The Annals of Probability
2011, Vol. 39, No. 6, 2043–2078
DOI: 10.1214/10-AOP601

BRUNET–DERRIDA PARTICLE SYSTEMS, FREE BOUNDARY PROBLEMS AND WIENER–HOPF EQUATIONS

BY RICK DURRETT[1,2] AND DANIEL REMENIK[1,2]

Duke University and University of Toronto

We consider a branching-selection system in $\mathbb{R}$ with N particles which give birth independently at rate 1 and where after each birth the leftmost particle is erased, keeping the number of particles constant. We show that, as $N \to \infty$, the empirical measure process associated to the system converges in distribution to a deterministic measure-valued process whose densities solve a free boundary integro-differential equation. We also show that this equation has a unique traveling wave solution traveling at speed c or no such solution depending on whether $c \geq a$ or $c < a$, where a is the asymptotic speed of the branching random walk obtained by ignoring the removal of the leftmost particles in our process. The traveling wave solutions correspond to solutions of Wiener–Hopf equations.

1. Introduction and statement of the results. We will consider the following branching-selection particle system. At any time t we have N particles on the real line with positions $\eta_t^N(1) \geq \cdots \geq \eta_t^N(N)$. Each one of the N particles gives birth at rate 1 to a new particle whose position is chosen, relative to the parent particle, using a given probability distribution ρ on $\mathbb{R}$. Whenever a new particle is born, we reorder the $N+1$ particles and erase the leftmost one (so the number of particles is always kept equal to N). We will denote by $\mathcal{X}_N = \{(\eta(1), \dots, \eta(N)) \in \mathbb{R}^N : \eta(1) \geq \cdots \geq \eta(N)\}$ the state space of our process.

We learned of this process through the work of Durrett and Mayberry (2010), who considered the special case in which ρ corresponds to a uniform random variable on $[-1, 1]$. However, our process is a member of a family of processes that first arose in work of Brunet and Derrida (1997), who studied a discrete analog of the Fisher–Kolmogorov PDE:

$$\frac{\partial h}{\partial t} = \frac{\partial^2 h}{\partial x^2} + h - h^3.$$

In a simpler version of their process, model A in Brunet et al. (2007), the discrete time dynamics occur in the following way: at each time step each of the N particles

Received July 2009; revised April 2010.

[1]This work was completed while R. Durrett was a Professor of Mathematics and D. Remenik was a graduate student at Cornell University.

[2]Supported in part by NSF Grant DMS-07-04996 from the probability program.

MSC2010 subject classifications. 60J99, 60J80, 35R35, 35C07, 60F99.

Key words and phrases. Branching-selection system, branching random walk, free boundary equation, Wiener–Hopf equation, traveling wave solutions.

is replaced by a fixed number k of particles whose displacements from the parent particle are chosen independently, and then only the N rightmost particles are kept. They conjectured that the system moves to the right with a deterministic asymptotic speed v_N which increases as $N \to \infty$ to some explicit maximal speed v at a rate of order $(\log N)^{-2}$. This slow rate of convergence was recently proved by Bérard and Gouéré (2010) in the case $k = 2$ under some assumptions on the distribution used to choose the locations of the new particles (one being that new particles are always sent to the right of the parent particle).

Although we will say something about the behavior of the system for fixed N, our main interest in this paper is to study the behavior of the empirical distribution of the process as $N \to \infty$. Before proceeding with this, let us specify some assumptions. When a particle at x gives birth, the new particle is sent to a location $x + y$ with y being chosen from an absolutely continuous probability distribution $\rho(y)\,dy$. We will assume that ρ is symmetric and that $\int_{-\infty}^{\infty} |x|\rho(x)\,dx < \infty$. The initial condition for our process will always be specified as follows: each particle starts at a location chosen independently from a probability measure $f_0(x)\,dx$, where $f_0(x) = 0$ for $x < 0$ and $f_0(x)$ is strictly positive and continuous for $x > 0$.

1.1. *Convergence to the solution of a free boundary problem.* Let

$$\nu_t^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\eta_t^N(i)}$$

be the empirical measure associated to η_t^N. Observe that the initial empirical measure $\nu_0^N(dx)$ converges in distribution to $f_0(x)\,dx$. We will show that, as $N \to \infty$, this empirical measure process converges to a deterministic measure-valued process whose densities are the solution of a certain free boundary problem.

Alternatively one could think of the following (weaker) version of the problem. It is not hard to see that the probability measure $\mathbb{E}(\nu_t^N(\cdot))$ on $\mathbb{R}$ is absolutely continuous. Let $\overline{f}^N(t,x)$ be its density. We want to study its limit as $N \to \infty$. We expect this limit $f(t,x)$ to correspond to the densities of the limiting measure-valued process mentioned above. Now observe that if ξ_t^N is a version of our process in which we do not erase the leftmost particle after births (i.e., ξ_t^N is a branching random walk), then we would expect that the density of the corresponding expected empirical measure converges to the solution $\widehat{f}(t,x)$ of the following integro-differential equation:

$$\frac{\partial \widehat{f}}{\partial t}(t,x) = \int_{-\infty}^{\infty} \widehat{f}(t,y)\rho(x-y)\,dy. \tag{1.1}$$

This is indeed the case, as we will see in Proposition 2.1. Observe that the total mass of $\widehat{f}(t,\cdot)$ grows exponentially [in fact $\int_{-\infty}^{\infty} \widehat{f}(t,y)\,dy = e^t$, see (2.5)].

By adding the selection step to our process we ensure that the limiting density always has mass 1, but otherwise the branching mechanism is still governed by the convolution term appearing in (1.1). Thus we expect that if the limit

$f(t,x) = \lim_{N\to\infty} \overline{f}^N(t,x)$ exists, then it has to satisfy the following: there exists a continuous increasing function $\gamma : [0,\infty) \longrightarrow \mathbb{R}$ with $\gamma(0) = 0$ such that $(f(t,x), \gamma(t))$ is the unique solution to the following free boundary problem (FB):

$$\text{(FB1)} \qquad \frac{\partial f}{\partial t}(t,x) = \int_{-\infty}^{\infty} f(t,y)\rho(x-y)\,dy \qquad \forall x > \gamma(t),$$

$$\text{(FB2)} \qquad \int_{\gamma(t)}^{\infty} f(t,y)\,dy = 1,$$

$$\text{(FB3)} \qquad f(t,x) = 0 \qquad \forall x \le \gamma(t)$$

with initial condition $f(0,x) = f_0(x)$ for all $x \in \mathbb{R}$. $\gamma(t)$ is a moving boundary which keeps the mass of $f(t,\cdot)$ at 1, but the speed at which it moves is not known in advance and depends in turn on f.

