



## A $N$ -branching random walk with random selection

Aser Cortines and Bastien Mallein

Faculty of Industrial Engineering and Management,  
Technion - Israel Institute of Technology,  
Technion City,  
Haifa 32000, Israel.

*E-mail address:* [cortines@technion.ac.il](mailto:cortines@technion.ac.il)

*URL:* <https://web.iem.technion.ac.il/en/people/userprofile/cortines.html>

Institut für Mathematik, Universität Zürich  
Winterthurststrasse 190,  
CH-8057, Zrich, Swiss.

*E-mail address:* [bastien.mallein@math.uzh.ch](mailto:bastien.mallein@math.uzh.ch)

*URL:* <http://www.math.ens.fr/~mallein>

**Abstract.** We consider an exactly solvable model of branching random walk with random selection, which describes the evolution of a population with  $N$  individuals on the real line. At each time step, every individual reproduces independently, and its offspring are positioned around its current locations. Among all children,  $N$  individuals are sampled at random without replacement to form the next generation, such that an individual at position  $x$  is chosen with probability proportional to  $e^{\beta x}$ . We compute the asymptotic speed and the genealogical behavior of the system.

### 1. Introduction

In a general sense, a *branching-selection particle system* is a Markovian process of particles on the real line evolving through the repeated application of the two steps:

**Branching step:** every individual currently alive in the system splits into new particles, with positions (with respect to their birth place) given by independent copies of a point process.

**Selection step:** some of the new-born individuals are selected to reproduce at the next *branching step*, while the other particles are “killed”.

---

*Received by the editors May 13, 2016; accepted January 4, 2017.*

*2010 Mathematics Subject Classification.* 60J80, 60K35.

*Key words and phrases.* Branching random walk, Coalescent process, Selection, Poisson-Dirichlet partition.

Research partially supported by the Israeli Science Foundation grant 1723/14.

We will often see the particles as individuals and their positions as their *fitness*, that is, their score of adaptation to the environment. From a biological perspective, branching-selection particle systems model the competition between individuals in an environment with limited resources.

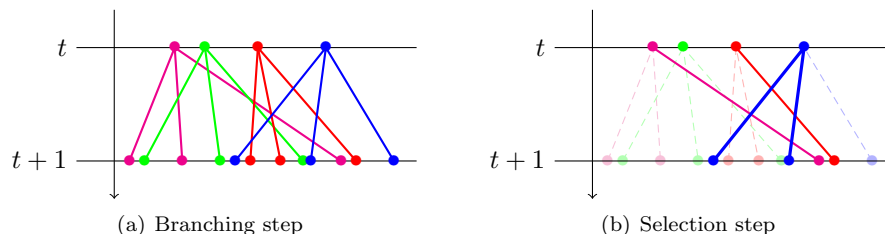


FIGURE 1.1. One time step of a branching-selection particle system

These models are of physical interest (Brunet and Derrida, 1997; Brunet et al., 2007) and can be related to reaction-diffusion phenomena and the F-KPP equation. Different methods can be used to select the individuals. For example, one can consider an *absorbing barrier*, below which particles are killed (Aïdékon and Jaffuel, 2011; Berestycki et al., 2013; Mallein, 2015a; Pain, 2016). Another example is the case where only the  $N$  rightmost individuals are chosen to survive (Brunet and Derrida, 1997; Brunet et al., 2007; Bérard and Gouéré, 2010; Durrett and Remenik, 2011), the so-called “ $N$ -branching random walk”. In this paper, we introduce a new selection mechanism, in which the individuals are randomly selected with probability depending on their positions.

Based on numerical simulations (Brunet and Derrida, 1997) and the study of solvable models (Brunet et al., 2007), it has been predicted that the dynamical and structural aspects of many branching selection particle systems satisfy universal properties. For example, the cloud of particles travels at speed  $v_N$ , which converges to a limit  $v$  as the size of the population  $N$  diverges. It has been conjectured in Brunet et al. (2007) that

$$v_N - v = -\varphi(\log N + 3 \log \log N + o(\log \log N))^{-2} \quad \text{as } N \rightarrow \infty,$$

for an explicit constant  $\varphi$  depending on the law of reproduction.

Some of these conjectures have been recently proved for the  $N$ -branching random walk (Bérard and Maillard, 2014; Bérard and Gouéré, 2010; Couronné and Gerin, 2014; Durrett and Remenik, 2011; Maillard, 2016; Mallein, 2015b). Bérard and Gouéré (2010) prove that  $v_N - v$  behaves like  $-\varphi(\log N)^{-2}$ . Nevertheless, several conjectures about this process remain open, such as the asymptotic behavior of the genealogy or the second-order expansion of the speed. Other examples in which the finite-size correction to the speed of a branching-selection particle system is explicitly computed can be found in Bérard and Maillard (2014); Comets and Cortines (2015); Couronné and Gerin (2014); Durrett and Remenik (2011); Mallein (2015b); Mueller et al. (2011).

To study the genealogical structure of such models we define the *ancestral partition process*  $\Pi_n^N(t)$  of a population choosing  $n \ll N$  individuals from a given generation  $T$  and tracing back their genealogical lineages. That is,  $\Pi_n^N(t)$  is a process in  $\mathcal{P}_n$  the set of partitions (or equivalence classes) of  $[n] := \{1, \dots, n\}$  such

that  $i$  and  $j$  belong to the same equivalence class if the individuals  $i$  and  $j$  have a common ancestor  $t$  generations backwards in time. Notice that the direction of time is the opposite of the direction of time for the natural evolution of the population, that is,  $t = 0$  is the current generation,  $t = 1$  brings us one generation backward in time and so on.

It has also been conjectured in [Brunet et al. \(2007\)](#) that the genealogical trees of branching selection particle systems converge to those of a Bolthausen-Sznitman coalescent and that the average coalescence times scale like a power of the logarithm of the population size. These conjectures contrast with classical results in neutral population models, such as Wright-Fisher and Moran's models, that lay in the Kingman coalescent universality class ([Möhle and Sagitov, 2001](#)). Mathematically, these conjectures are difficult to be verified and they have only been proved for some particular models, see [Berestycki et al. \(2013\)](#); [Cortines \(2016\)](#).

We define in this article a solvable model of branching selection particle system evolving in discrete time, and compute its asymptotic speed as well as its genealogical structure. Given  $N \in \mathbb{N}$  and  $\beta > 1$ , it consists in a population with a fixed number  $N$  of individuals. At each time step, the individuals die giving birth to offspring that are positioned according to independent Poisson point processes with intensity  $e^{-x} dx$  (that we write  $\text{PPP}(e^{-x} dx)$  for short). Then,  $N$  individuals are sampled (without replacement) to form the next generation, such that a child at the position  $x$  is sampled with probability proportional to  $e^{\beta x}$ .

We introduce the following notation. Let  $X_0^N(1), \dots, X_0^N(N) \in \mathbb{R}$  be the initial position of the particles and  $\{\mathcal{P}_t(j), j \leq N, t \in \mathbb{N}\}$  be a family of i.i.d.  $\text{PPP}(e^{-x} dx)$ . Given  $t \geq 1$  and  $X_{t-1}^N(1), \dots, X_{t-1}^N(N)$  the  $N$  positions at time  $t-1$ , we define the new positions as follows:

- i. Each individual  $X_{t-1}^N(j)$  gives birth to infinitely many children that are positioned according to the point process  $X_{t-1}^N(j) + \mathcal{P}_t(j)$ . Let  $\Delta_t := (\Delta_t(k); k \in \mathbb{N})$  be the sequence obtained by all positions ranked decreasingly, that is

$$(\Delta_t(k), k \in \mathbb{N}) = \text{Rank}(\{X_{t-1}^N(j) + p; p \in \mathcal{P}_t(j), j \leq N\}).$$

- ii. We sample successively  $N$  individuals  $X_t^N(1), \dots, X_t^N(N)$  composing the  $t$ th generation from  $\{\Delta_t(1), \Delta_t(2), \dots\}$  such that for all  $i \in \{1, 2, \dots, N\}$ :

$$\begin{aligned} \mathbf{P}(X_t^N(i) = \Delta_t(j) \mid \Delta_t, X_t^N(1), \dots, X_t^N(i-1)) \\ = \frac{e^{\beta \Delta_t(j)} \mathbf{1}_{\{\Delta_t(j) \notin \{X_t^N(1), \dots, X_t^N(i-1)\}\}}}{\sum_{k=1}^{+\infty} e^{\beta \Delta_t(k)} - \sum_{k=1}^{i-1} e^{\beta X_t^N(k)}}. \end{aligned} \quad (1.1)$$

To keep track of the genealogy of the process we define

$$A_t^N(i) = j \quad \text{if } X_t^N(i) \in \{X_{t-1}^N(j) + p, p \in \mathcal{P}_t(j)\}, \quad (1.2)$$

that is,  $A_n(i) = j$  if  $X_n^N(i)$  is an offspring of  $X_{n-1}^N(j)$ . We call this system the  $(N, \beta)$ -branching random walk or  $(N, \beta)$ -BRW for short.

It can be checked that the sum in the denominator of (1.1) is finite if  $\beta \in (1, \infty)$  and that it diverges as  $\beta \rightarrow 1$  (see [Proposition 1.3](#) below), thus the model is only defined for  $\beta \in (1, \infty)$ . Notice that as  $\beta \rightarrow \infty$  the sum in the denominator is dominated by the high values of  $\Delta_t$ . Precisely, it can be checked that the following limits hold a.s.

$$\lim_{\beta \rightarrow \infty} e^{-\beta \Delta_t(1)} \sum_{k=1}^{+\infty} e^{\beta \Delta_t(k)} = 1, \quad \lim_{\beta \rightarrow \infty} e^{-\beta \Delta_t(2)} \sum_{k=2}^{+\infty} e^{\beta \Delta_t(k)} = 1,$$

and so on. Therefore, the case “ $\beta = \infty$ ” is the “exponential model” from Brunet and Derrida (1997, 2012); Brunet et al. (2007), in which the  $N$  rightmost individuals are selected to form the next generation. In contrast with the examples already treated in the literature, when  $\beta < \infty$  one does not necessarily select the rightmost offspring. In this paper, we will take interest in the dynamical and genealogical aspects of the  $(N, \beta)$ -BRW, showing that it travels at a deterministic speed and that its genealogical trees converge in distribution. The next result concerns the speed of the  $(N, \beta)$ -BRW.

**Theorem 1.1.** *For all  $N \in \mathbb{N}$  and  $\beta \in (1, \infty]$ , there exists  $v_{N, \beta}$  such that*

$$\lim_{t \rightarrow +\infty} \frac{\max_{j \leq N} X_t^N(j)}{t} = \lim_{n \rightarrow +\infty} \frac{\min_{j \leq N} X_t^N(j)}{t} = v_{N, \beta} \quad a.s. \quad (1.3)$$

moreover,  $v_{N, \beta} = \log \log N + o(1)$  as  $N \rightarrow \infty$ .

