



**HAL**  
open science

# Coagulation-transport equations and the nested coalescents

Amaury Lambert, Emmanuel Schertzer

► **To cite this version:**

Amaury Lambert, Emmanuel Schertzer. Coagulation-transport equations and the nested coalescents. *Probability Theory and Related Fields*, 2020, 176 (1-2), pp.77-147. 10.1007/s00440-019-00914-4 . hal-02860464

**HAL Id: hal-02860464**

<https://hal.sorbonne-universite.fr/hal-02860464v1>

Submitted on 8 Jun 2020

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

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

# Coagulation-transport equations and the nested coalescents

Amaury Lambert <sup>†\*</sup>, Emmanuel Schertzer<sup>†\*</sup>

May 26, 2019

**Abstract.** The nested Kingman coalescent describes the dynamics of particles (called genes) contained in larger components (called species), where pairs of species coalesce at constant rate and pairs of genes coalesce at constant rate provided they lie within the same species. We prove that starting from  $rn$  species, the empirical distribution of species masses (numbers of genes/ $n$ ) at time  $t/n$  converges as  $n \rightarrow \infty$  to a solution of the deterministic coagulation-transport equation

$$\partial_t d = \partial_x(\psi d) + a(t)(d \star d - d),$$

where  $\psi(x) = cx^2$ ,  $\star$  denotes convolution and  $a(t) = 1/(t + \delta)$  with  $\delta = 2/r$ . The most interesting case when  $\delta = 0$  corresponds to an infinite initial number of species. This equation describes the evolution of the distribution of species of mass  $x$ , where pairs of species can coalesce and each species' mass evolves like  $\dot{x} = -\psi(x)$ . We provide two natural probabilistic solutions of the latter IPDE and address in detail the case when  $\delta = 0$ . The first solution is expressed in terms of a branching particle system where particles carry masses behaving as independent continuous-state branching processes. The second one is the law of the solution to the following McKean-Vlasov equation

$$dx_t = -\psi(x_t) dt + v_t \Delta J_t$$

where  $J$  is an inhomogeneous Poisson process with rate  $1/(t + \delta)$  and  $(v_t; t \geq 0)$  is a sequence of independent random variables such that  $\mathcal{L}(v_t) = \mathcal{L}(x_t)$ . We show that there is a unique solution to this equation and we construct this solution with the help of a marked Brownian coalescent point process. When  $\psi(x) = x^\gamma$ , we show the existence of a self-similar solution for the PDE which relates when  $\gamma = 2$  to the speed of coming down from infinity of the nested Kingman coalescent.

**Keywords and phrases.** Kingman coalescent; Smoluchowski equation; McKean-Vlasov equation; degenerate PDE; PDE probabilistic solution; hydrodynamic limit; entrance boundary; empirical measure; coalescent point process; continuous-state branching process; phylogenetics.

**MSC2010 Classification.** Primary 60K35; secondary 35Q91; 35R09; 60G09; 60B10; 60G55; 60G57; 60J25; 60J75; 60J80; 62G30; 92D15.

**Acknowledgements.** The authors thank *Center for Interdisciplinary Research in Biology* (CIRB, Collège de France) for funding.

## CONTENTS

1. Introduction and informal description of the main results	2
2. Weak solutions and branching CSBP	12
3. Finite population McKean–Vlasov equation	26
4. $\infty$ -population McKean–Vlasov equation	32
5. Main convergence results	48
6. Some useful estimates	51
7. Convergence of the empirical measure	54
8. Coming down from infinity in the nested Kingman coalescent	62
Appendix A.	64
References	66

### 1. INTRODUCTION AND INFORMAL DESCRIPTION OF THE MAIN RESULTS

**1.1. The nested Kingman coalescent.** The Kingman coalescent [19] is a stochastic process describing the dynamics of a system of coalescing particles, where each pair of particles independently merges at constant rate. It originates from population genetics, where it is used to model the dynamics of gene lineages in the backward direction of time, thus generating a random genealogy. This model can be enriched by embedding gene lineages into species (each gene belongs to a living organism which in turn belongs to some species). *Nested coalescents* were recently introduced [9] to model jointly the genealogy of the genes and the genealogy of the species. The nested Kingman coalescent is the simplest example of a nested coalescent, where both the gene tree and the species tree are given by (non-independent) Kingman coalescents:

- Each pair of species independently coalesces at rate 1, and when two species coalesce into one so-called mother species, all the genes they harbor are pooled together into the mother species (but do not merge). See Fig. 1.
- Conditional on the species tree, each pair of gene lineages *lying in the same species* independently coalesces at rate  $c$ .

It is easy to see (and well-known) that  $\infty$  is an entrance boundary for the number of particles in the Kingman coalescent, it is said that the Kingman coalescent *comes down from infinity*. It is further known that for the Kingman coalescent started at  $\infty$ , called the *standard* coalescent, the number of

particles  $K_t$  behaves as  $t \rightarrow 0$  like the solution to

$$(1.1) \quad \dot{x} = -x^2/2, \quad x_0 = +\infty$$

so that  $tK_t$  converges to 2 a.s. [6]. In the nested Kingman coalescent starting from infinitely many species containing infinitely many genes, the number of species behaves like  $2/t$  and the mass of each species decreases at the same speed, but is constantly replenished with the genes of other species upon species coalescences. However, because of the domination by the standard coalescent, it is reasonable to conjecture that each species at time  $t$  still carries of the order of  $1/t$  genes, so that the total number of genes at time  $t$  scales like  $1/t^2$ . This conjecture has been confirmed to hold in a recent work [10].

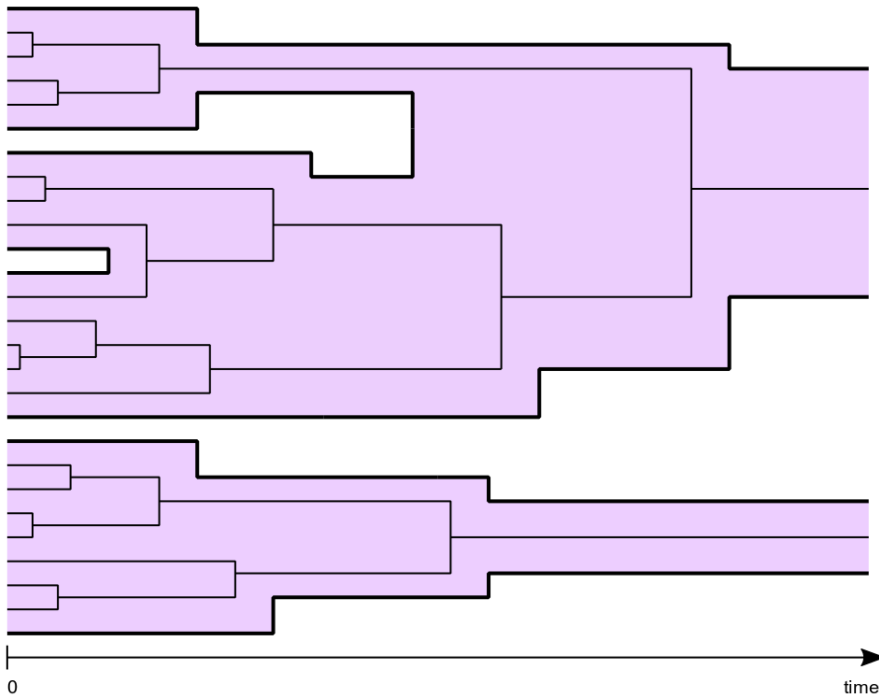


FIGURE 1. The nested coalescent.

The starting point of this paper was the study of the distribution (rather than the total mass) of species masses at small times in the nested Kingman coalescent, for arbitrary initial conditions. This led us to a journey through Smoluchowski coagulation-transport PDEs and McKean-Vlasov equations in which we develop new techniques that are interesting *per se* and could presumably be applied to more general nested coalescents than the nested Kingman coalescent. We now describe informally the main results of this work.

For a random variable  $X$ ,  $\mathcal{L}(X)$  denotes the law of  $X$ . We let  $M_F(\mathbb{R}^+)$  (resp.,  $M_P(\mathbb{R}^+)$ ) denote the set of finite measures (resp. probability measures) on  $\mathbb{R}^+$ .

**Definition 1.1.** We denote by  $\mathcal{H}$  the set of increasing homeomorphisms  $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ . In particular, for any  $\psi \in \mathcal{H}$ ,  $\psi$  is continuous,  $\psi(0) = 0$ ,  $\psi(x) > 0$  for all  $x > 0$  and  $\lim_{x \rightarrow \infty} \psi(x) = +\infty$ . The following (optional) condition

$$\int_1^\infty \frac{dx}{\psi(x)} < \infty,$$

will be called Grey's condition and sometimes abbreviated as  $\int^\infty 1/\psi < \infty$ .

**1.2. Convergence to the Smoluchowski equation.** We first consider a nested coalescent with a large but finite number of initial species. More formally, we consider a sequence of nested Kingman coalescents indexed by  $n$ . Let  $s_t^n$  be the number of species at time  $t$  (in the model indexed by  $n$ ). Let  $\Pi_t^n$  be the vector of size  $s_t^n$  recording the number of gene lineages in each species. We call this vector the genetic composition vector. We wish to investigate the dynamics of the distribution of species masses on a time scale  $O(1/n)$  and rescaling the number of gene lineages by  $1/n$ . Namely, we define

$$(1.2) \quad g_t^n = \frac{1}{s_t^n} \sum_{i=1}^{s_t^n} \delta_{\Pi_t(i)/n} \quad \text{and} \quad \tilde{g}_t^n = g_{t/n}^n$$

that is,  $g^n$  is the empirical distribution of the number of gene lineages per species renormalized by  $n$  and  $\tilde{g}^n$  is obtained from  $g^n$  by rescaling time by  $n$ .

**Result 1** (Theorem 5.1). Assume that there exist two deterministic quantities  $r \in (0, \infty)$  and  $\nu \in M_P(\mathbb{R}^+)$  such that

- (i)  $\frac{s_0^n}{n} \rightarrow r$  in  $L^{2+\epsilon}$
- (ii)  $\tilde{g}_0^n \implies \nu$  as  $n \rightarrow \infty$  in the weak topology for finite measures.

Then the sequence of rescaled empirical measures  $(\tilde{g}_t^n; t \geq 0)$  converges to the unique solution of the following IPDE (Integro Partial Differential Equation)

$$(1.3) \quad \partial_t d(t, x) = \partial_x(\psi d)(t, x) + \frac{1}{t + \delta} (d \star d(t, x) - d(t, x)) \quad t, x \geq 0$$

with initial condition  $d(0, x) dx = \nu(dx)$ . Here,  $d \star d(t, x) = \int_0^x d(t, x - y)d(t, y) dy$  denotes the convolution product,  $\psi(x) = \frac{c}{2}x^2$  and  $\delta = 2/r$ .

**Remark 1.2.** The solution displayed in Result 1 is the unique “weak” solution of (1.3). The notion of weak solution will be made precise in forthcoming Definition 1.8.

**Remark 1.3.** In [17], the authors studied a continuous time version of the Derrida-Retaux model [12] where another transport-coagulation equation also emerges.

The latter result provides a natural interpretation of the two terms on the RHS of (1.3). The transport term is interpreted as the number of gene lineages inside each species obeying the asymptotic dynamics (1.1) whereas the coagulation term is due to the coalescence of species lineages. Finally,  $\delta/2 = 1/r$  is the inverse of the initial population size. By a slight abuse of language, the parameter  $\delta$  will be referred to as the inverse population size in the rest of this manuscript.

**Remark 1.4.** *Consider the coagulation-transport equation*

$$\partial_t n = \partial_x(\psi n) + \frac{1}{2} n \star n - \hat{n}(t)n,$$

where  $\hat{n}(t) = \int_0^\infty n(t, x) dx$ . See e.g. [1, 30, 26] for studies on this class of equations. Here, pairs of clusters coalesce at rate 1 and  $n(t, x)$  is the amount (rather than the density) of clusters of mass  $x$  at time  $t$ . Then informal computations yield that  $\partial_t \hat{n} = -\frac{1}{2} \hat{n}^2$ , so that

$$\hat{n}(t) = \frac{2}{t + \delta} \quad \text{with } \delta := 2/\hat{n}(0).$$

Motivated by the previous heuristics, define  $d(t, x) = n(t, x) \frac{(t+\delta)}{2}$  so that the function  $d$  is interpreted as the density of clusters of mass  $x$  at time  $t$ . Straightforward computations then show that  $d$  satisfies (1.3). This gives an additional motivation for studying (1.3) and an additional justification for why  $\delta$  is called the inverse population size.

**1.3. Coming down from infinity.** Let us now motivate Equation (1.3) when  $\delta = 0$ . Since  $\delta$  is interpreted as the inverse population size, this equation will be referred to as the infinite population ( $\infty$ -population) Smoluchowski equation. In the previous section, we started with a finite but large population. Not surprisingly, the  $\infty$ -population Smoluchowski equation arises when the initial number of species is infinite.

**Result 2** (Theorem 5.5, Theorem 5.7). *Consider a nested coalescent with the following two properties.*

- (i)  $s_0 = \infty$
- (ii) *Each species contains at least one gene lineage.*

*Then the sequence of rescaled empirical measures  $(\tilde{g}_t^n; t \geq 0)$  converges to the unique weak  $\infty$ -population solution of (1.3).*

*As an application of this result, we can derive the speed of coming down from infinity in the nested coalescent. If  $\rho_t$  denotes the number of gene lineages at time  $t$ ,*

$$(1.4) \quad \frac{1}{n^2} \rho_{t/n} \implies \frac{2}{t} \int_0^\infty x \mu_t^{(0)}(dx) = \frac{2\mathbb{E}(\Upsilon)}{t^2} < \infty, \quad \text{as } n \rightarrow \infty,$$

*where  $\Upsilon$  is a random r.v. which is characterized in Result 4 below.*

**Remark 1.5.** *As a special case of the previous statement, that is when all species initially contain the same number of genes, we recover with (1.4) a special case of Theorem 1 in [10] (when  $s_j \equiv \infty$ ), where the constant  $\mathbb{E}(\Upsilon)$  is characterized otherwise.*

**Remark 1.6.** *The  $\infty$ -population solution displayed in Result 2 is the unique weak solution of (1.3) with  $\delta = 0$  ( $\infty$ -population), which is “proper”, in a sense that it has a “non-degenerate” initial condition, which will be made precise in forthcoming Definition 1.8.*

**Remark 1.7.** *Result 1 required us to “tune” the initial number of species and the number of gene lineages per species in order to get a nondegenerate scaling limit at time  $t/n$  as  $n \rightarrow \infty$ . Indeed, according to Result 1(i)(ii), both quantities must be of order  $n$ . In contrast, a fact that stands out in Result 2 is that the  $\infty$ -population equation arises without any delicate scaling. (Compare conditions (i)(ii) in Result 1 and in Result 2.)*

**1.4. General coagulation-transport equation and solution classes.** Motivated by the previous convergence results, we will consider (1.3) with a general depletion term  $\psi \in \mathcal{H}$ . As already discussed, Equation (1.3) describes the density of clusters of mass  $x$ , where pairs of clusters coalesce at rate 1 (coagulation) and each cluster’s mass evolves like  $\dot{x} = -\psi(x)$  (transport). It will be referred to as the *Smoluchowski equation* with (initial) *inverse population size*  $\delta$  (and depletion term  $\psi$ ). The case  $\delta = 0$  corresponds to an infinite initial number of species. Note that this specific case raises some important issues since (1.3) becomes degenerate at  $t = 0$ . (One of the main contributions of the present work is to make sense of such a degeneracy – see below.)

In the next two sections, we will define two types of solution of the IPDE (1.3). (with a general transport term.) Namely,

- (1) The notion of weak solution which is the usual framework in the PDE literature and which arose in our previous convergence results. See Definition 1.8;
- (2) A more restrictive class of solutions that we call McKean-Vlasov solutions, and that relates (1.3) to a natural McKean-Vlasov process describing the evolution of a ‘typical cluster’ in the population. See Definition 1.12.

**1.5. Weak solutions.** In this section, we assume that  $\psi$  is the *Laplace exponent of a spectrally positive and (sub)critical Lévy process*  $Y$ . Note that in particular  $\psi \in \mathcal{H}$ . This choice of  $\psi$  should encompass cases of interest regarding the descent from infinity of nested coalescents where the intra-species coalescence mechanism is more general than the Kingman coalescent. See Section 1.7(3) for a discussion on a natural conjecture extending the convergence results of the previous two sections to general nested  $\Lambda$ -coalescents.

Let us now proceed with the definition of weak solutions in the sense of measures. We are interested in solutions to (1.3) which have total mass 1 at all times, which implies that  $\int_0^\infty \partial_x(\psi d)(t, x) dx = 0$  or equivalently that  $\lim_{x \rightarrow \infty} \psi(x) d(t, x) = 0$  (since  $\psi(0) = 0$ ). To be more specific, we will say that  $f$  is a test-function if and only if  $f \in C^1(\mathbb{R}^+)$ , and further  $f$  and  $\psi f'$  are bounded. (Think of  $f(x) = \exp(-\lambda x)$  for  $\lambda \geq 0$ .) Integrating both sides of (1.3) with respect to such a test-function  $f$  and performing an integration by parts yields the following definition in the spirit of [23], which also follows the usual framework of the PDE literature.

**Definition 1.8.** *Let  $\delta > 0$  and  $\nu \in M_P(\mathbb{R}^+)$ . We say that a probability-valued process  $(\mu_t; t \geq 0)$  is a weak solution of (1.3) with initial condition  $\nu$  if for every test-function  $f$  and every  $t \geq 0$*

$$(1.5) \quad \langle \mu_t, f \rangle = \langle \nu, f \rangle - \int_0^t \langle \mu_s, \psi f' \rangle ds + \int_0^t \frac{1}{s + \delta} (\langle \mu_s \star \mu_s, f \rangle - \langle \mu_s, f \rangle) ds,$$

where we used the notation  $\langle \mu, f \rangle = \int_{\mathbb{R}^+} f(x) \mu(dx)$  for any finite measure  $\mu$ .

Let  $\delta = 0$ . We say that a probability-valued process  $(\mu_t; t > 0)$  is a weak solution of (1.3) if for every test-function  $f$  and every  $s, t > 0$ :

$$(1.6) \quad \langle \mu_t, f \rangle = \langle \mu_s, f \rangle - \int_s^t \langle \mu_u, \psi f' \rangle du + \int_s^t \frac{1}{u} (\langle \mu_u \star \mu_u, f \rangle - \langle \mu_u, f \rangle) du,$$

- We say that the solution is a dust solution if and only if  $\mu_t \rightarrow \delta_0$  in the weak topology as  $t \downarrow 0$ .
- We say that the solution is proper otherwise.

**Remark 1.9.** *The previous terminology is borrowed from fragmentation theory [7].*

Let  $(\mu_t; t \geq 0)$  be a weak solution to (1.3). In view of the convolution product in (1.3) and (1.5), it is natural to consider the Laplace transform  $u(t, \lambda)$  of  $\mu_t$ , namely

$$u(t, \lambda) = \int_{\mathbb{R}^+} e^{-\lambda x} \mu_t(dx) \quad \lambda, t \geq 0.$$

Let  $A^\psi$  denote the generator of the Lévy process  $Y$  with Laplace exponent  $\psi$ . Recall that by definition

$$\mathbb{E}_\lambda(\exp(-xY_t)) = e^{-\lambda x + t\psi(x)} \quad t, \lambda, x \geq 0.$$

Taking  $f(x) = \exp(-\lambda x)$  in (1.5), we get (by Dominated Convergence if  $\lambda > 0$ , the final result being trivial when  $\lambda = 0$ )

$$-\langle \mu_s, \psi f' \rangle = \lambda \int_{[0, \infty)} \psi(x) e^{-\lambda x} \mu_s(dx) = \lambda \frac{\partial}{\partial t} \Big|_{t=0} \int_{[0, \infty)} e^{-\lambda x + t\psi(x)} \mu_s(dx),$$

so that by Fubini–Tonelli,



$$\begin{aligned}
-\langle \mu_s, \psi f' \rangle &= \lambda \frac{\partial}{\partial t} \Big|_{t=0} \int_{[0, \infty)} \mathbb{E}_\lambda(e^{-xY_t}) \mu_s(dx) \\
&= \lambda \frac{\partial}{\partial t} \Big|_{t=0} \mathbb{E}_\lambda \int_{[0, \infty)} e^{-xY_t} \mu_s(dx) = \lambda \frac{\partial}{\partial t} \Big|_{t=0} \mathbb{E}_\lambda u(s, Y_t) = \lambda A^\psi u(s, \lambda),
\end{aligned}$$

where it is implicit that  $A^\psi$  acts on the second component of  $u$ .

Then this yields easily that the Laplace transform  $u(t, \lambda)$  satisfies the non-linear IPDE

$$(1.7) \quad \partial_t u = \lambda A^\psi u + a(t)(u^2 - u),$$

with initial condition  $u(0, \lambda) = \int_{\mathbb{R}^+} e^{-\lambda x} \nu(dx)$ , where  $a(t) = 1/(t+\delta)$ . In particular, it is crucial that  $\lambda A^\psi$  is the generator of the Continuous-State Branching Process (CSBP)  $Z$  with branching mechanism  $\psi$  [22]. When  $\gamma = 2$ ,  $A^\psi$  acting on  $\mathcal{C}^2(\mathbb{R}^+)$  functions coincide with the differential operator  $\partial^2/\partial^2\lambda$ ,  $Y$  is Brownian motion and  $Z$  is the Feller diffusion [14].

By a martingale approach, we can prove the uniqueness (under mild, natural conditions) of the solution of (1.7) when  $\lambda A^\psi$  is replaced with any generator  $A$  of a non-negative Feller process and  $a$  is a general continuous function.

**Result 3** (Theorem 2.4, Theorem 2.7, Theorem 4.2). *Assume that  $\psi$  is the Laplace exponent of a spectrally positive and (sub)critical Lévy process.*

*( $\delta > 0$ ) There exists a unique weak solution solution to (1.3) with initial condition  $\nu$ .*

*( $\delta = 0$ ) Assume that  $\psi$  satisfies Grey's condition, i.e., that  $\int^\infty 1/\psi < \infty$ . Then*

- *There exists a unique proper solution. See Theorem 2.7(ii) for a probabilistic expression of the solution.*
- *In the stable case (when  $\psi(x) = cx^\gamma$ ), there exist infinitely many dust solutions.*

**Remark 1.10.** *The existence of infinitely many dust solutions is reminiscent of a similar behavior for the Boltzmann equation [31].*

**Result 4** (Theorem 2.8). *Let  $\psi(x) = cx^\gamma$  with  $\gamma \in (1, 2]$ . (The case  $\gamma = 2$  corresponds to the nested Kingman coalescent.) The proper solution  $(\mu_t; t > 0)$  is self-similar in the sense that there exists a r.v.  $\Upsilon$  such that*

$$\forall t > 0, \mu_t = \mathcal{L}(t^{-\beta}\Upsilon), \quad \text{where } \beta := \frac{1}{\gamma - 1}.$$

Further,

- (1) *With  $h(x) = \mathbb{E}(\exp(-x\Upsilon))$  then  $h$  is the unique solution of the ODE described in (2.19).*

- (2)  $-h'(0) = \mathbb{E}(\Upsilon) < \infty$  can be expressed as the measure of mass extinction under the 0-entrance measure of a branching CSBP with branching mechanism  $\psi$ . See also Theorem 2.8(ii) for a detailed description of  $\Upsilon$  in terms of a Branching CSBP.

**Remark 1.11.** *The variable  $\Upsilon$  also arises in the study [10] of the speed of coming down from infinity of the nested Kingman coalescent. However,  $\Upsilon$  is not characterized in terms of an IDE or an excursion measure (as in items (1) and (2) of Result 4). Instead, it is characterized as the fixed point of a distributional transformation. (see Theorem 1 and Eq (1) in [10].)*

**1.6. McKean-Vlasov (MK-V) equation.** Recall that in Section 1.5, we have restricted our attention to the case where  $\psi$  is the Laplace exponent of a Lévy process. In this section, we only assume that  $\psi \in \mathcal{H}$ . (Additional assumptions will be needed when considering uniqueness in  $\infty$ -population regime.)

Let us now consider the McKean–Vlasov equation

$$(1.8) \quad dx_t = -\psi(x_t) dt + v_t \Delta J_t^{(\delta)}, \quad \mathcal{L}(x_0) = \nu$$

where  $J^{(\delta)}$  is an inhomogeneous Poisson process with rate  $1/(t + \delta)$  at time  $t$ , and  $(v_t; t \geq 0)$  is a sequence of independent random variables (indexed by  $\mathbb{R}^+$ ) with the property that  $\mathcal{L}(v_t) = \mathcal{L}(x_t)$ , and  $\nu \in M_P(\mathbb{R}^+)$ .

**Result 5** (Theorem 3.4, Corollary 3.5). *For any  $\psi \in \mathcal{H}$  and  $\delta > 0$ , there exists a unique solution to MK-V (1.8) with initial condition  $\nu \in M_P(\mathbb{R}^+)$ .*

Informally, one can think of  $(x_t; t > 0)$  as the evolution of the mass of a focal cluster in the population: mass is depleted at rate  $\psi(x_t)$  and upon coalescence, occurring at rate  $1/(t + \delta)$ , the cluster gains a mass with law  $\mathcal{L}(x_t)$ , i.e., the cluster (or species) gains a random mass whose law is the one of another ‘typical’ and independent cluster at time  $t$ .

It is straightforward to check from Itô’s formula that  $(\mu_t := \mathcal{L}(x_t); t \geq 0)$  is also a weak solution of (1.3) with inverse population size  $\delta$  and initial measure  $\nu$  (in the sense of Definition 1.8). See Lemma 3.2 for more details. This motivates the following definition.

**Definition 1.12.** *Let  $\delta > 0$  and  $\nu \in M_P(\mathbb{R}^+)$ . We say that  $(\mu_t; t > 0)$  is a MK-V solution of the Smoluchowski equation (1.3) if and only if  $(\mu_t; t \geq 0)$  is the law of a solution to the MK-V equation (1.8) (with the same parameters). Note that any MK-V solution is also a weak solution of the IPDE, and as a consequence, the class of MK-V solutions is smaller than its weak counterpart.*

In Section 3, we study the solutions to the MK-V equation when  $\delta > 0$  under minimal assumptions on  $\psi$ . In particular, we show that the solution can be constructed naturally from the Brownian

Coalescent Point Process (CPP) of Popovic [24]. This construction is reminiscent of the construction of the solution to the classical Smoluchowski equation when the coagulation kernel  $K$  is equal to 1 [3]. This explicit representation will allow us to develop some coupling techniques to investigate the  $\infty$ -population regime ( $\delta = 0$ ) and the long-time behavior of the solution to (1.3) when  $\delta > 0$ . (We also note that at the intuitive level, our representation of MK-V solutions in terms of the CPP can be understood in the light of [21] where it is shown that the Kingman coalescent at small scales can be described in terms of the CPP. In our setting the CPP describes the limiting species coalescent.)

Let us now consider the  $\infty$ -population MK-V equation in more detail. More precisely, we consider (1.8) when  $\delta = 0$  and *with no prescription of the initial condition*. Analogously to Definition 1.8, we can define proper and dust solutions for the  $\infty$ -population MK-V equation. By arguing as before, if  $(x_t, t > 0)$  is a proper (resp., dust) solution to MK-V, then  $(\mu_t := \mathcal{L}(x_t), t > 0)$  is a proper (resp., dust) solution to the *Smoluchowski* equation (1.3). In the same vein as Definition 1.12, the law of such processes will be referred to as solutions of the IPDE in the MK-V sense.

**Result 6** (Proposition 4.4, Theorem 4.1, Theorem 4.2). *Assume that  $\delta = 0$  and assume that  $\psi \in \mathcal{H}$  satisfies Grey's condition and is convex. Then*

- (1) *There exists at least one proper solution to the  $\infty$ -population MK-V equation.*
- (2) *In the stable case  $\psi(x) = cx^\gamma$  with  $\gamma > 1$ , there is a unique proper solution  $(x_t^{(0)}; t \geq 0)$ .*

*Further, this process is self-similar in the sense that*

$$\mathcal{L}\left(x_t^{(0)}\right) = \mathcal{L}\left(t^{-\beta}\Upsilon\right).$$

*where  $\Upsilon$  is a positive random variable (depending implicitly on  $c$  and  $\gamma$ ).*

- (3) *In the stable case, there exists infinitely many dust solutions.*

**Remark 1.13.** *Note that in the finite population case ( $\delta > 0$ ) the uniqueness of the MK-V solution advertised in Result 5 holds for any  $\psi \in \mathcal{H}$  while in the  $\infty$ -population case ( $\delta = 0$ ), we have only been able to show uniqueness of the proper solution for  $\psi(x) = cx^\gamma$  with  $\gamma > 1$  (Result 6). We nevertheless conjecture that it holds in the general case.*

**Remark 1.14.** *When  $\gamma \in (1, 2]$ , the definition of  $\Upsilon$  in Results 4 and 6 must coincide. This follows from the fact that any solution in the MK-V sense must be a solution in the weak sense, and the uniqueness of a weak proper solution when  $\gamma \in (1, 2]$ . See Result 3.*

**Result 7** (Theorem 4.13). *Let  $\gamma > 1$  and recall  $\beta = 1/(\gamma - 1)$ . Assume that  $\psi(x) = cx^\gamma$ . Let  $\nu \in M_P(\mathbb{R}^+)$  such that  $\nu \neq \delta_0$  and let  $x^{(\delta)}$  be the MK-V solution with inverse population size  $\delta > 0$  and initial probability measure  $\nu$ . Then*

$$\lim_{t \rightarrow \infty} t^\beta x_t^{(\delta)} = \Upsilon, \text{ in law}$$

where  $\Upsilon$  is defined in Result 6. In particular, this shows that as  $t \rightarrow \infty$ , the typical mass of a cluster goes to 0 as  $O(t^{-\beta})$ .