It is not a priori obvious that (FB) has a solution, let alone that such a solution is unique. We will prove the existence and uniqueness using arguments closely related to the ones we will use to prove the existence of the limiting density.

We will denote by $\mathcal{P}$ the space of probability measures on $\mathbb{R}$, which we endow with the topology of weak convergence, and by $D([0,T],\mathcal{P})$ the space of càdlàg functions from $[0,T]$ to $\mathcal{P}$ endowed with the Skorohod topology.

THEOREM 1. *For any fixed $T > 0$ the sequence of $\mathcal{P}$-valued processes ν_t^N on $[0,T]$ converges in distribution in $D([0,T],\mathcal{P})$ to a deterministic $\nu_t \in D([0,T],\mathcal{P})$. ν_t is absolutely continuous with respect to the Lebesgue measure for every $t \in [0,T]$ and the corresponding densities $f(t,\cdot)$ are characterized by the following: there exists a continuous, strictly increasing function $\gamma : [0,\infty) \longrightarrow \mathbb{R}$ with $\gamma(0) = 0$ such that $(f(t,x), \gamma(t))$ is the unique solution of the free boundary problem* (FB). *In particular for $x > \gamma(t)$, $f(t,x)$ is strictly positive, jointly continuous in t and x and differentiable in t.*

Let us remark that there are at least two other examples in the literature of particle systems converging to the solution of a free boundary equation, but in both cases the limiting equation is of a different type. Landim, Olla and Volchan (1998) study a tracer particle moving in a varying environment corresponding to the simple symmetric exclusion process, while Gravner and Quastel (2000) study an internal diffusion limited aggregation model. In both cases an hydrodynamic limit is proved with the limiting equation being closely related to the famous Stefan problem, which involves free boundary problems for the heat equation where the moving boundary separates a solid and a liquid phase [see Meirmanov (1992) and references therein for more on this problem].

1.2. *Behavior of the finite system.* To study the finite system it will be useful to introduce the *shifted* process Δ_t^N, which we define as follows:

$$\Delta_t^N = (\Delta_t^N(1), \dots, \Delta_t^N(N)) \qquad \text{with } \Delta_t^N(j) = \eta_t^N(j) - \eta_t^N(N).$$

Observe that $\Delta_t^N(N)$ is always 0. It is clear that Δ_t^N is also a Markov process, and its transitions are the same as those of η_t^N except that after erasing the leftmost particle the N remaining particles are shifted to the left so that the new leftmost one lies at the origin.

We will denote by $\min \eta_t^N = \eta_t^N(N)$ and $\max \eta_t^N = \eta_t^N(1)$ the locations of the leftmost and rightmost particles in η_t^N.

THEOREM 2. *For every fixed $N > 0$ the following hold*:

(a) *There is an $a_N > 0$ such that*

$$\lim_{t\to\infty} \frac{\min \eta_t^N}{t} = \lim_{t\to\infty} \frac{\max \eta_t^N}{t} = a_N$$

almost surely and in L^1. Moreover, the sequence $(a_N)_{N>0}$ is nondecreasing.

(b) *The process Δ_t^N has a unique stationary distribution μ_N, which is absolutely continuous.*

(c) *For any* (*random or deterministic*) *initial condition ν_0 we have*

$$\|\mathbb{P}^{\nu_0}(\Delta_t^N \in \cdot) - \mu_N(\cdot)\|_{\mathrm{TV}} \xrightarrow[t\to\infty]{} 0.$$

From this point on we will assume that the displacement distribution ρ has exponential decay. To be precise, we assume that there is an $\alpha > 0$ such that

$$\rho(x) \le Ce^{-\alpha|x|} \tag{1.2}$$

for some $C > 0$. We will write

$$\Theta = \sup\Big\{\alpha > 0 : \sup_{x\in\mathbb{R}}\big[e^{\alpha|x|}\rho(x)\big] < \infty\Big\}.$$

That is, $\Theta \in (0, \infty]$ is the maximal exponential rate of decay of ρ in the sense that $\rho(x) \le Ce^{-\alpha x}$ for some $C > 0$ when $\alpha < \Theta$ but not when $\alpha > \Theta$. Θ may be ∞ (as in the cases where ρ has compact support or ρ corresponds to a normal distribution), while $\Theta > 0$ is ensured by (1.2).

Now let

$$\phi(\theta) = \int_{-\infty}^{\infty} e^{\theta x}\rho(x)\,dx \tag{1.3}$$

be the moment generating function of the displacement distribution ρ. Equation (1.2) and our definition of Θ imply that $\phi(\theta) < \infty$ for $\theta \in (-\Theta, \Theta)$. To avoid unnecesary technical complications we will make the following extra assumption, which in particular implies that $\phi(\theta) = \infty$ for $|\theta| > \Theta$:

$$\frac{\phi(\theta)}{\theta} \xrightarrow[\theta \to \Theta^-]{} \infty. \tag{1.4}$$

This assumption always holds when $\Theta = \infty$: choosing $0 < l_1 < l_2$ so that $\rho(x) \geq M$ for some $M > 0$ and all $x \in [l_1, l_2]$ we get

$$\frac{1}{\theta} \int_{-\infty}^{\infty} e^{\theta x} \rho(x)\, dx \geq \frac{M(l_2 - l_1)}{\theta} e^{\theta l_1} \xrightarrow[\theta \to \infty]{} \infty.$$

Our next result will relate the asymptotic propagation speed a_N of our process η_t^N with the asympotic speed of the rightmost particle in the branching random walk ξ_t^N.

THEOREM 3.

$$\lim_{N \to \infty} a_N = a,$$

where a is the asymptotic speed of the rightmost particle in ξ_t^N, that is, in a branching random walk where particles branch at rate 1 *and their offspring are displaced by an amount chosen according to ρ.*

REMARK. If ρ is uniform on $[-1, 1]$ this follows from (8) in Durrett and Mayberry (2010). However, the couplings on which the proof is based extend easily to our more general setting, so we do not give the details of the proof.

The speed a has an explicit expression [see Biggins (1977)]: by standard results of the theory of large deviations, if S_t is a continuous time random walk jumping at rate 1 and with jump distribution ρ, then the limit

$$\Lambda(x) = \lim_{t \to \infty} \frac{1}{t} \log \mathbb{P}(S_t > xt) \tag{1.5}$$

exists and equals $-(\sup_{\theta > 0}\{x\theta - \phi(\theta)\} + 1)$; a is given then by the formula $\Lambda(a) = -1$ [see Durrett and Mayberry (2010) for more on this].

1.3. *Traveling wave solutions.* A *traveling wave solution* of (FB) is a solution of the form $f(t, x) = w(x - ct)$ and $\gamma(t) = ct$ for some $c > 0$ [with initial condition $f_0(x) = w(x)$].