The main result of the paper is the following theorem concerning the convergence in law of the ancestral partition process  $(\Pi_n^N(t); t \in \mathbb{N})$  of the  $(N, \beta)$ -BRW.

**Theorem 1.2.** *For all  $N \in \mathbb{N}$  and  $\beta \in (1, \infty]$ , let  $c_N$  be the probability that two individuals uniformly chosen at random have a common ancestor one generation backwards in time. Then, we have  $\lim_{N \rightarrow \infty} c_N \log N = 1$  and the rescaled coalescent process  $(\Pi^N(\lfloor t/c_N \rfloor), t \geq 0)$  converges in distribution toward the Bolthausen-Sznitman coalescent.*

The Bolthausen-Sznitman coalescent in Theorem 1.2 can be roughly explained by an individual going far ahead of the rest of the population, so that its offspring are more likely to be selected and overrun the next generation. Based on precise asymptotic of the coalescence time, Brunet et al. (2007) argue that the genealogical trees of the exponential model converge to the Bolthausen-Sznitman coalescent and conjecture that this behavior should be expected for a large class of models. The  $(N, \beta)$ -BRW can be thought as a finite temperature version of the exponential model from Brunet et al. (2007). In this sense, Theorem 1.2 attests for the robustness of their conjectures showing that even under weaker selection constrains this convergence occurs. It indicates that whenever the rightmost particles are likely to be selected, then the Bolthausen-Sznitman coalescent is to be expected.

Different coalescent behavior should be expected when the selection mechanism does not favor the rightmost particles, the classical example being the Wright-Fisher model. Another example can be obtained modifying the selection mechanism of the  $(N, \beta)$ -BRW. It can be checked using the techniques developed in this paper (see for example Theorem 3.3) that if we systematically eliminate the first individual sampled  $X_1^N(t)$ , so that it does not reproduce in the next generation, then this new branching-selection particle system lays in the Kingman’s coalescent universality class. Notice that this new selection procedure no longer favors the rightmost particles (in this case, the rightmost particle), which justifies this change of behavior.

**Notation.** In this article, we write

$$\begin{aligned}
 f(x) \sim g(x) \quad \text{as } x \rightarrow a \text{ if } \quad & \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1; \\
 f(x) = o(g(x)) \text{ as } x \rightarrow a \text{ if } \quad & \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0; \\
 \text{and } f(x) = \mathcal{O}(g(x)) \text{ as } x \rightarrow a \text{ if } \quad & \limsup_{x \rightarrow a} \frac{f(x)}{g(x)} < +\infty.
 \end{aligned}$$

1.1. *Preliminary results.* In this section, we prove that the  $(N, \beta)$ -BRW is well defined and provide elementary properties such as the existence of the speed  $v_{N, \beta}$ .

**Proposition 1.3.** *The  $(N, \beta)$ -BRW is well-defined for all  $N \in \mathbb{N}$  and  $\beta \in (1, \infty]$ . Moreover, setting  $X_t^N(\text{eq}) := \log \sum_{j=1}^N e^{X_t^N(j)}$ , the sequence*

$$\left( \sum_{k \in \mathbb{N}} \delta_{\Delta_k(t+1) - X_t^N(\text{eq})} : t \in \mathbb{N} \right)$$

*is an i.i.d. family of Poisson point processes with intensity measure  $e^{-x} dx$ .*

*Proof:* With  $N$  and  $\beta$  fixed, assume that the process has been constructed up to time  $t$  with  $X_t^N(1), \dots, X_t^N(N)$  denoting the positions of the  $N$  particles. Thanks to the invariance of superposition of independent PPP,  $\{X_t^N(j) + p; p \in \mathcal{P}_t(j), j \leq N\}$  is also a PPP with intensity measure

$$\sum_{i=1}^N e^{-(x - X_t^N(i))} dx = e^{-(x - X_t^N(\text{eq}))} dx.$$

Therefore, with probability one: all points have multiplicity one and the sequence  $(\Delta_k(t+1); k \in \mathbb{N})$  is uniquely defined. Since there are finitely many points  $\Delta_k(t+1)$  that are positive and  $\mathbf{E} \left( \sum e^{\beta \Delta_k(t+1)} \mathbf{1}_{\{\Delta_k(t+1) < 0\}} \right) < \infty$  we have that  $\sum e^{\beta \Delta_k(t+1)} < \infty$  a.s. As a consequence, the selection step is well-defined, proving the first claim. Moreover,  $(\Delta_k(t+1) - X_t^N(\text{eq}); k \in \mathbb{N})$  is a PPP( $e^{-x} dx$ ) independent from the  $t$  first steps of the  $(N, \beta)$ -BRW, proving the second claim.  $\square$

*Remark 1.4.* It is convenient to think  $X_t^N(\text{eq})$  as an “equivalent position” of the front at time  $t$ , in the sense that the particles positions in the  $(t+1)$ th generation are distributed as if they were generated by a unique individual positioned at  $X_t^N(\text{eq})$ .

We use Proposition 1.3 to prove the existence of the speed  $v_{N, \beta}$ , the study of its asymptotic behavior is postponed to Section 4.

**Lemma 1.5.** *With the notation of the previous proposition, (1.3) in Theorem 1.1 holds with*

$$v_{N, \beta} := \mathbf{E}(X_1^N(\text{eq}) - X_0^N(\text{eq})).$$

*Proof:* By Proposition 1.3,  $(X_{t+1}^N(\text{eq}) - X_t^N(\text{eq}) : t \in \mathbb{N})$  are i.i.d. random variables with finite mean, therefore  $\lim_{t \rightarrow +\infty} \frac{X_t^N(\text{eq})}{t} = v_{N, \beta}$  a.s. by the law of large numbers. Notice that both  $(\max X_t^N(j) - X_{t-1}^N(\text{eq}))_t$  and  $(\min X_t^N(j) - X_{t-1}^N(\text{eq}))_t$  are sequences of i.i.d. random variables with finite mean, which yields (1.3).  $\square$

In a similar way, we are able to obtain a simple structure for the genealogy of the process, and describe its law conditionally on the position of the particles.

**Lemma 1.6.** *The sequence  $(A_t^N)_{t \in \mathbb{N}}$  defined in (1.2) is i.i.d. Moreover, it remains independent conditionally on  $\mathcal{H} = \sigma(X_t^N(j), j \leq N, t \geq 0)$ , with the conditional probabilities*

$$\mathbf{P}(A_{t+1}^N = \bar{k} \mid \mathcal{H}) = \theta_t^N(k_1) \dots \theta_t^N(k_N), \quad \text{where } \bar{k} = (k_1, \dots, k_N) \in \{1, \dots, N\}^N;$$

$$\text{and } \theta_t^N(k) := \frac{e^{X_t^N(k)}}{\sum_{i=1}^N e^{X_t^N(i)}}. \quad (1.4)$$

*Proof:* To each point  $x \in \sum \delta_{\Delta_{t+1}(k) - X_t^N(\text{eq})}$  we associate the mark  $i$  if it is a point coming from  $X_t^N(i) + \mathcal{P}_t(i)$ . The invariance under superposition of independent PPP yields  $\mathbf{P}(x \in X_t^N(i) + \mathcal{P}_t(i) \mid \mathcal{H}) = \frac{e^{-(x - X_t^N(i))}}{\sum_{j=1}^N e^{-(x - X_t^N(j))}} = \frac{e^{X_t^N(i)}}{\sum_{j=1}^N e^{X_t^N(j)}}$ . By definition of  $A_t^N(i)$ , it is precisely the mark of  $X_{t+1}^N(i)$ , which yields (1.4). The independence between the  $A_t^N$  can be easily checked using Proposition 1.3.  $\square$

**Organization of the paper.** In Section 2, we obtain some technical lemmas concerning the Poisson-Dirichlet distributions. We focus in Section 3 on a class of coalescent processes generated by Poisson Dirichlet distributions and we prove a convergence criterion. Finally, in Section 4, we provide an alternative construction of the  $(N, \beta)$ -BRW in terms of a Poisson-Dirichlet distribution, and we use the results obtained in the previous sections to prove Theorems 1.1 and 1.2.

## 2. Poisson-Dirichlet distribution

In this section, we focus on the two-parameter Poisson-Dirichlet distribution denoted as PD( $\alpha, \theta$ ) distribution.

**Definition 2.1** (Definition 1 in Pitman and Yor, 1997). For  $\alpha \in (0, 1)$  and  $\theta > -\alpha$ , let  $(Y_j : j \in \mathbb{N})$  be a family of independent r.v. such that  $Y_j$  has Beta( $1 - \alpha, \theta + j\alpha$ ) distribution and write

$$V_1 = Y_1, \quad \text{and} \quad V_j = \prod_{i=1}^{j-1} (1 - Y_i) Y_j, \quad \text{if } j \geq 2.$$

Let  $U_1 \geq U_2 \geq \dots$  be the ranked values of  $(V_n)$ , we say that the sequence  $(U_n)$  is the *Poisson-Dirichlet distribution with parameters*  $(\alpha, \theta)$ .

Notice that for any  $k \in \mathbb{N}$  and  $n \in \mathbb{N}$ , we have

$$\mathbf{P}(V_n = U_k \mid (U_j, j \in \mathbb{N}), V_1, \dots, V_{n-1}) = \frac{U_k \mathbf{1}_{\{U_k \notin \{V_1, \dots, V_{n-1}\}\}}}{1 - V_1 - V_2 - \dots - V_{n-1}},$$

for this reason we say that  $(V_n)$  follows the *size-biased pick* from a PD( $\alpha, \theta$ ). It is well known that there exists a strong connexion between PD distributions and PPP, we recall some of these results in the proposition below.

**Proposition 2.2** (Proposition 10 in Pitman and Yor, 1997). *Let  $x_1 > x_2 > \dots$  be the points of a PPP( $e^{-x} dx$ ) and write  $L = \sum_{j=1}^{+\infty} e^{\beta x_j}$  and  $U_j = e^{\beta x_j} / L$ . Then  $(U_j, j \geq 1)$  has PD( $\beta^{-1}, 0$ ) distribution and  $\lim_{n \rightarrow +\infty} n^\beta U_n = 1/L$  a.s.*

Notice from Propositions 1.3 and 2.2 that  $(e^{\beta X_t^N(i)} / \sum_{j=1}^{+\infty} e^{\beta \Delta_n(j)})$  has the distribution of  $(V_i)$  the size-biased pick from PD( $\beta^{-1}, 0$ ), which makes the model solvable.

