A N-branching random walk with random selection

Aser Cortines, Bastien Mallein

A N-branching random walk with random selection

Aser Cortines* and Bastien Mallein†

May 27, 2016

Abstract

We consider an exactly solvable model of branching random walk with random selection, describing the evolution of a fixed number N of individuals in the real line. At each time step $t \to t+1$, the individuals reproduce independently at random making children around their positions and the N individuals that form the $(t+1)$th generation are chosen at random among these children according to the Gibbs measure at temperature β. We compute the asymptotic speed and the genealogical behaviour of the system.

1 Introduction

In a general sense, a *branching-selection particle system* is a stochastic process of particles evolving according to the two following steps.

Branching step: every individual currently alive in the system dies giving birth to new particles, that are positioned according to independent copies of a given point process, shifted by the position of their parent.

Selection step: some of new-born individuals are selected according to a given rule, to reproduce at the next step, while the other particles are "killed".

From a biological perspective these systems model the competition between individuals in an environment with limited resources. Different methods can be used to select individuals, for example, an *absorbing barrier* killing particles that go below it [1, 2, 11]. An other example is the case where only the N rightmost individuals are chosen to constitute the next generation [4, 6], the so-called N-branching random walk. Branching processes with selection of the N rightmost individuals have been the subject of recent studies [7, 10, 12], and several conjectures on this processes remain open, such as the asymptotic behaviour of the genealogy or the second order in the asymptotic behaviour of the speed. In this paper, we introduce a branching random walk, in which the individuals are randomly selected according to the Gibbs measure at temperature β.

Branching-selection particle systems as above are of physical interest [4, 6] and can be related to reaction-diffusion phenomena. Based in numerical simulations [4] and the study of solvable models [6], it has been predicted that the dynamical and structural aspects of these models satisfy universal properties depending on the behaviour of the right-most particles. These conjectures have been recently proved for some specific models, for example, the asymptotic velocity of the N-branching random walk converges to a limiting value at the slow rate $(\log N)^{-2}$. Its continuum analogue being the speed of travelling fronts of the noisy KPP equations [14]. An other example, in which the finite-size correction to the speed of a branching-selection particle system is explicitly computed can be found in [8].

The conjectures also comprise the genealogical structure of those 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 linages. That is, $\Pi^N_n(t)$ is a Markov process in $\mathcal{P}_n$ the set of partitions (or equivalent classes) of $[n] := \{1, \dots, n\}$ such that i and j belong the same equivalent class if the individual i and j have a common ancestor t generations backwards in time. Notice

*Supported by the Israeli Science Foundation grant 1723/14
†DMA, ENS

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 been conjectured [6] that the genealogical trees of these models converge to those of a Bolthausen-Sznitman coalescent and the average coalescence times scale like the logarithmic of the population's 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 [13]. Mathematically, it is difficult to verify such conjectures and they have only been proved for some particular models [2, 6, 9].

In this article, we study an exactly solvable model of discrete-time branching-selection particle system that we name by (N,β)-*branching random walk*, or (N,β)-BRW for short, where $N \in \mathbb{N}$ and $\beta > 1$. It consists in a fixed number N of particles, initially at the positions $\left(X_0^N(1), X_0^N(2), \dots, X_0^N(N)\right) \in \mathbb{R}^N$. Then, with $\{\mathcal{P}_t(j), j \leq N, t \in \mathbb{N}\}$ a family of i.i.d. Poisson point processes (PPP) with intensity measure $\mathrm{e}^{-x}\mathrm{d}x$, the process evolves as follows:

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

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

ii. Let $(X_t^N(i); i \in \mathbb{N})$ be a random shuffling of $(\Delta_t(k); k \in \mathbb{N})$ such that

$$\mathbf{P}\left(X_t^N(i) = \Delta_t(j) \middle| \left(\Delta_t(k)\right)_k, X_t^N(1), \dots, X_t^N(i-1)\right)$$
$$= \mathbf{1}_{\{\Delta_t(j) \not\in \{X_t^N(1), \dots X_t^N(i-1)\}\}} \frac{\mathrm{e}^{\beta \Delta_t(j)}}{\sum_{k=1}^{+\infty} \mathrm{e}^{\beta \Delta_t(k)} - \sum_{k=1}^{i-1} \mathrm{e}^{\beta X_t^N(k)}},$$

and write $A_t^N(i) = j$ if $X_t^N(i) \in \left\{X_{t-1}^N(j) + p, p \in \mathcal{P}_t(j)\right\}$, that is, if $X_t^N(i)$ is an offspring of $X_{t-1}^N(j)$. Then, the positions of the particles in generation t is given by $X_t^N(1), \dots, X_t^N(N)$ the N first terms in the sequence $(X_t^N(i); i \in \mathbb{N})$.

As we show in Proposition 1.3, the model can only be defined for $\beta > 1$. Nevertheless, one may interpret, in a loose sense, the case $\beta \leq 1$ to be the classical Wright-Fisher model, in which individuals are selected uniformly at random (regardless of their positions). In contrast with the examples already treated in the literature, when $1 < \beta < \infty$ one does not necessarily select the rightmost children. Observe that letting $\beta \to +\infty$, one recovers the model of the N-BRW from [6]. The next result concerns the speed of the (N,β)-BRW.

Theorem 1.1. *For any $N \in \mathbb{N}$ and $\beta > 1$, there exists $v_{N,\beta}$ such that*

$$\lim_{t \to +\infty} \frac{\max_{j \leq N} X_t^N(j)}{t} = \lim_{n \to +\infty} \frac{\min_{j \leq N} X_t^N(j)}{t} = v_{N,\beta} \qquad a.s. \tag{1.1}$$

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

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

Theorem 1.2. *Let $\beta > 1$ and c_N be the probability that two individuals uniformly chosen at random have a common ancestor one generation backwards in time. Then, $c_N \sim (\log N)^{-1}$ as $N \to +\infty$ and the rescaled coalescent process $\left(\Pi^N(\lfloor t/c_N \rfloor), t \geq 0\right)$ converges in distribution toward the* Bolthausen-Snitzmann coalescent.