If w is a traveling wave solution, then from (FB1) we get

$$cw'(z) = -\int_0^{\infty} w(y)\rho(z - y)\, dy \qquad \forall z > 0,$$

so integrating from $z = x$ to ∞ we deduce that w must solve the following system of equations (TW):

$$w(x) = \frac{1}{c}\int_0^\infty w(y)R(x-y)\,dy \qquad \forall x > 0, \tag{TW1}$$

$$\int_0^\infty w(x)\,dx = 1, \tag{TW2}$$

$$w(x) = 0 \qquad \forall x \le 0, \tag{TW3}$$

$$w(x) \ge 0 \qquad \forall x > 0, \tag{TW4}$$

where

$$R(x) = \int_x^\infty \rho(y)\,dy$$

is the tail distribution of ρ. On the other hand, it is easy to check that if w satisfies (TW) and $f_0(x) = w(x)$, then $(w(x - ct), ct)$ is the solution of (FB).

Equation (TW1) is known as a Wiener–Hopf equation. Equations of this type have been studied since at least the 1920s (at the time in relation with the theory of radiative energy transfer), and have since been extensively studied and found relevance in diverse problems in mathematical physics and probability. In general, these equations can be solved using the Wiener–Hopf method, which was introduced in Wiener and Hopf (1931) [see Chapter 4 of Paley and Wiener (1987) and also Kreĭn (1962)]. But the solutions provided by this method are not necessarily positive, so they are not useful in our setting. Instead, we will rely on the results of Spitzer (1957), who studied these equations via probabilistic methods in the case where $R(x)/c$ is a probability kernel.

To do that we need to convert our equation to one where the kernel with respect to which we integrate is a probability kernel. To that end, we need to make the following observation. In Lemma 4.1 we will show that there is a $\lambda^* \in (0, \Theta)$ such that

$$\frac{\phi(\lambda^*)}{\lambda^*} = \min_{\lambda \in (0,\Theta)} \frac{\phi(\lambda)}{\lambda} = a, \tag{1.6}$$

where ϕ is the moment generating function of ρ defined in (1.3) and a is the asymptotic speed introduced in Theorem 3. On the other hand, the function $\lambda \mapsto \phi(\lambda)/\lambda$ is continuous and goes to ∞ as $\lambda \to 0$ [and moreover, as we will show in Lemma 4.1, it is decreasing on $(0, \lambda^*)$]. Thus for every $c \ge a$ there is a $\lambda \in (0, \lambda^*]$ such that $\phi(\lambda)/\lambda = c$.

Observe that the tail distribution R of ρ has the same decay as ρ [see (1.2)]. That is, for every $0 < \alpha < \Theta$ there is a $C > 0$ so that

$$R(x) \le Ce^{-\alpha x} \qquad \forall x \ge 0. \tag{1.7}$$

Fix $c \geq a$ and use the above observation to pick a $\lambda \in (0, \lambda^*]$ such that $\phi(\lambda)/\lambda = c$. Equation (1.7) implies that the function $x \mapsto e^{\lambda x} R(x)$ is in $L^1(\mathbb{R})$. Moreover, integration by parts yields

$$\int_{-\infty}^{\infty} e^{\lambda x} R(x)\, dx = \frac{\phi(\lambda)}{\lambda}.$$

Therefore

$$k(x) = \frac{\lambda}{\phi(\lambda)} e^{\lambda x} R(x) \tag{1.8}$$

is a probability kernel. On the other hand, if w is a solution of (TW) with $c = \phi(\lambda)/\lambda$ then it is easy to check that $u(x) = e^{\lambda x} w(x)$ satisfies

$$u(x) = \int_0^{\infty} u(y) k(x-y)\, dy \qquad \forall x \geq 0. \tag{1.9}$$

Thus the idea will be to recover solutions of (TW) from positive solutions of (1.9).

Positive solutions of (1.9) can be regarded as densities of stationary measures for the following Markov chain. Let ξ_n be a sequence of i.i.d. random variables with distribution given by k, let $X_0 = 0$ and define

$$X_{n+1} = (X_n + \xi_n)^+,$$

where $x^+ = \max\{0, x\}$. This chain appears, for example, in the study of ladder variables for a random walk [see Chapter XII of Feller (1971)] and in the study of the GI/G/1 queue [see Chapter 5 of Durrett (2004)]. If u satisfies (1.9), it is 0 on the negative half-line and it is nonnegative on the positive half-line, then the measure (supported on $[0, \infty)$) having u as its density is invariant for X_n. Assuming that $\mathbb{E}(|\xi_1|) < \infty$, X_n is recurrent, null-recurrent or transient according to whether $\mathbb{E}(\xi_1)$ is negative, zero or positive. As we will see in Section 4, this expectation is negative in our case if and only if $\lambda < \lambda^*$ and it is zero for $\lambda = \lambda^*$. In both cases the theory of recurrent Harris chains suggests (and Theorem 4.2 will prove) that there exists a unique (up to multiplicative constant) invariant measure for X_n, although this measure may not be finite in the null-recurrent case. The difference between the recurrent and null-recurrent cases explains the difference between the cases $c > a$ and $c = a$ in Theorem 4 below. The fact that the chain is transient when $\lambda > \lambda^*$ suggests that there are no positive solutions of (1.9) for these values of λ (again see Theorem 4.2 for a proof). This in turn hints at the possibility that there are no solutions of (TW) for $c < a$. Our proof of this fact will not rely in seeing (TW) as a Wiener–Hopf equation, but instead will use explicitly the fact that its solutions are traveling wave solutions of (FB).

THEOREM 4. *Assume that* (1.2) *and* (1.4) *hold.*

(a) *If $c \geq a$ the equation* (TW) *has a unique solution w. This solution is differentiable except at the origin.*

(b) *When* $c > a$, *and letting* $\lambda \in (0, \lambda^*)$ *be such that* $\phi(\lambda)/\lambda = c$, *the solution satisfies* $\int_0^\infty e^{\lambda x} w(x)\,dx < \infty$ [*which, in particular, implies that* $\int_x^\infty w(y)\,dy = o(e^{-\lambda x})$]. *Moreover, if* $\widetilde{\lambda} > \lambda$ *then* $\sup_x e^{\widetilde{\lambda} x} \int_x^\infty w(y)\,dy = \infty$.
(c) *When* $c = a$ *the solution satisfies* $\int_x^\infty w(y)\,dy = O(e^{-\lambda^* x})$ *together with* $\int_0^x e^{\lambda^* y} w(y)\,dy = O(x)$. *The last integral goes to* ∞ *as* $x \to \infty$.
(d) *If* $c < a$ *the equation* (TW) *has no solution.*

We remark that, when $c > a$, the solution w given by the theorem can be obtained by the following limiting procedure. Take $0 < \lambda < \lambda^*$ as in the above statement and let u_0 be the density of any nonnegative random variable whose distribution is absolutely continuous. Now let $w_0(x) = e^{-\lambda x} u_0(x)$ and then for $n \geq 1$ let

$$w_{n+1}(x) = \frac{1}{c} \int_0^\infty w_n(y) R(x - y)\,dy \qquad \text{for } x \geq 0.$$

Then the limit $w_\infty(x) = \lim_{n\to\infty} w_n(x)$ exists and defines an integrable continuous function. The solution w is then given by $w(x) = K w_\infty(x)$ with $K > 0$ chosen so that w integrates to 1. The fact that w has this representation follows from the results of Spitzer (see Theorem 4.2).