*Remark 2.3* (Change of parameter). If  $V_1, V_2, \dots$  is a *size-biased pick* of  $\text{PD}(\alpha, \theta)$ , then  $\left(\frac{V_2}{1-V_1}, \frac{V_3}{1-V_1}, \dots\right)$  has the distribution of a size-biased pick from a  $\text{PD}(\alpha, \alpha + \theta)$ . Moreover, it is independent of  $V_1$ . That is, the sequence obtained from  $V_1, V_2, \dots$  after discarding the first sampled element  $V_1$  and re-normalizing is a *size-biased pick* from a  $\text{PD}(\alpha, \alpha + \theta)$ . Therefore, ordering this new sequence one obtains a  $\text{PD}(\alpha, \alpha + \theta)$  sequence.

In what follows, we fix  $\alpha \in (0, 1)$  and  $\theta > -\alpha$ , and let  $c$  and  $C$  be positive constants, that may change from line to line and implicitly depend on  $\alpha$  and  $\theta$ . We will focus attention on the convergence and the concentration properties of

$$\Sigma_n := \sum_{j=1}^n V_j^\alpha = \sum_{j=1}^n Y_j^\alpha \prod_{i=1}^{j-1} (1 - Y_i)^\alpha; \quad n \in \mathbb{N}. \quad (2.1)$$

**Lemma 2.4.** *Set  $M_n := \prod_{i=1}^n (1 - Y_i)$ , then, there exists a positive r.v.  $M_\infty$  such that*

$$\lim_{n \rightarrow +\infty} \left( n^{\frac{1-\alpha}{\alpha}} M_n \right)^\gamma = M_\infty^\gamma \quad \text{a.s. and in } \mathbb{L}^1 \text{ for all } \gamma > -(\theta + \alpha), \quad (2.2)$$

with  $\gamma$ -moment verifying  $\mathbf{E}(M_\infty^\gamma) = \Phi_{\theta, \alpha}(\gamma) := \alpha^\gamma \frac{\Gamma(\theta+1)\Gamma(\frac{\theta+\gamma}{\alpha}+1)}{\Gamma(\theta+\gamma+1)\Gamma(\frac{\theta}{\alpha}+1)}$ . Moreover, if  $0 < \gamma < \theta + \alpha$ , then there exists  $C_\gamma > 0$  such that

$$\mathbf{P} \left( \inf_{n \geq 0} n^{\frac{1-\alpha}{\alpha}} M_n \leq y \right) \leq C_\gamma y^\gamma, \quad \text{for all } n \geq 1 \text{ and } y \geq 0. \quad (2.3)$$

Notice that if  $\gamma > -\theta$ , then  $\Phi_{\theta, \alpha}(\gamma) = \alpha^\gamma \frac{\Gamma(\theta)\Gamma(\frac{\theta+\gamma}{\alpha})}{\Gamma(\theta+\gamma)\Gamma(\frac{\theta}{\alpha})}$ .

*Proof:* Fix  $\gamma > -(\theta + \alpha)$ , then  $(M_n^\gamma / \mathbf{E}(M_n^\gamma))$  is a non-negative martingale with respect to its natural filtration and

$$\begin{aligned} \mathbf{E}(M_n^\gamma) &= \frac{\Gamma(\theta + \gamma + n\alpha)}{\Gamma(\theta + n\alpha)} \frac{\Gamma(n + \frac{\theta}{\alpha})}{\Gamma(n + \frac{\theta+\gamma}{\alpha})} \frac{\Gamma(\theta + 1)\Gamma(\frac{\theta+\gamma}{\alpha} + 1)}{\Gamma(\theta + \gamma + 1)\Gamma(\frac{\theta}{\alpha} + 1)} \\ &\sim \Phi_{\theta, \alpha}(\gamma) n^{-\gamma \frac{1-\alpha}{\alpha}}, \quad \text{as } n \rightarrow +\infty. \end{aligned}$$

Since  $\lim_{n \rightarrow +\infty} \mathbf{E}(M_n^{\gamma/2}) \mathbf{E}(M_n^\gamma)^{-1/2} > 0$ , Kakutani's theorem yields  $M_n^\gamma / \mathbf{E}(M_n^\gamma)$  converges a.s. and in  $\mathbb{L}^1$  as  $n \rightarrow +\infty$ , implying (2.2) with  $M_\infty = \lim_{n \rightarrow +\infty} M_n n^{\frac{1-\alpha}{\alpha}}$ .

In particular, if  $0 < \gamma < \theta + \alpha$  we obtain from Doob's martingale inequality that

$$\begin{aligned} \mathbf{P} \left( \inf_{n \geq 0} n^{\frac{1-\alpha}{\alpha}} M_n \leq y \right) &= \mathbf{P} \left( \sup_{n \geq 0} M_n^{-\gamma} n^{-\gamma \frac{1-\alpha}{\alpha}} \geq y^{-\gamma} \right) \\ &\leq \mathbf{P} \left( \sup_{n \geq 0} \frac{M_n^{-\gamma}}{\mathbf{E}(M_n^{-\gamma})} \geq y^{-\gamma} / C_\gamma \right), \end{aligned}$$

with  $C_\gamma := \sup_{n \in \mathbb{N}} \mathbf{E}[M_n^{-\gamma}] / n^{\gamma \frac{1-\alpha}{\alpha}} < \infty$ , proving (2.3).  $\square$

We now focus on the convergence of the series  $\sum Y_j^\alpha j^{\alpha-1}$ .

**Lemma 2.5.** *Let  $S_n := \sum_{j=1}^n Y_j^\alpha j^{\alpha-1}$  and  $\Psi_\alpha := \alpha^{-\alpha} \Gamma(1 - \alpha)^{-1}$ , then, there exists a random variable  $S_\infty$  such that*

$$\lim_{n \rightarrow +\infty} S_n - \Psi_\alpha \log n = S_\infty \quad \text{a.s.}$$

Moreover, there exists  $C > 0$  such that for all  $n \in \mathbb{N}$  and  $y \geq 0$ ,

$$\mathbf{P}(|S_n - \mathbf{E}(S_n)| \geq y) \leq Ce^{-y^{2-\alpha}}.$$

*Proof:* Since  $Y_j$  has Beta( $1 - \alpha, \theta + j\alpha$ ) distribution, we have

$$\begin{aligned} \mathbf{E}((jY_j)^\alpha) &= \frac{1}{\alpha^\alpha \Gamma(1 - \alpha)} + \mathcal{O}(1/j) \\ \text{and } \mathbf{Var}((jY_j)^\alpha) &= \frac{\Gamma(1 + \alpha)\Gamma(1 - \alpha) - 1}{\alpha^{2\alpha}\Gamma(1 - \alpha)^2} + \mathcal{O}(1/j), \end{aligned}$$

which implies that  $\sum \mathbf{Var}(Y_j^{\alpha-1}) < +\infty$  and that  $\mathbf{E}(S_n) = \Psi_\alpha \log n + C_S + o(1)$  with  $C_S \in \mathbb{R}$ . Thanks to  $Y_j^\alpha - \mathbf{E}(Y_j^\alpha) \in (-1, 1)$  a.s. we deduce from Kolmogorov's three-series theorem that  $S_n - \mathbf{E}(S_n)$  and hence that  $S_n - \Psi_\alpha \log n$  converge a.s. To bound  $\mathbf{P}(S_n - \mathbf{E}(S_n) \geq y)$ , notice that

$$\mathbf{P}(S_n - \mathbf{E}(S_n) \geq y) \leq e^{-\lambda y} \mathbf{E} \left[ e^{\lambda(S_n - \mathbf{E}(S_n))} \right] \leq e^{-\lambda y} \prod_{j=1}^n \mathbf{E} \left( e^{\lambda j^{\alpha-1}(Y_j^\alpha - \mathbf{E}(Y_j^\alpha))} \right),$$

for all  $y \geq 0$  and  $\lambda > 0$ . Taking  $c > 0$  such that  $e^x \leq 1 + x + cx^2$  for  $x \in (-1, 1)$ , we obtain

$$\mathbf{E} \left[ e^{\lambda j^{\alpha-1}(Y_j^\alpha - \mathbf{E}(Y_j^\alpha))} \right] \leq \begin{cases} e^{\lambda j^{\alpha-1}} & \text{if } \lambda j^{\alpha-1} > 1; \\ 1 + c\lambda^2 j^{2(\alpha-1)} \mathbf{Var}(Y_j^\alpha) & \text{if } \lambda j^{\alpha-1} \leq 1. \end{cases}$$

Since  $\sum_{j^{1-\alpha} \leq \lambda} j^{\alpha-1} < \lambda^{\frac{1}{1-\alpha}}$  for all  $\alpha \in (0, 1)$  and  $\theta \geq 0$ , there exists  $c = c(\alpha, \theta)$  such that

$$\begin{aligned} \mathbf{P}(S_n - \mathbf{E}(S_n) \geq y) &\leq e^{-\lambda y} \prod_{j^{1-\alpha} \leq \lambda} e^{\lambda j^{\alpha-1}} \times \prod_{j^{1-\alpha} > \lambda} \left( 1 + c \frac{\lambda^2}{j^2} \right) \\ &\leq \exp \left( -\lambda y + \lambda^{\frac{1}{1-\alpha}} + c\lambda^2 \right), \quad \text{for all } n \in \mathbb{N} \text{ and } y \geq 0. \end{aligned}$$

Let  $\varrho := (2 - \alpha)/(1 - \alpha) > 2$ , then there exists  $C = C(\alpha, \theta) > 0$  such that

$$\mathbf{P}(S_n - \mathbf{E}(S_n) \geq y) \leq C \exp(-\lambda y + C\lambda^\varrho), \quad \text{for all } n \in \mathbb{N} \text{ and } y \geq 0.$$

Optimizing in  $\lambda > 0$  we obtain

$$\mathbf{P}(S_n - \mathbf{E}(S_n) \geq y) \leq C \exp \left( -y^{\varrho/(\varrho-1)} C^{1/(1-\varrho)} \left( \varrho^{1/(1-\varrho)} - \varrho^{\varrho/(1-\varrho)} \right) \right),$$

with  $C^{1/(1-\varrho)} [\varrho^{1/(1-\varrho)} - \varrho^{\varrho/(1-\varrho)}] > 0$ , as  $\varrho > 1$ . The same argument, with the obvious changes, holds for  $\mathbf{P}(S_n - \mathbf{E}(S_n) \leq -y)$ , therefore, there exists  $C > 0$  such that  $\mathbf{P}(|S_n - \mathbf{E}(S_n)| \geq y) \leq C \exp(-y^{\varrho/(\varrho-1)}/C)$ , proving the second statement.  $\square$

With the above results, we obtain the convergence of  $\Sigma_n = \sum V_j^\alpha$  as well as its tail probabilities.

