal{P}$  is a Poisson point process with intensity measure  $e^{-x}dx$ .

Next, we note that for all  $i, j, n \in \mathbb{N}$

$$\begin{aligned} \mathbf{P}(A_{n+1}(j) = i | \mathcal{H}) &= \mathbf{P}(X_{n+1}(j) - aX_n(i) \in \mathcal{P}_{n+1,i} | \mathcal{H}) \\ &= \frac{e^{-X_{n+1}(j) - aX_n(i)}}{\sum_{k=1}^N e^{-X_{n+1}(j) - aX_n(k)}} = \frac{e^{aX_n(j)}}{\sum_{k=1}^N e^{aX_n(k)}}. \end{aligned}$$

Moreover, the  $A_{n+1}(j)$ 's are (conditionally) independent, which concludes the proof.  $\square$

Using Lemma 2.1, we observe that for any  $n \in \mathbb{N}$ , we have the following recursion equation:

$$X_{n+1}(\text{eq}) - aX_n(\text{eq}) = \log \left( \sum_{j=1}^N e^{a(X_{n+1}(j) - X_n(\text{eq}))} \right) \stackrel{(d)}{=} \log \left( \sum_{j=1}^N e^{a\Delta_j} \right),$$

where  $(\Delta_j, j \in \mathbb{N})$  are the ranked atoms of a Poisson point process with intensity measure  $e^{-x}dx$ . As a result, writing  $(\xi_j, j \in \mathbb{N})$  for an i.i.d. sequence of random variables with the same law as  $X_1(\text{eq}) - aX_0(\text{eq})$ , we have

$$X_n(\text{eq}) \stackrel{(d)}{=} \sum_{j=0}^{n-1} a^j \xi_{n-j} + a^n X_0(\text{eq}). \quad (2.2)$$

In other words,  $(X_n(\text{eq}), n \in \mathbb{N})$  behaves as a perpetuity process.

*Proof of Proposition 1.2.* We recall that  $M_n$  and  $m_n$  denote respectively the positions of the rightmost and the leftmost particles at time  $n$ . Using Lemma 2.1, we observe that

$$M_n - X_{n-1}(\text{eq}) \stackrel{(d)}{=} p_1 \quad \text{and} \quad m_n - X_{n-1}(\text{eq}) \stackrel{(d)}{=} p_N,$$

where  $p_j$  is the  $j$ th largest atom in a Poisson point process with intensity  $e^{-x}dx$ . Hence, we have by standard Poisson computations that

$$\mathbf{E}(M_n) = \mathbf{E}(X_{n-1}(\text{eq})) + \gamma \quad \text{and} \quad \mathbf{E}(m_n) = \mathbf{E}(X_{n-1}(\text{eq})) - \psi(N) \quad (2.3)$$

where  $\psi(N) := \Gamma'(N)/\Gamma(N)$  is the digamma function, which satisfies  $\psi(N) = \log N + \mathcal{O}(N^{-1})$  as  $N \rightarrow \infty$ . It is therefore enough to compute the asymptotic behaviour of  $\mathbf{E}(X_n(\text{eq}))$  as  $n \rightarrow \infty$  and then  $N \rightarrow \infty$  to conclude the proof.

For what follows in the proof, we shall assume that  $a < 1$  and  $\mathbf{E}(X_0(\text{eq})) < \infty$ . Thus, taking the expected value in (2.2), one gets

$$\mathbf{E}(X_n(\text{eq})) = a^n \mathbf{E}(X_0(\text{eq})) + \frac{1-a^n}{1-a} \mathbf{E}(\xi_1), \quad \text{yielding} \quad \lim_{n \rightarrow \infty} \mathbf{E}(X_n(\text{eq})) = \frac{1}{1-a} \mathbf{E}(\xi_1). \quad (2.4)$$

It remains therefore to compute  $\mathbf{E}(\xi_1)$ . For this propose, we will first compute the Laplace transform  $\Lambda(\lambda)$  of  $\xi_1$ , then recover its mean via  $\mathbf{E}(\xi_1) = -(\log \Lambda)'(0)$ .