The rest of the paper is devoted to proofs, with one section devoted to each one of the proofs of Theorems 1, 2 and 4.

2. Proof of Theorem 1.

2.1. *Outline of the proof.* Most of the work in the proof of Theorem 1 will correspond to showing that for each fixed $t \geq 0$ the tail distribution of our process at that time, defined as

$$F^N(t, x) = \nu_t^N([x, \infty)),$$

converges (almost surely and in L^1) to a deterministic limit $F(t, x)$ corresponding to the tail distribution of a random variable and that, moreover, the limit $F(t, x)$ has a density $f(t, x)$ [i.e., $F(t, x) = \int_x^\infty f(t, y)\,dy$] which solves (FB).

To achieve this we will compare the process ν_t^N with two auxiliary measure-valued processes $\nu_t^{N,k}$ and ν_t^k. As we will see below, the first of the two will be a stochastic process, but the second one will be deterministic.

REMARK. To avoid confusion (and notational complications) we will use the following convention: upper-case superscripts, such as in ν_t^N or $F^N(t, x)$, refer to quantities associated with our stochastic process η_t^N, while lower-case superscripts, such as in ν_t^k or the function $F^k(t, x)$ which we will introduce below, refer to deterministic quantities associated with the deterministic process ν_t^k.

We begin by defining a process $\eta_t^{N,k}$ with values in $\bigcup_{M\geq N}\mathcal{X}_M$ inductively as follows. For each $m=0,\dots,2^k-1$, run the process with no killing on the interval $[\frac{m}{2^k},\frac{m+1}{2^k})$ and then at time $\frac{m+1}{2^k}$ repeatedly delete the leftmost particle until there are only N left. To make clear the distinction between the particles we erase in η_t^N and those we erase in this modified process at dyadic times, we will refer to this last procedure as *shaving off the extra mass* in $\eta_t^{N,k}$.

Having defined $\eta_t^{N,k}$ we now define $\nu_t^{N,k}$ as the empirical measure associated to it:

$$\nu_t^{N,k}=\frac{1}{N}\sum_{i=1}^{|\eta_t^{N,k}|}\delta_{\eta_t^{N,k}(i)},$$

where $|\eta_t^{N,k}|\geq N$ is the number of particles in $\eta_t^{N,k}$. In everything that follows we will consider the càdlàg version of $\eta_t^{N,k}$ and $\nu_t^{N,k}$, and we do the same for the other processes and functions defined below.

The first step in the proof of Theorem 1 will be to study the convergence of the tail distribution of $\nu_t^{N,k}$, defined by

$$F^{N,k}(t,x)=\nu_t^{N,k}([x,\infty)).$$

We will see in Proposition 2.2 that $F^{N,k}(t,x)$ converges in probability to $F^N(t,x)$ as $k\to\infty$, and a key fact will be that this convergence is uniform in N.

The second auxiliary process, ν_t^k, will turn out to be the limit of $\nu_t^{N,k}$ as $N\to\infty$. We define it in terms of its density, which is constructed inductively on each of the dyadic subintervals of $[0,1]$. We let $f^k(0,x)=f_0(x)$. If we have constructed f^k up to time $\frac{m}{2^k}$ for some $m\in\{0,\dots,2^k-1\}$ then for $t\in(\frac{m}{2^k},\frac{m+1}{2^k})$ we let $f^k(t,x)$ be the solution of

$$\frac{\partial f^k}{\partial t}(t,x)=\int_{-\infty}^{\infty}f^k(t,y)\rho(x-y)\,dy. \tag{2.1}$$

Then at time $\frac{m+1}{2^k}$ we let X_{m+1}^k be such that

$$\int_{X_{m+1}^k}^{\infty}f^k\left(\left(\frac{m+1}{2^k}\right)-,y\right)dy=1 \tag{2.2}$$

and define

$$f^k\left(\frac{m+1}{2^k},x\right)=f^k\left(\left(\frac{m+1}{2^k}\right)-,x\right)\mathbf{1}_{x>X_{m+1}^k}.$$

In other words, on each dyadic subinterval we let f^k evolve following (1.1) and then at each dyadic time we shave off the extra mass in f^k. The measure ν_t^k is defined as the measure having $f^k(t,\cdot)$ as its density. We also denote by F^k be the tail distribution of ν_t^k:

$$F^k(t,x)=\nu_t^k([x,\infty))=\int_x^{\infty}f^k(y)\,dy.$$

We will show in Lemma 2.5 that, for fixed t and x, $F^k(t,x)$ is decreasing in k, so we may define $F(t,x) = \lim_{k\to\infty} F^k(t,x)$. We will show in Proposition 2.2 that $F^{N,k}(t,x)$ converges in probability to $F^k(t,x)$ as $N \to \infty$ for fixed k. Since the convergence of $F^{N,k}(t,x)$ to $F^N(t,x)$ as $k \to \infty$ is uniform in N, we will be able to interchange limits to obtain the convergence of $F^N(t,x)$ to $F(t,x)$ as $N \to \infty$. The rest of the proof will consist of showing that $F(t,x)$ has a density which solves (FB) and then to extend the convergence of the tail distributions $F^N(t,x)$ for fixed t and x to the convergence of the measure-valued process $\nu_t^{N,k}$.

2.2. *Convergence of the auxiliary processes.* Our first result will allow us to deduce the limiting behavior of $F^{N,k}(t,x)$ as $N \to \infty$ inside each dyadic subinterval. Recall that in Section 1.1 we introduced the branching random walk ξ_t^N defined like η_t^N but with no killing. Let $\widehat{\nu}_t^N$ be the associated empirical measure. Let $\mathcal{M}$ be the space of finite measures on $\mathbb{R}$, endowed with the topology of weak convergence, and let $C([0,1],\mathcal{M})$ and $D([0,1],\mathcal{M})$ be the spaces of continuous and càdlàg functions from $[0,1]$ to $\mathcal{M}$ endowed, respectively, with the uniform and Skorohod topologies.