**Lemma 2.6.** *With the notation of Lemmas 2.4 and 2.5, we have*

$$\lim_{n \rightarrow +\infty} \frac{\Sigma_n}{\log n} = \Psi_\alpha M_\infty^\alpha \quad \text{a.s. and in } \mathbb{L}^1.$$

Moreover, for any  $0 < \gamma < \alpha + \theta$  there exists  $D_\gamma$  such that for all  $n \geq 1$  large enough and  $u > 0$  we have

$$\mathbf{P}(\Sigma_n \leq u \log n) \leq D_\gamma u^{\frac{\gamma}{\alpha}}.$$



*Proof:* Notice that  $\Sigma_n = \sum_{j=1}^n (S_j - S_{j-1})j^{1-\alpha}M_{j-1}^\alpha$  (since  $S_0 := 0$ ) and that

$$\lim_{n \rightarrow +\infty} (M_n(n+1)^{\frac{1-\alpha}{\alpha}})^\alpha = M_\infty^\alpha \quad \text{and} \quad \lim_{n \rightarrow +\infty} \frac{S_n}{\log n} = \Psi_\alpha \quad \text{a.s.}$$

by Lemmas 2.4 and 2.5 respectively. Then, Stolz-Cesàro theorem yields the a.s. convergence of  $\Sigma_n/\log n$  to  $\Psi_\alpha M_\infty^\alpha$  as  $n \rightarrow \infty$ . Expanding  $(\Sigma_n)^2$  we obtain

$$\begin{aligned} \mathbf{E}(\Sigma_n^2) &= \sum_{j=1}^n \mathbf{E}(V_j^{2\alpha}) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \mathbf{E}((V_i V_j)^\alpha) \\ &= \sum_{j=1}^n \mathbf{E}(M_{j-1}^{2\alpha}) \mathbf{E}(Y_j^{2\alpha}) \\ &\quad + 2 \sum_{i=1}^{n-1} \mathbf{E}(M_{i-1}^{2\alpha}) \mathbf{E}((Y_i(1-Y_i))^\alpha) \sum_{j=i+1}^n \mathbf{E}\left(\frac{M_{j-1}^\alpha}{M_i^\alpha}\right) \mathbf{E}(Y_j^\alpha) \\ &\leq C \sum_{j=1}^n j^{-2(1-\alpha)} j^{-2\alpha} + C \sum_{i=1}^{n-1} i^{-2(1-\alpha)} i^{-\alpha} \sum_{j=i+1}^n \frac{j^{-(1-\alpha)}}{i^{-(1-\alpha)}} j^{-\alpha} \leq C(\log n)^2. \end{aligned}$$

Therefore,  $\sup \mathbf{E}[(\Sigma_n/\log n)^2] < +\infty$  implying its  $\mathbb{L}^1$  convergence. To obtain bounds for  $\mathbf{P}(\Sigma_n \leq u \log n)$ , we study the two cases  $u \geq 1/n$  and  $u \leq 1/n$  separately. Assume first that  $u \geq 1/n$ , then

$$\Sigma_n = \sum_{j=1}^n \left( (j-1)^{\frac{1-\alpha}{\alpha}} M_{j-1} \right)^\alpha \frac{(jY_j)^\alpha}{j} \geq \left( \inf_{j \in \mathbb{N}} j^{\frac{1-\alpha}{\alpha}} M_j \right)^\alpha S_n.$$

For all  $\gamma' < \theta + \alpha$  and  $t > 0$  such that  $t < \mathbf{E}[S_n]$  we have

$$\begin{aligned} \mathbf{P}(\Sigma_n \leq u \log n) &\leq \mathbf{P}(S_n \leq t) + \mathbf{P}\left(\left(\inf j^{\frac{1-\alpha}{\alpha}} M_j\right)^\alpha \leq (u \log n)/t\right) \\ &\leq C \exp\left(-C^{-1}(\mathbf{E}[S_n] - t)^{\frac{\theta}{\theta-1}}\right) + C_{\gamma'} \left(\frac{u \log n}{t}\right)^{\gamma'/\alpha}. \end{aligned}$$

Let  $0 < \varepsilon < 1/2$  and set  $t = u^\varepsilon \log n$ , since  $\lim_{n \rightarrow +\infty} \frac{\mathbf{E}(S_n)}{\log n} = \Psi_\alpha$ , there exists a constant  $c > 0$  depending only on  $\alpha$  such that  $u^\varepsilon \log n \leq \mathbf{E}[S_n]$  for all  $u \leq c$ . Decreasing  $c$  if necessary, we can and will assume that

$$C^{-1}(\mathbf{E}[S_n] - u^\varepsilon \log n) < a \log n$$

and hence that

$$\mathbf{P}(\Sigma_n \leq u \log n) \leq C \exp\left(- (a \log n)^{\theta/(\theta-1)}\right) + C_{\gamma'} u^{(1-\varepsilon)\gamma'/\alpha}, \quad \text{for all } u \leq c,$$

where  $a > 0$  is to be chosen conveniently small. Observe that

$$C \exp\left(- (a \log n)^{\theta/(\theta-1)}\right) < C_{\gamma'} u^{(1-\varepsilon)\gamma'/\alpha},$$

for all  $u \in \left[\frac{C}{C_{\gamma'}} \exp\left(-\frac{\alpha}{\gamma'} (a \log n)^{\frac{\theta}{\theta-1}}\right), c\right]$  and that  $e^{-\frac{\alpha}{\gamma'} (a \log n)^{\frac{\theta}{\theta-1}}} \ll 1/n$ . Therefore, taking  $\gamma = (1-\varepsilon)\gamma' < \alpha + \theta$  there exists  $D_\gamma$  such that

$$\mathbf{P}(\Sigma_n \leq u \log n) \leq D_\gamma u^{\frac{\gamma}{\alpha}},$$

for all  $n$  large enough and  $u \in [\frac{1}{n}, +\infty)$ .

On the other hand if  $u \leq 1/n$ , let  $j^* \in \mathbb{N}$  be such that  $(1 - \alpha)j^* > \gamma$ , then

$$\mathbf{P}(\Sigma_n < u \log n) = \mathbf{P}\left(\sum_{j=1}^n V_j^\alpha \leq u \log n\right) \leq \mathbf{P}(V_j^\alpha < u \log n; \text{ for all } 1 \leq j \leq j^*).$$

Observe that if  $u < 1/n$  and  $V_j^\alpha < u \log n$  for all  $j \leq j^*$ , we have

$$Y_j^\alpha = \frac{V_j^\alpha}{((1-Y_1)(1-Y_2)\cdots(1-Y_{j-1}))^\alpha} \leq \frac{u \log n}{(1-Y_1^\alpha)\cdots(1-Y_{j-1}^\alpha)}.$$

We prove by recurrence that under the above hypothesis,  $Y_j^\alpha \leq \frac{u \log n}{(1 - \frac{(j-1) \log n}{n})}$ . The case  $j = 1$  holds by the assumption  $Y_1^\alpha \leq u \log n$ . Assuming that the statement holds for all  $i \leq j - 1$ , that is  $Y_i^\alpha \leq u \log n / (1 - \frac{(i-1) \log n}{n})$  then

$$\begin{aligned} (1 - Y_1^\alpha)(1 - Y_2^\alpha) \cdots (1 - Y_{j-1}^\alpha) &\geq \prod_{i=1}^{j-1} \left(1 - \frac{\frac{\log n}{n}}{1 - \frac{(i-1) \log n}{n}}\right) \\ &\geq \prod_{i=1}^{j-1} \frac{1 - \frac{i \log n}{n}}{1 - \frac{(i-1) \log n}{n}} = 1 - \frac{(j-1) \log n}{n}, \end{aligned}$$

yielding  $Y_j^\alpha \leq \log n / (1 - \frac{(j-1) \log n}{n})$ . As a consequence, for all  $j \leq j^*$  and  $n$  sufficiently large  $Y_j^\alpha < 2u \log n$  and hence  $\mathbf{P}(\Sigma_n < u \log n) \leq \prod_{j=1}^{j^*} \mathbf{P}(Y_j^\alpha < 2u \log n)$ .

Using crude estimate for the probability distribution function of the Beta distribution, we bound the product in the display by  $Cu^{\frac{\gamma}{\alpha}} \rho(u^{j^*(1-\alpha)-\gamma} (\log n)^{j^*(1-\alpha)})$ , with  $C$  an explicit constant. Since  $u < 1/n$  and  $j^*(1 - \alpha) - \gamma > 0$ , the term inside the parentheses tends to zero uniformly in  $u$ . Therefore, increasing  $D_\gamma > 0$  if necessary, the upper-bound  $\mathbf{P}(\Sigma_n < u \log n) \leq D_\gamma u^{\frac{\gamma}{\alpha}}$  holds for all  $n \geq 1$  and  $u \geq 0$  finishing the proof.  $\square$

In some cases, we are able to identify the random variable  $\Psi_\alpha M_\infty^\alpha$ .

**Corollary 2.7.** *Let  $(U_n)_n$  be a PD( $\alpha, 0$ ), then  $\Psi_\alpha M_\infty^\alpha = L^{-\alpha}$ , where we have set  $1/L = \lim_{n \rightarrow +\infty} n^{1/\alpha} U_n$ .*

*Proof:* By Proposition 2.2,  $L := \lim_{n \rightarrow +\infty} n^{-1/\alpha} / U_n$  exists a.s. and by Lemma 2.6 we have that  $\Psi_\alpha M_\infty^\alpha \sim \frac{1}{\log n} \sum_{j=1}^n V_j^\alpha \leq \frac{1}{\log n} \sum_{j=1}^n U_j^\alpha \sim L^{-\alpha}$  as  $n \rightarrow +\infty$ , thus  $\Psi_\alpha M_\infty^\alpha \leq L^{-\alpha}$  a.s. By Lemma 2.4, the  $p$ th moments of  $\Psi_\alpha M_\infty^\alpha$  are equal to

$$\mathbf{E}[(\Psi_\alpha M_\infty^\alpha)^p] = \frac{\Gamma(p+1)}{\Gamma(p\alpha+1)} \Gamma(1-\alpha)^{-p}, \quad \text{for all } p > -1.$$

By Pitman and Yor (1997, Equation (30)), it matches with the  $p$ th moments of  $L^{-\alpha}$ , which implies that the two random variables have the same distribution (the Mittag-Leffler ( $\alpha$ ) distribution), and hence that  $\Psi_\alpha M_\infty^\alpha = L^{-\alpha}$  a.s. by monotonicity.  $\square$

### 3. Convergence of discrete exchangeable coalescent processes

In this section, we study a family of coalescent processes with dynamics driven by PD-distributions and obtain a sufficient criterion for the convergence in distribution of these processes. For the sake of completeness, we include a brief introduction to coalescent theory with the main results we will use, for a detailed account we recommend the paper Berestycki (2009) from where we borrow the approach.