It was already proved in [6] that the rescaled genealogy of an (N,∞)-branching random walk converges toward the Bolthausen-Snitzmann coalescent. Theorem 1.2 proves the robustness of this result: when an individual goes far ahead of the rest of the population, its offspring overrun the next generation and the Bolthausen-Snitzmann coalescent is to be expected. Nevertheless, if the modifications made to the process are important enough, different coalescent behaviour might emerge.

For example, using the same reasoning as in Section 4, the genealogy of a process with a similar selection procedure as the (N,β)-branching random walk, except the first individual selected $X_t^N(1)$ is killed, and individuals $X_t^N(2)$ to $X_t^N(N+1)$ will reproduce, have a genealogy converging toward the Kingman coalescent. Moreover, in the same way as in [5], the explicit formulas we obtain in Section 4 suggest that conditioning the (N,β)-BRW to be faster (or slower) than expected might modify the asymptotic behaviour of the genealogy. A third direction would be to consider (N,β_N)-branching random walks, with $\beta_N \to 1$. Numerical simulations suggest that as long as β_N converges to 1 fast enough, the genealogy of the process behaves asymptotically as the genealogy of a classical Wright-Fisher model (Kingman coalescent). However, setting $\beta_N = 1 + \frac{\lambda}{\log N}$ different limit coalescent is to be expected, interpolating between the star-shaped coalescent for a given $\lambda_c > 0$ and the Kingman coalescent for $\lambda \to 0$.

The (N,β)-branching random walk and preliminary results

In this short section, we use Poisson point processes to obtain basic properties of the (N,β)-BRW, such as the existence of $v_{N,\beta}$. We first prove that for any $\beta > 1$ the model is well-defined and that the generations are independent.

Proposition 1.3. *The (N,β)-BRW is well-defined for all $N \in \mathbb{N}$ and $\beta > 1$. Moreover, if we set $X_t^N(\mathrm{eq}) := \log \sum_{j=1}^N e^{X_t^N(j)}$, then the sequence $\left(\sum_{k\in\mathbb{N}} \delta_{\Delta_k(t+1) - X_t^N(\mathrm{eq})} : t \in \mathbb{N}\right)$ is an i.i.d. family of Poisson point processes with intensity measure $e^{-x}dx$.*

Remark 1.4. We observe that $X_t^N(\mathrm{eq})$ is an "equivalent position" for the particles at generation t: the set of individuals belonging to the $(t+1)$th generation are distributed as if they were made by a unique particle positioned at $X_t^N(\mathrm{eq})$.

Proof. Fix $N \in \mathbb{N}$ and $\beta > 1$. We assume that the process has been constructed up to time t and let $X_t^N(1), \ldots, X_t^N(N)$ be the positions of the N particles. By the invariance of superposition of PPP, conditionally on $(X_t^N(j), j \le N)$, $\left\{X_t^N(j) + p; p \in \mathcal{P}_t(j), j \le N\right\}$ is also a PPP with intensity measure $\sum_{i=1}^N e^{-(x - X_t(i))}dx = e^{-(x - X_t^N(\mathrm{eq}))}dx$.

Therefore, with probability one: all points have multiplicity one, $(\Delta_k(t+1);\ k \in \mathbb{N})$ is uniquely defined and $\sum e^{\beta \Delta_k(t+1)} < \infty$. Thus, both the selection mechanism and $(A_{t+1}^N(i), i \le N)$ are well-defined, proving the first claim. Moreover, $(\Delta_k(t+1) - X_t(\mathrm{eq});\ k \in \mathbb{N})$ is a PPP$(e^{-x}dx)$, which is independent of the t first steps of the (N,β)-BRW, proving the second claim. $\square$

In the next lemma, we prove the existence of the *speed* $v_{N,\beta}$ of the (N,β)-branching random walk. The asymptotic behaviour of $v_{N,\beta}$ is obtained in Section 4.

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

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

Proof. By Proposition 1.3, $\left(X_{t+1}^N(\mathrm{eq}) - X_t^N(\mathrm{eq}) :\ t \in \mathbb{N}\right)$ are i.i.d. random variables with finite mean, therefore

$$\lim_{t\to+\infty} \frac{X_t^N(\mathrm{eq})}{t} = v_{N,\beta} \quad \text{a.s. by the law of large numbers.}$$

Notice that both $\left(\max X_t^N(j) - X_{t-1}^N(\mathrm{eq})\right)_t$ and $\left(\min X_t^N(j) - X_{t-1}^N(\mathrm{eq})\right)_t$ are sequences of i.i.d. random variables with finite mean as well, which yields (1.1). $\square$

In a similar way, we are able to obtain a simple structure for the genealogy of the process and to describe its law conditionally on the position of every particle $\mathcal{H} = \sigma(X_t^N(j), j \le N, t \ge 0)$.