**1.7. Discussion and conjectures.** We have called ‘infinite population regime’ the case where the coagulation term becomes degenerate at  $t = 0$ . The uniqueness of solutions to the Smoluchowski equation that we obtain in these cases in the apparent absence of an initial condition is actually due to the fact that this initial condition can be seen as a Dirac mass at  $\infty$ , which ‘comes down from infinity’, in the sense that  $\mu_t$  is a probability measure on  $[0, \infty)$  at any positive time  $t$ .

In this regard, Grey’s condition  $\int^\infty 1/\psi < \infty$  is not anecdotal. It ensures that  $\dot{x} = -\psi(x)$  comes down from infinity, so that in the Smoluchowski equation, the antagonism between mass transport towards 0 and the increase of mass by coagulation is dominated by the transport term, in such a way that the mass distribution over species converges to the Dirac mass at 0 as  $t \rightarrow \infty$ . We conjecture that when  $1/\psi$  is not integrable at  $\infty$ , coagulation overwhelms transport, in such a way that the mass distribution concentrates around large masses as  $t \rightarrow \infty$ . We further conjecture that in this case, our results concerning uniqueness of solutions in the infinite population regime will not hold any longer.

In addition to the previous remarks, and in view of our results, it is natural to make the additional following conjectures.

- (1) The notion of weak and MK-V solutions coincide. (We only showed that if  $(x_t; t \geq 0)$  is a solution to MK-V, then  $(\mathcal{L}(x_t); t \geq 0)$  is a weak solution.)
- (2) There exists a unique proper weak and MK-V solution for any convex  $\psi \in \mathcal{H}$  satisfying Grey’s condition.
- (3) Let  $\Lambda$  be a finite measure on  $[0, 1]$ . Results 1 and 2 concerning the convergence of the rescaled distribution of species masses at small times to the solution of the Smoluchowski equation extends to the case when the genes undergo a  $\Lambda$ -coalescent process coming down from infinity (and the species still undergo a Kingman coalescent) to the same Smoluchowski equation where  $\psi(\lambda) = c\lambda^2$  is replaced with

$$(1.9) \quad \psi(\lambda) = \int_{(0,1]} (e^{-\lambda r} - 1 + \lambda r) r^{-2} \Lambda(dr),$$

which is the Laplace exponent of a spectrally positive Lévy process. This is due to the fact that in the standard  $\Lambda$ -coalescent coming down from infinity, the number of lineages obeys the asymptotic dynamics  $\dot{x} = -\psi(x)$  [5, 4].

- (4) The entrance law at  $\infty$  (in the sense that the number of species at time  $t$  goes to  $\infty$  as  $t \downarrow 0$ ) of the nested Kingman coalescent is unique.

- (5) Recall the notion of dust solution of Definition 1.8. For a given dust solution, we conjecture the existence of a scaling such that the nested Kingman coalescent converges to this solution.
- (6) For the sake of simplicity, we only considered in the manuscript a Kingman species coalescent. For a general  $\Lambda$ -coalescent, we expect the same type of results to hold. More precisely, the coagulation term in the IPDE should be replaced the coagulation kernel described in [8] Proposition 3.

## 2. WEAK SOLUTIONS AND BRANCHING CSBP

**2.1. Main assumptions.** In this section, we consider the case where  $\psi$  is of the form

$$\psi(x) = ax + bx^2 + \int_{(0,\infty)} (\exp(-rx) - 1 + rx) \pi(dr)$$

where  $b \geq 0$ ,  $\pi$  is a  $\sigma$ -finite measure on  $(0, \infty)$  such that  $\int_{\mathbb{R}^+} (r \wedge r^2) \pi(dr) < \infty$ , and  $a = \psi'(0^+) \geq 0$ , so that  $\psi \in \mathcal{H}$  is the Laplace exponent of a spectrally positive (sub)critical Lévy process  $Y$  with

$$\mathbb{E}_0(\exp(-\lambda Y_t)) = e^{t\psi(\lambda)} \quad t, \lambda \geq 0.$$

In particular, we can recover the case  $\psi(x) = cx^2$  by taking  $\pi = 0$  and we can recover the cases  $\psi(x) = cx^\gamma$  for  $\gamma \in (1, 2)$  by taking  $b = 0$  and the jump measure of the form  $\pi(dx) = \frac{\bar{c}}{x^{\gamma+1}}$ , where  $\bar{c}$  is some positive constant. Note that under our assumptions, for any  $t_0 \in \mathbb{R}$  and  $x_0 > 0$ , there exists a unique solution  $(v(t); t \geq t_0)$  to the ODE

$$\dot{v} = -\psi(v), \quad v(t_0) = x_0,$$

and that  $\lim_{t \rightarrow \infty} v(t) = 0$ .

**2.2. Laplace transform of weak solutions.** We let  $A^\psi$  denote the generator of the Lévy process  $Y$ . There is a Feller process with generator  $A$  given by  $Af(\lambda) = \lambda A^\psi f(\lambda)$ , also known as the CSBP (Continuous State Branching Process) with branching mechanism  $\psi$ . For any  $\mu \in M_F(\mathbb{R}^+)$ ,  $f(\lambda) = \int_{\mathbb{R}^+} \exp(-\lambda x) \mu(dx)$  is in the domain of the generator  $A$ , and further

$$\forall \lambda \geq 0, \quad Af(\lambda) = \lambda \int_{\mathbb{R}^+} \psi(x) \exp(-\lambda x) \mu(dx).$$

Let us now consider  $(\mu_t; t \geq 0)$  a weak solution of the Smoluchowski equation and set

$$u(t, \lambda) = \int_{\mathbb{R}^+} e^{-\lambda x} \mu_t(dx) \quad \lambda, t \geq 0.$$

We have proved after Remark 1.9 that  $u$  satisfies (1.7) for all  $\lambda > 0$  and  $t \geq 0$ , that is,

$$\partial_t u = Au + a(t)(u^2 - u),$$

with initial condition  $u(0, \lambda) = \int_{\mathbb{R}^+} e^{-\lambda x} \nu(dx)$ , where  $a(t) = 1/(t + \delta)$ .

In the next subsection, we recall some well-known facts on the CSBP with branching mechanism  $\psi$ . Then we introduce a closely related object: the branching CSBP. In Subsection 2.5, we prove that there exists a unique weak solution to (1.3) which can be expressed in terms of a branching particle system. In Subsection 2.6, we focus on proper weak solutions in the infinite population of the Smoluchowski equation (1.3). In Subsection 2.7, we provide additional results in the stable case  $\psi(\lambda) = c\lambda^\gamma$  for  $\gamma \in (1, 2]$ .

**2.3. Continuous-state branching processes (CSBP).** We collect here known results about CSBP. See e.g., [13] or Section 2.2.3. in [20].

A CSBP  $Z = (Z_t; t \geq 0)$  is a Feller process with values in  $\mathbb{R}^+$  with the branching property, namely if  $P_x$  denotes the law of  $Z$  started at  $x \geq 0$ , then  $P_x \star P_y = P_{x+y}$ . It is well-known [22, 11], that CSBPs are in one-to-one correspondence with Lévy processes with no negative jumps via several different bijections, including a random time-change known as Lamperti's transform. If  $\psi$  is the Laplace exponent of such a Lévy process  $Y$  (assumed to be critical or subcritical), then the CSBP  $Z$  associated with  $Y$  is called the *CSBP with branching mechanism  $\psi$*  and the generator  $A$  of  $Z$  is given by

$$Af(\lambda) = \lambda A^\psi f(\lambda) \quad \lambda \geq 0,$$

for any  $f$  in the domain of the generator  $A^\psi$  of  $Y$ . Further,

$$(2.10) \quad \mathbb{E}_x(\exp(-\lambda Z_t)) = \exp(-xu_t(\lambda)), \quad \dot{u} = -\psi(u), \quad u_0 = \lambda.$$

where  $\dot{u} = -\psi(u)$  will be a shorthand for  $\partial_t u(t, \lambda) = -\psi(u(t, \lambda))$ . Note that by the branching property, 0 is an absorbing state for  $Z$ . It is accessible if and only if  $1/\psi$  is integrable at  $\infty$  (Grey's condition [16]). We then denote by  $T_0$  the first hitting time of 0 by  $Z$ .

By the branching property, the law of  $Z_t$  is infinitely divisible. More specifically, if Grey's condition is fulfilled, then under  $P_x$ ,  $Z_t$  is equal to the sum  $\sum_i Z_t^{(i)}$ , where  $(Z_t^{(i)})$  are the atoms of a Poisson point process with intensity measure  $xN$ , where  $N$  is a  $\sigma$ -finite measure on the set of càdlàg processes started at 0 and with non-negative values. We will call  $N$  the *entrance measure at 0 of the CSBP*. Exactly as  $E_x$ ,  $N$  is a measure on the space of non-negative càdlàg paths, except that it has infinite mass. As we will now see, for any  $\varepsilon > 0$ , the infinite mass comes from the subset of paths which return to 0 within times smaller than  $\varepsilon$ . For any  $\lambda \geq 0$ , by the exponential formula for Poisson point processes,

$$(2.11) \quad E_x(\exp(-\lambda Z_t)) = \exp(-xN(1 - e^{-\lambda Z_t})),$$

and so for any measurable functional of paths  $G$  such that  $|G| \leq KT_0$  for some  $K$ ,

$$(2.12) \quad \lim_{x \downarrow 0} x^{-1} E_x(G) = N(G) < \infty.$$

Note as announced that  $N(T_0 > t) < \infty$  for all  $t > 0$ .

**2.4. Duality between a branching and a coalescing particle system.** In this section, we fix  $\mathfrak{t}$  an ultrametric binary tree with  $n$  labelled leaves and depth  $T$  (i.e., the distance from the root to each leaf is  $T$ ). We let  $N_t(\mathfrak{t})$  denote the number of points in  $\mathfrak{t}$  at time  $t$ , i.e., at distance  $t$  from the root. We will now introduce two particle systems, a coalescing particle system initialized at the leaves of  $\mathfrak{t}$  and a branching particle system initialized at the root of  $\mathfrak{t}$ .

Let us first introduce the coalescing particle system. We start by assigning a mark  $\lambda_i \geq 0$  to the leaf of the tree  $\mathfrak{t}$  labelled by  $i$ . Then we let the marks propagate from the leaves to the root according to the following rules: (i) along each branch, the marking evolves according to the deterministic dynamics  $\dot{x} = -\psi(x)$ , and (ii) when two branches merge, we add the corresponding two marks. We call  $F(\mathfrak{t}, \lambda)$  the resulting mark at the root.

Now, we consider a system of branching particles with random mass running along the branches of the tree, from the root to the leaves, according to the following rules: (i) we start with one particle at the root, (ii) at each branching point the incoming particle, carrying say mass  $x$ , duplicates into two copies of itself, one copy for each branch, each with mass  $x$ , and (iii) along each branch, the mass of the particle on that branch evolves independently according to a CSBP with branching mechanism  $\psi$ . See Fig. 2.

Let  $Q_x^\mathfrak{t}$  denote the law of the branching particle system started with one particle with mass  $x$  at time 0 and let  $\mathcal{Z}_t^\mathfrak{t} = (Z_t^i)_{1 \leq i \leq N_t(\mathfrak{t})}$  denote the masses carried by the particles at time  $t$ . Finally, as a direct consequence of the branching property, it is not hard to see that  $\mathcal{Z}^\mathfrak{t}$  is infinitely divisible, and as for a simple CSBP, under Grey's condition,  $\mathcal{Z}^\mathfrak{t}$  under  $Q_x^\mathfrak{t}$  can be decomposed into a Poisson sum of elementary processes starting from 0 with intensity measure  $xM^\mathfrak{t}$ , where  $M^\mathfrak{t}$  is the *entrance measure at 0 of  $\mathcal{Z}^\mathfrak{t}$* . In view of (2.12), we have

$$M^\mathfrak{t}(G) = \lim_{x \downarrow 0} x^{-1} Q_x^\mathfrak{t}(G)$$

for any measurable functional of paths  $G$  such that  $|G| \leq KT_0$  for some  $K$  (where here  $T_0$  is the extinction time of the particle on the root edge of  $\mathfrak{t}$ , set to  $+\infty$  if the particle splits before going extinct).

Now analogously to (2.10), there exists a nice characterization of the Laplace transform of  $\mathcal{Z}_t^\mathfrak{t}$ .

**Proposition 2.1.** *For any non-negative numbers  $x$  and  $(\lambda_i)_{1 \leq i \leq n}$ ,*

$$(2.13) \quad \mathbb{E}_x \left( \exp(-\lambda \cdot \mathcal{Z}_T^\mathfrak{t}) \right) = \exp(-xF(\mathfrak{t}, \lambda)),$$

*with the notation  $\lambda \cdot \mathcal{Z}_T^\mathfrak{t} = \sum_{i=1}^n \lambda_i Z_T^i$ . Under Grey's condition, we additionally have*

$$(2.14) \quad M^\mathfrak{t}(1 - \exp(-\lambda \cdot \mathcal{Z}_T^\mathfrak{t})) = F(\mathfrak{t}, \lambda).$$

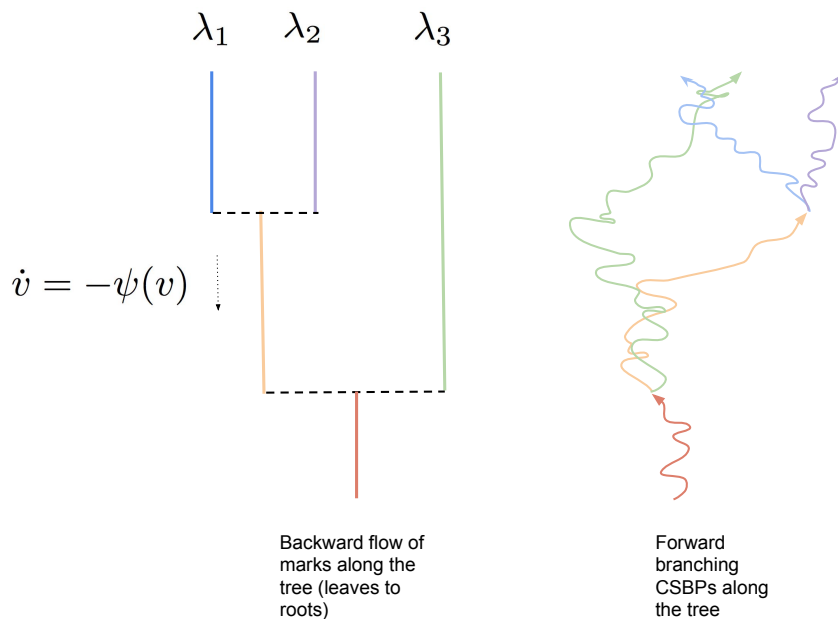


FIGURE 2. Duality between branching CSBPs and the coalescing particle system

*Proof.* Using (2.10), it is straightforward to get (2.13) by induction on the number of nodes of the tree. Under Grey's condition, analogously to (2.11), we have

$$\mathbb{E}_x \left( \exp(-\lambda \cdot Z_T^t) \right) = \exp \left\{ -xM^t (1 - \exp(-\lambda \cdot Z_T^t)) \right\},$$

which yields (2.14). □

**2.5. Smoluchowski equation in finite population.** Let  $P_t$  denote the semigroup and  $A$  the infinitesimal generator of a time-homogeneous Feller process  $Z$  with values in  $[0, \infty)$ . We are interested in the case when  $Af(\lambda) = \lambda A^\psi(\lambda)$  but do not immediately restrict our study to this case. Let  $\mathcal{E}_A$  be the space of continuous functions  $f : [0, \infty) \rightarrow [0, 1]$  such that  $Af$  is a well-defined continuous, bounded function on  $(0, \infty)$  and  $f(Z_t) - \int_0^t Af(Z_s) ds$  is a martingale (or equivalently  $t \mapsto P_t f(x)$  is differentiable at 0 with derivative  $Af(x)$  for all  $x \geq 0$ ). Further let  $\mathcal{E}'_A$  be the space of two-variable continuous functions  $f : [0, \infty) \times [0, \infty) \rightarrow [0, 1]$  such that  $f(t, \cdot) \in \mathcal{E}_A$  and  $f(\cdot, x) \in \mathcal{C}^1$ .

Let  $a : [0, \infty) \rightarrow [0, \infty)$  be a continuous map and fix  $T > 0$ . Let us consider a time-inhomogeneous system of branching particles, where each particle carries an individual mass evolving like the process  $Z$  and each particle independently gives birth at rate  $\tilde{a}(t) = a(T - t)$  (at time  $t$ ) to a copy of itself (i.e., a branching particle with mass  $x$  splits into two particles, each with mass  $x$ ). Let  $Z_t$  denote



the state of this system (e.g., its empirical measure) at time  $t$ , assumed to be càdlàg. We assume that the system starts at time 0 with one particle carrying mass  $x$ , and we then let  $Q_x^T$  denote the law of  $\mathcal{Z}$  up to time  $T$ . We also let  $N_t$  denote the number of particles at time  $t$  and by  $(Z_t^i)_{1 \leq i \leq N_t}$  the masses carried by these particles.

**Remark 2.2.** *Let  $\mathfrak{t}$  be an ultrametric tree with depth  $T$ . When  $Af(\lambda) = \lambda A^\psi(\lambda)$ , and if we condition on the genealogy of the process up to  $T$  to be  $\mathfrak{t}$ , the process  $\mathcal{Z}$  coincides with  $\mathcal{Z}^\mathfrak{t}$  as defined in the previous section.*

**Lemma 2.3.** *For any  $g \in \mathcal{E}_A$  there is at most one solution  $u \in \mathcal{E}'_A$  to the PDE (or IPDE)*

$$(2.15) \quad \partial_t u = Au + a(t)(u^2 - u),$$

with initial condition  $u(0, x) = g(x)$ . Further, if such a solution exists, it must be of the following form

$$u(T, x) = Q_x^T \left( \prod_{i=1}^{N_T} g(Z_T^i) \right) \quad T > 0, x \geq 0$$

Before proving this lemma, we wish to state the relevant corollary regarding weak solutions to (1.3), taking  $A$  in the previous lemma equal to the generator of the CSBP  $Z$  with branching mechanism  $\psi$ .

**Theorem 2.4.** *Assume that  $\psi$  is the Laplace exponent of a (sub)critical spectrally positive Lévy process.*

- (i) *There exists a unique weak solution  $(\mu_t; t \geq 0)$  to the Smoluchowski equation (1.3) with initial distribution  $\nu$  and inverse population  $\delta > 0$ .*
- (ii) *For any  $T \geq 0$ ,  $\mu_T$  is equal to the law of  $F(\mathbf{T}, (W_i); 1 \leq i \leq N_T(\mathbf{T}))$ , where  $\mathbf{T}$  is the time-inhomogeneous tree started at 0 with one particle, stopped at time  $T$  and with birth rate  $\tilde{a}(t) = 1/(T - t + \delta)$ , the  $(W_i)$  are iid with law  $\nu$ .*

*Proof.* Let us apply Lemma 2.3 with  $A$  defined by  $Af(\lambda) = \lambda A^\psi(\lambda)$  the generator of the CSBP  $Z$  with branching mechanism  $\psi$  and  $a(t) = \frac{1}{t+\delta}$ . Then Equation (2.15) is the same equation as (1.7), with initial condition  $g(\lambda) = \int_{\mathbb{R}^+} e^{-\lambda x} \nu(dx)$ , which we can write

$$g(\lambda) = \mathbb{E}(\exp(-\lambda W)),$$

where  $W$  denotes a random variable with law  $\nu$ . Note that  $g$  takes values in  $[0, 1]$ . Let us check that  $g \in \mathcal{E}_A$ . Recall that for any  $f$  of the form  $f(\lambda) = \int_{\mathbb{R}^+} \exp(-\lambda x) \mu(dx)$ ,  $f$  is in the domain of  $A$ , and further  $Af(\lambda) = \lambda \int_{\mathbb{R}^+} \psi(x) \exp(-\lambda x) \mu(dx)$ . This shows that

$$Ag(\lambda) = \lambda \int_{\mathbb{R}^+} \psi(x) \exp(-\lambda x) \nu(dx),$$

so that  $Ag$  is a well-defined continuous function on  $(0, \infty)$  (since  $\psi$  increases at most polynomially at  $\infty$ ). In addition, it is well-known (see e.g. [11]) that

$$e^{-xZ_t} - \psi(x) \int_0^t Z_s e^{-xZ_s} ds$$

is a martingale for any  $x \geq 0$ , so by integrating  $x$  with respect to the probability measure  $\nu$ , we get that  $g(Z_t) - \int_0^t Ag(Z_s) ds$  is also a martingale. By dominated convergence,  $Ag$  vanishes at  $\infty$  and so is bounded and we conclude that  $g \in \mathcal{E}_A$ . So by Lemma 2.3 there is at most one solution  $v \in \mathcal{E}'_A$  to (1.7) given by

$$v(T, \lambda) = Q_\lambda^T \left( \prod_{i=1}^{N_T} g(Z_T^i) \right).$$

Now let  $(\mu_t; t \geq 0)$  be a weak solution to the Smoluchowski equation (1.3) and set

$$u(t, \lambda) = \int_{\mathbb{R}^+} e^{-\lambda x} \mu_t(dx) \quad \lambda, t \geq 0.$$

Recall that  $u$  satisfies (1.7) with initial condition  $u(0, \lambda) = g(\lambda)$ . The exact same reasoning used to prove that  $g \in \mathcal{E}_A$  shows that  $u(t, \cdot) \in \mathcal{E}_A$  for all  $t$ , and because  $u$  satisfies (1.7),

$$\partial_t u(t, \lambda) - a(t)(u^2 - u)(t, \lambda) = \lambda A^\psi u(t, \lambda) = \lambda \int_{\mathbb{R}^+} \psi(x) \exp(-\lambda x) \mu_t(dx),$$

which is continuous in  $t$ . Since  $a$  is continuous, we get that  $u(\cdot, \lambda)$  is of class  $\mathcal{C}^1$  and so  $u \in \mathcal{E}'_A$ . This shows that  $u = v$ , so that

$$u(T, \lambda) = Q_\lambda^T \left( \prod_{i=1}^{N_T} \mathbb{E}(\exp(-Z_T^i W_i) | Z_T^i) \right) = \mathbb{E} \left[ Q_\lambda^T \left( \exp \left( - \sum_{i=1}^{N_T} W_i Z_T^i \right) \right) \right],$$

where  $\mathbb{E}$  is the expectation taken with respect to the  $(W_i)$ , which are independent copies of  $W$  (and we have applied Fubini–Tonelli Theorem). From (2.13), we get

$$u(T, \lambda) = \mathbb{E} \left[ Q_\lambda^T \left( \exp \left( - \sum_{i=1}^{N_T} W_i Z_T^i \right) \right) \right] = \mathbb{E}_T [\exp(-\lambda F(\mathbf{T}, (W_i); 1 \leq i \leq N_T(\mathbf{T})))],$$

where now  $\mathbb{E}_T$  is the expectation taken with respect to the Yule tree  $\mathbf{T}$  with branching parameter  $\tilde{a}(t) = 1/(T - t + \delta)$  stopped at time  $T$  and the iid random variables  $(W_i)$ . It follows that

$$\int_{\mathbb{R}^+} e^{-\lambda x} \mu_T(dx) = u(T, \lambda) = \mathbb{E} [\exp(-\lambda F(\mathbf{T}, (W_i); 1 \leq i \leq N_T(\mathbf{T})))],$$

and by the injectivity of the Laplace transform,  $\mu_T$  is the law of  $F(\mathbf{T}, (W_i); 1 \leq i \leq N_T(\mathbf{T}))$ .

For the sake of completeness, in Appendix A we check that  $\mu_T$  defined as  $F(\mathbf{T}, (W_i); 1 \leq i \leq N_T(\mathbf{T}))$  indeed is solution to (1.3). (As a matter of fact, the existence of a weak solution will be proved in the MK-V section (see Theorem 3.4) and thus checking that  $\mu$  is indeed a weak solution is not formally needed.)  $\square$

*Proof of Lemma 2.3.* Recall that  $P_t$  denotes the semigroup of  $Z$ . We will use the notation  $E_x$  to denote the expectation associated with its law when started from  $x$ . We extend the semigroup and generators by defining for any  $f \in \mathcal{E}'_A$ ,

$$\bar{P}_s f(t, x) = E_x(f(t + s, Z_s)) \quad \text{and} \quad \bar{A}f(t, x) = Af(t, x) + \partial_t f(t, x),$$

so that in particular,

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} (\bar{P}_\varepsilon f(t, x) - f(t, x)) = \bar{A}f(t, x).$$

Let  $u \in \mathcal{E}'_A$  be a solution to (2.15) with initial condition  $g$ . Fix  $T > 0$  and recall the system of branching particles defined before the statement of the lemma. Because the dynamics of the system are time-inhomogeneous, we will need to denote by  $Q_{t,x}^T$  the law of  $Z$  when started with one single particle with mass  $x$  at time  $t$ . In particular,  $Q_x^T = Q_{0,x}^T$ . Let  $\mathcal{F}_t$  denote the  $\sigma$ -field generated by  $Z_t$ . Fix  $T > 0$  and set  $Z^u$  the  $(\mathcal{F}_t)$ -adapted process given by

$$\forall t \in [0, T], \quad Z_t^u = \prod_{i=1}^{N_t} \tilde{u}(t, Z_t^i), \quad \text{where } \tilde{u}(t, x) := u(T - t, x),$$

$N_t$  is the number of particles present at time  $t$  and  $Z_t^i$  is the mass of particle  $i$  (note from the definition of  $Z_t^u$  that it does not depend on the labelling chosen).