Let  $\mathcal{P}_n$  be the set of partitions (or equivalence classes) of  $[n] := \{1, \dots, n\}$  and  $\mathcal{P}_\infty$  the set of partitions of  $\mathbb{N} = [\infty]$ . A partition  $\pi \in \mathcal{P}_n$  is represented by blocks  $\pi(1), \pi(2), \dots$  listed in the increasing order of their least elements, that is,  $\pi(1)$  is the block (class) containing 1,  $\pi(2)$  the block containing the smallest element not in  $\pi(1)$  and so on. There is a natural action of the symmetric group  $S_n$  on  $\mathcal{P}_n$  setting  $\pi^\sigma := \{\{\sigma(j), j \in \pi(i)\}, i \in [n]\}$  for  $\sigma \in S_n$ . If  $m < n$ , one can define the projection of  $\mathcal{P}_n$  onto  $\mathcal{P}_m$  by the restriction  $\pi|_m = \{\pi(j) \cap [m]\}$ . For  $\pi, \pi' \in \mathcal{P}_n$ , we define the *coagulation of  $\pi$  by  $\pi'$*  to be the partition  $\text{Coag}(\pi, \pi') = \{\cup_{i \in \pi'(j)} \pi(i); j \in \mathbb{N}\}$ .

With this notation, a coalescent process  $\Pi(t)$  is a discrete (or continuous) time Markov process in  $\mathcal{P}_n$  such that for any  $s, t \geq 0$ ,

$$\Pi(t+s) = \text{Coag}(\Pi(t), \tilde{\Pi}_s), \quad \text{with } \tilde{\Pi}_s \text{ independent of } \Pi(t).$$

We say that  $\Pi(t)$  is *exchangeable* if  $\Pi^\sigma(t)$  and  $\Pi(t)$  have the same distribution for all permutation  $\sigma$ .

An important class of continuous-time exchangeable coalescent processes in  $\mathcal{P}_\infty$  are the so-called  $\Lambda$ -coalescents (see [Pitman, 1999](#)), introduced independently by Pitman and Sagitov. They are constructed as follows: let  $\Pi_n(t)$  be the restriction of  $\Pi(t)$  to  $[n]$ , then  $(\Pi_n(t); t \geq 0)$  is a Markov jump process on  $\mathcal{P}_n$  with the property that whenever there are  $b$  blocks, each  $k$ -tuple ( $k \geq 2$ ) of blocks is merging to form a single block at the rate

$$\lambda_{b,k} = \int_0^1 x^{k-2} (1-x)^{b-k} \Lambda(dx), \quad \text{where } \Lambda \text{ is a finite measure on } [0, 1].$$

Among such, we distinguish the Beta( $2 - \lambda, \lambda$ )-coalescents obtained from  $\Lambda(dx) = \frac{x^{1-\lambda}(1-x)^{\lambda-1}}{\Gamma(\lambda)\Gamma(2-\lambda)} dx$ , where  $\lambda \in (0, 2)$ , the case  $\lambda = 1$  (uniform measure) being the celebrated *Bolthausen-Sznitman coalescent*.

The set  $\mathcal{P}_\infty$  can be endowed with a topology making it a Polish space, therefore, one can study the weak convergence of processes in  $\mathcal{D}([0, \infty), \mathcal{P}_\infty)$ , see [Berestycki \(2009\)](#) for the definitions. Without going into details, we say that a process  $\Pi^N(t) \in \mathcal{P}_\infty$  converges in the *Skorokhod sense* (or in distribution) to  $\Pi(t)$ , if for all  $n \in \mathbb{N}$  the projection  $\Pi^N(t)|_n$  converges in distribution to  $\Pi(t)|_n$  in  $\mathcal{D}([0, \infty) \mathcal{P}_n)$ .

**3.1. Coalescent processes obtained from multinomial distributions.** In this section, we define a family of discrete-time coalescent processes  $(\Pi^N(t); t \in \mathbb{N})$  and prove sufficient criteria for its convergence in distribution. Let  $(\eta_1^N, \dots, \eta_N^N)$  be an  $N$ -dimensional random vector satisfying

$$1 \geq \eta_1^N \geq \eta_2^N \geq \dots \geq \eta_N^N \geq 0 \quad \text{and} \quad \sum_{j=1}^N \eta_j^N = 1.$$

Conditionally on a realization of  $(\eta_j^N)$ , let  $\{\xi_j; j \leq N\}$  be i.i.d. random variables satisfying  $\mathbf{P}(\xi_j = k | \eta^N) = \eta_k^N$  and define the partition

$$\pi_N = \{\{j \leq N : \xi_j = k\}; k \leq N\}.$$

With  $(\pi_t; t \in \mathbb{N})$  i.i.d. copies of  $\pi_N$ , let  $\Pi^N(t)$  be the discrete time coalescent such that

$$\Pi^N(0) = \{\{1\}, \{2\}, \dots, \{n\}\} \quad \text{and} \quad \Pi^N(t+1) = \text{Coag}(\Pi^N(t), \pi_{t+1}).$$

The goal of this section is to obtain conditions under which  $\Pi^N(t)$  converges in distribution. First, we assume that there exist a sequence  $L_N$  and a function  $f : (0, 1) \rightarrow \mathbb{R}_+$  such that

$$\lim_{N \rightarrow +\infty} L_N = +\infty, \quad \lim_{N \rightarrow +\infty} L_N \mathbf{P}(\eta_1^N > x) = f(x) \quad \text{and} \quad \lim_{N \rightarrow +\infty} L_N \mathbf{E}(\eta_2^N) = 0. \quad (3.1)$$

Denote by  $c_N = \sum_{j=1}^N \mathbf{E}[(\eta_j^N)^2]$ , which corresponds to the probability that two individuals have a common ancestor one generation backward in time.

**Lemma 3.1.** *Assume that (3.1) holds and that*

$$\int_0^1 x \left( \sup_{N \in \mathbb{N}} L_N \mathbf{P}(\eta_1^N > x) \right) dx < +\infty. \quad (3.2)$$

Then,  $c_N \sim_{N \rightarrow \infty} L_N^{-1} \int_0^1 2xf(x)dx$  and the coalescent process  $(\Pi^N(t/c_N); t \in \mathbb{R}_+)$ , properly rescaled, converges in distribution to the  $\Lambda$ -coalescent, with  $\Lambda$  satisfying  $\int_x^1 \frac{\Lambda(dy)}{y^2} = f(x)$ .

*Proof:* Denote by  $\nu_k = \#\{j \leq N : \xi_j = k\}$ , then  $(\nu_1, \dots, \nu_N)$  has multinomial distribution with  $N$  trials and (random) probabilities outcomes  $\eta_i^N$ . By [Möhle and Sagitov \(2001, Theorem 2.1\)](#), the convergence of finite dimensional distribution of  $\Pi^N(t)$  is obtained from the convergence of the factorial moments of  $\nu$ , that is

$$\frac{1}{c_N(N)_b} \sum_{\substack{i_1, \dots, i_a=1 \\ \text{all distinct}}}^N \mathbf{E}[(\nu_{i_1})_{b_1} \dots (\nu_{i_a})_{b_a}], \quad \text{with } b_i \geq 2 \text{ and } b = b_1 + \dots + b_a,$$

where  $(n)_a := n(n-1)\dots(n-a+1)$ . Since  $(\nu_1, \dots, \nu_N)$  is multinomial distributed, we obtain that  $\mathbf{E}[(\nu_{i_1})_{b_1} \dots (\nu_{i_a})_{b_a}] = (N)_b \mathbf{E}[\eta_{i_1}^{b_1} \dots \eta_{i_a}^{b_a}]$ , see [Cortines \(2016, Lemma 4.1\)](#) for a rigorous a proof. Therefore, we only have to show that for all  $b$  and  $a \geq 2$

$$\begin{aligned} \lim_{N \rightarrow +\infty} c_N^{-1} \sum_{i_1=1}^N \mathbf{E}[(\eta_{i_1}^N)^b] &= \int_0^1 x^{b-2} \Lambda(dx) \\ \text{and } \lim_{N \rightarrow +\infty} c_N^{-1} \sum_{\substack{i_1, \dots, i_a=1 \\ \text{all distinct}}}^N \mathbf{E}[(\eta_{i_1}^N)^{b_1} \dots (\eta_{i_a}^N)^{b_a}] &= 0. \end{aligned}$$

We obtain by dominated convergence that

$$\begin{aligned} L_N \mathbf{E}[(\eta_1^N)^2] &= \int_0^1 2x L_N \mathbf{P}(\eta_1^N > x) dx \\ &\rightarrow \int_0^1 2xf(x)dx = \int_0^1 \Lambda(dx) < +\infty, \quad \text{as } N \rightarrow \infty. \end{aligned}$$

We also get  $\mathbf{E}[(\eta_2^N)^2 + \dots + (\eta_N^N)^2] \leq \mathbf{E}[\eta_2^N(1-\eta_1)]$ , with  $L_N \mathbf{E}(\eta_2^N)$  tending to zero as  $N \rightarrow \infty$ , as  $(\eta_i^N)$  is ordered and sums up to 1. In particular, it implies that  $L_N c_N = L_N \sum \mathbf{E}[(\eta_i^N)^2]$  tends to  $\int_0^1 2xf(x)dx$  as  $N \rightarrow \infty$ . A similar calculation shows that for any  $b \geq 2$

$$\lim_{N \rightarrow \infty} L_N \sum_{i=1}^N \mathbf{E}[(\eta_i^N)^b] = \int bx^{b-1} f(x)dx = \int x^{b-2} \Lambda(dx) = \lambda_{b,b},$$

where  $\lambda_{b,b}$  is the rate at which  $b$  blocks merge into one given that there are  $b$  blocks in total. The others  $\lambda_{b,k}$  can be easily obtained using the recursion formula  $\lambda_{b,k} = \lambda_{b+1,k} + \lambda_{b+1,k+1}$ .

We now consider the case  $a = 2$ , cases  $a > 2$  being treated in the same way. We have

$$\begin{aligned} & \sum_{\substack{i_1, i_2=1 \\ \text{distinct}}}^N \mathbf{E} \left[ (\eta_{i_1}^N)^{b_1} (\eta_{i_2}^N)^{b_2} \right] \\ & \leq \mathbf{E} \left[ (\eta_1^N)^{b_1} \eta_2^N \sum (\eta_i^N)^{b_2-1} \right] + \mathbf{E} \left[ (\eta_1^N)^{b_2} \eta_2^N \sum (\eta_i^N)^{b_1-1} \right] \\ & \quad + \mathbf{E} \left[ \sum_{i_1 \neq 1} (\eta_{i_1}^N)^{b_1} \eta_2^N \sum_{\substack{i_2 \neq 1 \\ i_2 \neq i_1}} (\eta_{i_2}^N)^{b_2-1} \right] \\ & \leq 3 \times \mathbf{E} [\eta_2^N], \end{aligned}$$

we recall that  $(\eta_2^N)^b + \dots + (\eta_N^N)^b \leq \eta_2^N + \dots + \eta_N^N = 1 - \eta_1^N < 1$ . Since  $L_N \mathbf{E} \eta_2^N \rightarrow 0$  as  $N \rightarrow \infty$ , the right hand side of the inequality tends to zero, concluding the proof.  $\square$

The next lemma gives sufficient conditions for the convergence to the Kingman's coalescent.