Lemma 1.6. *Denote by $A_t^N := \left(A_t^N(1), \ldots, A_t^N(N)\right) \in \{1, \ldots, N\}^N$, then the sequence $(A_t^N)_{t\in\mathbb{N}}$ is i.i.d. with common distribution determined by the conditional probabilities*

$$\mathbf{P}\left(A_{t+1}^N = \overline{k} \mid \mathcal{H}\right) = \theta_t^N(k_1)\ldots\theta_t^N(k_N), \quad \textit{where} \quad \overline{k} = (k_1, \ldots, k_N) \in \{1, \ldots, N\}^N;$$

$$\textit{and} \quad \theta_t^N(k) := \frac{e^{X_t^N(k)}}{\sum_{i=1}^N e^{X_t^N(i)}}. \tag{1.2}$$

Proof. We add marks $i \in \{1,\ldots,N\}$ to the points in $\sum \delta_{\Delta_{t+1}(k)-X_t^N(\mathrm{eq})}$ such that a point x has mark i if it is a point coming from $X_t^N(i)+\mathcal{P}_t(i)$. The invariance under superposition of independent PPP says that

$$\mathbf{P}\big(x \in X_t^N(i)+\mathcal{P}_t(i) \mid \mathcal{F}\big) = \frac{\mathrm{e}^{-(x-X_t^N(i))}}{\sum_{j=1}^N \mathrm{e}^{-(x-X_t^N(j))}} = \frac{\mathrm{e}^{X_t^N(i)}}{\sum_{j=1}^N \mathrm{e}^{X_t^N(j)}}.$$

By construction, $A_t^N(i)$ is given by the mark of $X_{t+1}^N(i)$, which yields (1.2). The independence between the A_t^N can be easily checked using Proposition 1.3. $\square$

Organisation of the paper. In Section 2, we obtain some technical lemmas concerning the Poisson-Dirichlet distributions. In Section 3, we focus 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,β)-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

We focus in this section on the two-parameters Poisson-Dirichlet distribution, with parameters $\alpha \in (0,1)$ and $\theta > -\alpha$ (shorten into $\mathrm{PD}(\alpha,\theta)$ distribution), as defined by Pitman and Yor [16].

Definition 2.1 (Definition 1 in [16])**.** 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 $\mathrm{Beta}(1-\alpha, \theta + j\alpha)$ distribution and write

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

Then, denote by $U_1 \geq U_2 \geq \cdots$ the values of (V_n) ranked in the decreasing order. The sequence (U_n) as above is called the *Poisson-Dirichlet distribution with parameters* (α,θ).

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

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

for this reason we say that (V_n) follows the *size-biased pick* from $\mathrm{PD}(\alpha,\theta)$. It is well known that there exists a strong connexion between PD distributions and PPP [16].

Proposition 2.2 (Proposition 10 in [16])**.** *Let $x_1 > x_2 > \ldots$ be the points of a* $\mathrm{PPP}(\mathrm{e}^{-x}\mathrm{d}x)$*, we write $L = \sum_{j=1}^{+\infty} \mathrm{e}^{\beta x_j}$ and $U_j = \mathrm{e}^{\beta x_j}/L$. Then $(U_j, j \geq 1)$ has distribution* $\mathrm{PD}(\beta^{-1},0)$ *and*

$$\lim_{n\to+\infty} n^\beta U_n = 1/L \quad \textit{a.s.}$$

Notice from Propositions 1.3 and 2.2 that $(e^{\beta X_n^N(i)}/\sum_{j=1}^{+\infty} e^{\beta \Delta_n(j)})$ has the distribution of (V_i) the size biased pick from $\mathrm{PD}(\beta^{-1},0)$, which makes the model solvable. We explore in Section 4 this connection to study the genealogical and dynamical properties of the (N,β)-BRW.

Remark 2.3 (Change of parameter)*.* Let (U_n) be a sequence with $\mathrm{PD}(\alpha,\theta)$ distribution. Set $(\widetilde{U}_n)$ for the values of $\{\frac{U_n}{1-V_1}, n \in \mathbb{N} : U_n \neq V_1\}$ ranked in the decreasing order, $(\widetilde{U}_n)$ has $\mathrm{PD}(\alpha, \alpha+\theta)$ distribution and is independent of V_1. In particular, $(\frac{V_j}{1-V_1}, j \geq 2)$ is a size-biased $\mathrm{PD}(\alpha,\alpha+\theta)$.

Henceforth in this section, we fix $\alpha \in (0,1)$ and $\theta > -\alpha$. Let also c and C be positive constants, that may change from line to line and implicitly depend on α and θ. We will now focus attention on the convergence and the concentration properties of

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

The next result concerns the convergence and concentration properties of the multiplicative martingale $\prod(1-Y_i)^\alpha$ appearing in Σ_n.

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

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

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

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

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

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

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

Observe also that $\lim_{n\to+\infty} \mathbf{E}(M_n^{\gamma/2})\,\mathbf{E}(M_n^\gamma)^{-1/2} > 0$, thus Kakutani's theorem yields the a.s. and $\mathbb{L}^1$ convergence of $(M_n^\gamma / \mathbf{E}(M_n^\gamma))$ as $n \to +\infty$. As a consequence, setting $M_\infty = \lim_{n\to+\infty} M_n n^{\frac{1-\alpha}{\alpha}}$, (2.2) holds. In particular, if $0 < \gamma < \theta+\alpha$, let $C_\gamma := \sup_{n\in\mathbb{N}} \mathbf{E}[M_n^{-\gamma}]/n^{\gamma\frac{1-\alpha}{\alpha}} < \infty$ then

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

by Doob's martingale inequality, proving (2.3). □

We now focus on the converge and concentration properties of the sum $\sum Y_j^\alpha j^{\alpha-1}$.

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

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

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

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

Proof. As Y_j has Beta$(1-\alpha, \theta+j\alpha)$ distribution, we have

$$\mathbf{E}((jY_j)^\alpha) = \frac{1}{\alpha^\alpha\Gamma(1-\alpha)} + \mathcal{O}(1/j) \quad \text{and} \quad \mathbf{V}\mathrm{ar}((jY_j)^\alpha) = \frac{\Gamma(1+\alpha)\Gamma(1-\alpha)-1}{\alpha^{2\alpha}\Gamma(1-\alpha)^2} + \mathcal{O}(1/j),$$

which implies that $\sum \mathbf{V}\mathrm{ar}(Y_j^{\alpha-1}) < +\infty$ and that $\mathbf{E}(S_n) = \Psi_\alpha \log n + C_S + o(1)$ with $C_S \in \mathbb{R}$. Since $Y_j^\alpha - \mathbf{E}(Y_j^\alpha) \in (-1,1)$ a.s. we deduce from Kolmogorov's three-series theorem that $S_n - \mathbf{E}(S_n)$, and hence $S_n - \Psi_\alpha \log n$, converge a.s. proving the first claim.