We aim at proving that  $Q_x^T(Z_t^u)$  is constant as a function of  $t$ . Let  $t, \varepsilon$  such that  $0 \leq t \leq t + \varepsilon \leq T$ . Conditional on  $\mathcal{F}_t$ , denote by  $\tau_i$  time when the  $i$ -th particle splits,  $i = 1, \dots, N_t$ . Then by the branching property,

$$\begin{aligned} Q^T(Z_{t+\varepsilon}^u | \mathcal{F}_t) &= \mathbb{P}(\tau_i > t + \varepsilon, \forall i) \prod_{i=1}^{N_t} E_{Z_t^i}(\tilde{u}(t + \varepsilon, Z_\varepsilon)) \\ &+ \sum_{j=1}^{N_t} \mathbb{P}(\tau_i > t + \varepsilon, \forall i \neq j) \int_t^{t+\varepsilon} \mathbb{P}(\tau_j \in dv) E_{Z_t^j}(Q_{v,Z_v}^T(Z_{t+\varepsilon}^u)^2) \prod_{i \neq j} E_{Z_t^i}(\tilde{u}(t + \varepsilon, Z_\varepsilon)) + C_\varepsilon, \end{aligned}$$

where  $C_\varepsilon \leq \mathbb{P}(B_\varepsilon \geq 2)$ , with

$$B_\varepsilon := \#\{i \leq N_t : t \leq \tau_i \leq t + \varepsilon\}.$$

Now  $B_\varepsilon$  is a binomial random variable with parameters  $N_t$  and  $u_\varepsilon := 1 - e^{-\int_t^{t+\varepsilon} \bar{a}(u) du}$ , so

$$\mathbb{P}(B_\varepsilon \geq 2) \leq \frac{N_t(N_t - 1)}{2} u_\varepsilon^2 \leq N_t^2 M^2 \varepsilon^2,$$

where  $M := \sup_{t \in [0, T]} a(t)$ . Re-arranging, we get

$$\begin{aligned} \frac{1}{\varepsilon} (Q^T(\mathcal{Z}_{t+\varepsilon}^u | \mathcal{F}_t) - \mathcal{Z}_t^u) &= \frac{1}{\varepsilon} \left( \prod_{i=1}^{N_t} E_{Z_t^i}(\tilde{u}(t + \varepsilon, Z_\varepsilon)) - \mathcal{Z}_t^u \right) \\ &\quad + \frac{1}{\varepsilon} (1 - e^{-N_t \int_t^{t+\varepsilon} \tilde{a}(u) du}) \prod_{i=1}^{N_t} E_{Z_t^i}(\tilde{u}(t + \varepsilon, Z_\varepsilon)) \\ &\quad + \frac{e^{-(N_t-1) \int_t^{t+\varepsilon} \tilde{a}(u) du}}{\varepsilon} \sum_{j=1}^{N_t} \int_t^{t+\varepsilon} \tilde{a}(v) dv e^{-\int_t^v \tilde{a}(u) du} E_{Z_t^j}(Q_{v, Z_v}^T(\mathcal{Z}_{t+\varepsilon}^u)^2) \prod_{i \neq j} E_{Z_t^i}(\tilde{u}(t + \varepsilon, Z_\varepsilon)) \\ &\quad + \frac{1}{\varepsilon} C_\varepsilon. \end{aligned}$$

Because  $u \in \mathcal{E}'_A$  and  $a$  is continuous, the right-hand side of the last equality converges as  $\varepsilon \downarrow 0$  to

$$\begin{aligned} &\sum_{i=1}^{N_t} (\bar{A}\tilde{u}(t, Z_t^i)) \prod_{j \neq i} \tilde{u}(t, Z_t^j) - \tilde{a}(t) N_t \mathcal{Z}_t^u + \tilde{a}(t) \sum_{j=1}^{N_t} \tilde{u}(t, Z_t^j) \mathcal{Z}_t^u \\ &= \sum_{i=1}^{N_t} \left( \prod_{j \neq i} \tilde{u}(t, Z_t^j) \right) [\bar{A}\tilde{u}(t, Z_t^i) - \tilde{a}(t) \tilde{u}(t, Z_t^i) (1 - \tilde{u}(t, Z_t^i))]. \end{aligned}$$

Now the last quantity is zero because for any  $x$

$$\begin{aligned} &\bar{A}\tilde{u}(t, x) - \tilde{a}(t) \tilde{u}(t, x) (1 - \tilde{u}(t, x)) \\ &= Au(T-t, x) - \partial_t u(T-t, x) - a(T-t) u(T-t, x) (1 - u(T-t, x)), \end{aligned}$$

which is zero by (2.15). So we have proved

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} (Q^T(\mathcal{Z}_{t+\varepsilon}^u | \mathcal{F}_t) - \mathcal{Z}_t^u) = 0.$$

We would now like to take expectations inside the limit. Since  $u$  and so  $\mathcal{Z}^u$  take values in  $[0, 1]$ , we first have

$$\left| \frac{1}{\varepsilon} (Q^T(\mathcal{Z}_{t+\varepsilon}^u | \mathcal{F}_t) - \mathcal{Z}_t^u) \right| \leq \left| \frac{1}{\varepsilon} \left( \prod_{i=1}^{N_t} \bar{P}_\varepsilon \tilde{u}(t, Z_t^i) - \mathcal{Z}_t^u \right) \right| + 2MN_t + N_t^2 M^2 \varepsilon.$$

Now because  $\bar{A}\tilde{u} = a(t)(\tilde{u} - \tilde{u}^2)$ ,  $\bar{A}\tilde{u}$  takes values in  $[0, M]$ , and since

$$\frac{\bar{P}_\varepsilon \tilde{u}(t, x) - \tilde{u}(t, x)}{\varepsilon} = \frac{1}{\varepsilon} \int_0^\varepsilon \bar{P}_s \bar{A}\tilde{u}(t, x) ds,$$

we get

$$0 \leq \frac{\bar{P}_\varepsilon \tilde{u}(t, x) - \tilde{u}(t, x)}{\varepsilon} \leq M.$$

So we can write

$$\frac{1}{\varepsilon} \left( \prod_{i=1}^{N_t} \bar{P}_\varepsilon \tilde{u}(t, Z_t^i) - \mathcal{Z}_t^u \right) = \frac{H(\varepsilon) - H(0)}{\varepsilon}$$

where

$$H(\varepsilon) = \prod_{i=1}^{N_t} (x_i + \varepsilon y_i),$$

with  $x_i = \tilde{u}(t, Z_t^i)$  and  $y_i = \frac{\bar{P}_\varepsilon \tilde{u}(t, Z_t^i) - \tilde{u}(t, Z_t^i)}{\varepsilon}$ , so that  $0 \leq x_i \leq 1$  and  $0 \leq y_i \leq M$ . This shows that for any  $z \in [0, \varepsilon]$

$$0 \leq H'(z) \leq H'(\varepsilon) = \sum_{i=1}^{N_t} y_i \prod_{j \neq i} (x_j + \varepsilon y_j) \leq N_t M (1 + \varepsilon M)^{N_t - 1}.$$

Then by the Mean Value Theorem

$$\left| \frac{1}{\varepsilon} \left( \prod_{i=1}^{N_t} \bar{P}_\varepsilon \tilde{u}(t, Z_t^i) - Z_t^u \right) \right| = \left| \frac{H(\varepsilon) - H(0)}{\varepsilon} \right| \leq N_t M (1 + \varepsilon M)^{N_t - 1}.$$

Finally we get

$$\left| \frac{1}{\varepsilon} (Q_x^T(Z_{t+\varepsilon}^u | \mathcal{F}_t) - Z_t^u) \right| \leq N_t M (1 + \varepsilon M)^{N_t - 1} + 2M N_t + N_t^2 M^2 \varepsilon =: S_t(\varepsilon).$$

Since under  $Q_x^T$ ,  $N_t$  is dominated by the number of lineages at time  $t$  in a Yule process with birth rate  $M$  started at 1, it is geometrically distributed and so there is  $\varepsilon_0$  such that for any  $\varepsilon \in [0, \varepsilon_0]$ ,  $S_t(\varepsilon) \leq S_t(\varepsilon_0)$  and  $\mathbb{E}(S_t(\varepsilon_0)) < \infty$ . Then the Dominated Convergence Theorem ensures that

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} (Q_x^T(Z_{t+\varepsilon}^u) - Q_x^T(Z_t^u)) = 0.$$

This proves that  $Q_x^T(Z_t^u)$  is constant as a function of  $t$ , so that

$$u(T, x) = Q_x^T(Z_0^u) = Q_x^T(Z_T^u) = Q_x^T \left( \prod_{i=1}^{N_T} u(0, Z_T^i) \right) = Q_x^T \left( \prod_{i=1}^{N_T} g(Z_T^i) \right) = v(T, x),$$

which yields the announced result.  $\square$

**Remark 2.5.** *By the branching property,*

$$\begin{aligned} v(T + \varepsilon, x) - v(T, x) &= -v(T, x) + Q_x^{T+\varepsilon} \left( \prod_{i=1}^{N_{T+\varepsilon}} g(Z_{T+\varepsilon}^i), N_\varepsilon = 1 \right) + a(T) \varepsilon v(T, x)^2 + o(\varepsilon) \\ &= -v(T, x) + (1 - a(T) \varepsilon) E_x(v(T, Z_\varepsilon)) + a(T) \varepsilon v(T, x)^2 + o(\varepsilon) \\ &= E_x(v(T, Z_\varepsilon)) - v(T, x) + a(T) \varepsilon (v(T, x)^2 - v(T, x)) + o(\varepsilon). \end{aligned}$$

So for any  $t, x \geq 0$ ,

$$\lim_{\varepsilon \downarrow 0} \frac{(v(t + \varepsilon, x) - v(t, x)) - (P_\varepsilon v(t, x) - v(t, x))}{\varepsilon} = a(t) (v(t, x)^2 - v(t, x)).$$

If we could prove that  $v(\cdot, x)$  is of class  $\mathcal{C}^1$  or that  $v(t, \cdot) \in \mathcal{E}_A$ , then the RHS would equal  $\partial_t v - Av$  and  $v$  would indeed be solution to (2.15).

**2.6. Proper solutions for the  $\infty$ -population Smoluchowski equation.** In addition to the assumptions of Theorem 2.4, we now assume Grey's condition. Under this assumption, 0 is accessible and since  $Y$  is assumed to be (sub)critical

$$P_x(T_0 < \infty) = 1 \quad x \geq 0,$$

where we recall that  $T_0 = \inf\{t \geq 0 : Z_t = 0\}$ .

As in the proof of Theorem 2.4, we start with a general lemma, which is the  $\infty$ -population analog of Lemma 2.3. We make the same general assumptions with the notable difference that we only assume that  $a$  is only defined on  $(0, T)$ . We denote by  $T_{\mathcal{M}}$  the time of mass extinction of the branching particle system with birth rate  $\tilde{a}$  defined on  $[0, T)$ , i.e., the first time when all particles carry zero mass. Under  $Q_x^T$ , we denote by  $\mathcal{M}$  the event  $\{T_{\mathcal{M}} < T\}$ .

**Lemma 2.6.** *Assume that  $\int_{(0, T)} a(t) dt = \infty$ . Then there exists a unique solution  $u \in \mathcal{E}'_A$  to the PDE (or IPDE)*

$$(2.16) \quad \partial_t u = Au + a(t)(u^2 - u)$$

defined for  $t > 0$ , such that  $\limsup_{t \downarrow 0} \sup_{y \in [x, \infty)} u(t, y) < 1$  for all  $x > 0$  and  $u(t, 0) = 1$  for all  $t > 0$ . This solution is given for any  $T > 0$  by

$$u(T, x) = Q_x^T(\mathcal{M}).$$

Before proving this lemma, we wish to state the relevant corollary regarding proper (weak) solutions to (1.3) when  $A = \lambda A^\psi$ . Recall  $M^\mathfrak{t}$  is the entrance measure at 0 of the branching particle system conditional on the genealogy  $\mathfrak{t}$ , where particle masses evolve like the CSBP  $Z$  with branching mechanism  $\psi$  (and at each branching event in  $\mathfrak{t}$ , the particle with mass  $x$  undergoing division, splits into two particles, each with mass  $x$ ).

**Theorem 2.7.** *Let  $\psi$  be the Laplace exponent of a spectrally positive and (sub)critical Lévy process such that  $1/\psi$  is integrable at  $\infty$ . Then*

- (i) *There exists a unique proper weak solution  $(\mu_t; t > 0)$  to the Smoluchowski equation (1.3).*
- (ii) *For any  $T > 0$ ,  $\mu_T$  is the law of  $M^{\mathbf{T}}$  ( $\mathcal{M}^c$ ), where  $\mathbf{T}$  is the time-inhomogeneous binary tree started at 0 with one particle, stopped at time  $T$  and with birth rate  $\tilde{a}(t) = 1/(T - t)$ .*

*Proof.* The proof follows the same lines as the proof of Theorem 2.4, but here in the absence of an initial condition. If  $(\mu_t; t > 0)$  is a weak solution to the Smoluchowski equation (1.3), then its Laplace transform  $u(t, \lambda)$  is solution of (2.15). Further, the fact that the solution is proper implies that

$$\limsup_{t \downarrow 0} \sup_{y \in [x, \infty)} u(t, y) = \limsup_{t \downarrow 0} u(t, x) < 1$$

and as a consequence of the previous lemma  $u(T, x) = Q_x^T(\mathcal{M})$ . The same application of the branching property as in the proof of Theorem 2.4 shows that

$$Q_x^T(\mathcal{M}) = \mathbb{E} [\exp(-xM^{\mathbf{T}}(\mathcal{M}^c))],$$

where the expectation is taken with respect to the binary tree  $\mathbf{T}$  with branching rate  $\tilde{a}(t) = 1/(T-t)$ . More specifically, we can write conditionally on  $\mathbf{T} = \mathfrak{t}$

$$\mathbb{1}_{\mathcal{M}} = \exp\left(-\sum_i \chi_i\right),$$

where the sum is taken over the atoms of the Poisson process of branching particles with intensity  $xM^{\mathfrak{t}}$  and  $\chi_i = 0$  if the  $i$ -th particle has no descendance alive at time  $T$  and  $+\infty$  otherwise. The result is obtained by an application of the exponential formula and taking expectation with respect to  $\mathbf{T}$ .

As a consequence, we get that  $\mu_T$  is the law of  $M^{\mathbf{T}}(\mathcal{M}^c)$ . Finally, it remains to show the existence of a proper solution. One option consists of checking that  $\mathcal{L}(M^{\mathbf{T}}(\mathcal{M}^c))$  is solution. Alternatively, we will provide a construction through the MK-V approach in the next section.  $\square$

*Proof of Lemma 2.6.* Let us first prove the uniqueness part of the statement. Let  $u$  be a solution to (2.15) with the requested properties. Following the proof of Lemma 2.3,  $Q_x^T(\mathcal{Z}_t^u)$  is constant as a function of  $t \in [0, T-\varepsilon]$  for any  $\varepsilon \in (0, T)$ , so that

$$(2.17) \quad u(T, x) = Q_x^T(\mathcal{Z}_0^u) = Q_x^T(\mathcal{Z}_{T-\varepsilon}^u) = Q_x^T\left(\prod_{i=1}^{N_{T-\varepsilon}} u(\varepsilon, Z_{T-\varepsilon}^i)\right).$$

If we set  $X_\varepsilon := \prod_{i=1}^{N_{T-\varepsilon}} u(\varepsilon, Z_{T-\varepsilon}^i)$ , we can write

$$u(T, x) = Q_x^T(X_\varepsilon \mathbb{1}_{T_{\mathcal{M}} < T}) + Q_x^T(X_\varepsilon \mathbb{1}_{T_{\mathcal{M}} \geq T}).$$

On  $\mathcal{M}$ , because  $u(\varepsilon, 0) = 1$ ,  $X_\varepsilon = 1$  for any  $\varepsilon$  such that  $T_{\mathcal{M}} < T - \varepsilon$ . By Dominated Convergence,

$$\lim_{\varepsilon \downarrow 0} Q_x^T(X_\varepsilon \mathbb{1}_{\mathcal{M}}) = Q_x^T(\mathcal{M}).$$

Then it only remains to show that

$$\lim_{\varepsilon \downarrow 0} Q_x^T(X_\varepsilon \mathbb{1}_{\mathcal{M}^c}) = 0.$$

First observe that because  $a$  is not integrable in the neighborhood of  $0+$ , the birth rate  $\tilde{a}$  is not integrable in the neighborhood of  $T-$ , so that a.s.  $\lim_{\varepsilon \downarrow 0} N_{T-\varepsilon} = +\infty$ . On  $\{T_{\mathcal{M}} \geq T\}$ , there exists at least one particle born before  $T$  which carries positive mass up until time  $T$ . The probability that the mass of this particle vanishes exactly at time  $T$  is zero. As a consequence, there is  $\eta, \epsilon > 0$  such that this focal particle has mass larger than  $\eta$  on  $[T-\epsilon, T]$ . Let  $t_n$  the times at which the focal particle gives birth, where  $(t_n)$  is an increasing sequence converging to  $T$ . Let  $\eta_n \geq \eta$  be the

mass carried by the focal particle at time  $t_n$ . Then conditional on  $(t_n)$  and  $(\eta_n)$ , let  $\alpha_n$  denote the probability that the particle born at time  $t_n$  with mass  $\eta_n$  carries mass always larger than  $\eta/2$  between  $t_n$  and  $T$ . Since  $Z$  is Feller, the sequence  $(\alpha_n)$  is bounded away from 0 and by independence of these particles conditional on  $(\eta_n)$ , infinitely many of them carry mass larger than  $\eta/2$  on  $[t_n, T]$ . As a consequence, a.s. on  $\{T_{\mathcal{M}} \geq T\}$ , there is  $\eta > 0$  such that

$$X_\varepsilon \leq \prod_{i=1}^{N_{T-\varepsilon}^\eta} u(\varepsilon, Z_{T-\varepsilon}^{\eta, i})$$

where  $N_t^\eta$  is the number of particles at time  $t$  which carry more than  $\eta/2$  on  $[t, T]$  and  $(Z_t^{\eta, i})$  are their masses, which satisfy

$$\lim_{\varepsilon \downarrow 0} N_{T-\varepsilon}^\eta = +\infty \quad \text{and} \quad Z_{T-\varepsilon}^{\eta, i} \geq \eta/2.$$

Since  $\limsup_{\varepsilon \downarrow 0} \sup_{[\eta/2, \infty)} u(\varepsilon, x) < 1$ ,  $\lim_{\varepsilon \downarrow 0} X_\varepsilon = 0$  a.s. on  $\mathcal{M}^c$ . The result follows by Dominated Convergence.

The fact that  $u$ , defined as in the statement by  $u(T, x) = Q_x^T(\mathcal{M})$ , satisfies (2.15) is due to the same reasoning as in Remark 2.5, except that here  $u(\cdot, x)$  is of class  $\mathcal{C}^1$  (since  $T_0$ , and so  $T_{\mathcal{M}}$ , has a continuous density), so we can conclude that  $u(t, \cdot) \in \mathcal{E}_A$  and that indeed  $\partial_t u = Au + a(t)(u^2 - u)$ .  $\square$

**2.7. The self-similar case.** Here, we assume that  $A$  is the generator of the stable CSBP  $Z$  with Laplace exponent  $\psi(x) = cx^\gamma$ , for  $\gamma \in (1, 2]$  and  $c > 0$ . For any real number  $r$ , we denote by  $Z^{(r)}$  the CSBP with branching mechanism  $\psi(\lambda) - r\lambda$ , which is the Feller process with generator  $A_r$  defined by

$$A_r f(x) = Af(x) + rxf'(x) \quad x \geq 0.$$

Set

$$\beta = \frac{1}{\gamma - 1}.$$

Here we denote by  $Q_x^{(\beta)}$  the law of a branching particle system started with a single particle with mass  $x$ , where particles branch at rate 1 and masses follow independent copies, *not of the original CSBP (with Laplace exponent  $cx^\gamma$ ), but of the CSBP  $Z^{(\beta)}$* . Similarly, for any infinite binary tree  $\mathfrak{t}$  embedded in continuous time,  $Q_x^{(\beta), \mathfrak{t}}$  now denotes the law of the particle system started with one particle with mass  $x$  at time 0, where particle masses evolve like independent copies of  $Z^{(\beta)}$  and where the genealogy of particles is given by  $\mathfrak{t}$ . Consistently, the entrance measure at 0 of the branching particle system with genealogy  $\mathfrak{t}$  is

$$M^{(\beta), \mathfrak{t}} = \lim_{x \downarrow 0} x^{-1} Q_x^{(\beta), \mathfrak{t}}.$$

Here  $\mathcal{M}$  denotes the event of total mass extinction, i.e., the event  $\mathcal{M} = \{T_{\mathcal{M}} < \infty\}$  that *after some finite time* all particles have mass 0.



**Theorem 2.8.** *If  $(\mu_t; t \geq 0)$  is the unique proper solution of the Smoluchowski equation (1.3) advertized in Theorem 2.7, then*

(i)  $\mu_t$  is the law of  $t^{-\beta}\Upsilon$ , where

$$(2.18) \quad \Upsilon = M^{(\beta), \mathbf{T}}(\mathcal{M}^c),$$

where  $\mathbf{T}$  denotes the Yule tree, i.e., the pure-birth tree with unit birth rate.

(ii) Let  $h(x) = \mathbb{E}(\exp(-x\Upsilon))$ . Then  $h(x) \leq \exp(-x(\beta/c)^\beta)$  and  $h$  is the unique solution in  $\mathcal{E}_{A_\beta}$  to

$$(2.19) \quad \begin{aligned} A_\beta h + h^2 - h &= 0 \\ h(0) = 1, \quad \limsup_{x \rightarrow \infty} h(x) &< 1. \end{aligned}$$

*Proof.* We first recall some known facts about  $Z^{(r)}$  (see for example Section 2.2.3 in [20], in particular Corollary 2.2.3.3). For any  $x \geq 0$ , the law of  $Z^{(r)}$  started at  $x$  is denoted  $P_x^{(r)}$ . It is well-known that 0 is absorbing for  $Z^{(r)}$  and that if  $T_0$  denotes the absorbing time of  $Z^{(r)}$  at 0, then

$$P_x^{(r)}(T_0 < t) = e^{-x\varphi_r(t)} \quad x, t \geq 0,$$

where  $\varphi_r$  is the inverse of

$$\phi_r(\lambda) = \int_\lambda^\infty \frac{dx}{\psi(x) - rx} = -\frac{\beta}{r} \ln \left( 1 - \frac{r}{c\lambda^{1/\beta}} \right)$$

when  $r \neq 0$ , and if  $r = 0$ ,

$$\phi_0(\lambda) = \frac{\beta/c}{\lambda^{1/\beta}}.$$

This yields

$$\varphi_r(t) = \left( \frac{r/c}{1 - e^{-rt/\beta}} \right)^\beta$$

when  $r \neq 0$ , and if  $r = 0$ ,

$$\varphi_0(t) = \left( \frac{\beta/c}{t} \right)^\beta.$$

In particular, the probability of extinction (of a non-branching particle) is

$$P_x^{(r)}(T_0 < \infty) = \exp(-x(r/c)^\beta)$$

when  $r > 0$  and is 1 if  $r \leq 0$ . More specifically,

$$E_x^{(r)}(e^{-\lambda Z_t^{(r)}}) = e^{-x\varphi_r(t + \phi_r(\lambda))},$$

where

$$(2.20) \quad \varphi_r(t + \phi_r(\lambda)) = \left( \frac{r/c}{1 - e^{-rt/\beta} \left( 1 - \frac{r}{c\lambda^{1/\beta}} \right)} \right)^\beta$$

when  $r \neq 0$ .

Now we wish to compare the two branching particle systems. We will refer to the  $(0, T)$ -system as the time-inhomogeneous branching particle system with inhomogeneous branching rate  $\tilde{a}(t) = a(T - t) = 1/(T - t)$  blowing up at time  $T$  and masses evolving as independent copies of  $Z^{(0)}$ . We will refer to the  $(\beta, \infty)$ -system as the time-homogeneous particle system where particles branch at rate 1 and masses evolve as independent copies of  $Z^{(\beta)}$ . For either system, the genealogy of particles can be represented by the infinite binary tree, and for each node  $v$  of the infinite binary tree, we record the corresponding branching time  $U_0(v)$  (resp.  $U_\beta(v)$ ) and the mass of the corresponding particle just before it splits  $M_0(v)$  (resp.  $M_\beta(v)$ ) in the  $(0, T)$ -system (resp. the  $(\beta, \infty)$ -system). Recall from Lemma 2.6 that the probability that all masses go extinct in the  $(0, T)$ -system is  $u(T, x)$ . We denote by  $h(x)$  this probability in the  $(\beta, \infty)$ -system. We claim that  $u(T, x) = h(x/T^\beta)$ .

First, it is straightforward to check that the first branching time  $U_0$  of the first particle in the  $(0, T)$ -system is uniformly distributed in  $(0, T)$ , so that using (2.20) with  $r = \beta$ , we get

$$\begin{aligned}
E_{T^\beta x}(e^{-\lambda Z_{U_0}^{(0)}/(T-U_0)^\beta}) &= \frac{1}{T} \int_0^T ds \exp \{ -T^\beta x \varphi_0(s + \phi_0(\lambda/(T-s)^\beta)) \} \\
&= \int_0^\infty du e^{-u} \exp \{ -T^\beta x \varphi_0(T(1-e^{-u}) + \phi_0(\lambda/T^\beta e^{-\beta u})) \} \\
&= \int_0^\infty du e^{-u} \exp \left\{ -T^\beta x \left( \frac{\beta/c}{T(1-e^{-u}) + \frac{\beta/c}{\lambda^{1/\beta}/T e^{-u}}} \right)^\beta \right\} \\
&= \int_0^\infty du e^{-u} \exp \left\{ -x \left( \frac{\beta/c}{1-e^{-u} + e^{-u} \frac{\beta/c}{\lambda^{1/\beta}}} \right)^\beta \right\} \\
&= \int_0^\infty du e^{-u} e^{-x \varphi_\beta(u + \phi_\beta(\lambda))} \\
&= E_x(e^{-\lambda Z_{U_\beta}^{(\beta)}}),
\end{aligned}$$

where  $U_\beta$  is an independent exponential variable with parameter 1. By an immediate induction, we see that if the  $(0, T)$ -system starts with mass  $T^\beta x$  and the  $(\beta, \infty)$ -system starts with mass  $x$ , then the sequence of rescaled masses  $(w^{(0)}(v)/(T - U^{(0)}(v))^\beta)_v$  indexed by the binary tree is equally distributed as the sequence  $(w^{(\beta)}(v))_v$ . Now by a similar argument as the one used in the proof of Lemma 2.6, it can be seen that in both cases, there is extinction of mass if and only if all masses are zero except in a finite number of nodes  $v$  of the genealogy. This shows that  $u(T, T^\beta x) = h(x)$ , so that  $u(T, x) = h(x/T^\beta)$ , as claimed earlier. Finally, the same application of the branching property as in the proof of Theorem 2.7 shows (2.18).

Now let us show the properties of  $h$  stated in (ii) of the theorem. Let  $\mathcal{Z}_t$  denote the state (e.g., the empirical measure) at time  $t$  of the  $(\beta, \infty)$ -system. Let  $(P_t)$  denote the semigroup of  $Z^{(\beta)}$ , that

we now simply denote  $Z$  (more generally, we will omit the  $\beta$  superscript until the end of the proof). Since the first branching time  $\tau$  of the initial particle is independent of the dynamics of its mass, by the branching property

$$h(x) = E_x(h(Z_t)) \mathbb{P}(\tau > t) + \int_0^t ds \mathbb{P}(\tau \in ds) E_x(h(Z_s)^2).$$

Re-arranging, we get for any  $t > 0$

$$\frac{1}{t}(P_t h(x) - h(x)) = \frac{1 - e^{-t}}{t} P_t h(x) - \frac{1}{t} \int_0^t ds e^{-s} P_s(h^2)(x).$$

Since  $h$  is bounded continuous and  $Z$  is a Feller process, the right-hand side converges as  $t \downarrow 0$  to  $h(x) - h(x)^2$ . Then  $h$  is in the domain of  $A_\beta$  and we have

$$A_\beta h(x) = h(x) - h(x)^2 \quad x \geq 0.$$

Note that  $h(x) \in [0, 1]$  and  $h(0) = 1$ . Also notice that  $\mathcal{M} \subset \mathcal{E}$ , where  $\mathcal{E}$  is the event that the mass of a single (randomly chosen, say) lineage goes to 0. Now  $Q_x(\mathcal{E}) = P_x^{(\beta)}(T_0 < \infty) = \exp(-x(\beta/c)^\beta)$ , which yields

$$h(x) = Q_x(\mathcal{M}) \leq Q_x(\mathcal{E}) = \exp(-x(\beta/c)^\beta).$$

As a consequence  $\lim_{x \rightarrow \infty} h(x) = 0$  and so  $h$  is a solution to (2.19), which yields the existence part of the statement.

For the uniqueness part, it is sufficient to note that if  $h$  is solution to (2.19), then  $u(t, x) = h(x/t^\beta)$  is solution of the IPDE (2.15). Since the solution of this equation is unique, the result follows.  $\square$

### 3. FINITE POPULATION MCKEAN–VLASOV EQUATION

In this section, we only assume that  $\psi \in \mathcal{H}$ , which notably encompasses the case studied in the previous section, i.e., when  $\psi$  is the Laplace transform of a spectrally positive (sub)critical Lévy process.

For any  $u > 0$ ,  $\theta_u$  will be the time shift operator by  $u$  so that  $\theta_u \circ f(t) = f(t + u)$  for any generic function  $f$  of time.

We fix  $\delta > 0$  and  $\nu \in M_P(\mathbb{R}^+)$ . (Recall that we think of  $\delta$  as the inverse population size.) We consider the McKean–Vlasov equation (1.8). As already mentioned in the introduction, one may think of  $(x_t; t \geq 0)$  as the evolution of the mass of a typical cluster in the population described by the Smoluchowski equation (1.3). We start by giving a more formal definition in terms of a fixed point problem (see Proposition 3.1).

Let us consider the Skorohod space  $D(\mathbb{R}^+, \mathbb{R}^+)$  (i.e., the space of càdlàg functions equipped with the Skorohod topology on every finite interval  $[0, T]$ ). For every probability measure  $m$  on

$D(\mathbb{R}^+, \mathbb{R}^+)$ , define  $\phi(m)$  to be the law of the process

$$dy_t = -\psi(y_t)dt + \Delta J_t^{(\delta)} v_t, \quad \mathcal{L}(y_0) = \nu$$

where  $(v_t; t \geq 0)$  is a family of independent random variables with  $v_t$  being distributed as  $z_t$  – the process with law  $m$  evaluated at time  $t$  – and  $J^{(\delta)}$  is an inhomogeneous Poisson process with rate  $1/(t + \delta)$ . (More precisely, conditioned on the jump times  $\{s_i\}$  of  $J^{(\delta)}$ ,  $\{v_{s_i}\}_i$  is a sequence of independent random variables with respective law  $\mathcal{L}(z_{s_i})$ .)

We will say that  $(x_t; t \geq 0)$  is solution of the McKean–Vlasov equation (1.8) if and only if the law of the process  $x$  is a fixed point for the map  $\phi$ .