**Lemma 3.2.** *Assume that (3.1) holds and that*

$$\int_0^1 x f(x) dx = +\infty \quad \text{and} \quad \exists n \geq 2 : \int_0^1 x^n \left( \sup_{N \in \mathbb{N}} L_N \mathbf{P}(\eta_1^N > x) \right) dx < +\infty. \quad (3.3)$$

Then,  $\lim_{N \rightarrow +\infty} c_N L_N = +\infty$  and the process  $(\Pi_n^N(\lfloor tc_N^{-1} \rfloor); t \in \mathbb{R}_+)$  converges in the Skorokhod sense to the Kingman's coalescent restricted to  $\mathcal{P}_n$ .

*Proof:* A similar argument to the one used in Lemma 3.1 shows that

$$L_N c_N \geq L_N \mathbf{E}[\eta_1^2] = \int_0^1 2x L_N \mathbf{P}(\eta_1 > x) dx.$$

Thus, Fatou lemma and (3.3) yield  $\liminf_{N \rightarrow +\infty} L_N c_N = +\infty$ , proving the first claim. By Berestycki (2009, Theorem 2.5), the convergence to the Kingman's coalescent follows from

$$\lim_{N \rightarrow \infty} \sum_{i=1}^N \mathbf{E}[(\nu_i)_3] / (N)_3 c_N = 0.$$

We rewrite the sum in the display as  $c_N^{-1} \sum_{i=1}^N \mathbf{E}[(\eta_i^N)^3]$  and apply Hölder inequality to obtain

$$\mathbf{E}[(\eta_1^N)^\lambda (\eta_1^N)^{3-\lambda}] \leq \mathbf{E}[(\eta_1^N)^2]^{\lambda/2} \mathbf{E}[(\eta_1^N)^{\frac{2(3-\lambda)}{2-\lambda}}]^{1-\lambda/2} \quad \text{for all } \lambda < 2.$$

Let  $\lambda \in (0, 2)$  be the unique solution of  $2(3-\lambda)/(2-\lambda) = n+4$ , then we obtain from (3.3) that

$$\frac{\mathbf{E}[(\eta_1^N)^3]}{c_N} \leq \frac{\mathbf{E}[(\eta_1^N)^3]}{\mathbf{E}[(\eta_1^N)^2]} \leq \left( \frac{\mathbf{E}[(\eta_1^N)^{n+1}]}{\mathbf{E}[(\eta_1^N)^2]} \right)^{1-\lambda/2} \xrightarrow{N \rightarrow +\infty} 0.$$

A similar argument to the one used in Lemma 3.1 shows that  $c_N^{-1} \sum_{i=2}^N \mathbf{E}[(\eta_i^N)^3]$  tends to zero as  $N \rightarrow \infty$ , which proves the statement.  $\square$

**3.2. The Poisson-Dirichlet distribution case.** In this section, we construct a coalescent using the PD distribution and obtain a criterion for its convergence in distribution. With  $(V_j, j \geq 1)$  a size-biased pick from a  $\text{PD}(\alpha, \theta)$  partition, define

$$\theta_j^N := \frac{V_j^\alpha}{\sum_{i=1}^N V_i^\alpha} \quad \text{and} \quad \theta_{(1)}^N \geq \theta_{(2)}^N \geq \dots \geq \theta_{(N)}^N, \quad \text{the order statistics of } (\theta_j^N).$$

In what follows,  $\theta_{(i)}^N$  will stand for the  $\eta_i^N$  from Section 3.1 and  $(\Pi_n^N(t); t \in \mathbb{N})$  for the coalescent with transition probabilities  $\Pi_n^N(t+1) = \text{Coag}(\Pi_n^N(t), \pi_t^n)$ , as defined there.

**Theorem 3.3.** *With the above notation, set  $\lambda = 1 + \theta/\alpha$  and*

$$L_N = c_{\alpha, \theta} (\log N)^\lambda, \quad \text{where} \quad c_{\alpha, \theta} = \left( \Gamma(1 - \theta/\alpha) \Gamma(1 - \alpha)^{\theta/\alpha} \Gamma(1 + \theta) \right)^{-1}.$$

- (1) *If  $\theta \in (-\alpha, \alpha)$ , then  $c_N \sim_{N \rightarrow +\infty} (1 - \theta/\alpha)/L_N$  and  $(\Pi^N(t/c_N), t \geq 0)$  converges weakly to the  $\text{Beta}(2 - \lambda, \lambda)$ -coalescent.*
- (2) *Otherwise,  $\lim_{N \rightarrow +\infty} c_N L_N = +\infty$  and  $(\Pi^N(t/c_N))$  converges weakly to the Kingman's coalescent.*

Before proving Theorem 3.3 we obtain a couple of technical results. The next lemma studies the asymptotic behavior of  $\theta_1^N$ .

**Lemma 3.4.** *With the notations of Theorem 3.3, we have*

$$\lim_{N \rightarrow +\infty} L_N \mathbf{P}(\theta_1^N > x) = \frac{1}{\lambda \Gamma(\lambda) \Gamma(2 - \lambda)} \left( \frac{1 - x}{x} \right)^\lambda = \int_x^1 \frac{\text{Beta}(2 - \lambda, \lambda)(dy)}{y^2}.$$

Moreover, there exists  $C > 0$  such that for all  $x \in (0, 1)$

$$\sup_{N \in \mathbb{N}} L_N \mathbf{P}(\theta_1^N > x) \leq C x^{-\lambda}.$$

*Proof:* Let  $\Sigma'_N := \sum_{j=2}^N \left( \frac{V_j}{1 - V_1} \right)^\alpha$ , then Remark 2.3 says that  $\Sigma'_N$  and  $V_1$  are independent and that  $\Sigma'_N$  has the distribution of  $V'_1 + \dots + V'_{N-1}$  with  $V'_i$  a size-biased pick from a  $\text{PD}(\alpha, \alpha + \theta)$  distribution. By Lemma 2.6, for all  $\varepsilon \in (0, 1)$  there exists  $C = C(\varepsilon)$  and  $N_0 \in \mathbb{N}$  such that

$$\sup_{N \geq N_0} \mathbf{P}(\Sigma'_N \leq u \log N) \leq \min(Cu^{\lambda+\varepsilon}, 1), \quad \text{for all } u \geq 0. \quad (3.4)$$

Writing  $\mathbf{P}(\theta_1^N > x)$  in terms of  $\Sigma'_N$  and  $V_1 = Y_1$ , we obtain

$$\begin{aligned} \mathbf{P}(\theta_1^N > x) &= \mathbf{P}(V_1^\alpha > x(V_1^\alpha + (1 - V_1)^\alpha \Sigma'_N)) = \mathbf{P}\left(\frac{V_1}{1 - V_1} > \left(\frac{x}{1 - x} \Sigma'_N\right)^{1/\alpha}\right) \\ &= \int_0^1 \mathbf{P}\left(1/y - 1 > \left(\frac{x}{1 - x} \Sigma'_N\right)^{1/\alpha}\right) \frac{\Gamma(1 + \theta)(1 - y)^{-\alpha} y^{\alpha + \theta - 1}}{\Gamma(1 - \alpha) \Gamma(\alpha + \theta)} dy. \end{aligned}$$

Making the change of variables  $u = (\frac{1-x}{x \log N})(1/y - 1)^\alpha$  the display reads

$$\begin{aligned} & \mathbf{P}(\theta_1^N > x) \\ &= \left(\frac{1-x}{x \log N}\right)^\lambda \frac{\Gamma(1+\theta)}{\alpha \Gamma(1-\alpha) \Gamma(\alpha+\theta)} \int_0^{+\infty} \frac{\mathbf{P}(\Sigma'_N < u \log N)}{u^{2-1/\alpha} \left(u^{1/\alpha} + \left(\frac{1-x}{x \log N}\right)^{1/\alpha}\right)^{1+\theta}} du. \end{aligned}$$

Then, we use (3.4) to bound the equation within the integral, obtaining

$$\frac{\mathbf{P}(\Sigma'_N < u \log N)}{u^{2-1/\alpha} \left(u^{1/\alpha} + \left(\frac{1-x}{x \log N}\right)^{1/\alpha}\right)^{1+\theta}} \leq \frac{\mathbf{P}(\Sigma'_N \leq u \log N)}{u^{2+\theta/\alpha}} \leq \min(Cu^{\varepsilon-1}, u^{-2}),$$

for all  $N$  large enough. In particular, there exists  $C > 0$  such that  $L_N \mathbf{P}(\eta_1^N > x) \leq Cx^{-\lambda}$  for all  $N \in \mathbb{N}$ . Moreover, by dominated convergence and Lemma 2.6, we obtain

$$\begin{aligned} & \lim_{N \rightarrow +\infty} (\log N)^\lambda \mathbf{P}(\theta_1^N > x) \\ &= \left(\frac{1-x}{x}\right)^\lambda \frac{\Gamma(1+\theta)}{\alpha \Gamma(1-\alpha) \Gamma(\alpha+\theta)} \int_0^{+\infty} \frac{\mathbf{P}(\Psi_\alpha(M'_\infty)^\alpha < u)}{u^{1+\lambda}} \\ &= \left(\frac{1-x}{x}\right)^\lambda \frac{\Gamma(1+\theta)}{\alpha \lambda \Gamma(1-\alpha) \Gamma(\alpha+\theta)} \mathbf{E}\left((\Psi_\alpha(M'_\infty)^\alpha)^{-\lambda}\right) \\ &= \left(\frac{1-x}{x}\right)^\lambda \frac{\alpha^{\alpha+\theta-1} \Gamma(1-\alpha)^{\theta/\alpha} \Gamma(1+\theta)}{\lambda \Gamma(\alpha+\theta)} \Phi_{\theta+\alpha, \alpha}(-(\theta+\alpha)), \end{aligned}$$

and hence  $\lim_{N \rightarrow +\infty} L_N \mathbf{P}(\theta_1^N > x) = \left(\frac{1-x}{x}\right)^\lambda \frac{1}{\lambda \Gamma(\lambda) \Gamma(2-\lambda)}$ , proving the statement.  $\square$

This result is used to study the asymptotic behavior of  $\theta_{(1)}^N = \max_{j \leq N} \theta_j^N$ .

**Lemma 3.5.** *For all  $\varepsilon \in (0, 1)$ , there exists  $C = C(\varepsilon)$  such that*

$$\left| \mathbf{P}(\theta_{(1)}^N > x) - \mathbf{P}(\theta_1^N > x) \right| \leq C(x \log N)^{\varepsilon-2-\theta/\alpha},$$

for all  $x \in (0, 1)$  and  $N$  large enough.