We now bound $\mathbf{P}(S_n - \mathbf{E}(S_n) \geq y)$, notice that for any $y \geq 0$ and $\lambda > 0$,

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

Take $c > 0$ such that $\mathrm{e}^x \leq 1 + x + cx^2$ for all $x \in (-1,1)$, then, since $Y_j^\alpha - \mathbf{E}(Y_j^\alpha) \in (-1,1)$ a.s. we obtain the inequalities

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

Consequently, for any $\alpha \in (0,1)$ and $\theta \geq 0$, there exists $c = c(\alpha,\theta)$ such that

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

where we use that $\sum_{j^{1-\alpha}\leq\lambda} j^{\alpha-1} < \lambda^{\frac{1}{1-\alpha}}$. Denote by $\varrho := (2-\alpha)/(1-\alpha) > 2$, then there exists $C = C(\alpha,\theta) > 0$ such that for any $n \in \mathbb{N}$ and $y \geq 0$,

$$\mathbf{P}\left(S_n - \mathbf{E}(S_n) \geq y\right) \leq C \exp\left(-\lambda y + C\lambda^{\varrho}\right).$$

We optimize this equation in $\lambda > 0$ to obtain

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

with $C^{1/(1-\varrho)}\left[\varrho^{1/(1-\varrho)} - \varrho^{\varrho/(1-\varrho)}\right] > 0$, because $\varrho > 1$. Since the same argument holds for $\mathbf{P}(S_n - \mathbf{E}(S_n) \leq -y)$, there exists $C > 0$ such that $\mathbf{P}\left(|S_n - \mathbf{E}(S_n)| \geq y\right) \leq C\exp\left(-y^{\varrho/(\varrho-1)}/C\right)$, proving the second statement. □

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*

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

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

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

Proof. We observe that $\Sigma_n = \sum_{j=1}^n (S_j - S_{j-1}) j^{1-\alpha} M_{j-1}^\alpha$, with $S_0 := 0$. Moreover, by Lemmas 2.4 and 2.5 we have

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

Consequently, by Stolz-Cesàro theorem, we obtain $\lim_{n\to+\infty} \frac{\Sigma_n}{\log n} = \Psi_\alpha M_\infty^\alpha$ for almost every event of the probability space. Moreover, expanding $\left(\Sigma_n\right)^2$ we obtain

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

Therefore, $\sup \mathbf{E}\left[(\Sigma_n/\log n)^2\right] < +\infty$, implying the $\mathbb{L}^1$ convergence of $\Sigma_n/\log n$ toward $\Psi_\alpha M_\infty^\alpha$.

We now bound $\mathbf{P}(\Sigma_n \leq u\log n)$, assuming first that $u \geq 1/n$. We write

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

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

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

Let $0 < \varepsilon < 1/2$ and set $t = u^\varepsilon \log n$. Since $\lim_{n\to+\infty} \frac{E(S_n)}{\log n} = \Psi_\alpha$, there exists a constant $c > 0$ depending only on α such that $u^\varepsilon \log n \le E[S_n]$ for any $u \le c$. By decreasing c we can also assume that $C^{-1}(\mathbf{E}[S_n] - y^\varepsilon \log n) < a \log n$ for all $\varepsilon < 1/2$, where $a > 0$ is chosen conveniently small, therefore

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

Observe that

$$C \exp\left(-(\eta \log n)^{\varrho/(\varrho-1)}\right) < C_{\gamma'} u^{(1-\varepsilon)\gamma'/\alpha}, \quad \text{for all } u \in \left[\frac{C}{C_{\gamma'}} \exp\left(-\frac{\alpha}{\gamma'}(\eta\log n)^{\frac{\varrho}{(\varrho-1)}}\right), c\right],$$

and that $\mathrm{e}^{-\frac{\alpha}{\gamma'}(\eta \log n)^{\frac{\varrho}{\varrho-1}}} \ll 1/n$. Therefore, taking $\gamma = (1-\varepsilon)\gamma' < \alpha + \theta$, there exists D_γ such that $\mathbf{P}\left(\Sigma_n \le u \log n\right) \le D_\gamma u^{\frac{\gamma}{\alpha}}$, for all n large enough and $u \in [\frac{1}{n}, +\infty)$.

Assume now that $u \le 1/n$ and let $j^* \in \mathbb{N}$ be such that $(1-\alpha)j^* > \gamma$. As $V_j > 0$, we have

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

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

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

Thus, we prove by recurrence that $Y_j^\alpha \le \frac{u\log n}{1 - \frac{(j-1)\log n}{n}}$. In effect, we have $Y_1^\alpha \le u \log n$. Moreover, if for any $i < j$, $Y_i^\alpha \le \frac{u \log n}{1-\frac{(i-1)\log n}{n}}$, we have

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

yielding $Y_j^\alpha \le \frac{u \log n}{1-\frac{(j-1)\log n}{n}}$. Thus, for any $j \le j^*$ and n large enough, $Y_j^\alpha < 2u\log n$, yielding

$$\mathbf{P}(\Sigma_n < u \log n) \le \prod_{j=1}^{j^*} \mathbf{P}\left(Y_j^\alpha < 2u \log n\right).$$

A crude estimate of the probability distribution function of the beta distribution, can be used to bound this quantity by $C u^{\frac{\gamma}{\alpha}} \varrho\big(u^{(j^*(1-\alpha)-\gamma}(\log n)^{j^*(1-\alpha)}\big)$, where C is an explicit constant. Since $u < 1/n$ and $j^*(1-\alpha) - \gamma > 0$, the term inside the parentheses goes to zero uniformly in u. Therefore, there exists $D_\gamma > 0$ such that for any $n \ge 1$ and $u \ge 0$, $\mathbf{P}\left(\Sigma_n < u \log n\right) \le D_\gamma u^{\frac{\gamma}{\alpha}}$, concluding the proof. □