**Proposition 3.1** (Uniqueness to MK-V). *The operator  $\phi$  has a unique fixed point. As a consequence, there is at most one solution to the MK-V equation (1.8).*

*Proof.* We give a contraction argument analogous to Theorem 1.1. in [27]. For every pair of measures  $m^1, m^2$  on  $D(\mathbb{R}^+, \mathbb{R}^+)$ , and every  $T \geq 0$ , we define the Wasserstein distance

$$D_T(m_1, m_2) = \inf \left\{ \mathbb{E} \left( \sup_{s \in [0, T]} |y_s^1 - y_s^2| \right) : \mathcal{L}(y^1) = m^1, \mathcal{L}(y^2) = m^2 \right\}$$

where the infimum is taken over every possible coupling between  $y^1, y^2$  under the constraint  $\mathcal{L}(y^1) = m^1$  and  $\mathcal{L}(y^2) = m^2$ . Consider  $z^1, z^2$  be two processes in  $D(\mathbb{R}^+, \mathbb{R}^+)$  with respective laws  $m^1$  and  $m^2$  and define

$$(3.21) \quad dx_t^i = -\psi(x_t^i)dt + \Delta J_t^{(\delta)} v_t^i, \quad \mathcal{L}(x_0^i) = \nu,$$

where  $x_0^1 = x_0^2$  and  $(v_s^1, v_s^2)$  are independent random variables with  $\mathcal{L}(v_s^i) = \mathcal{L}(z_s^i), i = 1, 2$  and  $(v_s^1, v_s^2)$  are coupled in a minimal way, i.e., for every  $s \geq 0$

$$\mathbb{E}(|v_s^1 - v_s^2|) = \inf \{ \mathbb{E}(|a - b|) : \mathcal{L}(a) = z_s^1, \mathcal{L}(b) = z_s^2 \}.$$

(One can show that the minimum is attained by considering an approximating subsequence and using a standard tightness argument.) Write  $\Delta x_t = x_t^2 - x_t^1$  and note that  $\Delta x_0 = 0$ . We have

$$\Delta x_t = \Delta x_0 - \int_0^t (\psi(x_s^2) - \psi(x_s^1)) ds + \sum_{s_i \leq t : \Delta J_{s_i}^{(\delta)} = 1} (v_{s_i}^2 - v_{s_i}^1).$$

Since  $\psi$  is positive and non-decreasing in  $x$ , the part of the dynamics  $(\psi(x_s^2) - \psi(x_s^1)) ds$  can only reduce the distance between  $x^1$  and  $x^2$ , it is not hard to see that

$$\sup_{s \leq t} |\Delta x_s| \leq \sum_{s_i \leq t : \Delta J_{s_i}^{(\delta)} = 1} |v_{s_i}^2 - v_{s_i}^1|$$

and thus

$$\begin{aligned} \mathbb{E}(\sup_{s \leq t} |\Delta x_s|) &\leq \int_0^t \frac{1}{\delta + s} \mathbb{E}(|v_s^2 - v_s^1|) ds \\ &\leq \frac{1}{\delta} \int_0^t D_s(m^1, m^2) ds \end{aligned}$$

where the last inequality follows from the choice of our coupling  $(v_s^1, v_s^2)$ . This implies that

$$D_t(\phi(m_1), \phi(m_2)) \leq \frac{1}{\delta} \int_0^t D_s(m_1, m_2) ds$$

By a simple induction, this implies that

$$D_t(\phi^k(m_1), \phi^k(m_2)) \leq \frac{t^k}{\delta^k k!} D_t(m_1, m_2).$$

Thus if  $m_1$  and  $m_2$  are two fixed points for  $\phi$ , letting  $k \rightarrow \infty$ , yields that  $m_1 = m_2$ .  $\square$

**Lemma 3.2.** *Let  $(x_t; t \geq 0)$  be a solution of (1.8), so that  $(\mu_t := \mathcal{L}(x_t); t \geq 0)$  is a (MK-V) solution of the Smoluchowski equation (1.3). Then  $(\mu_t; t \geq 0)$  is also a weak solution to the Smoluchowski equation (1.3) with initial condition  $\nu$  and inverse population size  $\delta$ .*

*Proof.* By definition of the process  $x$ , for any test function  $f$ , a direct application of Itô's formula yields

$$\mathbb{E}(f(x_t)) = \mathbb{E}(f(x_0)) - \int_0^t \mathbb{E}(\psi(x_s) f'(x_s)) ds + \int_0^t \frac{1}{\delta + s} \int \mathbb{E}(f(x_s + u) - f(x_s)) \mu_s(du) ds$$

which can be rewritten as (1.5).  $\square$

**3.1. The Brownian CPP.** Let us denote by  $\mathcal{P}$  a Poisson point process with intensity measure  $dl \times \frac{dt}{t^2}$  on  $\mathbb{R}_*^+ \times \mathbb{R}_*^+$ . We call Brownian Coalescent Point Process the ultrametric tree  $\mathcal{T}$  associated with  $\mathcal{P}$

$$\mathcal{T} = \{(l, s) \in \mathbb{R}^+ \times \mathbb{R}^+ : \exists t \geq s \text{ s.t. } (l, t) \in \mathcal{P}\} \cup \{(0, t); t \in \mathbb{R}^+\}$$

equipped with the distance

$$d_{\mathcal{T}}((l', s'), (l, s)) = \begin{cases} 2 \max\{\bar{t} : (\bar{l}, \bar{t}) \in \mathcal{P}, \text{ s.t. } \bar{l} \in [l' \wedge l, l' \vee l]\} - (s + s') & \text{if } l \neq l', \\ |s - s'| & \text{otherwise.} \end{cases}$$

$\{(0, t); t \in \mathbb{R}^+\} \subset \mathcal{T}$  will be referred to as the eternal branch of the tree. See Fig. 3 for a pictorial representation of the tree  $\mathcal{T}$  above level  $\delta > 0$ .



Fix  $\delta > 0$  and let  $\{\zeta_i^{(\delta)}\}_i$  be a (possibly random) sequence in  $\mathbb{R}^+$ . Let  $\{l_i, \delta\}_i$  be the elements in  $\mathcal{T}$  with time coordinate  $\delta$ , and assume that the  $l_i$ 's are listed in increasing order. For each point  $(l, s)$  of the tree  $\mathcal{T}$  with time coordinate  $s \geq \delta$ , we assign a mark  $m_l(s)$  such that (1)  $m_{l_i}(\delta) = \zeta_i^{(\delta)}$  for every  $i \in \mathbb{N}$ , and (2) the marks with higher time coordinates are deduced (deterministically conditioned on  $\mathcal{T}$  and the initial marking) from the differential relation

$$(3.22) \quad \forall t \geq \delta, \quad dm_l^{(\delta)}(t) = -\psi(m_l^{(\delta)}(t))dt + \sum_{(l', t) \in \mathcal{P}: l' > l} \sigma_{(l, l', t)} m_{l'}^{(\delta)}(t),$$

where  $\sigma_{(l, l', t)} = 1$  if  $\sup\{\bar{t} : \bar{l} \in (l, l'], (\bar{l}, \bar{t}) \in \mathcal{P}\} = t$ , 0 otherwise.

In words, we start by marking each point at level  $\delta$  from left to right with the sequence  $\zeta_i^{(\delta)}$ ; then the marks evolve according to the ODE  $\dot{x} = -\psi(x)$  along each branch, and when two branches merge, we simply add up the values of the marks. The previous procedure defines a marking of the tree  $\mathcal{T}$  above time horizon  $\delta$ .

**Remark 3.3.** *Let us assume here that  $\psi$  is again the Laplace exponent of (sub)critical spectrally positive Lévy process.*

*The previous marking is completely identical to the one involved in the definition of the function  $F$  in (2.13). This is obviously not coincidental.*

*Let  $(\mu_t; t \geq 0)$  be the weak solution of the Smoluchowski equation. Recall from Theorem 2.4 that  $\mu_T$  is the law of  $F(\mathbf{T}, (W_i); 1 \leq i \leq N_T(\mathbf{T}))$ , where  $\mathbf{T}$  is the time-inhomogeneous tree started at 0 with one particle, stopped at time  $T$  and with birth rate  $\tilde{a}(t) = 1/(T - t + \delta)$ , the  $(W_i)$  are iid with law  $\nu$ .*

*Now, it is not hard to see that the tree originated from  $(0, T)$  and stopped at time horizon  $\delta$  in the CPP is identical in law to  $\mathbf{T}$ . From Theorem 2.4, this implies that  $(\mathcal{L}(\theta_\delta \circ m_0^{(\delta)}(t)); t \geq 0)$  is a weak solution of (1.3). We will actually show more, namely that  $(\theta_\delta \circ m_0^{(\delta)}(t); t \geq 0)$  is solution to MK-V (and with no restriction on  $\psi$ ). See Theorem 3.4.*

*This provides a natural interpretation for the expression of  $\mu_t$  in terms of the MK-V equation.*

The marks  $(m_l^{(\delta)}(t); t \geq \delta, (l, t) \in \mathcal{T})$  will be referred to as the *partial* marking of  $\mathcal{T}$  above level  $\delta$  with initial condition  $\{\zeta_i^{(\delta)}\}_i$ . When the initial marks  $\{\zeta_i^{(\delta)}\}_i$  are distributed as independent copies with law  $\nu$ , the partial marking will be referred to as the partial marking above level  $\delta$  with initial condition  $\nu$ .

Finally, a *full* marking of  $\mathcal{T}$  will refer to marks  $(m_l(t); (l, t) \in \mathcal{T})$  defined on the whole tree  $\mathcal{T}$  such that for every  $\delta > 0$ ,  $(m_l(t); t \geq \delta, (l, t) \in \mathcal{T})$  is a partial marking above level  $\delta$  (with initial marking  $\{m_{l_i}(\delta), (l_i, \delta) \in \mathcal{T}\}$ ).

**Theorem 3.4.** *Let us consider  $m^{(\delta)}$  the partial marking above level  $\delta$  with initial law  $\nu$ . Then  $(\theta_\delta \circ m_0^{(\delta)}(t); t \geq 0)$  is solution to the MK-V equation (1.8) with initial condition  $\nu$  and inverse population  $\delta$ .*

*Proof.* It is enough to show that  $(m_0^{(\delta)}(t); t \geq \delta)$  solves the McKean–Vlasov

$$(3.23) \quad \forall t \geq \delta, \quad \bar{x}_t - \bar{x}_\delta = \underbrace{-\int_\delta^t \psi(\bar{x}_s) ds}_{\text{term I}} + \underbrace{\sum_{\delta \leq s_i \leq t : \Delta \bar{J}_{s_i}^{(0)} = 1} \bar{v}_{s_i}}_{\text{term II}}, \quad \text{and } \mathcal{L}(\bar{x}_\delta) = \nu$$

where  $\bar{J}^{(0)}$  is identical in law to a Poisson Point process with rate  $1/t$  and where conditional on  $\bar{J}^{(0)}$ ,  $\{\bar{v}_{s_i}\}$  is a collection of independent random variables with  $\mathcal{L}(v_s) = \mathcal{L}(\bar{x}_s)$ .

First, the initial condition is obviously satisfied.

Secondly, we note that in the absence of a coalescence event,  $m_0^{(\delta)}(t)$  decreases at rate  $-\psi(m_0(t))$  which exactly corresponds to term I in the dynamics (3.23). For term II, we note that  $m_0^{(\delta)}(t)$  experiences a jump upon a coalescence event (see the second term on the RHS of (3.22)). Recall that the Brownian CPP is defined as a Poisson Point process with intensity rate  $dl \times dt/t^2$ , and by definition of  $m_0^{(\delta)}$  such a coalescence event occurs whenever the left-most branch “alive” at time  $t$  with a strictly positive  $l$ -coordinate dies out (see Fig. 3). This occurs at a rate

$$\frac{1}{t^2} / \int_t^\infty \frac{ds}{s^2} = \frac{1}{t} \quad \text{at time } t$$

which exactly corresponds to the rate of  $\bar{J}^{(0)}$  in (3.23). Finally, by translation invariance and the independence structure in the Brownian CPP, the branch coalescing with the eternal branch  $\{0\} \times \mathbb{R}_*^+$  carries a mark that is identical in law to  $m_0^{(\delta)}(t)$ , and independent of  $m_0^{(\delta)}(t)$ . See again Fig. 3.  $\square$

As a corollary of Proposition 3.1 and Theorem 3.4, we get the following existence and uniqueness result.

**Corollary 3.5.** *For  $\delta > 0$  and  $\nu \in M_P(\mathbb{R}^+)$ , there exists a unique solution to the MK-V equation (1.8).*

**3.3. Scaling.** For every  $\tau > 0$ , define the scaling map

$$F_\tau : (l, t) \rightarrow (\tau l, \tau t)$$

Fix  $\gamma > 1$ . We set

$$\beta := \frac{1}{\gamma - 1}$$

and for any  $\nu \in M_F(\mathbb{R}^+)$ , we define  $\mathcal{S}^{\tau, \gamma}(\nu)$  as the push-forward of the measure  $\nu$  by the map

$$x \mapsto \tau^{-\beta} x.$$



**Proposition 3.6** (scaling). *For every  $\tau > 0$*

- (i)  $\tilde{\mathcal{T}} = F_\tau(\mathcal{T})$  *is identical in law to  $\mathcal{T}$ .*
- (ii) *Assume that  $\psi(x) = cx^\gamma$  for  $c > 0$  and  $\gamma > 0$ . Let  $m^{(\delta)}$  be a partial marking above level  $\delta$  with initial measure  $\nu$ . Define*

$$(l, t) \in \tilde{\mathcal{T}}, \quad t \geq \delta\tau, \quad \tilde{m}_i^{(\delta\tau)}(t) := \frac{1}{\tau^\beta} m_{i/\tau}^{(\delta)}(t/\tau)$$

*Then  $\tilde{m}^{(\delta\tau)}$  is a partial marking of  $\tilde{\mathcal{T}}$  above level  $\tau\delta$  with initial measure  $\mathcal{S}^{\tau,\gamma}(\nu)$ .*

*Proof.* (i) is a direct consequence of the definition of the Brownian CPP. (ii) is a consequence of the observation that (a) for every  $(l, \delta\tau) \in F_\tau(\mathcal{T})$ , we have  $\mathcal{L}\left(\tilde{m}_i^{(\delta\tau)}(\tau\delta)\right) = \mathcal{S}^{\tau,\gamma}(\nu)$ , and (b) along each branch of the tree  $\tilde{\mathcal{T}}$ , the marking evolve according to the dynamics

$$dx_t = -cx_t^\gamma dt$$

(because of the pre-factor  $\frac{1}{\tau^\beta}$  in the definition of  $\tilde{m}$ ) and at a coalescence point, marks add up. (So that  $\tilde{m}$  defines a partial marking on  $\tilde{\mathcal{T}}$  with initial marking  $\mathcal{S}^{\tau,\gamma}(\nu)$ .)  $\square$

#### 4. $\infty$ -POPULATION MCKEAN-VLASOV EQUATION

In this section, we assume that  $\psi \in \mathcal{H}$  and that Grey's condition holds. Then we can define the homeomorphism  $q : (0, \infty) \rightarrow (0, V)$  as

$$q : x \mapsto \int_x^\infty \frac{1}{\psi(s)} ds,$$

with  $V = q(0+) \in (0, +\infty]$ . Define  $\phi : (0, \infty) \rightarrow (0, \infty)$  as

$$\phi(t) = \begin{cases} q^{-1}(t) & \text{if } t < V \\ 0 & \text{if } t \geq V. \end{cases}$$

Then for any  $x_0 \in (0, +\infty]$ , the ODE

$$\dot{u} = -\psi(u), \quad u(0) = x_0$$

has a unique solution on  $\mathbb{R}^+$  given by

$$u(t) = \phi(t + q(x_0)),$$

with  $q(x_0) = 0$  if  $x_0 = \infty$ . Notice that the flow  $x_0 \mapsto \phi(t + q(x_0))$  is continuous and keep in mind that  $\phi$  is the unique solution to

$$\dot{u} = -\psi(u), \quad u_0 = \infty.$$

In this section, we will also make the extra assumption that  $\psi$  is also convex. (so that  $V = \infty$ .) Note that if  $\psi$  is the Laplace exponent of a spectrally positive Lévy process the latter condition holds.

We will say that  $(x_t; t > 0)$  is an  $\infty$ -population solution of (1.8) if it is a solution of (1.8) for  $\delta = 0$  (without prescribing the initial condition at  $t = 0$ ). More precisely,  $x$  will be an  $\infty$ -pop solution if and only if for every  $\tau > 0$ , conditional on  $x_\tau$ , the process  $(x_t; t \geq \tau)$  is identical in law to the solution of

$$(4.24) \quad d\bar{x}_t = -\psi(\bar{x}_t)dt + \Delta J_t^{(0)}\bar{v}(t); t \geq \tau; \quad \bar{x}_\tau = x_\tau$$

where  $(\bar{v}_t)_{t \geq 0}$  is a family of independent random variable's with  $\mathcal{L}(\bar{v}_t) = \mathcal{L}(\bar{x}_t)$ .

Recall the definition of dust and proper solutions for the  $\infty$ -population Smoluchowski equation. (See Definition 1.8.) There is a natural extension of this definition to the MK-V equation. We will say that  $(x_t; t > 0)$  is a dust solution if it satisfies the  $\infty$ -population MK-V equation with  $\delta = 0$  and  $\lim_{t \rightarrow 0} x_t = 0$  in probability. Solutions are said to be proper otherwise. By a direct application of Itô's formula (analogous to Lemma 3.2), if  $x$  is a dust (resp., proper) solution then  $(\mathcal{L}(x_t); t > 0)$  is a dust (resp., proper) solution of the Smoluchowski equation. We state the two main results of this section.

**Theorem 4.1** (Proper solutions). (i) *There exists at least one proper solution to the  $\infty$ -population MK-V equation and any such solution satisfies*

$$(4.25) \quad \forall t > 0, x_t \geq \phi(t) \quad a.s. \quad (\text{Growth condition})$$

(ii) *In the stable case*

$$\psi(x) = cx^\gamma \quad \text{with } \gamma > 1,$$

*there exists a unique proper solution  $(x_t^{(0)}; t > 0)$ . Further,*

(Self-similarity) *For every  $t > 0$ ,  $\mathcal{L}(x_t^{(0)}) = \mathcal{L}(\Upsilon/t^\beta)$  where  $\Upsilon := x_1^{(0)}$ .*

(Integrability) *For every  $t > 0$ ,  $\mathbb{E}(x_t^{(0)}) < \infty$ .*

(Measurability)  *$(x_t^{(0)}; t > 0)$  can be constructed on the same space as the Brownian CPP, and under this coupling, it is measurable with respect to the  $\sigma$ -field generated by the CPP.*

**Theorem 4.2** (Dust solutions). *In the stable case  $\psi(x) = cx^\gamma$  ( $c > 1$  and  $\gamma > 1$ ), there exist infinitely many dust solutions to the  $\infty$ -population MK-V equation. Since the law of a dust solution to MK-V is also a weak dust solution, there exist infinitely many dust solutions to the Smoluchowski equation.*

The rest of the section is mainly dedicated to the proof of those two theorems. In Section 4.1, we prove the existence of a proper solution and derive some of its properties (growth condition, measurability, self-similarity...). In Section 4.2 we show that any proper solution satisfies the growth condition (4.25). This is shown by introducing what we call the marking associated with an  $\infty$ -population solution. In Section 4.3, we use this full marking to prove uniqueness of a proper solution

in the stable case, and we show all the required properties of the solution. (This will show part (ii) of Theorem 4.1).

Finally, in Section 4.4, we show that the long-term behavior of finite population solutions can be described in terms of the  $\infty$ -population. Finally, we close this section with the proof of Theorem 4.2.

**4.1. Existence and construction of an  $\infty$ -population proper solution.** We start with an  $\infty$ -population analog of Theorem 3.4 that will be exploited repeatedly throughout this section.

**Theorem 4.3.** *Let  $m$  be a full marking of the CPP. Assume that for every  $t > 0$ ,  $\{m_l(t) : (l, t) \in \mathcal{T}\}$  (assuming that the points are ranked according to the values of  $l$  in increasing order) is a sequence of i.i.d. random variables distributed as  $m_0(t)$ . Then  $(m_0(t); t > 0)$  is an infinite solution to MK-V.*

*Proof.* By Theorem 3.4, for every  $t > 0$ ,  $(\theta_t \circ m_0(u); u \geq 0)$  is solution of

$$\forall u \geq 0, dy_u = -\psi(y_u)du + \Delta J_u^t v_u, \quad \text{with } \mathcal{L}(y_0) = m_0(t),$$

where  $\{v_u\}_{u>t}$  is an infinite collection of r.v. with  $\mathcal{L}(y_u) = \mathcal{L}(v_u)$  for every  $u \geq t$ . Equivalently,  $(m_0(u); u \geq t)$  is identical in law to

$$\forall u \geq t, dz_u = -\psi(z_u)du + \Delta J_u^0 v_u, \quad \text{with } \mathcal{L}(z_t) = m_0(t),$$

where  $\{v_u\}_{u>t}$  is an infinite collection of r.v. distributed with  $\mathcal{L}(z_u) = \mathcal{L}(v_u)$  for every  $u \geq 0$ . Since this holds for any  $t > 0$ , this  $m_0$  is an  $\infty$ -population solution to MK-V.  $\square$

In order to construct a non-trivial  $\infty$ -population solution out of the CPP, we will consider  $(m_l^{(\delta),+}(t); (l, t) \in \mathcal{T}, t \geq \delta)$  the partial marking starting from level  $\delta > 0$  with initial measure  $\nu^+(dx) = \delta_\infty(dx)$ , i.e, we start with the initial condition  $+\infty$  at level  $\{t = \delta\}$ .

**Proposition 4.4.** *Let us consider a positive non-increasing sequence  $(\delta_n)$  going to 0.*

- (i) *For almost every realization of the CPP  $\mathcal{T}$ , for every  $(l, t) \in \mathcal{T}$ , the sequence  $(m_l^{(\delta_n),+}(t))$  is non-increasing and if we define  $m_l^+(t)$  its limit, then*

$$0 < \phi(t) \leq m_l^+(t) < \infty.$$

- (ii)  *$m^+$  is a full marking of  $\mathcal{T}$  which is measurable with respect to the  $\sigma$ -field generated by  $\mathcal{T}$  and does not depend on the sequence  $(\delta_n)$ .*
- (iii) *This marking is maximal in the sense that for every full marking  $m^-$  defined on  $\mathcal{T}$ , for every  $(l, t) \in \mathcal{T}$ ,  $m_l^+(t) \geq m_l^-(t)$ .*
- (iv)  *$(m_0^+(t); t \geq 0)$  is an  $\infty$ -population proper solution to MK-V.*
- (v) *In the stable case,  $\psi(x) = cx^\gamma$  (for  $c > 0$ ,  $\gamma > 1$ ), for every  $u > 0$ ,*

$$\mathcal{L}(u^\beta m_0^+(u)) = \mathcal{L}(m_0^+(1)).$$

*Proof.* We start with a monotonicity property of our marking of the Brownian CPP that we will use repeatedly throughout this proof. Let us first consider two partial markings  $m, \bar{m}$  of the CPP above a given level  $\delta > 0$ . It is clear that if for every  $(l, \delta) \in \mathcal{T}$  we have  $m_l(\delta) \geq \bar{m}_l(\delta)$ , then for every  $(l, t) \in \mathcal{T}$  with  $t \geq \delta$ , we must have

$$m_l(t) \geq \bar{m}_l(t).$$

Now, since  $(\delta_n)$  is non-increasing, it follows that for every  $(l_n, \delta_n) \in \mathcal{T}$

$$m_{l_n}^{(\delta_{n+1}),+}(\delta_n) < m_{l_n}^{(\delta_n),+}(\delta_n) = \infty$$

which implies that the sequence of marks  $(m_l^{(\delta_n),+}(t))$  (for any  $(l, t) \in \mathcal{T}$ ) is non-increasing and converges to a mark  $m_l(t) < \infty$ . (Note that since there are only finitely many coalescences in each compact time interval of  $(0, \infty)$ , Grey's condition ensures that the limiting marking is finite.)

Next, in the absence of coalescence along the vertical branch  $[(0, t), (0, \delta_n)]$  in  $\mathcal{T}$  (for  $\delta_n < t$ ), we have

$$m_0^{(\delta_n),+}(t) = \phi(t - \delta_n).$$

Thus, ignoring all the coalescence events along the vertical branch  $[(0, t), (0, \delta_n)]$  ensures that  $m_0^{(\delta_n),+}(t)$  is bounded from below by the RHS of the latter identity. Since coalescence events can only add extra mass to the eternal branch (and by uniqueness of the solution to the ODE), the inequality in (i) follows after taking the limit  $n \rightarrow \infty$ .

Let us now show (ii) and (iii). By continuity of the flow, it is not hard to check that  $m^+$  defines a full marking of the tree  $\mathcal{T}$  in the sense prescribed in the beginning of Section 3.2. (In other words, the property “marks evolve according to  $\dot{x}_t = -\psi(x_t)$  along branches and marks add up upon coalescence” passes to the limit.) For details, we refer the reader to the proof of Lemma 4.15 below where we develop a similar argument in more detail.

Let us show that  $m^+$  is independent of the choice of the sequence  $(\delta_n)$ . Let  $\bar{\delta}_n$  be another non-increasing sequence going to 0, and let  $\bar{m}^+$  be the limit of  $m^{(\bar{\delta}_n),+}$  as  $n \rightarrow \infty$ . Going to a subsequence of  $(\bar{\delta}_n)$  if necessary, one can always assume that  $\bar{\delta}_n \leq \delta_n$  for every  $n$ . Under this assumption, one gets  $\bar{m}_{l_n}^{(\bar{\delta}_n),+}(\delta_n) \leq m_{l_n}^{(\delta_n),+}(\delta_n)$  for every  $(l_n, \delta_n) \in \mathcal{T}$  and by similar monotonicity arguments as above this ensures that  $\bar{m}^+ \leq m^+$ . Finally, assuming that  $\bar{\delta}_n \geq \delta_n$  yields the reverse inequality. This shows that the limiting marking  $m^+$  does not depend on the choice of the sequence  $(\delta_n)$ .

Further, by construction, for any  $\delta_n \leq t$  and any  $(l, t) \in \mathcal{T}$ , we must have  $m_l^-(t) \leq m_l^{(\delta_n),+}(t)$  since at time  $\delta_n$ , the initial condition of the the marking  $m^{(\delta_n),+}$  dominates the one of  $m^-$ . This completes the proof of the first part of (ii) and (iii), i.e., that  $m^+$  is a maximal full marking of the CPP. Finally, since the initial condition of  $m_0^{(\delta_n),+}$  is deterministic and the marking  $m^{(\delta_n),+}$

below level  $\delta_n$  is determined deterministically from the tree  $\mathcal{T}$  above level  $\delta$ , we deduce that the  $(m_0^+(t); t > 0)$  is measurable with respect to  $\sigma(\mathcal{T})$ .

Let us now proceed with the proof of (iv). By the branching structure of the CPP, at every  $t > 0$ , all the marks at level  $t$  are i.i.d. with law  $\mathcal{L}(m_0^+(t))$ . By Theorem 4.3,  $m_0^+$  is an  $\infty$ -population solution to MK-V. By the inequality in (i), the solution is proper (and goes to  $\infty$  as  $t \rightarrow 0$ ).

Let us now show the scaling identity (v).

From Proposition 3.6 and the invariance of the initial condition  $\delta_\infty(dx)$  under rescaling (i.e.,  $\mathcal{S}^{\tau, \gamma}(\delta_\infty) = \delta_\infty$  for every  $\tau > 0$ ), we get for every  $t \geq \delta_n \tau$

$$(4.26) \quad m_0^{(\delta_n \tau)^+, +}(t) =_{\mathcal{L}} \frac{1}{\tau^\beta} m_0^{(\delta_n)^+, +}(t/\tau).$$

Let  $\tilde{m}^+$  be the limit of  $m^{(\delta_n \tau)^+, +}$ . Since  $m^+$  does not depend on the choice of the sequence  $(\delta_n)$  (in particular if we replace  $(\delta_n)$  by  $(\tau \delta_n)$ ),

$$m_0^+(t) = \tilde{m}_0^+(t) =_{\mathcal{L}} \frac{1}{\tau^\beta} m_0^+(t/\tau),$$

where the latter identity follows from (4.26). This completes the proof of the scaling identity after taking  $\tau = t$ .  $\square$

**4.2. Full marking associated with an  $\infty$ -population solution.** In this subsection, we construct a full marking of the CPP using an  $\infty$ -population solution to MK-V.

**Proposition 4.5.** *Let  $(x_t; t > 0)$  be an  $\infty$ -population solution to MK-V. There exists a unique full marking  $(\mathcal{T}, m)$  such that*

$$(4.27) \quad \mathcal{L}(\{m_0(t); t > 0\}) = \mathcal{L}(\{x_t; t > 0\}).$$

and at every level  $s$ ,  $\{m_l(s) : (l, s) \in \mathcal{T}\}$  is a sequence with i.i.d. random variable's with law  $\mathcal{L}(x_s)$ . This marking  $m$  will be referred to as the marking associated with the solution  $x$ .

*Proof.* We start by proving existence. The proof goes along similar lines as Theorem 4.3. Let  $(\delta_n)$  a sequence decreasing to 0. For every  $n$ , we can consider the partial marking  $m^{(\delta_n)}$  above level  $\delta_n$  with initial marking  $\mathcal{L}(x_{\delta_n})$ . By Theorem 3.4,  $(\theta_{\delta_n} \circ m_0^{(\delta_n)}(t); t \geq 0)$  is identical in law to the solution of the MK-V equation with inverse population  $\delta_n$  and initial measure  $\mathcal{L}(x_{\delta_n})$ . Equivalently, this amounts to saying that  $(m_0^{(\delta_n)}(t); t \geq \delta_n)$  is identical in law to the solution of

$$t \geq \delta_n, \quad dy_t = -\psi(y_t)dt + \Delta J_t^0 v_t, \quad y_{\delta_n} = x_{\delta_n}, \quad \text{with } \mathcal{L}(v_t) = \mathcal{L}(y_t),$$

and by uniqueness of the solution to the latter equation,  $(m_0^{(\delta_n)}(t); t \geq \delta_n)$  is identical in law to  $(x_t; t \geq \delta_n)$ . Further, the independence of the marks at level  $\delta_n$  easily implies (by the branching

structure in the CPP) that for every  $s \geq \delta_n$ , the set of marks  $\{m_l^{(\delta_n)}(s) : (l, s) \in \mathcal{T}\}$  (labelled by increasing values of  $l$ ) is a sequence of i.i.d. marks with law  $\mathcal{L}(x_s)$ .

Let  $\mathcal{T}^{\delta_n}$  be the subset of the tree  $\mathcal{T}$  consisting of all the points above time level  $\delta_n$ . By the previous argument, the sequence  $\{(\mathcal{T}^{\delta_n}, m^{(\delta_n)})\}_n$  is consistent in the sense that if we consider the marked tree  $(\mathcal{T}^{\delta_{n+1}}, m^{(\delta_{n+1})})$  below level  $\delta_n$ , the resulting object is distributed as  $(\mathcal{T}^{\delta_n}, m^{(\delta_n)})$ .