By [9, Proposition 4.2], we remark that

$$e^{\xi_1} = \sum_{j=1}^N e^{a\Delta_j} \stackrel{(d)}{=} e^{aZ_n} \sum_{j=1}^n e^{aE_j},$$

where  $(E_j, j \geq 1)$  are i.i.d. exponential random variables with parameter 1 and  $Z_N$  is an independent random variable whose distribution has density  $(N!)^{-1}e^{-(N+1)x-e^{-x}}$  with respect to the Lebesgue measure. Therefore, we have

$$\Lambda(\lambda) := \mathbf{E}(e^{-\lambda\xi_1}) = \mathbf{E}\left(\left(\sum_{j=1}^N e^{aE_j}\right)^{-\lambda}\right) \mathbf{E}(e^{-\lambda a Z_N}).$$

By direct computations, we obtain  $\mathbf{E}(e^{-\lambda a Z_N}) = \Gamma(N+1+a\lambda)/\Gamma(N+1)$  and

$$\mathbf{E}\left(\left(\sum_{j=1}^N e^{aE_j}\right)^{-\lambda}\right) = \frac{1}{\Gamma(\lambda)} \int_0^\infty t^{\lambda-1} \mathbf{E}\left(e^{-t \sum_{j=1}^N e^{aE_j}}\right) dt = \frac{1}{\Gamma(\lambda)} \int_0^\infty t^{\lambda-1} I(t)^N dt, \quad (2.5)$$

where  $I(t)$  is the function defined by

$$I(t) := \mathbf{E}\left(e^{-te^{aE_1}}\right) = \int_0^\infty e^{-x} e^{-te^{ax}} dx = \frac{t^{1/a}}{a} \int_t^\infty u^{-(1+1/a)} e^{-u} du.$$

Making the change of variable  $t = x/N$  in the right-hand side of (2.5) one obtains

$$\mathbf{E}\left(\left(\sum_{j=1}^N e^{aE_j}\right)^{-\lambda}\right) = \frac{1}{N^\lambda} \int_0^\infty I(x/N)^N \frac{x^{\lambda-1} dx}{\Gamma(\lambda)} =: \frac{J_N(\lambda)}{N^\lambda},$$

where  $J_N(\lambda)$  is a smooth function such that  $J_N(0) = 1$ . Collecting all pieces, we obtain that  $\Lambda(\lambda) = J_N(\lambda) \frac{\Gamma(N+a\lambda+1)}{N^\lambda \Gamma(N+1)}$ , which yields

$$\mathbf{E}(\xi_1) = -(\log \Lambda)'(0) = \log N - a\psi(N+1) - J'_N(0), \quad (2.6)$$

where we recall that  $\psi(x) = \Gamma'(x)/\Gamma(x)$  is the digamma function and  $(\log J_N)'(0) = \frac{J'_N(0)}{J_N(0)} = J'_N(0)$ .

To compute  $J'_N(0)$ , we will take the  $\lambda \rightarrow 0$  limit of

$$\frac{J_N(\lambda) - 1}{\lambda} = \frac{1}{\lambda \Gamma(\lambda)} \int_0^\infty x^{\lambda-1} (I(x/N)^N - e^{-x}) dx = \frac{1}{\Gamma(\lambda+1)} \int_0^\infty x^{\lambda-1} (I(x/N)^N - e^{-x}) dx. \quad (2.7)$$

By definition,  $I(t) \leq e^{-t}$  for all  $t \in \mathbb{R}_+$  which implies

$$|x^{\lambda-1} (I(x/N)^N - e^{-x})| \leq 2e^{-x} \quad \text{for all } x \geq 1 \text{ and } \lambda \in (0, 1). \quad (2.8)$$

On the other hand, we observe that for all  $t \in (0, 1]$ , we have

$$\begin{aligned} I(t) &= \frac{t^{1/a}}{a} \left( \int_t^\infty u^{-(1+1/a)} (1-u) du + \int_t^\infty u^{-(1+1/a)} (e^{-u} - 1 + u) du \right) \\ &= \frac{t^{1/a}}{a} \left( at^{-1/a} + \frac{t^{1-1/a}}{1-1/a} + \int_t^\infty u^{-(1+1/a)} (e^{-u} - 1 + u) du \right) \\ &= 1 - \frac{t}{1-a} + \mathcal{O}(t^b) \quad \text{as } t \rightarrow 0, \end{aligned}$$