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

Corollary 2.7. *If (U_n) is a PD$(\alpha, 0)$, then $\Psi_\alpha M_\infty^\alpha = L^{-\alpha}$, where $1/L = \lim_{n\to+\infty} n^{1/\alpha} U_n$.*

Proof. By Proposition 2.2, $L := \lim_{n\to+\infty} \frac{1}{n^{1/\alpha}U_n}$ exists a.s. and by Lemma 2.6, we have

$$\Psi_\alpha M_\infty^\alpha \sim \frac{1}{\log n}\sum_{j=1}^n V_j^\alpha \le \frac{1}{\log n}\sum_{j=1}^n U_j^\alpha \sim L^{-\alpha} \qquad \text{as } n \to +\infty,$$

thus, $\Psi_\alpha M_\infty^\alpha \le L^{-\alpha}$ a.s. By Lemma 2.4, for any $p > -1$ we have

$$\mathbf{E}\left[(\Psi_\alpha M_\infty^\alpha)^p\right] = \frac{\Gamma(p+1)}{\Gamma(p\alpha+1)}\Gamma(1-\alpha)^{-p},$$

and $L^{-\alpha}$ have the same moments by [16, Equation (30)]. Consequently, these two random variables have the same distribution (the Mittag-Leffler (α) distribution, up to the same multiplicative constant), which implies that $\Psi_\alpha M_\infty^\alpha = L^{-\alpha}$ a.s. □

3 Convergence of discrete exchangeable coalescent processes

In this section, we focus attention on 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 now give a brief introduction to the coalescent theory stating the main results that will be used here. Nevertheless, we recommend [3] (from where we borrow the approach) for a detailed account and the proofs.

Let $\mathcal{P}_n$ be the set of partitions (or equivalent classes) of $[n] := \{1, \ldots, n\}$ and $\mathcal{P}_\infty$ the set of partitions of $\mathbb{N} = [\infty]$. A partition $\pi \in \mathcal{P}_n$ is represented by blocks $\pi(1), \pi(2), \ldots$ 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 $\pi^\sigma := \big\{\{\sigma(j), j \in \pi(i)\}, i \in [n]\big\}$, with $\sigma \in S_n$. If $m < n$, one can define the projection of $\mathcal{P}_n$ onto $\mathcal{P}_m$ by the restriction $\pi|_m = \{\pi(j) \cap [m]\}$. Finally, for $\pi, \pi' \in \mathcal{P}_n$, we define the *coagulation of* π *by* π' to be the partition $\mathrm{Coag}(\pi, \pi') = \big\{\cup_{i \in \pi'(j)} \pi(i);\ j \in \mathbb{N}\big\}$.

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

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

We say that $\Pi(t)$ is *exchangeable* if for all permutation σ, $\Pi^\sigma(t)$ and $\Pi(t)$ have the same distribution. There exists a natural metric in $\mathcal{P}_n$ and hence one can study the weak convergence of processes in $\mathcal{D}\big([0,\infty), \mathcal{P}_n\big)$, see [3] for the definitions. Without going deeply into details, we say a process $\Pi^N(t) \in \mathcal{P}_\infty$ converges in the *Skorokhod sense* (or in distribution) towards $\Pi(t)$, if $\Pi^N(t)|_n \in \mathcal{P}_n$ converge to $\Pi(t)|_n$ under the Skorokhod topology of $\mathcal{D}\big([0,\infty), \mathcal{P}_n\big)$, for any $n \in \mathbb{N}$.

An important class of continuous-time exchangeable coalescent processes in $\mathcal{P}_\infty$ are the so-called Λ-coalescents [15], 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

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

Among such, we distinguish the Beta$(2-\lambda, \lambda)$-coalescent obtained from $\Lambda(\mathrm{d}x) = \frac{x^{1-\lambda}(1-x)^{\lambda-1}}{\Gamma(\lambda)\Gamma(2-\lambda)}\mathrm{d}x$, where $\lambda \in (0,2)$. The case $\lambda = 1$ (uniform measure) being the celebrated *Bolthausen-Sznitman coalescent.* These coalescent appear in the coming results as the limit distribution of a family of ancestral partition processes.

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 a sufficient criteria for its convergence in distribution as $N \to \infty$. Let $(\eta_1^N, \ldots \eta_N^N)$ be an N-dimensional random vector satisfying

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

Conditionally on a realisation of (η_j^N), let $\big\{\xi_j; j \leq N\big\}$ be i.i.d. random variables satisfying $\mathbf{P}(\xi_j = k | \eta^N) = \eta_k^N$ and define the partition $\pi_N = \big\{\{j \leq N : \xi_j = k\};\ k \leq N\big\}$. With $(\pi_t;\ t \in \mathbb{N})$ i.i.d. copies of π_N, define the discrete time coalescent $\Pi^N(t)$ as follows:

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

We now prove a convergence criterion for these class of coalescent processes. We assume that there exists a sequence L_N and a function $f : (0,1) \to \mathbb{R}_+$ such that

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

We also denote by $c_N = \sum_{j=1}^N \mathbf{E}\left[(\eta_j^N)^2\right]$, from the point of view of ancestral linages it stands for the of probability that two individuals have a common ancestor one generation backward in time.

Lemma 3.1. *With $(\Pi_n^N(t); t \in \mathbb{N})$ defined as above, assume that (3.1) holds and that*

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

Then $c_N \sim_{N\to\infty} L_N^{-1} \int_0^1 2xf(x)\mathrm{d}x$ and the re-scaled coalescent process $\left(\Pi^N(t/c_N); t \in \mathbb{R}_+\right)$ converges in distribution to the Λ-coalescent, with Λ satisfying $\int_x^1 \frac{\Lambda(\mathrm{d}y)}{y^2} = f(x)$.