By the Kolmogorov extension theorem, we can then construct a unique full marking  $(\mathcal{T}, m)$  such that the restriction above level  $\delta_n$  coincides with  $(\mathcal{T}^{\delta_n}, m^{(\delta_n)})$ , and thus  $m$  is such that

$$\mathcal{L}(\{m_0(t); t > 0\}) = \mathcal{L}(\{x_t; t > 0\}).$$

and at every level  $s$ ,  $\{m_l(s) : (l, s) \in \mathcal{T}\}$  is a sequence with i.i.d. marks with law  $\mathcal{L}(x_s)$ .

For uniqueness, it is enough to note that the second property (i.e., the distribution of the marks at level  $s$ ) implies that any two full markings satisfying this property must be identical in law above level  $s$  for every  $s > 0$ .  $\square$

**Lemma 4.6.** *Let  $(x_t; t > 0)$  be a proper solution and let  $m$  be the associated full marking. Then  $(m_0(t); t > 0)$  (and thus  $(x_t; t > 0)$ ) satisfies the growth condition (4.25).*

*Proof.* We will first need the following preliminary step that will be also used later in this manuscript.

**Step 1.** Let us first consider a general rooted ultra-metric tree  $\mathbf{t}$  whose leaves are denoted by  $l_1, \dots, l_n$  and such that the distance between the root and the leaves is given by  $\tau$ . Consider a marking of the leaves  $M_{l_1}, \dots, M_{l_n} \in \mathbb{R}^+$ , and let us consider the marking  $M_0$  of the root obtained by propagating the marks according to the dynamics  $\dot{x} = -\psi(x)$  along the branches, and by adding the marks upon coalescence (as in the Brownian CPP). Recall that  $t \mapsto \phi(t + q(x_0))$  is the unique solution to  $\dot{u} = -\psi(u)$  with initial condition  $x_0$  at time 0. We claim that

$$(4.28) \quad \phi\left(\tau + q\left(\sum_{i=1}^n M_i\right)\right) < \tau \leq M_0 \leq \sum_{i=1}^n \phi(\tau + q(M_i)).$$

On the one hand, the RHS of the inequality corresponds to the extreme case of the star tree i.e., when all the branches coalesce simultaneously at the root of the tree (in this case the marks evolve independently along each branch, and then add up at the root). On the other hand, the LHS corresponds to the degenerate situation where all the leaves coalesce instantaneously (in which case, the marks add up to  $\sum_{i=1}^n M_i$  and then evolve along a single branch).

Since  $\psi \in \mathcal{H}$  and is convex, it is also super-linear in the sense that for  $x, y > 0$

$$\psi(x + y) \geq \psi(x) + \psi(y).$$

Recall that marks evolve according to the dynamics  $\dot{x} = -\psi(x)$  along each branch, so that the latter super-linearity assumption implies that the marking decreases faster if we collapse two branches into

a single branch, i.e., if we consider

$$\dot{r}_t = -\psi(r_t), \quad r_0 = a_0 + a_1, \quad \text{and } \forall i = 1, 2, \quad \dot{r}_t^i = -\psi(r_t^i), \quad r_0 = a_i.$$

then  $r_t \leq r_t^1 + r_t^2$ . (4.28) can easily be deduced from there and a simple induction on the number of nodes of the ultrametric tree.

**Step 2.** Let  $D_{\mathcal{T}}^{(\delta_k)}(l, t)$  be the set of descendants of  $(l, t)$  in the tree  $\mathcal{T}$  with time coordinate  $\delta_k$ . See Fig 3.  $|D_{\mathcal{T}}^{(\delta_k)}(l, t)|$  denotes the cardinality of the set  $D_{\mathcal{T}}^{(\delta_k)}(l, t)$ .

Let  $m$  be the full marking constructed from the solution  $x$ . By construction, at a given level  $\delta_k < t$ , the set of marks  $\{m_l(\delta_k)\}_{(l, \delta_k) \in D_{\mathcal{T}}^{(\delta_k)}(0, t)}$  is identical in law with  $\{y_n^k\}_{n \leq |D_{\mathcal{T}}^{(\delta_k)}(0, t)|}$  where  $\{y_n^k\}_n$  is a sequence of independent random variables with law  $\mathcal{L}(x_{\delta_k})$ , and independent of  $|D_{\mathcal{T}}^{(\delta_k)}(0, t)|$ . Since  $|D_{\mathcal{T}}^{(\delta_k)}(0, t)| \rightarrow \infty$  a.s. as  $k \rightarrow \infty$ , we have

$$M^k(0, t) := \sum_{(l, \delta_k) \in D_{\mathcal{T}}^{(\delta_k)}(0, t)} m_l(\delta_k) \rightarrow \infty \text{ in probability.}$$

Here, we used the fact that  $\mathcal{L}(m_l(\delta_k)) = \mathcal{L}(x_{\delta_k})$  and that the solution is proper. Indeed, since the solution is proper, there exist  $\varepsilon > 0$  and a sequence  $\delta'_k > 0$  converging to 0 such that  $x_{\delta'_k} \geq \varepsilon$  with probability larger than  $\varepsilon$ . Substituting  $\delta_k$  with  $\delta'_k$  yields the result.

Finally, by Step 1, we have

$$(4.29) \quad m_i(t) \geq \phi(t - \delta_k + q(M^k(0, t))),$$

and the result follows by letting  $k \rightarrow \infty$ . □

**Remark 4.7.** Note that we only used the LHS of (4.28) in the proof of the lemma. The RHS will be useful later.

**4.3. Uniqueness of a proper solution in the stable case.** In this section, we restrict our study to the stable case  $\psi(x) = cx^\gamma$  with  $\gamma > 1$ . Then  $q(x) = \beta x^{1-\gamma}/c$  and  $\phi(t) = (ct/\beta)^{-\beta}$ . As a consequence, the ODE  $\dot{z}_t = -\psi(z_t)$  with initial condition  $z_0$  at time 0 has the solution

$$(4.30) \quad \forall t \geq 0, \quad z_t = \left( c(\gamma - 1)t + z_0^{1-\gamma} \right)^{-\beta}.$$

**Proposition 4.8.** Let  $(x_t; t > 0)$  be a proper solution to MK-V. Let  $m^-$  be the full marking of the CPP  $\mathcal{T}$  associated with  $x$  (as defined in Proposition 4.5). Let  $m^+$  be the maximal full marking of  $\mathcal{T}$  (as defined in Proposition 4.4). Then  $(m_0^+(t); t > 0) = (m_0^-(t); t > 0)$  a.s.. As a consequence,  $x$  is identical in law to  $m_0^+$ .

Let  $T > 0$ . By right continuity of  $m_0^\pm$  it is enough to show that

$$(4.31) \quad m_0^+(T) = m_0^-(T) \text{ a.s..}$$

The rest of the proof is dedicated to this result. We note that in the following, we will use repeatedly the fact that  $m^\pm$  satisfy the growth condition

$$(4.32) \quad \forall t > 0, \quad m_0^\pm(t) \geq (c(\gamma - 1)t)^{-\beta},$$

(as a consequence of Lemma 4.6 and Proposition 4.4).

Let us consider the set of descendants of  $(0, T)$  belonging to the PPP  $\mathcal{P} \subset \mathcal{T}$  (i.e. the branching points in the descendance of  $(0, T)$ ) and let us denote this set by  $\mathcal{D}_{0,T}$ .  $\mathcal{D}_{0,T}$  is a set of points endowed with a natural (a.s.) binary tree structure – See Fig. 3. We can index every point in  $\mathcal{D}_{0,T}$  by a point in  $\mathbf{t} = \cup_{n \in \mathbb{N}^*} \{0\} \otimes \{0, 1\}^n$ , i.e., we construct a bijection  $G$  from  $\mathbf{t}$  to  $\mathcal{D}_{0,T}$  in such a way that

- $G(0) = (0, T)$
- If  $\kappa \in \mathbf{t}$ ,  $G(\kappa, 0)$  (resp.,  $G(\kappa, 1)$ ) is the left-child (resp., right-child) of  $G(\kappa)$  in  $\mathcal{D}_{0,T}$ .

The binary (planar) tree  $\mathbf{t}$  is naturally equipped with a triplet  $(g_\kappa^+, g_\kappa^-, d_\kappa)_{\kappa \in \mathbf{t}}$  where  $g_\kappa^\pm$  is the mark  $m_{\bar{l}}^\pm(\bar{l}-)$  where  $(\bar{l}, \bar{t}) = G(\kappa)$  and  $d_\kappa$  (the depth of the point  $\kappa$ ) is the time coordinate of the point  $G(\kappa)$  in  $\mathcal{P}$ . Note that for  $\kappa \neq 0$ , the point  $G(\kappa)$  is a branching point of the CPP, and corresponds to a discontinuity point in the marking. In our notation, the marks  $g_\kappa^\pm$  are considered right before the occurrence of the discontinuity (i.e., before we add up the marks at the branching point).

We now fix  $\kappa \in \mathbf{t}$ . Our first goal is to show the “passage formula” (4.41) below, which will be achieved through Lemmas 4.9 and 4.10. This formula will allow us to go from an arbitrary mark  $\kappa$  to the marks of its children. The desired formula (4.31) will be achieved from there by an induction on the nodes of the tree.

First, from the definition of the marking  $m^\pm$ , we can deduce the marking at  $\kappa$  from the marking of its two children  $((\kappa, 0), (\kappa, 1))$ , namely, if we consider the dynamics

$$(4.33) \quad \dot{z}_u^\pm = -\psi(z_u^\pm), \quad z_0^\pm = \sum_{i=0}^1 g_{(\kappa, i)}^\pm$$

then  $g_\kappa^\pm = z_{d_\kappa - d_{(\kappa, 0)}}^\pm = z_{d_\kappa - d_{(\kappa, 1)}}^\pm$  (since  $d_{(\kappa, 0)} = d_{(\kappa, 1)}$ ). In other words, we sum up the marks carried by the two children of  $G(\kappa)$  and let the marking evolve according to the dynamics  $\dot{x} = -\psi(x)$  along the branch connecting  $\kappa$  to its children  $(\kappa, 0)$  and  $(\kappa, 1)$ .

Alternatively, we will consider the dynamics

$$(4.34) \quad \text{for } i = 0, 1 \quad \dot{z}_u^{\pm, i} = -\psi(z_u^{\pm, i}), \quad z_0^{\pm, i} = g_{(\kappa, i)}^\pm$$

In words, instead of merging the two branches at  $G(\kappa, 0) = G(\kappa, 1)$ , we treat the two branches as if the merging had not occurred.



**Lemma 4.9.** *Let  $z^\pm$  and  $z^{\pm,i}$  be defined as above, then for every  $u \geq 0$ ,*

$$(4.35) \quad \Delta z_u \leq \sum_{i=0}^1 \Delta z_u^i, \quad \text{where } \Delta z_u := z_u^+ - z_u^-, \quad \text{and for } i = 0, 1, \quad \Delta z_u^i := z_u^{+,i} - z_u^{-,i}.$$

*Proof.* Define

$$f_t(x, y) := (c(\gamma - 1)t + (x + y)^{1-\gamma})^{-\beta} - (c(\gamma - 1)t + x^{1-\gamma})^{-\beta} - (c(\gamma - 1)t + y^{1-\gamma})^{-\beta}$$

so that according to (4.30)

$$z_t^\pm - \sum_{i=0}^1 z_t^{\pm,i} = f_t(z_0^{\pm,0}, z_0^{\pm,1}).$$

We need to prove that  $f_t(z_0^{+,0}, z_0^{+,1}) \leq f_t(z_0^{-,0}, z_0^{-,1})$ . Since  $z_0^{+,i} \geq z_0^{-,i}$  for  $i = 0, 1$ , the problem boils down to proving that the coordinates of the gradient of  $f_t$  are non-positive for  $x, y > 0$ . We have

$$\partial_x f_t(x, y) = g_t(1/(x + y)) - g_t(1/x), \quad \text{where } g_t(u) = (c(\gamma - 1)t + u^{\gamma-1})^{-\frac{\gamma}{\gamma-1}} u^\gamma,$$

and one can easily check that  $g$  is increasing in  $u$  (for  $\gamma > 1$ ), thus showing that  $\partial_x f_t(x, y) \leq 0$ . An analogous argument shows that the gradient is non-positive along the  $y$  coordinate. This completes the proof of the lemma.  $\square$

**Lemma 4.10.** *For any  $\kappa \in \mathbf{t}$ , and  $i = 0, 1$*

$$(4.36) \quad \Delta z_{d_\kappa - d_{(\kappa,i)}}^i \leq \Delta g_{(\kappa,i)} \left( \frac{d_{(\kappa,i)}}{d_\kappa} \right)^{\frac{\gamma}{\gamma-1}} \left( 1 + F\left(\frac{d_{(\kappa,i),*}}{d_\kappa}\right) \right)$$

where  $* = 0, 1$  (the value of  $d_{(\kappa,i),*}$  does not change the value of  $*$ ) and

$$\forall x \in [0, 1], \quad F(x) = \frac{(1 - 2^{1-\gamma})x^2}{1 - (1 - 2^{1-\gamma})x} \geq 0$$

*Proof. Step 1.* Recall from (4.34) that  $\Delta z_0^i = \Delta g_{(\kappa,i)}$  where  $\Delta g_{(\kappa,i)} = g_{(\kappa,i)}^+ - g_{(\kappa,i)}^-$  and

$$(4.37) \quad d\Delta z_u^i = -(\psi(z_u^{+,i}) - \psi(z_u^{-,i}))du.$$

The strategy will consist of bounding from above the RHS of the differential relation and then using Gronwall's lemma. First, by convexity of  $\psi$  and since  $z_u^{+,i} \geq z_u^{-,i}$

$$\psi(z_t^{+,i}) - \psi(z_t^{-,i}) \geq \psi'(z_u^{-,i})\Delta z_t^i = c\gamma(z_u^{-,i})^{\gamma-1}\Delta z_t^i.$$

We will now bound  $z_u^{-,i}$  on the interval  $[0, d_\kappa - d_{(\kappa,i)}]$ . Let  $v$  be such that

$$z_0^{-,i} = g_{(\kappa,i)}^- = \left( \frac{v}{c(\gamma - 1)d_{(\kappa,i)}} \right)^\beta$$

Note that since the depth of  $g_{(\kappa,i)}$  is given by  $d_{(\kappa,i)}$  and since we have  $m_t^\pm(u) \geq (\frac{1}{c(\gamma-1)u})^\beta$  (see (4.32)), we have  $v \geq 1$ . Solving the equation (4.34), we find that for any  $u \in [0, d_\kappa - d_{(\kappa,i)}]$

$$\begin{aligned} z_u^{-,i} &= \left( c(\gamma-1)u + c(\gamma-1)\frac{d_{(\kappa,i)}}{v} \right)^{-\beta} \\ &= (c(\gamma-1)(u + d_{(\kappa,i)}))^{-\beta} \left( 1 + \frac{d_{(\kappa,i)}(1-1/v)}{(u + d_{(\kappa,i)}) - d_{(\kappa,i)}(1-\frac{1}{v})} \right)^\beta \\ &\geq (c(\gamma-1)(u + d_{(\kappa,i)}))^{-\beta} \left( 1 + \frac{1-\frac{1}{v}}{\frac{d_\kappa}{d_{(\kappa,i)}} - (1-\frac{1}{v})} \right)^\beta \end{aligned}$$

where the LHS and RHS of the inequality are equal when  $u = d_\kappa - d_{(\kappa,i)}$ . Putting everything together, we get

$$(4.38) \quad \forall u \in [0, d_\kappa - d_{(\kappa,i)}], \quad \psi(z_u^{+,i}) - \psi(z_u^{-,i}) \geq \frac{\Delta z_t^i}{u + d_{(\kappa,i)}} \frac{\gamma}{\gamma-1} \left( 1 + \frac{1-\frac{1}{v}}{\frac{d_\kappa}{d_{(\kappa,i)}} - (1-\frac{1}{v})} \right).$$

**Step 2.** Since the RHS of the latter inequality increases in  $v$ , we produce a lower bound for  $v$ . By construction, the mark  $g_{(\kappa,i)}^-$  is obtained by considering the dynamics

$$(4.39) \quad dy_t = -\psi(y_t)dt, \quad y_0 = \sum_{j=0}^1 g_{(\kappa,i,j)}^-$$

evaluated at time  $d_{(\kappa,i)} - d_{(\kappa,i,*)}$  and where  $*$  = 0 or 1. Since (4.32) implies that  $g_{(\kappa,i,j)}^- \geq (c(\gamma-1)d_{(\kappa,i,*)})^{-\beta}$ , we get

$$\begin{aligned} g_{(\kappa,i)}^- &\geq \left( c(\gamma-1)(d_{(\kappa,i)} - d_{(\kappa,i,*)}) + \frac{1}{2^{\gamma-1}}c(\gamma-1)d_{(\kappa,i,*)} \right)^{-\beta} \\ &= (c(\gamma-1)d_{(\kappa,i)})^{-\beta} \left( 1 - \left(1 - \frac{1}{2^{\gamma-1}}\right) \frac{d_{(\kappa,i,*)}}{d_{(\kappa,i)}} \right)^{-\beta} \end{aligned}$$

where the RHS and LHS of the inequality are equal when the initial condition of  $(y_t; t \geq 0)$  is taken to be  $2(c(\gamma-1)d_{(\kappa,i,*)})^{-\beta}$  (where the factor 2 comes from the sum in the initial condition of (4.39)). This implies that

$$(4.40) \quad v \geq \frac{1}{1 - \left(1 - \frac{1}{2^{\gamma-1}}\right) \frac{d_{(\kappa,i,*)}}{d_{(\kappa,i)}}}.$$

**Step 3.** Combining the two previous steps yields that

$$\forall u \in [0, d_\kappa - d_{(\kappa,i)}], \quad \psi(z_u^{+,i}) - \psi(z_u^{-,i}) \geq \frac{\Delta z_t^i}{u + d_{(\kappa,i)}} \frac{\gamma}{\gamma-1} \left( 1 + G\left(\frac{d_{(\kappa,i)}}{d_\kappa}, \frac{d_{(\kappa,i,*)}}{d_{(\kappa,i)}}\right) \right).$$

where  $G(x, y) = x \frac{(1-2^{1-\gamma})y}{1-(1-2^{1-\gamma})y}$ . Further since  $G$  is increasing in  $x$  and  $y$  (since  $\gamma > 1$ ), and  $\frac{d_{(\kappa, i)}}{d_\kappa}, \frac{d_{(\kappa, i, *)}}{d_{(\kappa, i)}} \geq \frac{d_{(\kappa, i, *)}}{d_\kappa}$  we have

$$\forall u \in [0, d_\kappa - d_{(\kappa, i)}], \quad \psi(z_u^{+, i}) - \psi(z_u^{-, i}) \geq \frac{\Delta z_t^i}{u + d_{(\kappa, i)}} \frac{\gamma}{\gamma - 1} \left( 1 + F\left(\frac{d_{(\kappa, i, *)}}{d_\kappa}\right) \right).$$

Next, if we solve

$$d\bar{y}_u = -\frac{\bar{y}_u}{u + d_{(\kappa, i)}} \frac{\gamma}{\gamma - 1} \left( 1 + F\left(\frac{d_{(\kappa, i, *)}}{d_\kappa}\right) \right) du, \quad \bar{y}_0 = \Delta g_{(\kappa, i)}$$

and evaluate the solution at  $u = d_\kappa - d_{(\kappa, i)}$ , one finds the RHS of (4.36). The proof of the lemma is achieved by a direct application of Gronwall's lemma.  $\square$

Recall that  $\Delta g_\kappa = \Delta z_{d_\kappa - d_{\kappa, *}}$  and  $\Delta z = z^+ - z^-$  where  $z^+$  and  $z^-$  follow the dynamics defined in (4.33). From Lemma 4.9, we have  $\Delta g_\kappa \leq \sum_{i=0}^1 \Delta z_{d_\kappa - d_{(\kappa, i)}}^i$ . By the previous lemma, this implies that

$$\Delta g_\kappa \leq \sum_{i=0}^1 \Delta g_{(\kappa, i)} \left( \frac{d_{(\kappa, i)}}{d_\kappa} \right)^{\frac{\gamma}{\gamma-1}} \left( 1 + F\left(\frac{d_{(\kappa, i, *)}}{d_\kappa}\right) \right)$$

or equivalently that

$$(4.41) \quad d_\kappa^\beta \Delta g_\kappa \leq \sum_{i=0}^1 d_{(\kappa, i)}^\beta \Delta g_{(\kappa, i)} \left( \frac{d_{(\kappa, i)}}{d_\kappa} \right)^{\left( 1 + \frac{\gamma}{\gamma-1} F\left(\frac{d_{(\kappa, i, *)}}{d_\kappa}\right) \right)}$$

where  $F$  is defined in the previous lemma.

Let us now proceed with the rest of the proof. For  $n \geq 1$ , let  $\mathbf{t}_n = \{0\} \otimes \{0, 1\}^n$  that could be thought of as the vertices in  $\mathbf{t}$  which are at a distance  $n$  from the root. For any  $v \in \mathbf{t}_n$ , let  $v(i) \in \mathbf{t}_i$  the vertex obtained by only considering the first  $i$  coordinates of  $v$  (i.e.,  $v(i)$  is the ancestor of  $v$  at a distance  $i$  from the root). After a simple induction, we can generalize the previous inequality to the following one

$$\forall n \geq 1, \quad T^\beta \Delta g_0 \leq \sum_{v \in \mathbf{t}_{2n}} d_v^\beta \Delta g_v \prod_{i=0}^{2n-1} \left( \frac{d_{v(i+1)}}{d_{v(i)}} \right)^{1 + \frac{\gamma}{\gamma-1} F\left(\frac{d_{v(i+2)}}{d_{v(i)}}\right)}$$

(using the fact that  $d_{v(0)} = d_0 = T$ ). Define  $\epsilon(x, y) := y (x^{\gamma/(\gamma-1)F(y)})$ . The previous equation implies that (using the fact that  $F(x) \geq 0$  on  $[0, 1]^2$ )

$$(4.42) \quad \begin{aligned} \forall n \geq 1, \quad T^\beta \Delta g_0 &\leq \sum_{v \in \mathbf{t}_{2n}} d_v^\beta \Delta g_v \prod_{i=0}^{n-1} \left( \frac{d_{v(2i+1)}}{d_{v(2i)}} \right)^{\left( 1 + \frac{\gamma}{\gamma-1} F\left(\frac{d_{v(2i+2)}}{d_{v(2i)}}\right) \right)} \frac{d_{v(2i+2)}}{d_{v(2i+1)}} \\ &= \sum_{v \in \mathbf{t}_{2n}} d_v^\beta \Delta g_v \prod_{i=0}^{n-1} \epsilon\left(\frac{d_{v(2i+1)}}{d_{v(2i)}}, \frac{d_{v(2i+2)}}{d_{v(2i)}}\right) \\ &\leq \sum_{v \in \mathbf{t}_{2n}} d_v^\beta g_v^+ \prod_{i=0}^{n-1} \epsilon\left(\frac{d_{v(2i+1)}}{d_{v(2i)}}, \frac{d_{v(2i+2)}}{d_{v(2i)}}\right) \end{aligned}$$

We will now take advantage of the self-similarity in the Brownian CPP.

**Lemma 4.11.** *Let  $v \in \mathbf{t}_n$ . Then*

- (1)  $\{d_{v(i+1)}/d_{v(i)}\}_{i=0}^{n-1}$  is a sequence of i.i.d. random variables uniformly distributed on  $[0, 1]$ .
- (2)  $d_v^\beta g_v^+ = T^\beta g_0^+ = T^\beta m_0^+(T)$  in law.
- (3)  $\{d_{v(i+1)}/d_{v(i)}\}_{i=0}^{n-1}$  and  $d_v^\beta g_v^+$  are independent.

*Proof.* W.l.o.g. we take  $v = 0_n$  and  $T = 1$ , where  $0_n$  is the vector of length  $n$  filled up with 0. Let  $\tau > 0$  be random or deterministic. Define the scaling operator  $F_{1/\tau}(l, t) = (l/\tau, t/\tau)$ . Finally,

$$l_\tau := \inf\{l > 0 : (l, t) \in \mathcal{P}, t \geq \tau\}, \quad \mathcal{P}_\tau := \{(l, t) \in \mathcal{P} : l \leq l_\tau\},$$

so that  $\{l_\tau\} \times \mathbb{R}_*^+$  is the left-most branch in the tree  $\mathcal{T}$  alive at time  $\tau$ . In particular, we note that  $m_0(\tau)$  is measurable with respect to  $\mathcal{P}_\tau$  since the vertical branches of the CPP  $l$ -coordinates such that  $l \geq l_\tau$  will not coalesce with the branch  $\{0\} \times \mathbb{R}^+$  before time  $\tau$ .

From the scale invariance properties of the CPP, it is not hard to show that (i)  $\{d_{0_n(i+1)}/d_{0_n(i)}\}_{i=0}^{n-1}$  is a sequence of uniform random variables on  $[0, 1]$ , and that if we take  $\tau = d_{0_n}$  then (ii)  $F_{1/\tau}(\mathcal{P}_\tau)$  is identical in law to  $\mathcal{P}_1$ , and (iii) that  $\{d_{0_n(i+1)}/d_{0_n(i)}\}_{i=0}^{n-1}$  and  $F_{1/\tau}(\mathcal{P}_\tau)$  are independent. Further, by reasoning along the exact same lines of Proposition 4.4(v), one can show that  $\tau^\beta m_\cdot^+(\cdot\tau)$  coincides with the marking  $m^+$  on the rescaled CPP  $F_{1/\tau}(\mathcal{P}_\tau)$ . This completes the proof of the lemma.  $\square$

Passing to the expectation on both sides of (4.42) and using the previous lemma, we get that

$$\begin{aligned} \forall n \geq 1, \quad \mathbb{E}(T^\beta \Delta g_0) &\leq 2 \sum_{v \in \mathbf{t}_{2n}} \mathbb{E}(d_v^\beta g_v^+) \prod_{i=0}^{n-1} \mathbb{E} \left( \epsilon \left( \frac{d_{v(2i+1)}}{d_{v(2i)}}, \frac{d_{v(2i+2)}}{d_{v(2i)}} \right) \right) \\ (4.43) \quad &= 2 \times 4^n \times T^\beta \mathbb{E}(g_0^+) \mathbb{E}(\epsilon(U_1, U_1 U_2))^n \end{aligned}$$

where  $U_1$  and  $U_2$  are two uniform random variables on  $[0, 1]$ . From the definition of  $\epsilon$ , we have

$$\epsilon(U_1, U_1 U_2) < U_1 U_2 \text{ a.s.}$$

Since  $\mathbb{E}(U_1 U_2) = 1/4$ , we have  $\mathbb{E}(\epsilon(U_1, U_1 U_2)) < 1/4$  and thus that the lim sup of the RHS of (4.43) goes to 0 as  $n \rightarrow \infty$ . Finally, the proof of (4.31) (and thus of Proposition 4.8) is completed by the following lemma.

**Lemma 4.12.**  $\mathbb{E}(g_0^+) = \mathbb{E}(m_0^+(T)) < \infty$ .

*Proof.* Let  $n \geq 1$ . For every  $v \in \mathbf{t}_n$  we consider the dynamics

$$(4.44) \quad \dot{r}_t^v = -\psi(r_t^v), \quad r_0^v = g_v.$$

From the RHS of (4.28), we get that

$$\mathbb{E}(m_0^+(T)) \leq 2^n \mathbb{E}\left(r_{d_0-d_{0_n}}^{0_n}\right) \leq 2^n \mathbb{E}(\bar{r}_{T-d_{0_n}})$$

where  $\bar{r}$  follows the dynamics  $d\bar{r} = -\psi(\bar{r})dt$  with initial condition  $+\infty$ . Solving for  $\bar{r}$  we get

$$\bar{r}_{T-d_{0_n}} = (c(\gamma-1)(d_0-d_{0_n}))^{-\beta} = \left(c(\gamma-1)T\left(1 - \prod_{i=0}^{n-1} \frac{d_{0_n(i+1)}}{d_{0_n(i)}}\right)\right)^{-\beta}$$

and from Lemma 4.11, it remains to show that the following integral is finite for  $n$  large enough

$$I'_n = \int_{[0,1]^n} (1 - \prod_{i=1}^n u_i)^{-\beta} du_1 \cdots du_n.$$

Since the singularity of the integral is at  $(1, \dots, 1)$  we can consider the integral

$$I_n = \int_{[\frac{1}{2}, 1]^n} (1 - \prod_{i=1}^n u_i)^{-\beta} du_1 \cdots du_n.$$

Let us now make the change of variable  $\forall i \in [n]$ ,  $w_i = \prod_{j=i}^n u_j$  so that

$$\begin{aligned} I_n &= \int_{w_1 \leq \dots \leq w_n, \forall i \in [n], w_i \in [\frac{1}{2^{n-i-1}}, 1]} (1 - w_1)^{-\beta} \frac{1}{\prod_{i=2}^n w_i} dw_1 \cdots dw_n \\ &\leq \frac{1}{2^{n-1}} \int_{w_1 \leq \dots \leq w_n, \forall i \in [n], w_i \in [\frac{1}{2^{n-i-1}}, 1]} (1 - w_1)^{-\beta} dw_1 \cdots dw_n \\ &= \frac{1}{2^{n-1}} \frac{1}{\beta-1} \int_{w_2 \leq \dots \leq w_n, \forall i \in \{2, \dots, n\}, w_i \in [\frac{1}{2^{n-i-1}}, 1]} (1 - w_2)^{1-\beta} dw_2 \cdots dw_n + K_{n,\gamma} \end{aligned}$$

where  $K_{n,\gamma}$  is a finite constant. Iterating the same calculation, we get

$$I_n \leq C_{n,\gamma} \int_{w_n \in [\frac{1}{2}, 1]} (1 - w_n)^{n-1-\beta} dw_n + \bar{C}_{n,\gamma}$$

where  $C_{\gamma,n}, \bar{C}_{\gamma,n} < \infty$ . Taking  $n > \beta$  makes the integral  $I_n$  finite. This completes the proof of the lemma.  $\square$

*Proof of Theorem 4.1.* The existence of a proper MK-V solution is provided by Proposition 4.4. We proved the uniqueness of the solution in the stable case in Section 4.3. The scaling and measurability properties follow directly from Proposition 4.4. The integrability property follows from Lemma 4.12.  $\square$

**4.4. Asymptotic behavior of finite population models.** In this section, we use Theorem 4.1 to deduce some asymptotic results on the MK-V equation (1.3).