for some  $b > 1$ . Indeed, we have  $t^{1/a} \int_1^\infty u^{-(1+1/a)} (e^{-u} - 1 + u) du = \mathcal{O}(t^{1/a})$  as  $t \rightarrow 0$ , and relying on the fact that  $(e^{-u} - 1 + u) = \mathcal{O}(u^2)$  for all  $u \leq 1$  one gets

$$t^{1/a} \int_t^1 u^{-(1+1/a)} (e^{-u} - 1 + u) du = \begin{cases} \mathcal{O}(t^{1/a}) & \text{if } a > 1/2; \\ \mathcal{O}(t^{1/a}) \log t & \text{if } a = 1/2; \\ \mathcal{O}(t^2) & \text{if } a < 1/2. \end{cases}$$

As a result, we obtain  $|I_N(x/N)^N - e^{-x}| \leq \frac{a}{a-1}x + C \left( \frac{x^b}{N^{b-1}} + x^2 \right)$  for all  $x \leq 1$ , where  $C > 0$  is a constant not depending on  $N$ . Thus, one can find a (possibly larger) constant  $\tilde{C} > 0$  such that

$$|x^{\lambda-1}(I(x/N)^N - e^{-x})| \leq \tilde{C}, \quad \text{for all } x < 1. \quad (2.9)$$

Thanks to (2.8) and (2.9), we can apply dominated convergence in (2.7), to obtain

$$J'_N(0) = \lim_{\lambda \rightarrow \infty} \frac{J_N(\lambda) - 1}{\lambda} = \int_0^\infty x^{-1}(I_N(x/N)^N - e^{-x})dx.$$

Now, we plug the above in (2.6) and use the fact that  $\lim_{N \rightarrow \infty} \psi(N+1) - \log N = 0$  to get

$$\mathbf{E}(\xi_1) = (1-a) \log N + \int_0^\infty x^{-1}(e^{-x} - I_N(x/N)^N)dx + o(1) \quad \text{as } N \rightarrow \infty. \quad (2.10)$$

Finally, we notice that  $I_N(x/N)^N$  tends to  $e^{-\frac{x}{1-a}}$  as  $N \rightarrow \infty$ . Therefore, we can rely again on (2.8) and (2.9), to apply dominated convergence thereby obtaining

$$\begin{aligned} \lim_{N \rightarrow \infty} \int_0^\infty x^{-1}(e^{-x} - I_N(x/N)^N)dx &= \int_0^\infty x^{-1}(e^{-x} - e^{-\frac{x}{1-a}})dx \\ &= \int_0^\infty \int_1^{\frac{1}{1-a}} e^{-ux} du dx = \int_1^{\frac{1}{1-a}} \frac{du}{u} = -\log(1-a), \end{aligned}$$

which, in sight of (2.3), (2.4) and (2.10) concludes the proof.  $\square$

We now turn to the proof of Theorem 1.1. We first use Lemma 2.1 to construct a bi-infinite version  $(\bar{X}_n, \bar{A}_n, n \in \mathbb{Z})$  of the  $(N, a)$ -exponential model.

*Remark 2.2.* Let  $(\mathcal{P}_n, n \in \mathbb{Z})$  be i.i.d. Poisson point processes with intensity  $e^{-x}dx$  and write  $(p_{n,j}, j \in \mathbb{N})$  for the sequence of atoms in  $\mathcal{P}_n$  ranked in the decreasing order. We set  $\bar{X}_0(\text{eq}) = 0$ , and for  $n \in \mathbb{Z} \setminus \{0\}$

$$\bar{X}_n(\text{eq}) = \begin{cases} \sum_{j=1}^n \zeta_j a^{n-j} & \text{if } n \geq 1; \\ -\sum_{j=-1}^n \zeta_j a^{n-j} & \text{if } n \leq -1, \end{cases}$$

where  $\zeta_n := \log \sum_{j=1}^N e^{ap_{n,j}}$ . The position of the  $N$  particles  $\bar{X}_n(1) > \dots > \bar{X}_n(N)$  in generation  $n \in \mathbb{Z}$  is then defined as

$$\bar{X}_n(k) = \bar{X}_{n-1}(\text{eq}) + p_{n,k}; \quad k = 1, \dots, N.$$

Next, we assign labels to the particles as follows: the particle  $\bar{X}_n(k)$  bears the label  $\bar{A}_n(k) \in \{1, \dots, N\}$  such that

$$\mathbf{P}(\bar{A}_n(k) = j \mid \mathcal{P}_{n-1}) = e^{ap_{n-1,j}} / \sum_{i=1}^N e^{ap_{n-1,i}}.$$