Proof. Denote by $\nu_k = \#\{j \leq N : \xi_j = k\}$, then $(\nu_1, \ldots, \nu_N)$ has multinomial distribution with N trials and (random) probabilities outcomes η_i^N. By [13, Theorem 2.1], the convergence of finite dimensional distribution of $\Pi^N(t)$ can be obtained from the convergence of the factorial moments of ν, that is

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

where $(n)_a := n(n-1)\ldots(n-a+1)$. We use that $(\nu_1, \ldots, \nu_N)$ is multinomial distributed to obtain $\mathbf{E}\left[(\nu_{i_1})_{b_1} \ldots (\nu_{i_a})_{b_a}\right] = (N)_b \, \mathbf{E}\left[\eta_{i_1}^{b_1} \ldots \eta_{i_a}^{b_a}\right]$, see [9, Lemma 4.1] for rigorous a proof. Therefore, we only have to show that for all b and $a \geq 2$

$$\lim_{N\to+\infty} c_N^{-1} \sum_{i_1=1}^N \mathbf{E}\left[(\eta_{i_1}^N)^b\right] = \int_0^1 x^{b-2}\Lambda(\mathrm{d}x) \quad \text{and} \quad \lim_{N\to+\infty} c_N^{-1} \sum_{\substack{i_1,\ldots,i_a=1 \\ \text{all distinct}}}^N \mathbf{E}\left[(\eta_{i_1}^N)^{b_1} \ldots (\eta_{i_a}^N)^{b_a}\right] = 0.$$

We first compute the asymptotic behaviour of $c_N = \sum \mathbf{E}[(\eta_i^N)^2]$, by dominated convergence,

$$L_N \, \mathbf{E}\left[\left(\eta_1^N\right)^2\right] = \int_0^1 2x L_N \mathbf{P}\left(\eta_1^N > x\right) \mathrm{d}x = \int_0^1 2xf(x)\mathrm{d}x = \int_0^1 \Lambda(\mathrm{d}x) < +\infty.$$

Since η_i^N are ordered and sum up to 1, we also get $\mathbf{E}\left[\left(\eta_2^N\right)^2 + \ldots + \left(\eta_N^N\right)^2\right] \leq \mathbf{E}\left[\eta_2^N(1-\eta_1)\right]$. Moreover, $L_N \, \mathbf{E}\left(\eta_2^N\right) \to 0$ as $N \to \infty$, which implies that $\lim_{N\to+\infty} L_N c_N = \int_0^1 2xf(x)\mathrm{d}x$. A similar calculation shows that for $b \geq 2$

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

where $\lambda_{b,b}$ is the rate at which b blocks merge given that there are b blocs 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

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

where we use that $\left(\eta_2^N\right)^b + \ldots + \left(\eta_N^N\right)^b \leq \eta_2^N + \ldots + \eta_N^N = 1 - \eta_1^N < 1$ for every $b \geq 1$. Since $L_N \, \mathbf{E}\, \eta_2^N \to 0$ as $N \to \infty$, we conclude the proof. $\square$

The next lemma provides sufficient conditions for the convergence towards the Kingman's coalescent.

Lemma 3.2. *With $(\Pi_n^N(t); t \in \mathbb{N})$ defined as above, assume that (3.1) holds,*

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

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

Proof. We start checking the first claim. Using the same arguments as in Lemma 3.1 one gets

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

Thus, by Fatou lemma, we have $\liminf_{N \to +\infty} L_N c_N = +\infty$, which proves the first claim. To check the second claim, we show that $\sum_{i=1}^N \mathbf{E}[(\nu_i)_3]/(N)_3 c_N \to 0$ as $N \to \infty$, *see e.g.* [3, Theorem 2.5], which is equivalent to $\sum_{i=1}^N \mathbf{E}[(\eta_i^N)^3]/c_N \to 0$. Applying Hölder inequality, we have that

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

Choosing $\lambda \in (0,2)$ to be the unique solution of $2(3-\lambda)/(2-\lambda) = n+4$, we have

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

by (3.3). To estimate the remaining terms in $\sum_{i=1}^N \mathbf{E}[(\eta_i^N)^3]/c_N$, it suffices to proceed as in the proof of Lemma 3.1 and show that they converge to zero, which proves the statement. □

3.2 The Poisson-Dirichlet distribution case

In this section, we construct a coalescent using the PD distribution to govern the coefficients η_i^N defined in Section 3.1. With $\alpha \in (0,1)$ and $\theta > -\alpha$, let $(V_j, j \geq 1)$ be the size biased pick from a $\mathrm{PD}(\alpha, \theta)$ partition. Then, define

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

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

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

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

- *If $\theta \in (-\alpha, \alpha)$, then $c_N \sim_{N \to +\infty} (1-\theta/\alpha)/L_N$ and the process $(\Pi^N(t/c_N), t \geq 0)$ converges toward a Beta$(2-\lambda, \lambda)$-coalescent.*
- *Otherwise, $\lim_{N \to +\infty} c_N L_N = +\infty$ and $(\Pi^N(t/c_N))$ converges toward a Kingman coalescent.*

To prove this result, we first study the asymptotic behaviour of θ_1^N.

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

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

Moreover, there exists $C > 0$ such that for all $x \in (0,1)$, $\sup_{N \in \mathbb{N}} L_N \mathbf{P}\left(\theta_1^N > x\right) \leq C x^{-\lambda}$.