**Theorem 4.13.** *Again, we assume that  $\psi(x) = cx^\gamma$  with  $c > 0$  and  $\gamma > 1$ . Let  $\nu \in M_P(\mathbb{R}^+)$ , with  $\nu(\{0\}) < 1$ . For every  $\delta > 0$ , let  $(x_t^{(\delta)}; t \geq 0)$  be the MK-V solution with inverse population size  $\delta$  and initial measure  $\nu$ . Finally, let  $(x_t^{(0)}; t > 0)$  be the unique proper solution to MK-V.*

- (Convergence to the  $\infty$ . population solution) For every  $t > 0$

$$\lim_{\delta \downarrow 0} x_t^{(\delta)} = x_t^{(0)} \text{ in law.}$$

- (Long time behavior) For every  $\delta > 0$ ,

$$\lim_{t \uparrow \infty} x_t^{(\delta)} t^\beta = \Upsilon = x_1^{(0)} \text{ in law.}$$

**Remark 4.14.** Let  $\gamma \in (1, 2]$  and let  $(\mu_t^{(0)}, t \geq 0)$  be the unique proper weak solution to the Smoluchowski equation (1.3).  $(\mathcal{L}(x_t^{(0)}); t > 0)$  coincides with the measure valued process  $\mu^{(0)}$ . (Using the fact that there is a unique proper Smoluchowski solution, and that  $(\mathcal{L}(x_t^{(0)}); t > 0)$  is a proper weak to the Smoluchowski solution.) As a direct corollary of Theorem 4.13, we obtain the following PDE result:  $S^{t,\gamma}(\mu_t^{(\delta)}) \implies \mu_1^0$ , where the convergence is meant in the weak topology.

The proof of the previous theorem relies on the following lemma, which is a corollary of the work carried out in the previous section.

**Lemma 4.15.** Let  $\nu \in M_P(\mathbb{R}^+)$  and consider a sequence  $\{\nu^{(\delta)}\}_\delta$  in  $M_P(\mathbb{R}_+)$  with  $\nu^{(\delta)} \geq \nu$ , where the domination is meant in the stochastic sense.

Let  $(X_t^{(\delta)}; t \geq 0)$  be a solution to (1.8) with  $\mathcal{L}(X_0^{(\delta)}) = \nu^{(\delta)}$ . Assume that  $\nu(\{0\}) < 1$ . Then for every fixed  $t > 0$ ,  $\{X_t^{(\delta)}\}_\delta$  converges in law to the proper solution  $x_t^{(0)}$  as  $\delta \rightarrow 0$ .

*Proof.* Let  $\{\delta_n\}$  be a sequence of positive numbers decreasing to 0. According to Theorem 3.4,  $\mathcal{L}(X^{(\delta_n)}) = \mathcal{L}(\theta_{\delta_n} \circ m_0^{(\delta_n)})$  where  $m^{(\delta_n)}$  is the partial marking above level  $\delta_n$  with initial condition  $\nu^{(\delta_n)}$ . The strategy of the proof will consist in showing that the sequence of partial markings converges (up to a subsequence and in a sense specified below) to a full marking  $\mathbf{m}$  (Step 1). Then, we show that  $\mathbf{m}_0$  must be the (unique) proper solution of MK-V due to the condition  $\nu^{(\delta)} \geq \nu$  (Step 2).

**Step 1.** For  $j < n$ , define  $a_{ij}^n$  to be the  $i$ th mark of  $m^{(\delta_n)}$  at level  $\delta_j$  (where marks are ranked from left to right). By the branching structure of the CPP, the marks  $\{a_{ij}^n\}_i$  are i.i.d..

We first claim that for every fixed  $i, j \in \mathbb{N}$ , the sequence  $\{a_{ij}^n\}_n$  is tight. In order to see that, we first consider the case where  $\nu^{(\delta_n)}(dx) = \delta_\infty(dx)$  for every  $n$ . This exactly corresponds to the marking  $m^{(\delta_n),+}$  introduced in Proposition 4.4, for which we showed that  $\{a_{ij}^n\}_n$  converges. Since in general  $\{a_{ij}^n\}_n$  is dominated by the previous case, it follows that  $\{a_{ij}^n\}_n$  is tight.

Now,  $\{(a_{ij}^n; i, j \in \mathbb{N})\}_n$  seen as a random infinite array (equipped with the product topology) is also tight. Going to a subsequence if necessary, there exists a limiting array  $(a_{ij}^\infty; i, j \in \mathbb{N})$  such that

$$\{(\mathcal{T}, (a_{ij}^n; i, j \in \mathbb{N}))\}_n \implies (\mathcal{T}, (a_{ij}^\infty; i, j \in \mathbb{N})) \text{ as } n \rightarrow \infty,$$

where the convergence is meant in the product topology. Note that since for every fixed  $j, n$ ,  $\{a_{ij}^n\}_i$  is a sequence of i.i.d. random variables, the same holds for the sequence  $\{a_{ij}^\infty\}_i$  for every  $j$  (the marks at level  $\delta_j$  are independent).

Using the Skorohod representation theorem, one can assume w.l.o.g. that the convergence holds a.s.. Let us now fix  $j$  and let us consider  $\mathbf{m}^j$  the marking above level  $\delta_j$  with initial marks  $(a_{ij}^\infty; i \in \mathbb{N})$  (where the initial marks are assigned from left to right). By continuity of the flow, for almost every realization of the CPP and the markings, for every  $t > \delta_j$  and  $(l, t) \in \mathcal{T}$ ,  $\{m_l^{(\delta_n)}(t)\}_{n>j}$  converges a.s. to  $\mathbf{m}_l^j(t)$  under our coupling. (In other words, the convergence of the initial conditions induce the convergence of the partial marking with the limiting mark.)

Let us now take  $j > j'$  so that  $\delta_j < \delta_{j'}$ . The previous result at time  $t = \delta_{j'}$ , together with the fact that  $\{(a_{ij'}^n; i \in \mathbb{N})\}_n$  converges to  $\{(a_{ij'}^\infty; i \in \mathbb{N})\}_n$  readily implies that the marks of  $\mathbf{m}^j$  at level  $\delta_{j'}$  coincides with  $(a_{ij'}^\infty; i \in \mathbb{N})$  – the initial marking of  $\mathbf{m}^{j'}$  at level  $\delta_{j'}$ . Equivalently, this guarantees that the sequence of markings  $\{\mathbf{m}^j\}_j$  is consistent in the sense that for  $j > j'$ , the marking  $\mathbf{m}^j$  restricted to  $\{t \geq \delta_{j'}\}$  coincides  $\mathbf{m}^{j'}$ . Thus, there exists a full marking  $\mathbf{m}$  of the CPP such that  $m$  coincides with  $\mathbf{m}^j$  on  $\{t \geq \delta_j\}$ .

Gathering the previous results, we showed that (i) for every  $s > 0$ ,  $\{\mathbf{m}_l(s) : (l, s) \in \mathcal{T}\}$  is a sequence of i.i.d. random variable's; and (ii) for every realization of the CPP and the markings,

$$(4.45) \quad \forall t > 0, \quad \{m_0^{(\delta_n)}(t)\}_n \rightarrow \mathbf{m}_0(t) \quad \text{a.s..}$$

As a consequence of (i),  $(\mathbf{m}_0(t); t > 0)$  is an  $\infty$ -population solution of MK-V by virtue of Theorem 4.3.

**Step 2.** Next, by reasoning along the same lines as Proposition 4.5 (see Step 2 therein), the condition  $\nu^{(\delta)} \geq \nu$  implies that this solution satisfies the growth condition (4.25), and  $(\mathbf{m}_0(t); t > 0)$  must be proper. By the uniqueness result of Theorem 4.1, we get that  $\mathcal{L}(\mathbf{m}_0(t)) = \mathcal{L}(x_t^{(0)})$ . Finally, (4.45) together with the fact that  $\mathcal{L}(X^{(\delta_n)}) = \mathcal{L}(\theta_{\delta_n} \circ m_0^{(\delta_n)})$  (see Theorem 3.4) completes the proof the lemma.  $\square$

*Proof of Theorem 4.13.* The first item follows directly from Lemma 4.15 after taking  $\nu^{(\delta)} = \nu$ . For the second item, as a direct consequence of Proposition 3.6,

$$(4.46) \quad t^\beta x_t^{(\delta)} = \bar{x}_1^{(\delta/t)} \text{ in law}$$

where  $x^{(\delta)}$  is the solution of (1.8) with initial measure  $\nu$  and inverse population size  $\delta$ , and where  $\bar{x}^{\delta/t}$  is the solution of (1.8) with initial measure  $\mathcal{S}^{1/t, \gamma}(\nu)$  and inverse population size  $\delta/t$ . Since for  $t \geq 1$ ,  $\mathcal{S}^{1/t, \gamma}(\nu) \geq \nu$  (in the stochastic sense), the second item follows again by a direct application of Lemma 4.15.

□

#### 4.5. Dust solutions.

*Proof of Theorem 4.2.* The proof is very similar to the one of Lemma 4.15. Let  $X$  be a positive random variable with finite mean and let  $\{\delta_k\}$  be a positive sequence going to 0. Let  $m^{(\delta_k)}$  be the partial marking above level  $\delta_k$  with initial law  $\delta_k \mathcal{L}(X)$ .

**Step 1.** Up to a subsequence, one can construct a full marking  $\mathbf{m}$  of the CPP and coupling between the  $(\mathcal{T}, m^{(\delta_k)})$ 's and  $(\mathcal{T}, \mathbf{m})$  such that for every  $(l, t) \in \mathcal{T}$

$$m_l^{(\delta_k)}(t) \rightarrow \mathbf{m}_l(t) \quad \text{a.s.},$$

and such that  $\mathbf{m}_0$  is an  $\infty$ -pop solution to MK-V. The proof goes along the exact same lines as Lemma 4.15 (Step 1).

**Step 2.** To prove that  $\mathbf{m}_0$  is a dust solution, we show that  $\mathbf{m}_0(t) \implies 0$  as  $t \rightarrow 0$ . Recall the definition of  $|D^{(\delta_k)}(0, t)|$ , the number of descendants of  $(0, t)$  at time  $\delta_k$  in the CPP. From (4.28) we get the following stochastic bounds

$$\left( c(\gamma - 1)t + \left( \sum_{i=1}^{|D^{(\delta_k)}(0, t)|} \delta_k X_i \right)^{1-\gamma} \right)^{-\beta} \leq m_0^{(\delta_k)}(t) \leq \sum_{i=1}^{|D^{(\delta_k)}(0, t)|} \left( c(\gamma - 1)t + (\delta_k X_i)^{1-\gamma} \right)^{-\beta}$$

where the  $\{X_i\}$  is an infinite sequence of random variable's distributed as  $X$  and independent of the CPP. From the definition of the CPP, one can show that

$$\delta_k |D^{(\delta_k)}(0, t)| \implies \mathcal{E}(t)$$

where  $\mathcal{E}(t)$  is an exponential random variable with mean  $t$ . From here, a direct application of the Law of Large Numbers shows that the LHS (resp., RHS) of the latter inequality converges (in law) to

$$\left( c(\gamma - 1)t + (\mathbb{E}(X)\mathcal{E}(t))^{1-\gamma} \right)^{-\beta} \quad (\text{resp.}, \mathbb{E}(X)\mathcal{E}(t)).$$

As a consequence,

$$\left( c(\gamma - 1)t + (\mathbb{E}(X)\mathcal{E}(t))^{1-\gamma} \right)^{-\beta} \leq \mathbf{m}_0(t) \leq \mathbb{E}(X)\mathcal{E}(t).$$

The RHS shows that  $\mathbf{m}_0(t) \implies 0$  in law as  $t \rightarrow 0$ . Finally, the LHS of the inequality ensures that  $\mathbb{P}(\mathbf{m}_0(t) > 0) = 1$  for  $t > 0$ . This shows that  $\mathbf{m}_0$  is a non-trivial dust solution of the  $\infty$ -population MK-V equation.

Following the same approach, we can construct  $m^1, m^2, \dots, m^k$ ,  $k$  dust MK-V solutions using distinct positive random variable's  $X^1, X^2, \dots, X^k$  to initialize the underlying marking. Next, if we



choose those random variables in such a way that for every  $j < k$ ,

$$\mathbb{E}\left(m_0^j(t)\right) \leq \mathbb{E}(X^j)t < \mathbb{E}\left((c(\gamma-1)t + (\mathbb{E}(X^{j+1})\mathcal{E}(t))^{1-\gamma})^{-\beta}\right) \leq \mathbb{E}(m_0^{j+1}(t))$$

so that the law of the  $m^k$ 's must be distinct at time  $t$ . This shows that one can construct infinitely many dust solutions to MK-V.  $\square$

## 5. MAIN CONVERGENCE RESULTS

**5.1. Notation.** We will typically consider a sequence of nested coalescents indexed by  $n$ , in such way that the initial number of species increase linearly with  $n$  (see Theorem 5.1).

Let  $s_t^n$  be the number of blocks in the species coalescent (for the model indexed by  $n$ ); Further  $\tilde{s}_t^n = s_{t/n}^n$ . We order the blocks of the species coalescent at any given time by their least element.  $B_t^n(i)$  will denote the  $i^{\text{th}}$  block at time  $t$ , and  $V_t^n(i)$  the number of elements in the block;  $\tilde{B}_t^n = B_{t/n}^n$ ;  $\tilde{V}_t^n = V_{t/n}^n$ .

Define  $R^n$  to be the scaling operator acting on measure-valued process such for every  $(\nu_t; t \geq 0)$  valued in  $M_F(\mathbb{R}^+)$ , the process  $(R^n(\nu)_t; t \geq 0)$  is the only measure valued process such that for every bounded and continuous function  $f$

$$\forall t > 0, \int_{\mathbb{R}^+} f(x)R^n(\nu)_t(dx) = \int_{\mathbb{R}^+} f(x/n)\nu_{t/n}(dx).$$

In words, space and time are both rescaled by  $1/n$ .

For  $i \leq s_t^n$ , let us denote by  $\Pi_t^n(i)$  the number of gene lineages in the species block with index  $i$ ; further let  $\tilde{\Pi}_t^n := \frac{1}{n}\Pi_{t/n}^n$ . We define

$$g_t^n = \sum_{i=1}^{s_t^n} \delta_{\Pi_t^n(i)}, \text{ and } \tilde{g}_t^n = R^n \circ g_t^n = \frac{1}{\tilde{s}_t^n} \sum_{i=1}^{\tilde{s}_t^n} \delta_{\tilde{\Pi}_t^n(i)}$$

For a formal definition of  $(\Pi_t^n(i); t \geq 0)$  seen as a Markov process, we refer the reader to the generator written in (7.54).

$\mathcal{C}_0(\mathbb{R}^+)$  will denote the set of continuous functions vanishing  $+\infty$ ;  $\mathcal{C}_b^\infty(\mathbb{R}^+)$  will denote the set of infinitely differentiable functions with bounded derivatives. The Stone-Weierstrass Theorem ensures that  $\mathcal{C}_b^\infty(\mathbb{R}^+)$  is dense in  $\mathcal{C}_0(\mathbb{R}^+)$ , and is also dense in the set of test functions (i.e., the  $\mathcal{C}^1$  functions  $f$  so that  $f$  and  $f'\psi$  remain bounded).

$(M_F(\mathbb{R}^+), v)$  will refer to the set of Radon measure on  $\mathbb{R}^+$  endowed with the vague topology (i.e., the smallest topology making the map  $g \rightarrow \langle g, f \rangle$  continuous for every function  $f \in \mathcal{C}_0(\mathbb{R}^+)$ ); whereas  $(M_F(\mathbb{R}^+), w)$  will refer to  $M_F(\mathbb{R}^+)$  equipped with the weak topology (i.e., the smallest topology making the map  $g \rightarrow \langle g, f \rangle$  continuous for every function  $f \in \mathcal{C}_b(\mathbb{R}^+)$  – the set of continuous bounded functions).

## 5.2. Statement of the main results.

**Theorem 5.1.** *Consider a sequence of nested Kingman coalescents  $\{\Pi^n, s^n\}_n$ . Let  $\{X_i^n\}_{i,n}$  be an infinite array of independent random variables such that*

$$\forall i, j, n, \quad \mathcal{L}(X_i^n) = \mathcal{L}(X_j^n), \quad \limsup_n \mathbb{E}((X_1^n)^2) < \infty,$$

and independent of the species coalescent  $(\tilde{s}_t^n; t \geq 0)$ . Assume that

$$\tilde{\Pi}_0^n = \left( X_1^n, \dots, X_{\tilde{s}_0^n}^n \right)$$

Assume further that

- (1)  $s_0^n/n$  converges to  $r \in (0, \infty)$  in  $L^{2+\epsilon}$  for some  $\epsilon > 0$ .
- (2) There exists  $\nu \in M_P(\mathbb{R}^+)$  such that  $\tilde{g}_0^n \implies \nu$  in  $(M_F(\mathbb{R}^+), w)$ .

Then

$$(5.47) \quad ((\tilde{g}_t^n, \tilde{s}_t^n); t \geq 0) \implies \left( (\mu_t, \frac{2}{\frac{2}{r} + t}); t \geq 0 \right)$$

where  $(\mu_t; t \geq 0)$  is the unique weak solution of (1.3) with initial condition  $\nu$  and inverse population  $\delta = \frac{2}{r}$ . The convergence of  $\{\tilde{g}_t^n\}_n$  is meant in the Skorohod topology on  $D([0, T], (M_F(\mathbb{R}^+), w))$  for every finite interval  $[0, T]$ .

The proof of this result is postponed until Section 7.3.

In the previous theorem, we considered a sequence of nested Kingman coalescents (indexed by  $n$ ) with finite initial populations going to  $\infty$  with  $n$ . Next, we aim at investigating the case of a (single) nested Kingman coalescent where the size of the population at time  $t = 0$  is infinite.

**Definition 5.2.** *We say that  $((\Pi_t, s_t); t > 0)$  is an  $\infty$ -population nested coalescent if and only if*

- (i) *For every  $t > 0$ , conditioned on  $(\Pi_t, s_t)$ , the shifted process  $(\theta_t \circ (\Pi_u, s_u); u \geq 0)$  is a nested coalescent with initial condition  $(\Pi_t, s_t)$ .*
- (ii) *For every  $t > 0$ ,  $\Pi_t(i) \geq 1$  for every  $i \in [s_t]$ .*
- (iii)  *$s_t \rightarrow \infty$  a.s. as  $t \downarrow 0$ .*

Note that properties (i) and (iii) immediately imply that  $(s_t; t > 0)$  is distributed as the block counting process of a standard Kingman coalescent coming down from  $\infty$ . In particular,  $\frac{t}{2}s_t \rightarrow 1$  a.s. as  $t \rightarrow 0$ .

**Lemma 5.3.** *There exists an  $\infty$ -population nested Kingman coalescent.*

*Proof.* The idea is very similar to the existence of a proper solution to the  $\infty$ -population MK-V equation, as described in Section 4.1. We omit the details and only give a brief outline of the construction.

Let  $\{\delta_n\}$  be a sequence of positive numbers decreasing to 0 and let  $(s_t; t > 0)$  be a species coalescent (i.e., a standard coalescent with rate 1) coming down from infinity. For every  $n$ , define  $(\Pi_t^{(\delta_n),+}, \theta_{\delta_n} \circ s_t; t > 0)$  be the nested coalescent starting with  $s_{\delta_n}$  species and infinitely many genes per species. Finally, for every  $t > \delta_n$ , define  $\Pi_t^{n,+} = \theta_{-\delta_n} \circ \Pi_t^{(\delta_n),+}$ . It is easy to find a coupling such that for every  $t > 0$ , the sequence  $\{\Pi_t^{n,+}\}_n$  is decreasing (i.e., each of the coordinates decreases in  $n$ ) and converges to a limit  $\Pi_t^+$  a.s., whereas  $\theta_{\delta_n} \circ s_t$  obviously converges to  $s_t$ . Finally, one can easily check that  $(\Pi_t^+, s_t)$  is an  $\infty$ -population nested Kingman coalescent at  $\infty$ .  $\square$

**Remark 5.4.** *The  $\infty$ -population nested Kingman coalescent relates to the regime of Theorem 1 in [10].*

**Theorem 5.5.** *Let  $(g_t; t > 0)$  be the empirical measure of an  $\infty$ -population nested Kingman coalescent. Then*

$$\left\{ \left( (R^n \circ g_t, \frac{1}{n} s_{t/n}); t > 0 \right) \right\}_n \implies \left( \left( \mu_t^{(0)}, \frac{2}{t} \right); t > 0 \right)$$

where  $\mu^{(0)}$  is the unique proper  $\infty$ -population solution of the Smoluchowski equation (as described in Theorem 2.7) for  $\psi(x) = \frac{\varepsilon}{2} x^2$  and the convergence is meant in  $D([\tau, T], (M_F(\mathbb{R}^+), w))$  for any pair such that  $0 < \tau < T < \infty$ .

In particular, for every  $t > 0$ ,  $\mu_t^{(0)}$  is the law of  $\frac{1}{t} \Upsilon$  where  $\Upsilon$  is the r.v. defined in Theorem 2.8.

**Remark 5.6.** *In the proof of Lemma 5.3, we outlined the construction of the “maximal”  $\infty$ -population nested coalescent by starting from the  $\infty$ -initial condition at level  $\delta_n$  (and letting  $\delta_n \rightarrow 0$ ). A “minimal”  $\infty$ -population nested Kingman coalescent would consist in setting the initial condition to 1 for every species.*

*Our next result suggests that those two extremal nested coalescents are actually identical since their asymptotic empirical measures are indistinguishable. By a simple coupling argument, this should also imply that all the nested coalescents coming down from  $\infty$  are identical. In other words, we conjecture that there is a single entrance law for the nested Kingman coalescent.*

As a corollary of the previous result, we will deduce the speed of coming down from infinity in the nested Kingman coalescent.

**Theorem 5.7** (Speed of coming down from  $\infty$ ). *Let  $\rho_t := s_t \langle g_t, x \rangle$  be the number of gene lineages at time  $t$ . Then*

$$(5.48) \quad \frac{1}{n^2} \rho_{t/n} \implies_n \frac{2}{t} \int_0^\infty x \mu_t^0(dx) = \frac{2}{t^2} \mathbb{E}(\Upsilon) < \infty.$$

The proof of the two previous statements is postponed until Section 8.

## 6. SOME USEFUL ESTIMATES

In this section, we establish some estimates which will be useful in due time. This section can be skipped at first reading.

**Lemma 6.1** (Large deviations). *Let  $z^n$  be the block counting process associated with a Kingman coalescent with rate  $c > 0$  such that  $\{z_0^n\}$  is a deterministic sequence in  $\mathbb{R}^+ \cup \{+\infty\}$  such that  $z_0^n/n \rightarrow r \in (0, +\infty]$ .*

*There exist two functions  $I_+$  and  $I_-$  and two constants  $K, \bar{K}$*

$$\begin{aligned} \forall \gamma > 0, \quad \mathbb{P} \left( \frac{1}{n} z(t/n) > (1 + \gamma) \frac{2r}{2 + rct} \right) &\leq K_+ \exp(-nI_+(\gamma)) \\ \forall \gamma \in (0, 1), \quad \mathbb{P} \left( \frac{1}{n} z(t/n) < (1 - \gamma) \frac{2r}{2 + rct} \right) &\leq \bar{K}_- \exp(-nI_-(\gamma)) \end{aligned}$$

*such that  $I_\pm(\gamma) > 0$  for  $\gamma > 0$  and  $\liminf_{\gamma \rightarrow \infty} I_+(\gamma)/\gamma > 0$ .*

*Proof.* The case  $z_0^n = +\infty$  was treated in [21]; see proof of Lemma 4.6 therein. The general case can be handled by a straightforward extension of our method.  $\square$

**Corollary 6.2.** *Let  $k \in \mathbb{N}$  and  $r \in (0, \infty)$ . Let us assume that  $\{s_0^n/n\}$  converges to a deterministic  $r \in (0, \infty)$  in  $L^{k+2+\epsilon}$  for some  $\epsilon > 0$ . Then*

$$\limsup_n \mathbb{E} \left( \left( \frac{\tilde{s}_0^n}{n} \right)^k \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i)^2 \right) < \infty.$$

*Proof.* Take  $\alpha = \frac{1}{2+rcT}$ . Then

$$\begin{aligned} \mathbb{E} \left( \left( \frac{\tilde{s}_0^n}{n} \right)^k \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i)^2 \mid \tilde{s}_0^n \right) &\leq \frac{(\tilde{s}_0^n)^{k+2}}{n^k} \mathbb{P} \left( \frac{\tilde{s}_T^n}{\tilde{s}_0^n} < \alpha \mid \tilde{s}_0^n \right) + \frac{(\tilde{s}_0^n)^{k+1}}{\alpha n^k} \mathbb{E} \left( \sum_{i=1}^{\tilde{s}_T^n} \frac{\tilde{V}_T^n(i)^2}{(\tilde{s}_0^n)^2} \mid \tilde{s}_0^n \right) \\ &= \frac{(\tilde{s}_0^n)^{k+2}}{n^k} \mathbb{P} \left( \frac{\tilde{s}_T^n}{\tilde{s}_0^n} < \alpha \mid \tilde{s}_0^n \right) + \frac{(\tilde{s}_0^n)^{k+1}}{\alpha n^k} \left( (1 - \exp(-\frac{T}{n})) (1 - \frac{1}{\tilde{s}_0^n}) + \frac{1}{\tilde{s}_0^n} \right), \end{aligned}$$

where the equality simply comes from the fact that the expectation on the RHS of the inequality is the probability that two elements sampled in  $\{1, \dots, \tilde{s}_0^n\}$  (with replacement) are in the same block of a standard Kingman coalescent at time  $T/n$ . (Note that our choice of  $\alpha = \frac{1}{2+rcT}$  is motivated

by the previous large deviation estimates, so that when  $s_0^n/n \sim r$  the first term on the RHS will be negligible.)

First, since  $s_0^n/n$  converges in  $L^{k+1}$ , the second term of the RHS remains bounded in  $L^1$ . Let us now deal with the first term and show that it remains bounded in  $L^1$ . Seeking a contradiction, we assume that

$$(6.49) \quad \limsup_n \mathbb{E} \left( \frac{(\tilde{s}_0^n)^{k+2}}{n^k} 1_{\frac{s_T^n}{\tilde{s}_0^n} < \alpha} \right) = \infty.$$

Then, up to a subsequence, the latter expectation goes to  $\infty$ . Our aim is to extract a further bounded subsequence in  $L^1$ , thus yielding a contradiction. Let us now consider  $p, q > 1$  such that  $1/p + 1/q = 1$  and  $(2+k)p < 2+k+\epsilon$ . By Lemma 6.1 and our choice of  $\alpha$ , we can take  $\gamma$  small enough such that

$$\mathbb{P}(\tilde{s}_T^n < \alpha r(1+\gamma)n \mid s_0^n = [r(1-\gamma)n])$$

goes to 0 exponentially fast in  $n$ . For this choice of  $\gamma \in (0, 1)$ , one can extract a further subsequence such

$$(6.50) \quad \limsup_n n^{2q} \mathbb{P}(s_0^n/n \notin [r(1-\gamma), r(1+\gamma)]) < \infty.$$

For the rest of the proof, we will work under this subsequence. Next,

$$\begin{aligned} \mathbb{E} \left( \frac{(\tilde{s}_0^n)^{k+2}}{n^k} 1_{\frac{s_T^n}{\tilde{s}_0^n} < \alpha} \right) &\leq \mathbb{E} \left( \frac{(\tilde{s}_0^n)^{k+2}}{n^k}, \frac{\tilde{s}_T^n}{\tilde{s}_0^n} < \alpha \text{ and } s_0^n/n \in [r(1-\gamma), r(1+\gamma)] \right) \\ &\quad + \mathbb{E} \left( \frac{(\tilde{s}_0^n)^{k+2}}{n^k}, s_0^n/n \notin [r(1-\gamma), r(1+\gamma)] \right). \end{aligned}$$

We first deal with the first term on the RHS of the inequality that we call (i).

$$\begin{aligned} (i) &\leq r^{2+k}(1+\gamma)^{2+k} n^2 \mathbb{P}(\tilde{s}_T^n < \alpha r(1+\gamma)n, s_0^n/n \in [r(1-\gamma), r(1+\gamma)]) \\ &\leq r^{2+k}(1+\gamma)^{2+k} n^2 \mathbb{P}(\tilde{s}_T^n < \alpha r(1+\gamma)n \mid s_0^n = [r(1-\gamma)n]), \end{aligned}$$

where the RHS goes to 0 by our choice of  $\alpha$  and  $\gamma$ . We now deal with the second term. By Hölder's inequality, we have

$$\mathbb{E} \left( \frac{(\tilde{s}_0^n)^{k+2}}{n^k}, s_0^n \notin [r(1-\gamma), r(1+\gamma)] \right) \leq \mathbb{E} \left( \left( \frac{s_0^n}{n} \right)^{(k+2)p} \right)^{\frac{1}{p}} (n^{2q} \mathbb{P}(s_0^n \notin [r(1-\gamma), r(1+\gamma)]))^{1/q}.$$

Since  $\frac{s_0^n}{n}$  remains bounded in  $L^{2+k+\epsilon}$ , and because of (6.50), this shows that  $\limsup \mathbb{E} \left( \frac{(\tilde{s}_0^n)^{k+2}}{n^k} 1_{\frac{s_T^n}{\tilde{s}_0^n} < \alpha} \right) < \infty$ , which is the desired contradiction. This completes the proof of the lemma.  $\square$

**Lemma 6.3.** *Assume that  $\{s_0^n/n\}$  converges to a deterministic  $r \in (0, \infty)$  in  $L^{3+\epsilon}$  for some  $\epsilon > 0$ . Let  $(\xi_i; i \in \mathbb{N})$  be i.i.d. block counting processes of Kingman coalescent with rate  $c > 0$  coming down*

from infinity and independent of the species coalescent. Then for every  $0 < \tau < T$ :

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left( \sup_{[\tau, T]} \beta_t^n \right) < \infty \quad \text{where} \quad \beta_t^n := \frac{1}{\tilde{s}_t^n} \sum_{i=1}^{\tilde{s}_t^n} \left( \sum_{j \in \tilde{B}_t^n(i)} \frac{1}{n} \xi_j(t/n) \right)^2.$$

and where  $\tilde{B}_t^n$  was defined in Section 5.1.