In the bi-infinite version  $(\bar{X}, \bar{A})$  of the  $(N, a)$ -exponential model, the family  $(\bar{A}_n, n \in \mathbb{Z})$  is i.i.d. We can now use the  $\bar{A}$ 's to reconstruct the *ancestral partition process* of the process as follows: for every  $n \in \mathbb{N}$ , we say that  $i$  and  $j$  belong to the same block of  $\Pi_n^N$  if

$$\bar{A}_{-n}(\bar{A}_{-(n-1)}(\dots \bar{A}_{-1}(i))) = \bar{A}_{-n}(\bar{A}_{-(n-1)}(\dots \bar{A}_{-1}(j))).$$

This allows us to express the law of  $\Pi^N$  in terms of a population dynamics with independent generation. Precisely, let

$$\theta_n(j) = \frac{e^{aX_{-n}(j)}}{\sum_{k=1}^N e^{aX_{-n}(k)}} = \frac{e^{ap_{n-1,j}}}{\sum_{i=1}^N e^{ap_{n-1,i}}}.$$

Then conditionally on  $(\theta_n(j), -n \in \mathbb{N}, j \leq N)$ , each individual at generation  $n \leq -1$  chooses its parent at generation  $n-1$  independently at random, selecting the parent  $j$  with probability  $\theta_{n-1}(j)$ . This is often called a Cannings model defined by a multinomial distribution with  $N$  independent trials and (random) probabilities outcomes  $(\theta_1(j), j \leq N)$  (see [4, Section 2.2.3] for a definition of such processes).

Thanks to this last observation, we finally prove Theorem 1.1.

*Proof of Theorem 1.1.* Thinking of  $\Pi^N$  as the *ancestral partition process* of a Canning model, Lemma 2.1 together with [9, Proposition 4.2] yields

$$\{\theta_1(i), i \leq N\} := \left\{ \frac{e^{a\bar{X}_{-1}(i)}}{\sum_{k=1}^N e^{a\bar{X}_{-1}(k)}}, i \leq N \right\} \stackrel{(d)}{=} \left\{ \frac{e^{aE_i}}{\sum_{k=1}^N e^{aE_k}}, i \leq N \right\},$$

where  $(E_j, j \in \mathbb{N})$  i.i.d. exponential random variables with mean 1. Moreover, for any  $y \geq 1$ , we have  $\mathbf{P}(e^{aE_j} \geq y) = y^{-1/a}$ . Therefore, applying [8, Theorem 1.2], we conclude the proof of Theorem 1.1.  $\square$

### 3 The branching Ornstein-Uhlenbeck processes

In this section, we draw a parallel between branching Ornstein-Uhlenbeck processes and the  $(N, a)$ -exponential model we have introduced here. We first recall that an Ornstein-Uhlenbeck process is a continuous time diffusion that solves the stochastic differential equation:

$$dX_t = -\mu X_t dt + \sigma dW_t, \quad (3.1)$$

where  $\mu \geq 0$  is the *pulling strength* of the process,  $\sigma > 0$  is the diffusion coefficient and  $W$  is a standard Brownian motion. Therefore, the branching Ornstein-Uhlenbeck process with parameters  $(\beta, \mu, \sigma)$  is a continuous-time branching process, whose underlying motion is governed by  $X_t$ . It starts with a unique particle at 0. This particle (or individual) evolves according to an Ornstein-Uhlenbeck process with parameters  $(\mu, \sigma)$  for an independent exponential time with mean  $\beta$ . At that time, the individual splits into two children, that start independent branching Ornstein-Uhlenbeck processes from the position of their parents. Up to a space-time linear transforms, we may assume without loss of generality that  $\beta = \sigma = 1$ .