Proof. Let $\Sigma'_N := \sum_{j=2}^N \left(\frac{V_j}{1-Y_1}\right)^\alpha$, then Σ'_N and Y_1 are independent and, by Remark 2.3, Σ'_N has the law of Σ_N in Lemma 2.6, with a $\mathrm{PD}(\alpha, \alpha+\theta)$ distribution. Therefore, by Lemma 2.6, for all $\varepsilon \in (0,1)$, there exists $C = C(\varepsilon)$ and $P \in \mathbb{N}$ such that

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

We observe that

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

Consequently, setting $u = (\frac{1-x}{x \log N})(1/y-1)^\alpha$, we have

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

Using (3.4), for any $N \geq 1$ large enough, we have

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

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

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

and hence $\lim_{N\to+\infty} L_N \mathbf{P}\left(\theta_1^N > x\right) = \left(\dfrac{1-x}{x}\right)^\lambda \dfrac{1}{\lambda\Gamma(\lambda)\Gamma(2-\lambda)}$. ☐

This result is used to study the asymptotic behaviour 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*

$$\left|\mathbf{P}\left(\theta_{(1)}^N > x\right) - \mathbf{P}\left(\theta_1^N > x\right)\right| \leq C(x \log N)^{\varepsilon - 2 - \theta/\alpha}, \qquad \textit{for all } x \in (0,1) \textit{ and } N \textit{ large enough.}$$

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

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

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

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

where $\Sigma'_N = \sum V_j^\alpha/(1-Y_1)^\alpha$. Applying Lemma 2.6 once again we obtain

$$\mathbf{P}\left(\theta_{(1)}^N > x, V_1 \leq 1/2\right) \leq C(x \log N)^{\varepsilon - 1 - \frac{\alpha+\theta}{\alpha}}, \qquad \text{for all } N \text{ sufficiently large,}$$

finishing the proof. ☐

We now study the asymptotic behaviour of the second maximal value $\theta^N_{(2)}$.

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

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

Proof. We use the same method as in the previous lemma. We observe that

$$\begin{aligned}
&\mathbf{P}\big(\theta^N_{(2)} > x\big) \\
&\quad = \mathbf{P}\big(\theta^N_{(2)} > x, V_1 \leq 1/2\big) + \mathbf{P}\big(\theta^N_{(2)} > x, V_1 > 1/2, V_2 < 1/3\big) + \mathbf{P}\big(\theta^N_{(2)} > x, V_1 > 1/2, V_2 > 1/3\big) \\
&\quad \leq \mathbf{P}\big(\theta^N_{(1)} > x, V_1 \leq 1/2\big) + \mathbf{P}\big(\theta^N_{(2)} > x, V_1 > 1/2, V_2 < 1/3\big) + \mathbf{P}\big(\theta^N_2 > x\big).
\end{aligned}$$

By Lemma 3.5, we have $\mathbf{P}\big(\theta^N_{(1)} > x, V_1 \leq 1/2\big) \leq C(x \log N)^{\varepsilon-2-\theta/\alpha}$, moreover, proceeding as in Lemma 3.4, we obtain

$$\mathbf{P}\big(\theta^N_2 > x\big) = \mathbf{P}\big(V_2^\alpha(1-x) - xV_1^\alpha > x(1 - V_1 - V_2)^\alpha \Sigma''_N\big) \leq C(x \log N)^{\varepsilon-2-\theta/\alpha},$$

with $\Sigma''_N := (1 - V_1 - V_2)^{-\alpha} \sum_{j=3}^N V_j^\alpha$ and independent of (V_1, V_2). By a similar argument, we get

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

concluding the proof. ∎

Proof of Theorem 3.3. Let $\alpha \in (0,1)$, $\theta > -\alpha$ and $\varepsilon \in (0,1)$, we recall that $L_N = c_{\alpha,\theta}(\log N)^\lambda$. Using Lemma 3.5, there exists $C > 0$ such that for any $x \in (0,1)$,

$$L_N \mathbf{P}\big(\theta^N_{(1)} > x\big) - L_N \mathbf{P}\left(\theta^N_1 > x\right) \leq C(\log N)^{\varepsilon-1} x^{-\lambda}.$$

Therefore, by Lemma 3.4 we have

$$\lim_{N \to +\infty} L_N \mathbf{P}\big(\theta^N_{(1)} > x\big) = \frac{1}{\lambda\Gamma(\lambda)\Gamma(2-\lambda)} \left(\frac{1-x}{x}\right)^\lambda \quad \text{and} \quad \sup_{N \in \mathbb{N}} L_N \mathbf{P}\big(\theta^N_{(1)} > x\big) \leq Cx^{-\lambda}. \tag{3.5}$$

Moreover, Lemma 3.6 implies

$$L_N\, \mathbf{E}\left[\theta^N_{(2)}\right] = \int_0^1 L_N \mathbf{P}\big(\theta^N_{(2)} > x\big) \mathrm{d}x \leq (\log N)^{\varepsilon-1} \int_0^1 x^{\varepsilon-1-\lambda} \mathrm{d}x \underset{N \to +\infty}{\longrightarrow} 0, \tag{3.6}$$

and hence $\Pi^N(t)$ satisfies (3.1). Assume now that $\theta \in (-\alpha, \alpha)$, which implies that $\lambda \in (0,2)$. By (3.5), we have that

$$\int_0^1 x \sup_{N \in \mathbb{N}} L_N \mathbf{P}\big(\theta^N_{(1)} > x\big) \mathrm{d}x \leq C \int_0^1 x^{1-\lambda} \mathrm{d}x < +\infty.$$

Therefore, the assumptions of Lemma 3.1 are satisfied, impling that $c_N L_N \sim_{N\to\infty} (1 - \theta/\alpha)$ and that $\Pi^N(t/c_N)$ converges in distribution to the Beta$(2-\lambda, \lambda)$-coalescent.