PROPOSITION 2.1. *The empirical process $\widehat{\nu}_t^N$ associated to the branching random walk ξ_t^N converges in distribution in $D([0,1],\mathcal{M})$ to a deterministic $\widehat{\nu}_t$ in $C([0,1],\mathcal{M})$ which for each $t \in [0,1]$ is absolutely continuous with respect to the Lebesgue measure. If we denote the density of $\widehat{\nu}_t$ by $\widehat{f}(t,x)$ then $\widehat{f}(t,x)$ is the unique solution to the integro-differential equation*

$$\frac{\partial \widehat{f}}{\partial t}(t,x) = \int_{-\infty}^{\infty} \widehat{f}(t,y)\rho(x-y)\,dy \tag{2.3}$$

on $[0,1]$ *with initial condition* $\widehat{f}(0,x) = f_0(x)$.

PROOF. By Theorem 5.3 of Fournier and Méléard (2004) we have that $\widehat{\nu}_t^N$ converges in distribution to a deterministic $\widehat{\nu}_t$ in $C([0,1],\mathcal{M})$ which is the unique solution of the following system: for all bounded and measurable φ,

$$\begin{aligned}\int_{-\infty}^{\infty}\varphi(x)\widehat{\nu}_t(dx) &= \int_{-\infty}^{\infty}\varphi(x)\widehat{\nu}_0(dx)\\ &\quad + \int_0^t\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(y)\rho(x-y)\,dy\,\widehat{\nu}_s(dx)\,ds.\end{aligned} \tag{2.4}$$

Moreover, by Proposition 5.4 of the same paper, $\widehat{\nu}_t$ is absolutely continuous for all $t \in [0,1]$, and hence its density $\widehat{f}(t,x)$ must satisfy $\widehat{f}(0,x) = f_0(x)$ and

$$\frac{d}{dt}\int_{-\infty}^{\infty}\varphi(x)\widehat{f}(t,x)\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(y)\rho(x-y)\widehat{f}(t,x)\,dy\,dx.$$

Taking $\varphi = \mathbf{1}_{[z,z+h]}$, dividing by h, taking $h \to 0$ and using the symmetry of ρ we deduce that $\widehat{f}$ satisfies (FB1) at z and the result follows. $\square$

Recall that $f_0(x) > 0$ for $x \geq 0$. It is clear that if $\widehat{f}$ solves (2.3) then $\widehat{f}(t,x) > 0$ for all $t \in [0,1]$ and $x \in \mathbb{R}$.

PROPOSITION 2.2. *For every fixed $k \geq 1$ and every $t \in [0,1]$ and $x \in \mathbb{R}$, $F^{N,k}(t,x)$ converges in probability to $F^k(t,x)$.*

PROOF. Proposition 2.1 implies that $\nu_t^{N,k}([x,\infty))$ converges in probability to $\nu_t^k([x,\infty))$ for all $t \in [0, \frac{1}{2^k})$ and all $x \in \mathbb{R}$. Using the partial order in $\mathcal{M}$ given by $\mu \preceq \nu$ if and only if $\mu([x,\infty)) \leq \nu([x,\infty))$ for all $x \in \mathbb{R}$, it is clear that the mappings $t \mapsto \nu_t^k$ and $t \mapsto \nu_t^{N,k}$ are increasing on $[0, \frac{1}{2^k})$, and thus the limits $\lim_{t\uparrow 1/2^k} \nu_t^k = \nu_-^k$ and $\lim_{t\uparrow 1/2^k} \nu_t^{N,k} = \nu_-^{N,k}$ exist (for $\nu_t^{N,k}$ this statement holds almost surely). On the interval $[0, \frac{1}{2^k})$ the process ν_t^k is the same as the process $\widehat{\nu}_t$ defined in Proposition 2.1. On the other hand, by (2.3) we have

$$\int_{-\infty}^{\infty} \widehat{f}(t,x)\,dx = 1 + \int_0^t \int_{-\infty}^{\infty} \widehat{f}(s,x)\,dx\,ds,$$

whence it is easy to see that

$$\int_{-\infty}^{\infty} \widehat{f}(t,x)\,dx = e^t. \tag{2.5}$$

Therefore for $0 \leq s \leq t \leq \frac{1}{2^k}$ we have

$$(\nu_t^k - \nu_s^k)(\mathbb{R}) = e^t - e^s.$$

Similarly, $\nu_t^{N,k}$ corresponds to the branching random walk ξ_t^N on $[0, \frac{1}{2^k})$, and thus using (5.4) of Fournier and Méléard (2004) we can see that $\mathbb{E}(\nu_t^{N,k}(\cdot))$ equals $\widehat{\nu}_t(\cdot)$ on this interval, so for $0 \leq s \leq t \leq \frac{1}{2^k}$ we have

$$\mathbb{E}\big((\nu_t^{N,k} - \nu_s^{N,k})(\mathbb{R})\big) = e^t - e^s.$$

Using these two equalities together with the fact that ν_-^k is absolutely continuous it is easy to see that $\nu_-^{N,k}([x,\infty))$ converges in probability to $\nu_-^k([x,\infty))$ for all x. Since $x \mapsto \nu_-^k((x,\infty))$ is strictly decreasing at $x = X_1^k$, the location of the point at which the mass of $\nu_t^{N,k}$ is shaved off at time $\frac{1}{2^k}$ converges in probability to X_1^k. The result for $t = \frac{1}{2^k}$ follows from this, and induction gives the desired result. □

Next we turn to the convergence of $F^{N,k}(t,x)$ as $k \to \infty$.

PROPOSITION 2.3. *For every given $t \in [0,1]$ and $x \in \mathbb{R}$, $F^{N,k}(t,x)$ converges in probability to $F^N(t,x)$ as $k \to \infty$, uniformly in N.*

The proof depends on the following lemma:

LEMMA 2.4. *We can couple $\eta_t^{N,k}$ and η_t^N (starting with the same initial configuration) in such a way that the following holds: for every $N \geq 1$, $k \geq 1$ and $t \in [0, 1]$,*

$$\mathbb{E}(|\eta_t^N \Delta \eta_t^{N,k}|) \leq N \frac{e-1}{e^{2^{-k}} - 1} 2^{-2k+1} e^{2^{-k}} + N 2^{-k+2},$$

where $A \Delta B = A \setminus B \cup B \setminus A$.