To the best of our knowledge, there is a reduced number of rigorous results about this branching process. The authors in [18, 19] study the behaviour of particles in the bulk and show that a CLT type result holds when the branching rate is weak (or analogously the pulling force is small), whereas a completely different behaviour takes place in the *large branching rate* case. In [21], a branching Ornstein-Uhlenbeck type process with infinite branching rate is introduced. Nevertheless, the behaviour of extreme particles and the genealogy associated to the process with selection remain open. We believe there exists a straight connection between branching Ornstein-Uhlenbeck processes with selection and the  $(N, a)$ -exponential models. In both models particles are subjected to a pulling strength that depends linearly on the position of particles. In view of this connection and the results from the previous section, one can expect that the genealogy of branching Ornstein-Uhlenbeck processes with selection has non-trivial behaviour.

As noticed in [6], the time step  $n \mapsto n + 1$  in the  $(N, 1)$ -exponential model is associated to a  $\mathcal{O}((\log N)^2)$  time step for the associated  $N$ -branching random walk. Hence, as the genealogical tree of the  $(N, 1)$ -exponential model converges in the  $(\log N)$  scale, the genealogical tree of the  $N$ -branching random walk should converge in the  $(\log N)^3$  scale. These conjecture were verified in [3] to the case of branching Brownian motion with quasi-critical absorption. Similarly, a time step in the  $(N, a)$ -exponential model should correspond to a time interval of the order  $(\log N)^2$  in the evolution of the branching Ornstein-Uhlenbeck process with selection. If this connection is precise, it is reasonable to choose a pulling force  $\mu_N$  depending on the size of the system. A reasonable choice would then be  $\gamma/(\log N)^2$  with  $\gamma > 0$  a parameter. We notice that an Ornstein-Uhlenbeck process  $X$  with pulling strength  $\mu_N$  and starting position  $x$  satisfies  $\mathbf{E}(X_{(\log N)^2}) = xe^{-\gamma}$ .

We ran a few simulations that reinforce the above heuristics. We considered discrete-time/space branching-selection particles systems which mimic branching Ornstein-Uhlenbeck processes. The results are displayed in Figure 3. We observe that when the pulling strength has the order  $(\log N)^{-2}$ , then the average coalescent time of two individuals chosen uniformly at random seems to grow polynomial with  $N$ . This type of behaviour is compatible with Beta coalescent. This leads us to the following conjecture.

**Conjecture 3.1.** *Let  $\gamma > 0$  and  $N \in \mathbb{N}$ , there exists a sequence  $(\rho_N)$  which is  $c_\gamma$ -regularly varying, with  $c_\gamma \in [0, 1]$  such that  $(\Pi_{[\rho_N t]}^N, t \geq 0)$  converges toward a  $\text{Beta}(1 - c_\gamma, 1 + c_\gamma)$ -coalescent.*



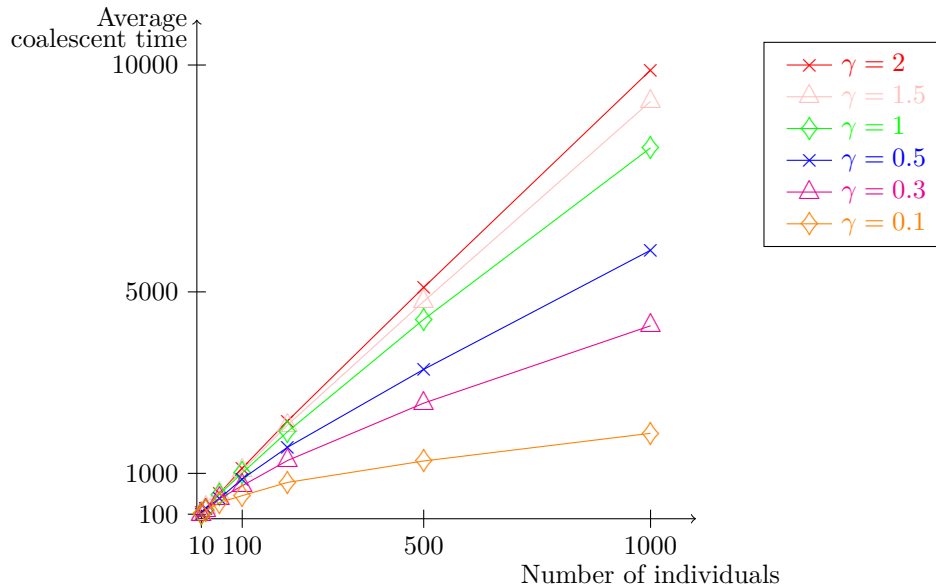


Figure 1: Average age of the most recent common ancestor of two individuals selected at random

Roughly speaking, this conjectures state that a branching Ornstein-Uhlenbeck process with pulling strength  $\gamma(\log N)^{-2}$  and selection of the  $N$  rightmost individuals can be associated to a  $(N, (c_\gamma + 1)^{-1})$ -exponential model. Even though we were not able to push the simulations far enough to guess the function  $c_\gamma$ , it is worth noting that it seems to be close to 1 for  $\gamma > 1$ , and that  $c_{0.1} \approx 0.5$ .