*Proof:* Notice that  $\mathbf{P}(\theta_1^N > x) \leq \mathbf{P}(\theta_{(1)}^N > x)$  and that  $\theta_{(1)}^N = \theta_1^N$  if  $V_1 > 1/2$ , thanks to  $\sum V_i \equiv 1$ . Therefore, splitting the events according to  $V_1 > 1/2$  and  $V_1 < 1/2$  we obtain

$$\begin{aligned} \mathbf{P}(\theta_{(1)}^N > x) - \mathbf{P}(\theta_1^N > x) &= \mathbf{P}\left(\theta_{(1)}^N > x; V_1 \leq \frac{1}{2}\right) - \mathbf{P}\left(\theta_1^N > x; V_1 \leq \frac{1}{2}\right) \\ &\leq \mathbf{P}\left(\theta_{(1)}^N > x; \leq \frac{1}{2}\right). \end{aligned}$$

Since  $0 < V_j < 1$  and  $V_1 = Y_1$ , we have

$$\begin{aligned} \mathbf{P}\left(\theta_{(1)}^N > x, V_1 \leq 1/2\right) &= \mathbf{P}\left(\max_{j \leq N} V_j^\alpha > x \sum_{j=1}^N V_j^\alpha; V_1 \leq 1/2\right) \\ &\leq \mathbf{P}\left(x^{-1} > Y_1^\alpha + (1-Y_1)^\alpha \Sigma'_N, Y_1 \leq 1/2\right) \\ &\leq \mathbf{P}(\Sigma'_N < 2^\alpha/x), \end{aligned}$$

where  $\Sigma'_N = \sum V_j^\alpha / (1 - Y_1)^\alpha$ . By Lemma 2.6, we obtain that for all  $N$  sufficiently large

$$\mathbf{P}(\theta_{(1)}^N > x, V_1 \leq 1/2) \leq C(x \log N)^{\varepsilon - 1 - \frac{\alpha + \theta}{\alpha}},$$

which finishes the proof.  $\square$

We now study the asymptotic behavior of the second maxima in  $\{\theta_1^N, \dots, \theta_N^N\}$ , written  $\theta_{(2)}^N$ .

**Lemma 3.6.** *For all  $\varepsilon \in (0, 1]$ , there exists  $C > 0$  such that for any  $x \in (0, 1)$  and  $N \in \mathbb{N}$ ,*

$$\mathbf{P}(\theta_{(2)}^N > x) < C(x \log N)^{\varepsilon - 2 - \theta/\alpha}.$$

*Proof:* We basically use the same method as in the previous lemma

$$\begin{aligned} \mathbf{P}(\theta_{(2)}^N > x) &= \mathbf{P}(\theta_{(2)}^N > x, V_1 \leq 1/2) + \mathbf{P}(\theta_{(2)}^N > x, V_1 > 1/2, V_2 < 1/3) \\ &\quad + \mathbf{P}(\theta_{(2)}^N > x, V_1 > 1/2, V_2 > 1/3) \\ &\leq \mathbf{P}(\theta_{(1)}^N > x, V_1 \leq 1/2) + \mathbf{P}(\theta_{(2)}^N > x, V_1 > 1/2, V_2 < 1/3) + \mathbf{P}(\theta_2^N > x). \end{aligned}$$

By Lemma 3.5, we have that  $\mathbf{P}(\theta_{(1)}^N > x, V_1 \leq 1/2) \leq C(x \log N)^{\varepsilon - 2 - \theta/\alpha}$ , so the same arguments used in Lemma 3.4 yield

$$\mathbf{P}(\theta_2^N > x) = \mathbf{P}(V_2^\alpha(1 - x) - xV_1^\alpha > x(1 - V_1 - V_2)^\alpha \Sigma_N'') \leq C(x \log N)^{\varepsilon - 2 - \theta/\alpha},$$

with  $\Sigma_N'' := (1 - V_1 - V_2)^{-\alpha} \sum_{j=3}^N V_j^\alpha$ . Moreover,  $\Sigma_N''$  is independent of  $(V_1, V_2)$  and

$$\begin{aligned} &\mathbf{P}(\theta_{(2)}^N > x, V_1 > 1/2, V_2 < 1/3) \\ &= \mathbf{P}\left(\max_{2 \leq j \leq N} V_j^\alpha > x(V_1^\alpha + V_2^\alpha + (1 - V_1 - V_2)^\alpha \Sigma_N'')\right) \\ &\leq \mathbf{P}(\Sigma_N'' \leq C/x) \leq C(x \log N)^{\varepsilon - 2 - \theta/\alpha}, \end{aligned}$$

concluding the proof.  $\square$

*Proof of Theorem 3.3:* Given  $\varepsilon > 0$ , Lemma 3.5 says that there exists  $C > 0$  such that

$$L_N \mathbf{P}(\theta_{(1)}^N > x) - L_N \mathbf{P}(\theta_1^N > x) \leq C(\log N)^{\varepsilon - 1} x^{-\lambda}, \quad \text{for all } x \in (0, 1).$$

Therefore, by Lemma 3.4 we have that

$$\begin{aligned} \lim_{N \rightarrow +\infty} L_N \mathbf{P}(\theta_{(1)}^N > x) &= \frac{1}{\lambda \Gamma(\lambda) \Gamma(2 - \lambda)} \left(\frac{1 - x}{x}\right)^\lambda \quad \text{and} \\ &\sup_{N \in \mathbb{N}} L_N \mathbf{P}(\theta_{(1)}^N > x) \leq Cx^{-\lambda}. \end{aligned} \tag{3.5}$$

We obtain from Lemma 3.6 that

$$L_N \mathbf{E}[\theta_{(2)}^N] = \int_0^1 L_N \mathbf{P}(\theta_{(2)}^N > x) dx \leq (\log N)^{\varepsilon - 1} \int_0^1 x^{\varepsilon - 1 - \lambda} dx \xrightarrow{N \rightarrow +\infty} 0, \tag{3.6}$$

which implies that  $\Pi^N(t)$  satisfies (3.1).



Assume now that  $\theta \in (-\alpha, \alpha)$ , so that  $\lambda \in (0, 2)$ , then (3.5) yields

$$\int_0^1 x \sup_{N \in \mathbb{N}} L_N \mathbf{P}(\theta_{(1)}^N > x) dx \leq C \int_0^1 x^{1-\lambda} dx < +\infty.$$

Therefore, the assumptions of Lemma 3.1 are satisfied implying that  $\Pi^N(t/c_N)$  converges in distribution to the Beta( $2-\lambda, \lambda$ )-coalescent and that  $c_N L_N \sim (1-\theta/\alpha)$  as  $N \rightarrow \infty$ . On the other hand if  $\theta \geq \alpha$ , we have that  $\int_0^1 x \left(\frac{1-x}{x}\right)^\lambda dx = +\infty$ . With  $k \geq \lambda$ , we obtain from (3.5) that

$$\int_0^1 x^k \sup_{N \in \mathbb{N}} L_N \mathbf{P}(\theta_{(1)}^N > x) dx \leq C \int_0^1 x^{k-\lambda} dx < +\infty.$$

We apply Lemma 3.2 to conclude that  $\Pi^N(t/c_N)$  converges weakly to the Kingman's coalescent.  $\square$

#### 4. Poisson-Dirichlet representation of the $(N, \beta)$ -branching random walk

We explore the relations between the  $(N, \beta)$ -BRW and the PD( $\beta^{-1}, 0$ ) distribution to show Theorems 1.1 and 1.2 in the case where  $\beta < \infty$ . The case  $\beta = \infty$  is also studied, but using different methods.

**Proposition 4.1.** *Let  $\beta \in (1, \infty)$  and  $(V_n)_n$  be its size-biased pick, and set  $(U_n)_n$  be a PD( $\beta^{-1}, 0$ ),  $L = \lim_{n \rightarrow +\infty} n^{-\beta} U_n^{-1}$  then*

$$(X_1^N(j) - X_0^N(\text{eq}), j \leq N) \stackrel{(d)}{=} \left( \frac{1}{\beta} \log V_j + \frac{1}{\beta} \log L \right),$$

in particular,  $X_1^N(\text{eq}) - X_0^N(\text{eq}) \stackrel{(d)}{=} \log \sum_{j=1}^N V_j^{1/\beta} + \frac{1}{\beta} \log L$ .

*Proof:* By Proposition 1.3,

$$(x_k, k \geq 1) := \text{Rank}(\{X_0^N(j) + p - X_0^N(\text{eq}), p \in \mathcal{P}_1(j), j \leq N\})$$

is the ordered points of a PPP( $e^{-x} dx$ ). With  $L = \sum_{j=1}^{+\infty} e^{\beta x_j}$  and  $U_j = e^{\beta x_j} / L$ , we know from Proposition 2.2 that  $(U_j, j \geq 1)$  is a PD( $\beta^{-1}, 0$ ) and that

$$\lim_{n \rightarrow +\infty} n^\beta U_n = L^{-1}.$$

By the definition of the  $(N, \beta)$ -BRW,  $V_j := e^{\beta(X_1^N(j) - X_0^N(\text{eq}))} / L$  is the  $j$ th particle sampled in the size-biased pick from  $(U_n)_n$ . Inversing the equation, we conclude that

$$X_1^N(j) - X_0^N(\text{eq}) = \frac{1}{\beta} (\log V_j + \log L),$$

proving the first statement. The second comes from the definition of  $X_1^N(\text{eq})$ .  $\square$

To study the case  $\beta = +\infty$ , we use the following representation of  $N$  rightmost points of a PPP( $e^{-x} dx$ ).

**Proposition 4.2.** *Let  $x_1 > x_2 > \dots > x_N$  be the  $N$  rightmost points of a PPP( $e^{-x} dx$ ), then*

$$(x_1, \dots, x_N) \stackrel{(d)}{=} \text{Rank}\{Z_N + e_1, \dots, Z_N + e_N\},$$

where  $(e_j)$  are i.i.d exponential random variables with mean 1 and  $Z_N$  is an independent random variable satisfying  $\mathbf{P}(Z_N \in dx) = \frac{1}{N!} \exp(-(N+1)x - e^{-x}) dx$ .

*Proof:* It is an elementary result about PPP that

$$\sum_{j=1}^{+\infty} \delta_{e^{-x_j}} \stackrel{(d)}{=} \text{PPP}(dx) \text{ on } \mathbb{R}_+ \quad \text{and that} \quad e^{-x_{N+1}} \stackrel{(d)}{=} \text{Gamma}(N+1, 1),$$

moreover, conditionally on  $x_{N+1}$ ,  $(e^{-x_1}, \dots, e^{-x_N})$  are the ranked values of  $N$  i.i.d. uniform random variables on the interval  $[0, e^{-x_{N+1}}]$ . Setting  $Z_N = x_{N+1}$  and  $U_1, \dots, U_N$  i.i.d. uniform random variables

$$(e^{-x_1}, \dots, e^{-x_N}) \stackrel{(d)}{=} \text{Rank}\{e^{-Z_N} U_1, \dots, e^{-Z_N} U_N\}.$$

It is straightforward that  $Z_N$  and  $(x_1, \dots, x_N)$  satisfy the desired properties.  $\square$

We first use these results to compute the asymptotic behavior of the speed of the  $(N, \beta)$ -BRW.