Before proving the lemma we need to give an explicit construction of the process η_t^N. Consider an i.i.d. family $(U_i^N)_{i \geq 1}$ with uniform distribution on $\{1, \ldots, N\}$ and an i.i.d. family $(R_i)_{i \geq 1}$ with distribution ρ and let $(T_i^N)_{i \geq 1}$ be the jump times of a Poisson process with rate N. To construct η_t^N we proceed as follows: at each time T_i^N we branch $\eta_{T_i^N}^N(U_i^N)$ using R_i for the displacement, erase the leftmost particle, and then relabel the particles to keep the ordering. The reader can check that the resulting process η_t^N has the desired distribution.

PROOF OF LEMMA 2.4. The coupling will be constructed inductively on each dyadic subinterval of $[0, 1]$. We start both processes with the same initial configuration. The idea will be to use the same branching times and displacements whenever possible. To do this we will decompose $\eta_t^{N,k}$ in the following way (for convenience we regard $\eta_t^{N,k}$ and η_t^N here as sets)

$$\eta_t^{N,k} = G_t^{N,k} \cup D_t^{N,k} \cup B_t^{N,k}, \tag{2.6}$$

where the unions are disjoint and:

- $G_t^{N,k} \subseteq \eta_t^N$ are "good particles," that is, particles which are coupled, in the sense that $G_t^{N,k} = \eta_t^{N,k} \cap \eta_t^N$;
- $B_t^{N,k}$ are "bad particles," that is, particles which are not coupled;
- $D_t^{N,k}$ are "dangerous particles," that is, particles which will become bad if not erased at the next dyadic time.

The basic idea of our coupling is the following. Good particles, which are present in both processes, evolve together using the same branching times and locations. When a good particle branches, a particle is erased from η_t^N but not from $\eta_t^{N,k}$. If the particle erased from η_t^N is not a good particle then the coupling is not affected. Otherwise, if the erased particle is good, we relabel it as dangerous in $\eta_t^{N,k}$. Observe that if this particle does not branch before the next dyadic time then it will not affect the coupling since it will surely get erased (by definition it is to the left of every good particle). When dangerous or bad particles give birth in $\eta_t^{N,k}$ we label their offspring as bad. Our goal will be to bound the number of bad particles.

Now we define the coupling more precisely. The first step is to construct η_t^N using the sequences U_i^N, R_i and T_i^N as described in the paragraph preceding this proof. Now we need to explain how to construct $\eta_t^{N,k}$ and decompose it into good, dangerous and bad particles. For the initial condition we choose $G_0^{N,k} = \eta_0^N$ and $D_0^{N,k} = B_0^{N,k} = \varnothing$.

We assume that we have constructed the coupling until time $\frac{m}{2^k}$ for some $0 \leq m \leq 2^k - 1$ and that the following holds:

$$G_{m/2^k}^{N,k} = \eta_{m/2^k}^{N,k} \cap \eta_{m/2^k}^N \quad \text{and} \quad D_{m/2^k}^{N,k} = \varnothing. \tag{2.7}$$

Observe that this condition holds trivially for $m = 0$. Observe also that $\eta_{m/2^k}^N \setminus G_{m/2^k}^{N,k}$ and $B_{m/2^k}^{N,k}$ both have $N - |G_{m/2^k}^{N,k}|$ particles, and thus we may identify particles in each set in a one-to-one fashion by, for example, going from left to right in each set.

Next we define the coupling on the interval $(\frac{m}{2^k}, \frac{m+1}{2^k}]$. Let $\frac{m}{2^k} \leq T_{I_m}^N < T_{I_m+1}^N < \cdots \leq T_{J_m}^N \leq \frac{m+1}{2^k}$ be the sequence of branching times for particles in η_t^N on this time interval (there are almost surely a finite number $J_m - I_m + 1$ of such times). We remark that after each branching event we will still have each particle in $\eta_t^N \setminus G_t^{N,k}$ identified with one particle in $B_t^{N,k}$ (see the second and last bullet below). For each $I_m \leq i \leq J_m$ we do the following:

- If the branching at time T_i^N occurs at a particle which is in $G_{T_i^N}^{N,k}$, we add the new particle to $G_{T_i^N}^{N,k}$.
- Otherwise, if the particle that is undergoing a branching in η_t^N at time T_i^N is not a good particle (and therefore it is not in $\eta_{T_i^N}^{N,k}$), we use the branching time and displacement to branch the particle in $B_{m/2^k}^{N,k}$ which is identified with it, and we identify this new bad particle with the new particle born in η_t^N at this branching event.
- If the particle erased from $\eta_{T_i^N}^N$ after the branching at time T_i^N is good (i.e., it is also in $G_{T_i^N}^{N,k}$), we relabel it as dangerous by moving it from $G_{T_i^N}^{N,k}$ to $D_{T_i^N}^{N,k}$. This dangerous particle in $\eta_t^{N,k}$ will not have an associated particle in η_t^N, so we use independent branching times and displacements for it and all its offspring, and label all its offspring as bad.
- Otherwise, if the particle erased from $\eta_{T_i^N}^N$ after the branching at time T_i^N is not in $G_{T_i^N}^{N,k}$ then there is a particle in $B_{T_i^N}^{N,k}$ identified with it; after the branching we remove this identification and use independent branching times and displacements for this bad particle and its offspring.

We remark that the offspring of dangerous and bad particles in $\eta_t^{N,k}$ is always labeled as bad and that whenever one of these particles has no associated particle in η_t^N it uses independent branching times and displacements.

The rules used to identify particles in $\eta_t^N \setminus G_t^{N,k}$ with particles in $B_t^{N,k}$ are not particularly important, the main point is that every branching event in η_t^N corresponds to a branching event in $\eta_t^{N,k}$ (though not the other way around, as some bad particles in $\eta_t^{N,k}$ branch independently of η_t^N).

At time $\frac{m+1}{2^k}$ we need to shave off the extra mass in $\eta_t^{N,k}$. Observe that, by our construction, we may erase particles of each of the three types. After erasing we relabel all remaining dangerous particles as bad by settting

$$G_{(m+1)/2^k}^{N,k} = \eta_{(m+1)/2^k}^{N,k} \cap \eta_{(m+1)/2^k}^{N}, \qquad B_{(m+1)/2^k}^{N,k} = \eta_{(m+1)/2^k}^{N,k} \setminus \eta_{(m+1)/2^k}^{N}$$

and

$$D_{(m+1)/2^k}^{N} = \varnothing.$$

In particular we see that the condition (2.7) holds at time $\frac{m+1}{2^k}$, allowing us to continue our inductive coupling.