**Acknowledgments:** We would like to thank Julien Berestycki and Éric Brunet for their comments, advices and interest in this work.

The work of A.C. is supported by the Swiss National Science Foundation 200021\_163170.

## References

- [1] J. Bérard and J.-B. Gouéré. Brunet-Derrida behavior of branching-selection particle systems on the line. *Comm. Math. Phys.*, 298(2):323–342, 2010.
- [2] J. Bérard and P. Maillard. The limiting process of  $N$ -particle branching random walk with polynomial tails. *Electron. J. Probab.*, 19:no. 22, 17, 2014.
- [3] J. Berestycki, N. Berestycki, and J. Schweinsberg. The genealogy of branching Brownian motion with absorption. *Ann. Probab.*, 41(2):527–618, 2013.
- [4] N. Berestycki. *Recent progress in coalescent theory*, volume 16 of *Ensaos Matemáticos [Mathematical Surveys]*. Sociedade Brasileira de Matemática, Rio de Janeiro, 2009.
- [5] E. Brunet and B. Derrida. Shift in the velocity of a front due to a cut-off. *Phys. Rev. E*, 56:2597–2604, 1997.
- [6] E. Brunet, B. Derrida, A.H. Mueller, and S. Munier. Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization. *Phys. Rev. E*, 76(4):041104, 2007.
- [7] F. Comets and A. Cortines. Finite-size corrections to the speed of a branching-selection process. *Braz. J. Probab. Stat.*, 31(3):476–501, 2017.
- [8] A. Cortines. The genealogy of a solvable population model under selection with dynamics related to directed polymers. *Bernoulli*, 22(4):2209–2236, 2016.

- [9] A. Cortines and B. Mallein. A  $N$ -branching random walk with random selection. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(1):117–137, 2017.
- [10] O. Couroonné and L. Gerin. A branching-selection process related to censored Galton-Walton processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 50(1):84–94, 2014.
- [11] A. De Masi, P.A. Ferrari, E. Presutti, and N. Soprano-Loto. Non-local branching Brownians with annihilation and free boundary problems. 2017. arXiv:1711.06390.
- [12] R. Durrett and D. Remenik. Brunet-Derrida particle systems, free boundary problems and Wiener-Hopf equations. *Ann. Probab.*, 39(6):2043–2078, 2011.
- [13] P. Maillard. Speed and fluctuations of  $N$ -particle branching Brownian motion with spatial selection. *Probab. Theory Related Fields*, 166(3-4):1061–1173, 2016.
- [14] B. Mallein. Branching random walk with selection at critical rate. *Bernoulli*, 23(3):1784–1821, 2017.
- [15] B. Mallein.  $N$ -branching random walk with  $\alpha$ -stable spine. *Teor. Veroyatn. Primen.*, 62(2):365–392, 2017.
- [16] M. Möhle and S. Sagitov. A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.*, 29(4):1547–1562, 2001.
- [17] M. Pain. Velocity of the  $L$ -branching Brownian motion. *Electron. J. Probab.*, 21:Paper No. 28, 28, 2016.
- [18] A. Radosław and P. Miłoś.  $U$ -statistics of Ornstein-Uhlenbeck branching particle system. *J. Theoret. Probab.*, 27(4):1071–1111, 2014.
- [19] A. Radosław and P. Miłoś. CLT for Ornstein-Uhlenbeck branching particle system. *Electron. J. Probab.*, 20:no. 42, 35, 2015.
- [20] J. Schweinsberg. Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Process. Appl.*, 106(1):107–139, 2003.
- [21] Q. Shi. A growth-fragmentation model related to Ornstein-Uhlenbeck type processes. 2017. arXiv:1702.01091.