We now assume that $\theta \geq \alpha$, in which case $\int_0^1 x\left(\frac{1-x}{x}\right)^\lambda \mathrm{d}x = +\infty$. However, taking $k \geq \lambda$ we have, by (3.5), that

$$\int_0^1 x^k \sup_{N \in \mathbb{N}} L_N \mathbf{P}\big(\theta^N_{(1)} > x\big) \mathrm{d}x \leq C \int_0^1 x^{k-\lambda} \mathrm{d}x < +\infty.$$

Applying Lemma 3.2, we conclude that $\Pi^N(t/c_N)$ converges in the Skorokhod sense to the Kingman coalescent. ∎

4 Poisson-Dirichlet representation of the (N,β)-branching random walk

In this section, we use Proposition 2.2 to construct the (N,β)-BRW, using i.i.d. size-biased $\mathrm{PD}(\beta^{-1},0)$ partitions. We first provide a coupling between the first generation of the (N,β)-BRW and a Poisson-Dirichlet partition.

Proposition 4.1. *Let (U_n) be a $PD(\beta^{-1},0)$, $L = \lim_{n\to+\infty} 1/(n^\beta U_n)$ and (V_n) its size-biased reordering. We have*

$$\left(X_1^N(j) - X_0^N(\mathrm{eq}), j \leq N\right) \overset{(d)}{=} \left(\tfrac{1}{\beta}\log V_j + \tfrac{1}{\beta}\log L\right).$$

In particular $X_1^N(\mathrm{eq}) - X_0^N(\mathrm{eq}) \overset{(d)}{=} \log \sum_{j=1}^N V_j^{1/\beta} + \frac{1}{\beta}\log L$.

Proof. We denote by $(x_k, k\geq 1) = \mathrm{Rank}\left(\left\{X_0^N(j) + p - X_0^N(\mathrm{eq}), p \in \mathcal{P}_1(j), j \leq N\right\}\right)$. By superposition property of the Poisson point processes, $(x_k, k \geq 1)$ is a $\mathrm{PPP}(\mathrm{e}^{-x}\mathrm{d}x)$. We set

$$L = \sum_{j=1}^{+\infty} e^{\beta x_j} \quad \text{and} \quad U_j = e^{\beta x_j}/L.$$

By Proposition 2.2, we observe that $(U_j, j \geq 1)$ is a $\mathrm{PD}(\beta^{-1},0)$ and $\lim_{n\to+\infty} n^\beta U_n = 1/L$. We then write $V_j = e^{\beta(X_1^N(j) - X_0^N(\mathrm{eq}))}/L$. By definition of the (N,β)-BRW, $V_1,\ldots,V_N$ are the N first elements in the size-biased reordering of (U_j). Reversing the equation, we conclude that

$$X_1^N(j) - X_0^N(\mathrm{eq}) = \tfrac{1}{\beta}\left(\log V_j + \log L\right).$$

□

We use this result to compute the asymptotic behaviour of the speed of the (N,β)-BRW.

Proof of Theorem 1.1. Recall from Lemma 1.5 that $v_{N,\beta} = \mathbf{E}\left(X_1^N(\mathrm{eq}) - X_0^N(\mathrm{eq})\right)$. Consequently, we have

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

We observe that by Lemma 2.6, $\log\left(\frac{\Sigma_N}{\log N}\right)$ is uniformly integrable and converges a.s. (and therefore in $\mathbb{L}^1$) toward $\log\left(\Psi_{\beta^{-1}} M_\infty^{1/\beta}\right)$. Using Corollary 2.7, we conclude that

$$\lim_{N\to+\infty} \log \sum_{j=1}^N V_j^{1/\beta} - \log\log N = -\frac{1}{\beta}\log L \quad \text{a.s. and in } \mathbb{L}^1.$$

This yields $\lim_{N\to+\infty} v_{N,\beta} - \log\log N = 0$. □

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

Proof of Theorem 1.2. We observe, by Proposition 4.1 and Lemma 1.6, that the genealogy of the (N,β) BRW can be described as in Theorem 3.3, with parameters $\alpha = \frac{1}{\beta}$ and $\theta = 0$. As a consequence, the genealogy of this process converges toward the Bolthausen-Snitzmann coalescent.

□

References

[1] E. Aïdékon and B. Jaffuel. Survival of branching random walks with absorption. *Sotchastic Process. Appl.*, 121(9):1901–1937, 2011.

[2] J. Berestycki, N. Berestycki, and J. Schweinsberg. The genealogy of branching Brownian motion with absorption. *Ann. Probab.*, 41(2):527–618, 2013.

[3] N. Berestycki. *Recent progress in coalescent theory.* Ensaios Matemáticos [Mathematical Surveys], 16, 2009.

[4] E. Brunet and B. Derrida. Shift in the velocity of a front due to a cut-off. *Phys. Rev. E*, 56:2597–2604, 1997.

[5] E. Brunet and B. Derrida. How genealogies are affected by the speed of evolution. *Philosophical Magazine*, 92(1–3):255–271, 2012.

[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] J. Bérard and P. Maillard. The limiting process of N-particle branching random walk with polynomial tails. *Elec. J. Probab.*, 19(22):1–17, 2014.

[8] F. Comets and A. Cortines. Finite-size corrections to the speed of a branching-selection process. *arXiv:1505.04971*, 2015.

[9] A. Cortines. The genealogy of a solvable population model under selection with dynamics related to directed polymers. *Bernoulli*, 22(4):2209–2236, 2016.

[10] P. Maillard. Speed and fluctuations of N-particle branching brownian motion with spatial selection. *arXiv:1304.0562*, 2013.

[11] B. Mallein. Branching random walk with selection at critical rate. *arXiv:1502.07390*, 2015.

[12] B. Mallein. N-branching random walk with α-stable spine. *arXiv:1503.03762*, 2015.

[13] M. Möhle and S. Sagitov. A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.*, 29(4):1547–1562, 2001.

[14] 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.

[15] J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902, 1999.

[16] J. Pitman and M. Yor. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.*, 25(2):855–900, 1997.