*Proof of Theorem 1.1:* Lemma 1.5 says that (1.3) holds with

$$v_{N,\beta} = \mathbf{E} (X_1^N(\text{eq}) - X_0^N(\text{eq})).$$

Thus, if  $\beta < \infty$  Proposition 4.1 yields

$$v_{N,\beta} = \mathbf{E} (X_1^N(\text{eq}) - X_0^N(\text{eq})) = \mathbf{E} \left( \log \left( \sum_{j=1}^N V_j^{1/\beta} \right) \right) + \frac{1}{\beta} \mathbf{E} (\log L).$$

By Lemma 2.6,  $(\log N)^{-1} \sum_{j=1}^N V_j^{1/\beta}$  converges to  $\Psi_{\beta-1} M_\infty^{1/\beta}$  a.s. and in  $\mathbb{L}^1$  as  $N \rightarrow \infty$ . Therefore, the logarithm of this quantity converges a.s. as well. We notice from Lemma 2.6 that for all  $u > 0$ ,

$$\mathbf{P} \left( \log \left( (\log N)^{-1} \sum_{j=1}^N V_j^{1/\beta} \right) \leq -u \right) = \mathbf{P} \left( \sum_{j=1}^N V_j^{1/\beta} \leq e^{-u} \log N \right) \leq D_{2/\beta} e^{-2u}.$$

The  $\mathbb{L}^1$  convergence of  $(\log N)^{-1} \sum_{j=1}^N V_j^{1/\beta}$  implies the existence of a constant  $K$  such that for all  $u > 0$ ,

$$\mathbf{P} \left( \log \left( (\log N)^{-1} \sum_{j=1}^N V_j^{1/\beta} \right) \geq u \right) \leq \mathbf{P} \left( \sum_{j=1}^N V_j^{1/\beta} \geq e^u \log N \right) \leq K e^{-u}.$$

In particular,  $(\log \sum_{j=1}^N V_j^{1/\beta} - \log \log N)$  is uniformly integrable, which implies its  $\mathbb{L}^1$  convergence. We know from Corollary 2.7 that  $\Psi_{\beta-1} M_\infty^{1/\beta} = L^{1/\beta}$ , and hence that

$$\begin{aligned} \lim_{N \rightarrow \infty} v_{N,\beta} - \log \log N &= \lim_{N \rightarrow \infty} \mathbf{E} \left[ \log \frac{\sum_{j=1}^N V_j^{1/\beta}}{\log N} \right] + \frac{\mathbf{E} [\log L]}{\beta} \\ &= \mathbf{E} \left[ \log \left( \Psi_{\beta-1} M_\infty^{1/\beta} \right) \right] + \frac{\mathbf{E} [\log L]}{\beta} = 0. \end{aligned}$$

For  $\beta = \infty$  we follow the ideas from Brunet et al. (2007) and use Laplace methods to estimate the asymptotic mean of  $X_1^N(\text{eq})$ . Let  $(x_1, \dots, x_N)$  be the  $N$  largest atoms of a  $\text{PPP}(e^{-x} dx)$ , and define for  $\lambda > 0$

$$\Lambda(\lambda) := \mathbf{E} \left( \exp \left( -\lambda \log \sum_{k=1}^N e^{x_k} \right) \right) = \mathbf{E} \left( \left( \sum_{k=1}^N e^{x_k} \right)^{-\lambda} \right).$$

By Proposition 4.2, we have  $\Lambda(\lambda) = \mathbf{E}(e^{-\lambda Z}) \mathbf{E}\left(\left(\sum_{k=1}^N e^{e_k}\right)^{-\lambda}\right)$ , where  $(e_k)$  are i.i.d exponential random variables and  $\exp(-Z_N)$  has  $\text{Gamma}(N+1, 1)$  distribution. Notice that the following equalities hold:  $\mathbf{E}(e^{-\lambda Z}) = \frac{\Gamma(N+1+\lambda)}{\Gamma(N+1)}$  and

$$\begin{aligned} \mathbf{E}\left(\left(\sum_{k=1}^N e^{e_k}\right)^{-\lambda}\right) &= \frac{1}{\Gamma(\lambda)} \int_0^{+\infty} t^{\lambda-1} \mathbf{E}\left(e^{-t \sum_{k=1}^N e^{e_k}}\right) dt \\ &= \frac{1}{\Gamma(\lambda)} \int_0^{+\infty} t^{\lambda-1} I_0(t)^N dt, \end{aligned} \tag{4.1}$$

with  $I_0(t) = \mathbf{E}(e^{-te^{e_1}})$  the Laplace transform of  $e^{e_1}$ . The function  $I_0$  can be represented using the exponential integral  $\text{Ei}$  we have  $I_0(x) = x\text{Ei}(-x) + e^{-x}$ . Therefore, there exists  $K > 0$  such that

$$|I_0(x) - 1 - x \log x| \leq Kx, \quad \text{for any } x \geq 0.$$

In particular for  $x = t/(N \log N)$  we have  $|I_0(t/(N \log N)) - 1 - t/N| \leq \frac{Kt}{N \log N}$ . Thus, (4.1) yields

$$\begin{aligned} \int_0^{+\infty} t^{\lambda-1} I_0(t)^N dt &= \frac{1}{(N \log N)^\lambda} \int_0^{+\infty} t^{\lambda-1} I_0(t/(N \log N))^N dt \\ &\leq \frac{1}{(N \log N)^\lambda} \int_0^{+\infty} t^{\lambda-1} e^{-t(1-\frac{K}{\log N})} dt \\ &\leq \frac{\Gamma(\lambda)}{(N \log N)^\lambda} \left(1 - \frac{K}{\log N}\right)^{-\lambda}. \end{aligned}$$

The same argument with the obvious change gives a similar lower bound, which implies

$$\Lambda(\lambda) = \frac{\Gamma(N+1+\lambda)}{(N \log N)^\lambda \Gamma(N+1)} (1 + O((\log N)^{-1})),$$

uniformly in  $\lambda \in [0, 1]$ , therefore  $\log \Lambda(\lambda) = \lambda \log \log N + O((\log N)^{-1})$ . As a consequence

$$\mathbf{E}\left(\log \sum_{k=1}^N e^{x_k}\right) = \lim_{\lambda \rightarrow 0} \frac{\log \Lambda(\lambda)}{\lambda} = \log \log N + o(1),$$

which concludes the proof. □

In a similar way, we obtain the genealogy of the  $(N, \beta)$ -branching random walk.

*Proof of Theorem 1.2:* If  $\beta \in (1, \infty)$ , then Lemma 1.6 and Proposition 4.1 say that the genealogy of the  $(N, \beta)$ -BRW can be described by 1. in Theorem 3.3, with  $\alpha = \frac{1}{\beta}$  and  $\theta = 0$ . Therefore, it converges to the Bolthausen-Sznitman coalescent.

On the other hand if  $\beta = \infty$ , the genealogy of the  $(N, \infty)$ -BRW is again described by a coalescent process obtained from multinomial random variables. In this case, by Proposition 4.2 we can rewrite the coefficients  $\eta_j^N$  as

$$\eta_j^N = \frac{e^{e_j}}{\sum_{i=1}^N e^{e_i}}, \quad \text{with } e_1, \dots, e_N \text{ i.i.d. exponential random variables.}$$

Thanks to  $\mathbf{P}(e^{e_j} \geq x) = x^{-1}$ , Cortines (2016, Theorem 1.2 (c)) says that the genealogy of the  $(N, \infty)$ -BRW converges to the Bolthausen-Sznitman coalescent with  $c_N \sim N$  as  $N \rightarrow \infty$ .  $\square$

### Acknowledgements

The authors would like to thank the anonymous referee for pointing out several mistakes in a preliminary version.

### References

- E. Aïdékon and B. Jaffuel. Survival of branching random walks with absorption. *Stochastic Process. Appl.* **121** (9), 1901–1937 (2011). [MR2819234](#).
- 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). [MR2669438](#).
- 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). [MR3167886](#).
- J. Berestycki, N. Berestycki and J. Schweinsberg. The genealogy of branching Brownian motion with absorption. *Ann. Probab.* **41** (2), 527–618 (2013). [MR3077519](#).
- N. Berestycki. *Recent progress in coalescent theory*, volume 16 of *Ensaïos Matemáticos [Mathematical Surveys]*. Sociedade Brasileira de Matemática, Rio de Janeiro (2009). ISBN 978-85-85818-40-1. [MR2574323](#).
- E. Brunet and B. Derrida. Shift in the velocity of a front due to a cutoff. *Phys. Rev. E* (3) **56** (3, part A), 2597–2604 (1997). [MR1473413](#).
- E. Brunet and B. Derrida. How genealogies are affected by the speed of evolution. *Philosophical Magazine* **92** (1-3), 255–271 (2012). [DOI: 10.1080/14786435.2011.620028](#).
- 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* (3) **76** (4), 041104, 20 (2007). [MR2365627](#).
- F. Comets and A. Cortines. Finite-size corrections to the speed of a branching-selection process. *ArXiv Mathematics e-prints* (2015). [arXiv: 1505.04971](#).
- A. Cortines. The genealogy of a solvable population model under selection with dynamics related to directed polymers. *Bernoulli* **22** (4), 2209–2236 (2016). [MR3498028](#).
- O. Couronné and L. Gerin. A branching-selection process related to censored Galton-Walton processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** (1), 84–94 (2014). [MR3161523](#).
- R. Durrett and D. Remenik. Brunet-Derrida particle systems, free boundary problems and Wiener-Hopf equations. *Ann. Probab.* **39** (6), 2043–2078 (2011). [MR2932664](#).
- P. Maillard. Speed and fluctuations of  $N$ -particle branching Brownian motion with spatial selection. *Probab. Theory Related Fields* **166** (3-4), 1061–1173 (2016). [MR3568046](#).
- B. Mallein. Branching random walk with selection at critical rate. *ArXiv Mathematics e-prints* (2015a). [arXiv: 1502.07390](#).
- B. Mallein.  $N$ -branching random walk with  $\alpha$ -stable spine. *ArXiv Mathematics e-prints* (2015b). [arXiv: 1503.03762](#).

- M. Möhle and S. Sagitov. A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.* **29** (4), 1547–1562 (2001). [MR1880231](#).
- C. Mueller, L. Mytnik and J. Quastel. Effect of noise on front propagation in reaction-diffusion equations of KPP type. *Invent. Math.* **184** (2), 405–453 (2011). [MR2793860](#).
- M. Pain. Velocity of the  $L$ -branching Brownian motion. *Electron. J. Probab.* **21**, Paper No. 28, 28 (2016). [MR3492932](#).
- J. Pitman. Coalescents with multiple collisions. *Ann. Probab.* **27** (4), 1870–1902 (1999). [MR1742892](#).
- J. Pitman and M. Yor. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25** (2), 855–900 (1997). [MR1434129](#).