We claim that the total number of bad particles after shaving and relabeling is bounded by the number of bad particles right before shaving:

$$|B_{(m+1)/2^k}^{N,k}| \le |B_{(m+1)/2^k-}^{N,k}|. \tag{2.8}$$

To see where the inequality comes from observe first that

$$B_{(m+1)/2^k}^{N,k} \subseteq B_{(m+1)/2^k-}^{N,k} \cup D_{(m+1)/2^k-}^{N,k}.$$

Now each particle in $D_{(m+1)/2^k-}^{N,k}$ is associated with a branching time T_i^N at which the number of particles in $\eta_{T_i^N}^{N,k}$ was increased by one. Therefore to each dangerous particle there corresponds some particle which will be erased when shaving; the corresponding particle to be erased is possibly the dangerous particle itself (in which case this particle will disappear after shaving so it will not be in $B_{(m+1)/2^k}^{N,k}$ after relabeling), and otherwise it has to be a bad particle because all good particles are to the right of any dangerous particle. In this way we continue the coupling until time 1.

Fix a dyadic subinterval $[\frac{m}{2^k}, \frac{m+1}{2^k})$. We claim that on this time interval the pair $(|D_t^{N,k}|, |B_t^{N,k}|)$ is stochastically dominated by a process (d_t^k, b_t^k) which evolves as follows:

$$\begin{aligned} d_t^k &\longrightarrow d_t^k + 1 \qquad \text{at rate } N, \\ b_t^k &\longrightarrow b_t^k + 1 \qquad \text{at rate } d_t^k + b_t^k \end{aligned}$$

with initial conditions $d_{m/2^k}^k = 0$ and $b_{m/2^k}^k = |B_{m/2^k}^{N,k}|$. In fact, bad particles increase by one when either a dangerous or a bad particle branches (so the second rate is actually the correct one), while dangerous particles are created as a consequence of some (but generally not all) of the branchings in η_t^N, which occur at rate N. An elementary calculation then shows that, for $h > 0$,

$$\mathbb{E}(d_{t+h}^k - d_t^k) = Nh + o(h), \qquad \mathbb{E}(b_{t+h}^k - b_t^k) = \mathbb{E}(d_t^k + b_t^k)h + o(h).$$

Then $\mathbb{E}(d_t^k) = N(t - \frac{m}{2^k})$ and thus dividing by h and taking $h \to 0$ we see that $\mathbb{E}(b_t^k)$ must solve

$$\frac{d\mathbb{E}(b_t^k)}{dt} = N\left(t - \frac{m}{2^k}\right) + \mathbb{E}(b_t^k) \tag{2.9}$$

for $t \in [\frac{m}{2^k}, \frac{m+1}{2^k})$. The solution of this ODE satisfies

$$\begin{aligned}\mathbb{E}(b_t^k) &= \big(\mathbb{E}(b_{m/2^k}^k) + N\big)e^{t-m/2^k} - N\left(t - \frac{m}{2^k} + 1\right)\\ &\le \mathbb{E}(b_{m/2^k}^k)e^{2^{-k}} + 2^{-2k}N,\end{aligned} \tag{2.10}$$

where we have used the inequality $e^x - 1 - x \le x^2$ for $x \in [0, 1]$ and the fact that $t - \frac{m}{2^k} \le \frac{1}{2^k}$. Since $b_0^k = 0$ we deduce that $\mathbb{E}(b_{(1/2^k)-}^k) \le N2^{-2k}$. At time $\frac{1}{2^k}$ we need to shave off the extra mass in $\eta_{(1/2^k)-}^{N,k}$ and this leaves us with $d_{1/2^k}^k = 0$ and $b_{1/2^k}^k \le b_{(1/2^k)-}^k$ by (2.8). Repeating this argument we get $\mathbb{E}(b_{2/2^k}^k) \le \mathbb{E}(b_{(2/2^k)-}^k) \le N[2^{-2k}e^{2^{-k}} + 2^{-2k}]$, and inductively we deduce that

$$\mathbb{E}(b_{m/2^k}^k) \le N\left[\sum_{j=0}^{m-1} e^{j2^{-k}}\right]2^{-2k} = N\frac{1 - e^{m2^{-k}}}{1 - e^{2^{-k}}}2^{-2k} \le N\frac{e - 1}{e^{2^{-k}} - 1}2^{-2k}, \tag{2.11}$$

where we used the fact that $m \le 2^k - 1$. Therefore, for $t \in [\frac{m}{2^k}, \frac{m+1}{2^k})$ we have, using (2.10),

$$\mathbb{E}(b_t^k) \le N\frac{e - 1}{e^{2^{-k}} - 1}2^{-2k}e^{2^{-k}} + N2^{-2k},$$

while, we recall, we also have $\mathbb{E}(d_t^k) = N(t - \frac{m}{2^k}) \le N2^{-k}$. Since $|\eta_t^N \Delta \eta_t^{N,k}| \le 2(d_t^k + b_t^k)$, the result follows. $\square$

PROOF OF PROPOSITION 2.3. Fix $k > 0$ for a moment and assume that $t \in [\frac{m}{2^k}, \frac{m+1}{2^k})$. Using the coupling introduced in Lemma 2.4 we have

$$\begin{aligned}\mathbb{E}\big(|F^{N,k}(t, x) - F^N(t, x)|\big) &= \mathbb{E}\left(\left|\frac{1}{N}\sum_{i=1}^N \big(\mathbf{1}_{\eta_t^{N,k}(i)\ge x} - \mathbf{1}_{\eta_t^N(i)\ge x}\big)\right|\right)\\ &\le \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\eta_t^{N,k}(i)\ne\eta_t^N(i)}\right)\\ &\le \frac{e - 1}{e^{2^{-k}} - 1}2^{-2k+1}e^{2^{-k}} + 2^{-k+2},\end{aligned}$$

so we deduce by Markov's inequality that

$$\mathbb{P}\big(|F^{N,k}(t,x) - F^N(t,x)| > \varepsilon\big) \le \frac{1}{\varepsilon}\left[\frac{e-1}{e^{2^{-k}}-1}2^{-2k+1}e^{2^{-k}} + 2^{-k+2}\right] \xrightarrow[k\to\infty]{} 0$$

uniformly in N. □

LEMMA 2.5.

$$F^k(t,x) \ge F^{k+1}(t,x) \qquad \textit{for all } t \in [0,1], x \in \mathbb{R}, k \ge 1.$$