*Proof.* For every  $k \in \mathbb{N}$ , we define the events  $\tilde{A}_t^{k,n} = \{ \frac{2kn}{ct} \leq \max_{i \in [\tilde{s}_0^n]} \xi_i(t/n) \leq \frac{2(1+k)n}{ct} \}$  and  $\bar{A}_{t,T}^{k,n} = \{ \frac{2kn}{cT} \leq \max_{i \in [\tilde{s}_0^n]} \xi_i(t/n) \}$ . Take  $k_0 \in \mathbb{N}$  such that  $k_0 \frac{\tau}{T} > 1$

$$\begin{aligned} \beta_t^n &\leq \left( \left( \frac{2k_0}{ct} \right)^2 + \sum_{k=k_0}^{\infty} \left( \frac{2(k+1)}{ct} \right)^2 1_{\tilde{A}_t^{k,n}} \right) \frac{1}{\tilde{s}_t^n} \sum_{i=1}^{\tilde{s}_t^n} \tilde{V}_t^n(i)^2 \\ &\leq \left( \left( \frac{2k_0}{ct} \right)^2 + \sum_{k=k_0}^{\infty} \left( \frac{2(k+1)}{ct} \right)^2 1_{\bar{A}_{t,T}^{k,n}} \right) \frac{1}{\tilde{s}_t^n} \sum_{i=1}^{\tilde{s}_t^n} \tilde{V}_t^n(i)^2. \end{aligned}$$

Next, for every  $t \leq T$ , let us denote by  $\tilde{C}_{t,T}^n(i)$  the indices of the blocks at time  $t/n$  partitioning the block  $i$  at time  $T/n$ . (In particular,  $\tilde{C}_{0,T}^n(i) = \tilde{B}_T^n(i)$ .) Then

$$\begin{aligned} \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i)^2 &= \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \left( \sum_{j \in \tilde{C}_{t,T}^n(i)} \tilde{V}_t^n(j) \right)^2 \\ &\geq \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \sum_{j \in \tilde{C}_{t,T}^n(i)} \tilde{V}_t^n(j)^2 = \frac{1}{\tilde{s}_T^n} \sum_{k=1}^{\tilde{s}_t^n} \tilde{V}_t^n(k)^2 \\ (6.51) \quad &\geq \frac{1}{\tilde{s}_t^n} \sum_{k=1}^{\tilde{s}_t^n} \tilde{V}_t^n(k)^2 \end{aligned}$$

which implies that  $\sup_{[\tau, T]} \frac{1}{\tilde{s}_t^n} \sum_{i=1}^{\tilde{s}_t^n} \tilde{V}_t^n(i)^2 = \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i)^2$ . Further, since  $\xi_i$  is non-increasing

$$\sup_{[\tau, T]} \beta_t^n \leq \left( \left( \frac{2k_0}{c\tau} \right)^2 + \sum_{k=k_0}^{\infty} \left( \frac{2(k+1)}{c\tau} \right)^2 1_{\bar{A}_{\tau, T}^{k,n}} \right) \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i)^2.$$

Since the  $\xi_i$ 's are independent of the coalescent, this yields

$$\begin{aligned} \mathbb{E}(\sup_{[\tau, T]} \beta_t^n \mid \tilde{s}_0^n) &\leq \left( \left( \frac{2k_0}{c\tau} \right)^2 + \sum_{k=k_0}^{\infty} \left( \frac{2(k+1)}{c\tau} \right)^2 \mathbb{P}(\bar{A}_{\tau, T}^{k,n} \mid \tilde{s}_0^n) \right) \mathbb{E} \left( \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i)^2 \mid \tilde{s}_0^n \right) \\ &\leq \left( \left( \frac{2k_0}{c\tau} \right)^2 + \sum_{k=k_0}^{\infty} \left( \frac{2(k+1)}{c\tau} \right)^2 \tilde{s}_0^n \mathbb{P} \left( \frac{1}{n} \xi(\tau/n) \geq \frac{2k}{cT} \right) \right) \mathbb{E} \left( \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i)^2 \mid \tilde{s}_0^n \right) \end{aligned}$$

Using Lemma 6.1

$$\mathbb{E}(\sup_{[\tau, T]} \beta_t^n \mid \tilde{s}_0^n) \leq \left( \left( \frac{2k_0}{c\tau} \right)^2 + \frac{\tilde{s}_0^n}{n} \sum_{k=k_0}^{\infty} \left( \frac{2(k+1)}{c\tau} \right)^2 n K_+ \exp(-n I_+(k \frac{\tau}{T} - 1)) \right) \mathbb{E} \left( \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i)^2 \mid \tilde{s}_0^n \right)$$

Recall that for every  $k \geq k_0$ , we have  $k\frac{\tau}{T} > 1$ . Since  $\liminf_{\gamma \rightarrow \infty} I_+(\gamma)/\gamma > 0$ , a straightforward application of the dominated convergence theorem shows that the sum on the RHS of the inequality goes to 0 as  $n \rightarrow \infty$ . Thus, there exists a constant  $C$  (independent of  $n$ ) such that

$$\mathbb{E}(\sup_{[\tau, T]} \beta_t^n \mid \tilde{s}_0^n) \leq \left( \left( \frac{2k_0}{c\tau} \right)^2 + C \frac{\tilde{s}_0^n}{n} \right) \mathbb{E} \left( \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i)^2 \mid \tilde{s}_0^n \right)$$

The fact that the RHS of the inequality is uniformly bounded in  $L^1$  is handled by Corollary 6.2 (with  $k = 0, 1$ ).  $\square$

## 7. CONVERGENCE OF THE EMPIRICAL MEASURE

In this section, we assume a sequence of nested coalescents  $\{(\Pi_t^n, s_t^n); t \geq 0\}$  indexed by  $n$ . Assume that there exists  $r \in (0, \infty)$  such that

$$(7.52) \quad \tilde{s}_0^n/n \rightarrow r \text{ in } L^1.$$

We will also assume that

$$(7.53) \quad \forall T > 0, \quad \limsup_n \mathbb{E} \left( \max_{[0, T]} \langle \tilde{g}_t^n, x^2 \rangle \right) < \infty.$$

As we shall see later, this condition will be satisfied under the initial conditions specified in Theorem 5.1, and will also appear naturally in the infinite population nested Kingman coalescent.

**7.1. Generators.** We start with some definition. The process of genetic composition  $(\tilde{\Pi}_t^n; t \geq 0)$  defines a Markov process valued in the space

$$E := \cup_{k \in \mathbb{N}_*} \mathbb{R}_+^k.$$

We define  $E_n$  to be the subspace of  $E$  such that every coordinate of  $\Pi \in E$  is such that  $n\Pi(i) \in \mathbb{N}_*$ .

For every  $\Pi \in E$ , we define  $|\Pi|$  as the number of entries in  $\Pi$ . In particular, we have  $|\tilde{\Pi}_t^n| = \tilde{s}_t^n$ . Finally, when  $|\Pi| > 1$ , for every  $i < j \leq |\Pi|$ ,  $\theta_{ij}(\Pi)$  is the only element  $\Pi' \in E$  such that  $\Pi'$  is obtained by coagulating coordinates  $i$  and  $j$ . More precisely,  $\theta_{ij}(\Pi)$  is a vector of size  $|\Pi| - 1$  with coordinates

$$\forall k < |\Pi|, \quad \theta_{i,j}(\Pi)(k) = \begin{cases} \Pi(i) + \Pi(j) & \text{if } k = i \\ \Pi(k+1) & \text{if } k \geq j \\ \Pi(k) & \text{otherwise} \end{cases}$$

For instance,

$$\text{if } \Pi = (X^1, X^2, X^3, X^4), \text{ then } \theta_{1,3}(\Pi) = (X^1 + X^3, X^2, X^4).$$

Finally, for every  $\Pi \in E$ , we define  $g_\Pi = \frac{1}{|\Pi|} \sum_{i=1}^{|\Pi|} \delta_{\Pi(i)}$ , the empirical measure associated to the genetic composition  $\Pi$ .

Let us now describe the generator of  $(\tilde{\Pi}_t^n; t \geq 0)$  describing the evolution of the number of gene lineages per species. Define  $e_i^k$  to be the vector of size  $k$  filled with zeros except for the  $i^{\text{th}}$  coordinate which is equal to 1. Then for every bounded function  $h$  from  $E$  to  $\mathbb{R}$ , and every  $\Pi \in E_n$  we have

$$(7.54) \quad G^n h(\Pi) := \underbrace{\frac{c}{n} \sum_{i=1}^n \frac{n\Pi(i)(n\Pi(i)-1)}{2} \left( h\left(\Pi - \frac{1}{n} e_i^{|\Pi|}\right) - h(\Pi) \right)}_{\text{term I}} + \underbrace{\frac{1}{n} 1_{|\Pi|>1} \sum_{i<j \leq |\Pi|} (h \circ \theta_{ij}(\Pi) - h(\Pi))}_{\text{term II}}$$

where the first term corresponds to a coalescence of two gene lineages (belonging to the same species), and the second term corresponds to a coalescence of two species lineages. Finally,

$$(7.55) \quad \tilde{g}_t^n = \frac{1}{|\tilde{\Pi}_t^n|} \sum_{i=1}^{|\tilde{\Pi}_t^n|} \delta_{\tilde{\Pi}_t^n(i)},$$

corresponds to the empirical measure of the block masses, where the mass of a block is measured in terms of its (renormalized) number of gene lineages.

Before going to the convergence of the empirical measure  $\tilde{g}^n$ , we will need to establish a few technical lemmas related to the generator of the process  $(\tilde{\Pi}_t^n; t \geq 0)$ . For every  $f \in C_b^\infty(\mathbb{R}^+)$ , we define

$$(7.56) \quad X_t^{n,f} := \langle \tilde{g}_t^n, f \rangle$$

Note that  $X_t^{n,f}$  can be regarded as a function of  $\tilde{\Pi}_t^n$ . We call  $h^f$  this function ( $h^f(\Pi) = \langle g_\Pi, f \rangle$ ), in such a way that  $X_t^{n,f} := h^f(\tilde{\Pi}_t^n)$ ,

**Lemma 7.1** (Generator approximation). *Assume that conditions (7.52) and (7.53) hold. For every  $f \in C_b^\infty(\mathbb{R}^+)$ ,  $\Pi \in E_n$ , define*

$$\bar{G}h^f(\Pi) = - \left\langle g_\Pi, \frac{cx^2}{2} f' \right\rangle + \frac{r}{2+rs} \int_{\mathbb{R}^2} g_\Pi(dx) g_\Pi(dy) (f(x+y) - f(x)).$$

Then for every  $t \geq 0$

$$\mathbb{E} \left( \int_0^t (G^n - \bar{G}) h^f(\tilde{\Pi}_s^n) ds \right) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

*Proof.* We first note that

$$\mathbb{E} \left( \int_0^t \left| \frac{|\tilde{\Pi}_s^n|}{2n} - \frac{r}{2+rs} \right| \int_{\mathbb{R}^2} g_{\tilde{\Pi}_s^n}(dx) g_{\tilde{\Pi}_s^n}(dy) (f(x+y) - f(x)) ds \right) \leq 2\|f\|_\infty \mathbb{E} \left( \int_0^t \left| \frac{|\tilde{\Pi}_s^n|}{2n} - \frac{r}{2+rs} \right| ds \right).$$

From Lemma 6.1 and using the fact that  $(|\tilde{\Pi}_t^n|/n; t \geq 0)$  is non-increasing

$$(|\tilde{\Pi}_t^n|/n; t \geq 0) \rightarrow \left( \frac{2r}{2+rt}; t \geq 0 \right) \quad \text{in probability}$$

(where the convergence is meant in the Skorohod topology on every interval  $[0, T]$ ) so that the integrand on the RHS of the latter inequality goes to 0 in probability. Further, by assumption



$\{|\tilde{\Pi}_0^n|/n\}$  is uniformly integrable, and since  $|\tilde{\Pi}_0^n| \geq |\tilde{\Pi}_t^n|$  it easily follows (by uniform integrability in  $(\Omega \times [0, t], \mathbb{P} \times dt)$ ) that

$$\mathbb{E}\left(\int_0^t \left| \frac{|\tilde{\Pi}_s^n|}{2n} - \frac{r}{2+rs} \right| ds\right) \rightarrow 0$$

so that the LHS of the latter inequality vanishes. From our assumptions, it is then sufficient to show the existence of a constant  $K$  such that for every  $\Pi \in E_n$  and  $f \in C_b^\infty(\mathbb{R}^+)$

$$(7.57) \quad \left| G^n h^f(\Pi) - \left( -\left\langle g_\Pi, \frac{cx^2}{2} f' \right\rangle + \frac{|\Pi|}{2n} \int_{\mathbb{R}^2} g_\Pi(dx) g_\Pi(dy) (f(x+y) - f(x)) \right) \right| \\ \leq \frac{K}{n} (\|f'\|_\infty \langle g_\Pi, x \rangle + \|f''\|_\infty \langle g_\Pi, x^2 \rangle) + K \times 1_{|\Pi|>1} \frac{1}{|\Pi|-1} \|f\|_\infty.$$

(Indeed, the first term on the RHS goes to 0 as  $n \rightarrow \infty$ . For the second term, it is enough to notice that  $|\tilde{\Pi}_t^n| \rightarrow \infty$  in probability and direct application of the Dominated Convergence Theorem implies that this term converges to 0 in  $L^1([0, 1] \times \Omega, dt \otimes d\mathbb{P})$ ). We start by approximating term I of the generator  $G^n$  (as defined in (7.54)). For every  $\Pi \in E_n$ , we have

$$I = \frac{c}{n} \sum_{i=1}^{|\Pi|} \frac{n\Pi(i)(n\Pi(i)-1)}{2} \left( h^f\left(\Pi - \frac{1}{n} e_i^{|\Pi|}\right) - h^f(\Pi) \right) \\ = \frac{c}{|\Pi|n} \sum_{i=1}^{|\Pi|} \frac{n\Pi(i)(n\Pi(i)-1)}{2} \left( f\left(\Pi(i) - \frac{1}{n}\right) - f(\Pi(i)) \right)$$

and by a simple Taylor expansion, it follows that there exists a constant  $c_1$  such that

$$|I + \left\langle g_\Pi, \frac{cx^2}{2} f' \right\rangle| \leq c_1 \left( \frac{\|f'\|_\infty}{n} \langle g_\Pi, x \rangle + \frac{\|f''\|_\infty}{n} \langle g_\Pi, x^2 \rangle \right).$$

Let us now deal with term  $II$  of the generator  $G^n$  (again as defined in (7.54)). For  $\Pi$  such that  $|\Pi| > 1$ , consider the measure

$$(7.58) \quad \nu_\Pi(dxdy) = g_\Pi(dx) \frac{|\Pi|}{|\Pi|-1} \left( g_\Pi(dy) - \frac{1}{|\Pi|} \delta_x(dy) \right),$$

i.e.,  $\nu_\Pi$  is the measure that consists of sampling two elements with no replacement according to the measure  $g_\Pi$ . Then

$$II = 1_{|\Pi|>1} \frac{|\Pi|(|\Pi|-1)}{2n} \int_{(\mathbb{R}^+)^2} \nu_\Pi(dxdy) \\ \times \left( \frac{|\Pi|}{|\Pi|-1} \langle g_\Pi, f \rangle + \frac{1}{|\Pi|-1} f(x+y) - \frac{1}{|\Pi|-1} f(x) - \frac{1}{|\Pi|-1} f(y) - \langle g_\Pi, f \rangle \right) \\ = 1_{|\Pi|>1} \frac{|\Pi|(|\Pi|-1)}{2n} \int_{(\mathbb{R}^+)^2} \nu_\Pi(dxdy) \left( \frac{1}{|\Pi|-1} \langle g_\Pi, f \rangle + \frac{1}{|\Pi|-1} f(x+y) - \frac{1}{|\Pi|-1} f(x) - \frac{1}{|\Pi|-1} f(y) \right) \\ = 1_{|\Pi|>1} \frac{|\Pi|}{2n} \int_{(\mathbb{R}^+)^2} \nu_\Pi(dxdy) (\langle g_\Pi, f \rangle + f(x+y) - f(x) - f(y))$$

In words, we coalesce two species lineages at rate  $\frac{|\Pi|(|\Pi|-1)}{2n}$ . Conditional on a coalescence event, we pick two species lineages according to the measure  $\nu_\Pi(dx dy)$ . If we pick two lineages with coordinate  $x$  and  $y$  respectively, then the change in  $\langle g_\Pi, f \rangle$  is readily given by the term between parenthesis.

By using (7.58), one gets the existence of a constant  $c_2$  such that

$$|II - 1_{|\Pi|>1} \frac{|\Pi|}{2n} \int_{(\mathbb{R}^+)^2} g_\Pi(dx) g_\Pi(dy) (\langle g_\Pi, f \rangle + f(x+y) - f(x) - f(y))| \leq 1_{|\Pi|>1} c_2 \frac{\|f\|_\infty}{|\Pi| - 1}$$

Using the fact that

$$\int_{(\mathbb{R}^+)^2} g_\Pi(dx) g_\Pi(dy) (\langle g_\Pi, f \rangle + f(x+y) - f(x) - f(y)) = \int_{(\mathbb{R}^+)^2} g_\Pi(dx) g_\Pi(dy) (f(x+y) - f(x))$$

this yields

$$|II - 1_{|\Pi|>1} \frac{|\Pi|}{2n} \int_{(\mathbb{R}^+)^2} g_\Pi(dx) g_\Pi(dy) (f(x+y) - f(x))| \leq c_2 1_{|\Pi|>1} \frac{\|f\|_\infty}{|\Pi| - 1}$$

which is the desired inequality (7.57).  $\square$

**Lemma 7.2.** *For every  $0 \leq u \leq t \leq T$*

$$|\langle X^{n,f} \rangle_t - \langle X^{n,f} \rangle_u| \leq \frac{C}{n} \int_u^t (\|f'\|_\infty^2 \langle \tilde{g}_t^n, x^2 \rangle + \|f\|_\infty^2) ds$$

*Proof.*  $h^f(\tilde{\Pi}^n)$  is a pure jump process and its bracket term  $\langle X^{n,f} \rangle_t - \langle X^{n,f} \rangle_u$  can be decomposed into two terms, i.e.  $\langle X^{n,f} \rangle_t - \langle X^{n,f} \rangle_u = \int_u^t I'_s ds + \int_u^t II'_s ds$  where

$$\begin{aligned} I'_t &= \frac{c}{n} \sum_{i=1}^{\tilde{s}_t^n} \frac{n\tilde{\Pi}_t^n(i)(n\tilde{\Pi}_t^n(i) - 1)}{2} \left( h^f(\tilde{\Pi}_t^n - \frac{1}{n} e_i^{\tilde{s}_t^n}) - h^f(\tilde{\Pi}_t^n) \right)^2 \\ &= \frac{c}{n|\tilde{\Pi}_t^n|^2} \sum_{i=1}^{|\tilde{\Pi}_t^n|} \frac{n\tilde{\Pi}_t^n(i)(n\tilde{\Pi}_t^n(i) - 1)}{2} \left( f(\tilde{\Pi}_t^n(i) - 1/n) - f(\tilde{\Pi}_t^n(i)) \right)^2 \end{aligned}$$

and

$$\begin{aligned} II'_t &= 1_{|\tilde{\Pi}_t^n|>1} \\ &\times \frac{|\tilde{\Pi}_t^n|(|\tilde{\Pi}_t^n| - 1)}{2n} \int_{(\mathbb{R}^+)^2} \nu_{\tilde{\Pi}_t^n}(dx dy) \left( \frac{1}{|\tilde{\Pi}_t^n| - 1} \langle g_{\tilde{\Pi}_t^n}, f \rangle + \frac{1}{|\tilde{\Pi}_t^n| - 1} f(x+y) - \frac{1}{|\tilde{\Pi}_t^n| - 1} f(x) - \frac{1}{|\tilde{\Pi}_t^n| - 1} f(y) \right)^2 \end{aligned}$$

where the sampling measure  $\nu_\Pi$  is defined as in (7.58), and the expression of  $II'_t$  is obtained by an argument analogous to the one for obtaining (7.59). Straightforward estimates yield that

$$|I'_t| \leq \frac{c\|f'\|_\infty^2}{n} \langle \tilde{g}_t^n, x^2 \rangle$$

and

$$|II'_t| \leq 16 \times 1_{|\tilde{\Pi}_t^n|>1} \|f\|_\infty \frac{|\tilde{\Pi}_t^n|}{n(|\tilde{\Pi}_t^n| - 1)} \leq 32 \frac{\|f\|_\infty}{n}$$

which is the desired result.  $\square$

**7.2. Tightness result.** The aim of this section is to show the following tightness result. This will be the key ingredient in the proof of our convergence results. In the following, we use the definition of weak ( $w$ ) and vague ( $v$ ) convergence as defined in Section 5.1.

**Proposition 7.3.** *Assume that conditions (7.52) and (7.53) hold. For every  $T > 0$ , the sequence processes  $\{\tilde{g}^n\}_{n \geq 0}$  is tight in  $D([0, T], (M_F(\mathbb{R}^+), w))$  and any converging subsequence belongs to  $\mathcal{C}([0, T], (M_F(\mathbb{R}^+), w))$ , the space of continuous functions from  $[0, T]$  to  $(M_F(\mathbb{R}^+), w)$ . Further*

- (i) *Any accumulation point  $g^\infty$  is a weak solution of the Smoluchowski (1.3) with inverse population  $2/r$  in the sense that for every  $t \geq 0$  and every test function  $f$*

$$0 = \langle g_t^\infty, f \rangle - \langle g_0^\infty, f \rangle + \int_0^t \left\langle g_s^\infty, c \frac{x^2}{2} f' \right\rangle ds - \int_0^t \frac{1}{s + \frac{2}{r}} \int_{(\mathbb{R}^+)^2} g_s^\infty(dx) g_s^\infty(dy) (f(x+y) - f(x)) ds$$

- (ii) *For every  $t \in [0, T]$ ,  $\langle g_t^\infty, x \rangle < \infty$  and*

$$\langle \tilde{g}_t^n, x \rangle \rightarrow \langle g_t^\infty, x \rangle \quad \text{in probability.}$$

In order to prove tightness, we follow a standard line of thoughts (see e.g., [15, 28, 29]). The approach is summarized in the statement of Theorem 7.4 which is cited from Tran [29] (Theorem 1.1.8). This Theorem can be obtained by concatenating the so-called Roelly criterium [25] (which states that the tightness of  $D([0, T], (M_F(\mathbb{R}^+), v))$  boils down to proving that  $\{\langle g_n, f \rangle\}_{n \geq 0}$  for  $f \in \mathcal{C}_b^\infty(\mathbb{R}^+)$  is tight in  $D([0, T], \mathbb{R})$ ), and a criterium due to Roelly [25], allowing us to go from vague to weak convergence by checking that no mass is lost at  $\infty$ .

**Theorem 7.4.** *Let  $\{\tilde{g}^n\}$  be a sequence in  $D([0, T], (M_F(\mathbb{R}^+), w))$ . Then the three following conditions are sufficient for the tightness of  $\{\tilde{g}^n\}$  in  $D([0, T], (M_F(\mathbb{R}^+), w))$ .*

- (i) *For every  $f \in \mathcal{C}_b^\infty(\mathbb{R}^+)$ , the sequence  $\{\langle g^n, f \rangle\}_n$  is tight in  $D([0, T], \mathbb{R})$ .*  
(ii)  $\limsup_n \mathbb{E}(\sup_{[0, T]} \langle \tilde{g}_t^n, x^2 \rangle) < \infty$ .  
(iii) *Any accumulation point  $g^\infty$  of  $\{\tilde{g}^n\}$  (in  $D([0, T], (M_F(\mathbb{R}^+), v))$ ) belongs to  $\mathcal{C}([0, T], (M_F(\mathbb{R}^+), w))$ .*

*Proof of Proposition 7.3. Step 1.* We first show that for every  $f \in \mathcal{C}_b^\infty(\mathbb{R}^+)$ , the sequence of processes  $\{X^{n, f}\}_n$  (as defined in (7.56)) is tight. In order to do so, we use the classical Aldous and Rebolledo criterium [2, 18]. We first note that for every  $t \geq 0$

$$|X_t^{n, f}| \leq \|f\|_\infty$$

so that the first requirement of Aldous criterion (i.e., for every deterministic  $t$ ,  $\{X_t^{n, f}\}_n$  is tight) is satisfied. Next, let  $\gamma > 0$  be an arbitrary small number and let us consider two stopping times  $(\tau, \sigma)$  such that

$$0 \leq \tau \leq \sigma \leq \tau + \gamma \leq T.$$

First, we decompose the semi-martingale  $X^{n,f}$  into its martingale part and its drift part, namely,

(7.59)

$$X_t^{n,f} = M_t^{n,f} + B_t^{n,f}, \text{ where } B_t^{n,f} := \int_0^t G^n h^f(\tilde{\Pi}_s^n) ds \text{ and } M_t^{n,f} := X_t^{n,f} - \int_0^t G^n h^f(\tilde{\Pi}_s^n) ds.$$

It remains to show that the quantities

$$\mathbb{E}(|B_\sigma^{n,f} - B_\tau^{n,f}|) \text{ and } \mathbb{E}(|M_\sigma^{n,f} - M_\tau^{n,f}|)$$

are bounded from above by a function of  $\gamma$  (uniformly in the choice of the two stopping times  $\tau$  and  $\sigma$  and  $n$ ) going to 0 as  $\gamma$  goes to 0. (This is the second part of Aldous and Robolledo criterium). In order to prove this result, we now make use of some of the technical results established earlier.

First, from (7.54), we note that there exists a constant  $\bar{K}$  such that for every  $\Pi \in E_n$  and  $f \in C_b^\infty(\mathbb{R}^+)$

$$|G^n h^f(\Pi)| \leq \bar{K} \left( \|f'\|_\infty \langle g_\Pi, x^2 \rangle + \frac{|\Pi|}{n} \|f\|_\infty \right).$$

This implies that

$$\begin{aligned} \mathbb{E}(|B_\sigma^{n,f} - B_\tau^{n,f}|) &\leq \mathbb{E} \left( \int_\tau^\sigma |G^n h^f(\tilde{\Pi}_s^n)| ds \right) \\ &\leq \bar{K} \mathbb{E} \left( \int_\tau^\sigma \left( \|f'\|_\infty \langle g_{\tilde{\Pi}_s^n}, x^2 \rangle + \frac{s_0^n}{n} \|f\|_\infty \right) ds \right) \\ &\leq \bar{K} \gamma \left( \|f'\|_\infty \mathbb{E} \left( \sup_{[0,T]} \langle g_{\tilde{\Pi}_s^n}, x^2 \rangle \right) + \frac{s_0^n}{n} \|f\|_\infty \right). \end{aligned}$$

Further,

$$\begin{aligned} \mathbb{E}(|M_\sigma^{n,f} - M_\tau^{n,f}|)^2 &\leq \mathbb{E}(|M_\sigma^{n,f} - M_\tau^{n,f}|^2) \\ &= \mathbb{E}(\langle X^{n,f} \rangle_\sigma - \langle X^{n,f} \rangle_\tau) \\ &\leq \frac{C}{n} \mathbb{E} \left( \int_\tau^\sigma \left( \|f'\|_\infty^2 \langle g_{\tilde{\Pi}_s^n}, x^2 \rangle + \|f\|_\infty^2 \right) ds \right) \\ &\leq \frac{C}{n} \gamma \left( \|f'\|_\infty^2 \mathbb{E} \left( \sup_{[0,T]} \langle g_{\tilde{\Pi}_s^n}, x^2 \rangle \right) + \|f\|_\infty^2 \right) \end{aligned}$$

where the second inequality follows from Lemma 7.2. Combining the two previous inequalities with (7.52) and (7.53) shows the tightness of  $\{X^{n,f}\}_{n \geq 0}$ .

**Step 2.** Let  $g^\infty$  be an accumulation point of the sequence  $\{\tilde{g}^n\}$  in  $D([0, T], (M_F(\mathbb{R}^+), \nu))$ . Since  $\tilde{g}_t^n = \frac{1}{\tilde{s}_t^n} \sum_{i=1}^{\tilde{s}_t^n} f(\tilde{\Pi}_t^n)$  and a transition can only affect two coordinates of  $\tilde{\Pi}^n$  at a time, it is not hard to show that

$$\sup_{t \in [0, T]} \sup_{f \in L^\infty([0, T]), \|f\|_\infty \leq 1} |\langle \tilde{g}_t^n, f \rangle - \langle \tilde{g}_{t-}^n, f \rangle| \leq \frac{4}{\tilde{s}_T^n}$$

where  $L^\infty([0, T])$  is the set of bounded functions from  $[0, T]$  to  $\mathbb{R}$ . Since  $\tilde{s}_T^n$  goes to  $\infty$  (in probability) as  $n \rightarrow \infty$ , this implies that  $g^\infty$  belongs to  $\mathcal{C}([0, T], (M_F([0, T]), w))$ . The tightness of  $\{\tilde{g}^n\}$  in

$D([0, T], (M_F(\mathbb{R}^+), w))$  then follows by a direct application of Theorem 7.4 (using the second moment assumption  $\limsup_n \mathbb{E} \left( \sup_{[0, T]} \langle \tilde{g}^n, x^2 \rangle \right) < \infty$ ).

**Step 3.** Next, let  $f$  be an arbitrary test function in  $\mathcal{C}_b^\infty(\mathbb{R}^+)$ . For every  $m \in D([0, T], (M_F(\mathbb{R}^+), w))$ , define

$$\varphi_{f,t}(m) = \langle m_t, f \rangle - \langle m_0, f \rangle + \int_0^t \left\langle m_s, c \frac{x^2}{2} f' \right\rangle ds - \int_0^t \frac{1}{s + \frac{2}{r}} \int_{(\mathbb{R}^+)^2} m_s(dx) m_s(dy) (f(x+y) - f(x)) ds.$$

In this step, we show that  $\varphi_{f,t}(g^\infty) = 0$ , for every  $t \in [0, T]$  and any choice of test function  $f$  in  $\mathcal{C}_b^\infty(\mathbb{R}^+)$ . We first observe

$$\mathbb{E}(|\varphi_{f,t}(\tilde{g}^n)|) \leq \mathbb{E}(|M_t^{n,f} - M_0^{n,f}|) + \mathbb{E} \left( \int_0^t |(G^n - \bar{G})h^f(\tilde{\Pi}_s^n)| ds \right)$$

where  $M^{n,f}$  is the martingale defined in (7.59) and  $\bar{G}$  is the generator approximation defined in Lemma 7.1. When we let  $n \rightarrow \infty$ , the second term vanishes by Lemma 7.1. For the first term, we have

$$\begin{aligned} \left( \mathbb{E}|M_t^{n,f} - M_0^{n,f}| \right)^2 &\leq \mathbb{E} \left( M_t^{n,f} - M_0^{n,f} \right)^2 \\ &= \mathbb{E} \left( \langle X^{n,f} \rangle_t - \langle X^{n,f} \rangle_0 \right) \end{aligned}$$

and the RHS can be handled by Lemma 7.2 and our second moment assumption (7.53). This implies

$$\lim_{n \rightarrow \infty} \mathbb{E}(|\varphi_{f,t}(\tilde{g}^n)|) = 0.$$

On the other hand, since any accumulation point  $g^\infty$  must be in  $\mathcal{C}([0, T], (M_F(\mathbb{R}^+), w))$  and since  $f$  and its derivative  $f'$  are continuous, we must have

$$\varphi_{f,t}(\tilde{g}^n) \implies \varphi_{f,t}(g^\infty).$$

(Here we use the fact that  $f$  is a test function so that  $f$  and  $\psi f'$  remain bounded, and further, if  $\{(\tilde{m}_t^n; t \geq 0)\}$  converges to a continuous  $(m_t^\infty; t \geq 0)$ , then for every continuous and bounded in  $u$ , the process  $(\langle m_t^n, u \rangle; t \geq 0)$  converges to  $(\langle m_t, u \rangle; t \geq 0)$  in the uniform norm on every finite interval) we get that  $\mathbb{E}(|\varphi_{f,t}(g^\infty)|) = 0$  by a direct application of the Dominated Convergence Theorem.

**Step 4.** In the previous step, we showed that  $\varphi_{f,t}(g^\infty) = 0$  for any test function in  $\mathcal{C}_b^\infty(\mathbb{R}^+)$ . By a standard density argument, the result also holds for any test function, thus showing that  $g^\infty$  is a weak solution of the Smoluchowski equation (1.3) with inverse population  $\delta$ .

**Step 5.** Let us now show the convergence of the mean. The argument is quite standard and goes by approximating the function  $x$  by a bounded and continuous function to make use of the weak convergence. Define

$$f^{(k)}(x) = x \text{ if } x \leq k, \quad f^{(k)}(x) = k \text{ otherwise,}$$

and note that

$$(7.60) \quad \langle \tilde{g}_t^n, x \rangle = \langle \tilde{g}_t^n, f^{(k)}(x) \rangle + \langle \tilde{g}_t^n, x - f^{(k)}(x) \rangle.$$

We now let  $n$  and then  $k$  go to 0 sequentially. By using the Cauchy-Schwarz and Markov inequalities, for any  $k \geq 1$ , we get

$$(7.61) \quad \begin{aligned} \langle \tilde{g}_t^n, x - f^{(k)}(x) \rangle^2 &\leq \langle \tilde{g}_t^n, 1_{x \geq k} \rangle \langle \tilde{g}_t^n, ((x - k)^+)^2 \rangle \\ &\leq \frac{1}{k^2} \langle \tilde{g}_t^n, x^2 \rangle^{3/2} \end{aligned}$$

and using (7.53), the RHS of the inequality goes to 0 (in probability) as  $n$  and then  $k$ , go sequentially to 0. On the other hand, since  $\{\tilde{g}_t^n\}_n$  converges to  $g_t^\infty$  as  $n \rightarrow \infty$  in the weak topology, the first term on the RHS of (7.60) converges to  $\langle g_t^\infty, f^{(k)} \rangle$ . Finally, as  $k \rightarrow \infty$ ,  $\langle g_t^\infty, f^{(k)} \rangle$  goes to  $\langle g_t^\infty, x \rangle$  by the monotone convergence theorem. This completes the proof for the convergence of the mean.  $\square$

**7.3. Proof of Theorem 5.1.** We start by showing the convergence of  $\{(\tilde{g}_t^n; t \in [0, T])\}_n$ . By Proposition 7.3 (and the unicity of (1.3) with initial condition  $\nu$  and  $\delta = 2/r$ ), it is enough to show that  $\limsup_n \mathbb{E} \left( \sup_{[0, T]} \langle \tilde{g}_t^n, x^2 \rangle \right) < \infty$ .

For every  $t \leq T$ , denote by  $\tilde{C}_{t, T}^n(i)$  be the indices of the blocks at time  $t/n$  partitioning the block  $i$  at time  $T/n$ . (In particular,  $\tilde{C}_{0, T}^n(i) = \tilde{B}_T^n(i)$ .) We have

$$(7.62) \quad \begin{aligned} \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \left( \sum_{k \in \tilde{B}_T^n(i)} \tilde{\Pi}_0^n(k) \right)^2 &= \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \left( \sum_{j \in \tilde{C}_{t, T}^n(i)} \sum_{k \in \tilde{B}_t^n(j)} \tilde{\Pi}_0^n(k) \right)^2 \\ &\geq \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \sum_{j \in \tilde{C}_{t, T}^n(i)} \left( \sum_{k \in \tilde{B}_t^n(j)} \tilde{\Pi}_0^n(k) \right)^2 \\ &= \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \left( \sum_{k \in \tilde{B}_t^n(i)} \tilde{\Pi}_0^n(k) \right)^2 \\ &\geq \frac{1}{\tilde{s}_t^n} \sum_{i=1}^{\tilde{s}_t^n} \left( \sum_{k \in \tilde{B}_t^n(i)} \tilde{\Pi}_0^n(k) \right)^2. \end{aligned}$$

Thus, for every  $t \leq T$ , this yields

$$\langle \tilde{g}_t^n, x^2 \rangle \leq \frac{1}{\tilde{s}_t^n} \sum_{i=1}^{\tilde{s}_t^n} \left( \sum_{j \in \tilde{B}_t^n(i)} \tilde{\Pi}_0^n(j) \right)^2 \leq \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \left( \sum_{j \in \tilde{B}_T^n(i)} \tilde{\Pi}_0^n(j) \right)^2$$

where the first inequality is obtained by ignoring the coalescence events between gene lineages (in particular, the first inequality becomes an equality when  $c = 0$ ). Let  $\{\mathcal{G}_t; t \geq 0\}$  be the natural

filtration generated by the species coalescent. From the previous arguments, we get that

$$\begin{aligned}
\mathbb{E} \left( \sup_{[0,T]} \langle \tilde{g}_t^n, x^2 \rangle \right) &\leq \mathbb{E} \left( \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \left( \sum_{j \in \tilde{B}_T^n(i)} \tilde{\Pi}_0^n(j) \right)^2 \right) \\
&\leq \mathbb{E} \left( \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i) \sum_{j \in \tilde{B}_T^n(i)} (\tilde{\Pi}_0^n(i))^2 \right) \\
&= \mathbb{E} \left( \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i) \sum_{j \in \tilde{B}_T^n(i)} \mathbb{E}((X_i^n)^2 \mid \mathcal{G}_t) \right) \\
&= \mathbb{E}((X_1^n)^2) \mathbb{E} \left( \frac{1}{\tilde{s}_T^n} \sum_{i=1}^{\tilde{s}_T^n} \tilde{V}_T^n(i)^2 \right),
\end{aligned}$$

and the RHS remains bounded by assumptions and Corollary 6.2.

It remains to show the joint convergence statement (5.47). The convergence of  $\{\tilde{s}^n\}$  was already stated in Lemma 6.1. The joint convergence follows from the fact that the limit of the marginals are both deterministic.

## 8. COMING DOWN FROM INFINITY IN THE NESTED KINGMAN COALESCENT

In the following  $(s_t, \Pi_t)$  will denote an  $\infty$ -population nested Kingman coalescent, and  $g_t$  will denote the associated empirical measure.

**Proposition 8.1.** *For every  $t > 0$ , we have  $\liminf_n R^n \circ g_t \geq \delta_{\frac{2}{ct}}$  in the sense that for every continuous, bounded and non-decreasing function  $f$*

$$\mathbb{E} \left( \liminf_n \langle R^n \circ g_t, f \rangle \right) \geq f(2/ct).$$

*Proof.* First, note that  $g_t$  stochastically dominates the case where each species lineage carries a single gene lineage at time 0. Hence, we can assume w.l.o.g. this particular initial condition. Secondly, since the species constraint forbids coalescence events between gene lineages belonging to different species,

$(g_t; t \geq 0)$  dominates  $(\delta_{K_t}; t \geq 0)$ , where  $K$  is the block counting process of a Kingman coalescent with rate  $c$ . (In other words, in  $K$ , we allow gene lineages to coalesce even if they belong to different species.) Finally, since  $\frac{ct}{2} K_t \rightarrow 1$  a.s., the result follows.  $\square$

Our next aim is to show the following result.

**Proposition 8.2.** *For every  $0 < \tau < T$ ,*

$$\limsup_n \mathbb{E} \left( \sup_{[\tau, T]} \langle R^n \circ g_t, x^2 \rangle \right) < \infty$$

*Proof.* We have

$$(8.63) \quad \sup_{t \in [\tau, T]} \langle R^n \circ g_t, x^2 \rangle = \sup_{t \in [\frac{\tau}{2}, T - \frac{\tau}{2}]} \langle R^n \circ \hat{g}_t^n, x^2 \rangle \quad \text{in law,}$$

where  $\hat{g}^n$  is the empirical measure associated with the nested coalescent with the initial number of species being equal to  $\hat{s}_0^n = s_{\tau/2n}$  and genetic composition vector  $\Pi_{\tau/2n}$ . Further, using the large deviation estimates of Lemma 6.1, we get

$$(8.64) \quad \hat{s}_0^n/n = s_{\tau/2n}/n \rightarrow \frac{4}{\tau} \in (0, \infty) \quad \text{in } L^p, \quad \forall p > 1.$$

The RHS of (8.63) is always bounded from above by the same quantity if we replace  $\hat{g}_t^n$  by the empirical measure associated with the nested coalescent starting with  $\hat{s}_0^n$  species and infinitely many gene lineages in each species. In turn, the latter model is bounded by the model starting from the infinite initial condition, but where gene lineages can only coalesce if they belong to the same species at time 0, i.e., even if species 1 and 2 coalesce, their respective gene lineages are forbidden to merge afterwards. The empirical measure associated with the process is identical in law to

$$m_t^n := \frac{1}{\hat{s}_t^n} \sum_{i=1}^{\hat{s}_t^n} \delta_{\sum_{j \in \hat{B}_t^n(i)} \xi_j(t)},$$

where  $\hat{B}_t^n(i)$  (with respect to the species coalescent  $\hat{s}^n$ ) and  $\xi_j^t$ s are defined analogously to Lemma 6.3. This yields

$$\forall t \in [\tau/2, T - \tau/2], \quad \langle R^n \circ \hat{g}_t^n, x^2 \rangle \leq \langle R^n \circ m_t^n, x^2 \rangle =_{\mathcal{L}} \beta_t^n$$

(where the domination is meant in the stochastic sense and  $\beta_t^n$  is defined as in Lemma 6.3) and thus

$$\mathbb{E} \left( \sup_{t \in [\tau, T]} \langle R^n \circ g_t, x^2 \rangle \right) \leq \mathbb{E} \left( \sup_{t \in [\tau/2, T - \tau/2]} \beta_t^n \right)$$

Proposition 8.2 then follows by a direct application of Lemma 6.3 (and (8.64)).  $\square$

*Proof of Theorem 5.5. Step 1.* Let us fix  $\tau > 0$ . Define  $\hat{g}^{n,(\tau)} = \theta_\tau \circ R^n \circ g_t$  and let  $\hat{s}^{n,(\tau)} = \theta_\tau \circ s_t^n$ . Proposition 8.2 implies that

$$\limsup_n \mathbb{E} \left( \sup_{t \in [0, T]} \langle \hat{g}_t^{n,(\tau)}, x^2 \rangle \right) = \mathbb{E} \left( \sup_{t \in [\tau, T + \tau]} \langle R^n \circ g_t, x^2 \rangle \right) < \infty$$

Further, by Lemma 6.1,

$$\underbrace{\frac{1}{n}}_{\text{time scaling}} \times \underbrace{\hat{s}_0^{n,(\tau)} (= s_{\tau/n})}_{\text{number of blocks in the species coalescent at time 0}} \rightarrow r = \frac{2}{\tau} \quad \text{in } L^p \text{ for every } p > 1.$$

By Proposition 7.3, it follows that the sequence  $\{\hat{g}^{n,(\tau)}\}_n$  is tight and that any sub-sequential limit  $\hat{g}^{\infty,(\tau)}$  is a weak solution of the Smoluchowski equation with inverse population size  $\tau/2$ . The



continuous mapping theorem implies that  $\{R^n \circ g_t = \theta_{-\tau} \circ \hat{g}^{n,(\tau)}; t \geq \tau\}$  converges to the limit  $(\theta_{-\tau} \circ g^{\infty,(\tau)}; t \geq \tau)$  where the latter process has a Laplace process satisfying the equation

$$(8.65) \quad \forall t \geq \tau, \langle \nu_t, f \rangle - \langle \nu_\tau, f \rangle + \int_\tau^t \left\langle \nu_s, c \frac{x^2}{2} f' \right\rangle ds - \int_\tau^t \frac{1}{s} \int_{(\mathbb{R}^+)^2} \nu_s(dx) \nu_s(dy) (f(x+y) - f(x)) ds = 0$$

Note that the coefficients of the IPDE do not depend on the value of  $\tau$ .

**Step 2.** Let us now take a sequence of positive numbers  $\{\tau_m\}_m$  going to 0. For every  $m$ , there exists a subsequence of  $\{(R^n \circ g_t; t \geq \tau_m)\}_n$  converging to  $\mu^{(0),m}$  satisfying (8.65). By a standard diagonalization argument, this ensures the existence of a subsequence of  $\{R^n \circ g\}$  converging to a process  $\mu^{(0)}$  defined on  $(0, \infty)$  (in comparison with step 1 where the process was defined on  $[\tau, \infty)$ ) and satisfying

$$(8.66) \quad \forall t, \tau > 0, \langle \nu_t, f \rangle - \langle \nu_\tau, f \rangle + \int_\tau^t \left\langle \nu_s, c \frac{x^2}{2} f' \right\rangle ds - \int_\tau^t \frac{1}{s} \int_{(\mathbb{R}^+)^2} \nu_s(dx) \nu_s(dy) (f(x+y) - f(x)) ds = 0,$$

i.e.,  $\mu^{(0)}$  is a weak solution of the  $\infty$ -population Smoluchowski equation. In order to prove Theorem 5.5, it remains to show that  $\mu^{(0)}$  is the only proper solution. This follows directly from the stochastic domination of Proposition 8.1. Finally, the joint convergence with  $(\tilde{g}^n, \tilde{s})$  follows from the fact that both marginals are deterministic in the limit.  $\square$

*Proof of Theorem 5.7.* Let  $\rho_t$  be the number of gene lineages at time  $t$ . We need to show that

$$\frac{1}{n^2} \rho_{t/n} \implies \frac{2}{t^2} \int_0^\infty x \mu_t^{(0)}(x) dx$$

where  $\mu^{(0)}$  is the proper solution of the Smoluchowski equation. By applying Proposition 7.3 and Theorem 5.5,  $(\frac{1}{n} s_{t/n}, \langle R^n \circ g_t, x \rangle)$  converges to  $(\frac{2}{t}; \frac{1}{t} \int_0^\infty x \mu_t^{(0)}(x) dx)$ . The result follows from the observation that

$$\frac{1}{n^2} \rho_{t/n} = \frac{1}{n} s_{t/n} \langle R^n \circ g_t, x \rangle.$$

$\square$

## APPENDIX A

Here, we complete the proof of Theorem 2.4 (ii) by showing that  $\mu_T$  defined as  $F(\mathbf{T}, (W_i); 1 \leq i \leq N_T(\mathbf{T}))$  indeed is a solution to (1.3).

Let  $f$  be a test-function as defined before Definition 1.8, i.e.,  $f \in \mathcal{C}^1(\mathbb{R}^+)$  such that  $f$  and  $f'\psi$  are bounded. Hereafter, we continue to denote by  $\mathbb{P}_T$  the joint law of the pure-birth tree  $\mathbf{T}$  started with one particle at time 0, birth rate  $a(T-t)$ , stopped at time  $T$ , and of the iid random variables

$(W_i; 1 \leq i \leq N_T(\mathbf{T}))$  with law  $\nu$ . We will abbreviate  $F(\mathbf{T}, (W_i); 1 \leq i \leq N_T(\mathbf{T}))$  into  $F(\mathbf{T})$ . In particular, denoting  $\mu_T$  as the law of  $F(\mathbf{T}, (W_i); 1 \leq i \leq N_T(\mathbf{T}))$  under  $\mathbb{P}_T$ , we have

$$\mu_T(f) := \int_{\mathbb{R}^+} f(x) \mu_T(dx) = \mathbb{E}_T(f \circ F(\mathbf{T})),$$

so that

$$\mu_{T+\varepsilon}(f) = \mathbb{E}_{T+\varepsilon}(f \circ F(\mathbf{T}), N_\varepsilon = 1) + a(T)\varepsilon \mathbb{E}_T^{\otimes 2}(f \circ F(\mathbf{T} + \mathbf{T}')) + o(\varepsilon),$$

where  $\mathbf{T}'$  is an independent copy of  $\mathbf{T}$  and  $\mathbf{T} + \mathbf{T}'$  denotes the tree splitting at time 0 into the two subtrees  $\mathbf{T}$  and  $\mathbf{T}'$ . First note that by construction,  $F(\mathfrak{t} + \mathfrak{t}') = F(\mathfrak{t}) + F(\mathfrak{t}')$ . Second, if we denote by  $\mathfrak{t} + \varepsilon$  the tree obtained from  $\mathfrak{t}$  by merely adding a length  $\varepsilon$  to its root edge, then by the Markov property of the entrance measure of the CSBP at 0,

$$\begin{aligned} F(\mathfrak{t} + \varepsilon) &= M^{\mathfrak{t}+\varepsilon} \left( 1 - \exp \left( - \sum_{i=1}^{N_T(\mathfrak{t})} w_i Z_{T+\varepsilon}^i \right) \right) \\ &= \int_{(0, \infty)} N(Z_\varepsilon \in dx) Q_x^\mathfrak{t} \left( 1 - \exp \left( - \sum_{i=1}^{N_T(\mathfrak{t})} w_i Z_T^i \right) \right) \\ &= \int_{(0, \infty)} N(Z_\varepsilon \in dx) \left( 1 - \exp \left( -x M^\mathfrak{t} \left( 1 - \exp \left( - \sum_{i=1}^{N_T(\mathfrak{t})} w_i Z_T^i \right) \right) \right) \right) \\ &= N(1 - \exp(-Z_\varepsilon F(\mathfrak{t}))). \end{aligned}$$

Now as specified at the end of Subsection 2.3, for each fixed  $\lambda$ ,  $N(1 - \exp(-\lambda Z_t))$  is solution to  $\dot{x} = -\psi(x)$  with initial condition  $x(0) = \lambda$ . As a consequence,

$$\lim_{\varepsilon \downarrow 0} \varepsilon^{-1} (F(\mathfrak{t} + \varepsilon) - F(\mathfrak{t})) = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1} (N(1 - \exp(-Z_\varepsilon F(\mathfrak{t}))) - F(\mathfrak{t})) = -\psi(F(\mathfrak{t})).$$

Combining the last two results, we obtain

$$\begin{aligned} \mu_{T+\varepsilon}(f) &= \mathbb{E}_{T+\varepsilon}(f \circ F(\mathbf{T}), N_\varepsilon = 1) + a(T)\varepsilon \mathbb{E}_T^{\otimes 2}(f \circ F(\mathbf{T} + \mathbf{T}')) + o(\varepsilon) \\ &= (1 - a(T)\varepsilon) \mathbb{E}_T(f \circ F(\mathbf{T} + \varepsilon)) + a(T)\varepsilon \mathbb{E}_T^{\otimes 2}(f \circ (F(\mathbf{T}) + F(\mathbf{T}'))) + o(\varepsilon) \\ &= \mu_T(f) + (1 - a(T)\varepsilon) \mathbb{E}_T(f \circ F(\mathbf{T} + \varepsilon) - f \circ F(\mathbf{T})) + a(T)\varepsilon (\mu_T^{\star 2}(f) - \mu_T(f)) + o(\varepsilon). \end{aligned}$$

Next, since  $\psi f'$  is bounded, by dominated convergence, we get

$$\lim_{\varepsilon \downarrow 0} \varepsilon^{-1} \mathbb{E}_T(f \circ F(\mathbf{T} + \varepsilon) - f \circ F(\mathbf{T})) = -\mathbb{E}_T(\psi(F(\mathbf{T})) f'(F(\mathbf{T}))) = -\mu_T(\psi f').$$

As a consequence,

$$\lim_{\varepsilon \downarrow 0} \varepsilon^{-1} (\mu_{T+\varepsilon}(f) - \mu_T(f)) = -\mu_T(\psi f') + a(T) (\mu_T^{\star 2}(f) - \mu_T(f)).$$

So  $t \mapsto \mu_t(f)$  is right-differentiable with continuous right-derivative equal to

$$\partial_t \mu_t(f) = -\mu_t(\psi f') + a(t) (\mu_t^{\star 2}(f) - \mu_t(f)) \quad t \geq 0.$$

Also note that

$$F(\varnothing + \varepsilon) = M^{\varnothing+\varepsilon} (1 - \exp(-w_1 Z_\varepsilon)) = N(1 - \exp(-w_1 Z_\varepsilon)),$$

so that  $F(\varnothing) = w_1$  and  $\mu_0(f) = \mathbb{E}_0(f(F(\mathbf{T}))) = \mathbb{E}_0(f(W)) = \nu(f)$ . This shows that  $\mu_0 = \nu$  so that  $(\mu_t(f); t \geq 0)$  satisfies (1.5).

**Acknowledgments.** We would like to thank the referee for her/his thorough review. We highly appreciated her/his comments and suggestions.

#### REFERENCES

- [1] E. Abulwafa, M. Abdou, and A. Mahmoud. The solution of nonlinear coagulation problem with mass loss. *Chaos, Solitons & Fractals*, 29(2):313–330, 2006.
- [2] D. Aldous. Stopping times and tightness. *The Annals of Probability*, 6(2):335–340, 1978.
- [3] D. J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 1999.
- [4] J. Berestycki, N. Berestycki, and V. Limic. The  $\Lambda$ -coalescent speed of coming down from infinity. *The Annals of Probability*, 38(1):207–233, 2010.
- [5] J. Berestycki, N. Berestycki, and J. Schweinsberg. Beta-coalescents and continuous stable random trees. *The Annals of Probability*, pages 1835–1887, 2007.
- [6] N. Berestycki. Recent progress in coalescent theory. *Ensaïos Matemáticos*, 16(1):1–193, 2009.
- [7] J. Bertoin. *Random fragmentation and coagulation processes*, volume 102. Cambridge University Press, 2006.
- [8] J. Bertoin and J.-F. Le Gall. Stochastic flows associated to coalescent processes. III. Limit theorems. *Illinois Journal of Mathematics*, 50(1-4):147–181, 2006.
- [9] A. Blancas, J.-J. Duchamps, A. Lambert, and A. Siri-Jégousse. Trees within trees: Simple nested coalescents. *Elec. J. Probab.*, 23(94):1–27, 2018.
- [10] A. Blancas, T. Rogers, J. Schweinsberg, and A. Siri-Jégousse. The nested kingman coalescent: Speed of coming down from infinity. *Ann. Appl. Probab.*, 29(3):1808–1836, 2019.
- [11] M. E. Caballero, A. Lambert, and G. Uribe Bravo. Proof (s) of the Lamperti representation of continuous-state branching processes. *Probability Surveys*, 6:62–89, 2009.
- [12] B. Derrida and M. Retaux. The depinning transition in presence of disorder: a toy model. *Journal of Statistical Physics*, 156(2):268–290, 2014.
- [13] T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281):vi+147, 2002.
- [14] W. Feller. Two singular diffusion problems. *Annals of mathematics*, pages 173–182, 1951.
- [15] N. Fournier and S. Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Annals of applied probability*, pages 1880–1919, 2004.
- [16] D. Grey. Asymptotic behaviour of continuous time, continuous state-space branching processes. *Journal of Applied Probability*, 11(4):669–677, 1974.
- [17] Y. Hu, B. Mallein, and M. Pain. An exactly solvable continuous-time Derrida–Retaux model. *arXiv preprint arXiv:1811.08749*, 2018.

- [18] A. Joffe and M. Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Advances in Applied Probability*, 18(1):20–65, 1986.
- [19] J. F. C. Kingman. The coalescent. *Stochastic processes and their applications*, 13(3):235–248, 1982.
- [20] A. Lambert. Population dynamics and random genealogies. *Stoch. Models*, 24(suppl. 1):45–163, 2008.
- [21] A. Lambert and E. Schertzer. Recovering the Brownian coalescent point process from the Kingman coalescent by conditional sampling. *To appear in Bernoulli*, 2016.
- [22] J. Lamperti. Continuous state branching processes. *Bulletin of the American Mathematical Society*, 73(3):382–386, 1967.
- [23] J. R. Norris. Smoluchowski’s coagulation equation: Uniqueness, nonuniqueness and a hydrodynamic limit for the stochastic coalescent. *Annals of Applied Probability*, pages 78–109, 1999.
- [24] L. Popovic. Asymptotic genealogy of a critical branching process. *Annals of Applied Probability*, 14(4):2120–2148, 2004.
- [25] S. Roelly-Coppoletta. A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics: An International Journal of Probability and Stochastic Processes*, 17(1-2):43–65, 1986.
- [26] P. Singh and G. Rodgers. Coagulation processes with mass loss. *Journal of Physics A: Mathematical and General*, 29(2):437, 1996.
- [27] A.-S. Sznitman. Topics in propagation of chaos. In *Ecole d’été de probabilités de Saint-Flour XIX—1989*, pages 165–251. Springer, 1991.
- [28] V. C. Tran. Large population limit and time behaviour of a stochastic particle model describing an age-structured population. *ESAIM: Probability and Statistics*, 12:345–386, 2008.
- [29] V. C. Tran. Une ballade en forêts aléatoires. Technical report, 2014.
- [30] J. A. Wattis, D. G. McCartney, and T. Gudmundsson. Coagulation equations with mass loss. *Journal of engineering mathematics*, 49(2):113–131, 2004.
- [31] B. Wennberg. An example of non-uniqueness for solutions to the homogeneous Boltzmann equation. *Journal of statistical physics*, 95(1-2):469–477, 1999.