PROOF. Fix $k \ge 1$ and $x \in \mathbb{R}$. The result is trivial at $t = 0$. We will work inductively on the intervals $(\frac{m}{2^k}, \frac{m+1}{2^k}]$. Take $0 \le m \le 2^k - 1$ and assume that

$$F^k\left(\frac{m}{2^k}, x\right) \ge F^{k+1}\left(\frac{m}{2^k}, x\right).$$

Then writing $H = F^k - F^{k+1}$ we have for $\frac{m}{2^k} < t < \frac{m}{2^k} + \frac{1}{2^{k+1}} = \frac{2m+1}{2^{k+1}}$ that

$$\begin{aligned}
\frac{\partial H}{\partial t}(t,x) &= \int_x^\infty \int_{-\infty}^\infty [f^k(t,u) - f^{k+1}(t,u)]\rho(y-u)\,du\,dy \\
&= \int_{-\infty}^\infty \int_x^\infty [f^k(t,y-v) - f^{k+1}(t,y-v)]\rho(v)\,dy\,dv \qquad (2.12)\\
&= \int_{-\infty}^\infty H(t,x-v)\rho(v)\,dv = \int_{-\infty}^\infty H(t,z)\rho(x-z)\,dz,
\end{aligned}$$

so H satisfies (FB1) on this interval and thus, since $H(\frac{m}{2^k},\cdot) \ge 0$, we get $H(t,\cdot) \ge 0$ for $t \in (\frac{m}{2^k}, \frac{2m+1}{2^{k+1}})$. At time $\frac{2m+1}{2^{k+1}}$ the density f^{k+1} is shaved off, leaving $F^k(\frac{2m+1}{2^{k+1}}, x) \ge F^{k+1}(\frac{2m+1}{2^{k+1}}, x)$. Repeating the above argument we get

$$F^k(t,x) \ge F^{k+1}(t,x) \qquad \text{for all } t \in \left(\frac{m}{2^k}, \frac{m+1}{2^k}\right). \tag{2.13}$$

Now at time $\frac{m+1}{2^{k+1}}$ both densities f^k and f^{k+1} are shaved off, say at points x_k and x_{k+1}, respectively. Then by (2.13), $x_k \ge x_{k+1}$, and thus (2.13) holds at $t = \frac{m+1}{2^{k+1}}$ as well. □

Since $F^k(t,x)$ is decreasing and positive, we can define

$$F(t,x) = \lim_{k\to\infty} F^k(t,x). \tag{2.14}$$

It is obvious that for each given t, $F(t,\cdot)$ is nonincreasing and its range is [0, 1].

PROPOSITION 2.6. *For every $t \in [0,1]$ and $x \in \mathbb{R}$, $F^N(t,x)$ converges almost surely and in L^1 as $N \to \infty$ to $F(t,x)$.*

PROOF. First observe that, for fixed $t \in [0,1]$ and $x \in \mathbb{R}$ and since $F^N(t,x) \leq 1$, the sequence of random variables $(F^N(t,x))_{N>0}$ is uniformly integrable, so it is enough to show that $F^N(t,x) \to F(t,x)$ in probability.

Fix $\varepsilon > 0$. Use Proposition 2.2 to choose, for each $k > 0$, an $N_k > 0$ so that

$$\mathbb{P}\left(|F^{N,k}(t,x) - F^k(t,x)| > \frac{\varepsilon}{2}\right) < \frac{1}{k}$$

for every $N \geq N_k$ and $N_k \uparrow \infty$. Define k_N as follows: $k_N = 1$ for $N < N_1$ and $k_N = k$ for $N_k \leq N \leq N_{k+1}$. We have

$$\begin{aligned}
&\mathbb{P}\big(|F^{N,k_N}(t,x) - F(t,x)| > \varepsilon\big)\\
&\quad \leq \mathbb{P}\left(|F^{N,k_N}(t,x) - F^{k_N}(t,x)| > \frac{\varepsilon}{2}\right)\\
&\qquad + \mathbb{P}\left(|F^{k_N}(t,x) - F(t,x)| > \frac{\varepsilon}{2}\right).
\end{aligned}$$

By the definition of k_N, the first term on the right-hand side is less than $1/k_N$, while by (2.14) the second one is 0 when k_N is large enough. We deduce that $F^{N,k_N}(t,x)$ converges in probability as $N \to \infty$ to $F(t,x)$.

To finish the proof write

$$\begin{aligned}
&\mathbb{P}\big(|F^{N}(t,x) - F(t,x)| > \varepsilon\big)\\
&\quad \leq \mathbb{P}\left(|F^{N}(t,x) - F^{N,k_N}(t,x)| > \frac{\varepsilon}{2}\right)\\
&\qquad + \mathbb{P}\left(|F^{N,k_N}(t,x) - F(t,x)| > \frac{\varepsilon}{2}\right).
\end{aligned}$$

We already know that the second term on the right-hand side goes to 0, while the first one goes to 0 thanks to Proposition 2.3 (here we use the fact that the convergence is uniform in N). $\square$

Recall the definition in (2.2) of the shaving points X_m^k and let $X^k : [0,1] \longrightarrow \mathbb{R}$ be the corresponding linear interpolation, that is,

$$X^k(t) = X_m^k + \frac{X_{m+1}^k - X_m^k}{2^k}\left(t - \frac{m}{2^k}\right) \qquad \text{for } \frac{m}{2^k} < t \leq \frac{m+1}{2^k}.$$

LEMMA 2.7. *$X^k(t)$ converges uniformly in $[0,1]$ to a continuous function $\gamma(t)$.*

PROOF. We will start by showing that the sequence of functions $(X^k)_{k>0}$ is relatively compact. By the Arzelà–Ascoli theorem, we only need to show that our sequence is uniformly bounded and equicontinuous.

Observe that, for each given k, $X^k(t)$ is increasing. Indeed, it is enough to show that $X_m^k \leq X_{m+1}^k$ for $0 \leq m < 2^k$, and this follows from the fact that, if $f^k(\frac{m}{2^k},\cdot)\geq 0$, then (FB1) implies that $f^k(t,\cdot)\geq f^k(\frac{m}{2^k},\cdot)$ for $\frac{m}{2^k}<t<\frac{m+1}{2^k}$. Therefore

$$\sup_{k>0}\sup_{t\in[0,1]} X^k(t)=\sup_{k>0} X^k(1).$$

To show that this last supremum is finite, observe that $\widehat{f}(t,x)$ (which was defined in Proposition 2.1) satisfies $\widehat{f}(t,x)\geq f^k(t,x)$ for all k. On the other hand we know by (2.5) that $\int_{-\infty}^{\infty}\widehat{f}(1,x)\,dx=e$. Therefore if we let $M>0$ be such that $\int_M^{\infty}\widehat{f}(1,x)\,dx<1$ we deduce that $X^k(1)\leq M$ for all k and the uniform boundedness follows.

For the equicontinuity we need to show that given any $\varepsilon>0$ there is a $\delta>0$ such that

$$\sup_{k>0}|X^k(t)-X^k(s)|<\varepsilon,$$

whenever $|t-s|<\delta$. Assume that $s<t$, fix k for a moment and let $\frac{l}{2^k}$ and $\frac{m}{2^k}$ be the dyadic numbers immediately to the right of s and t, respectively (here we assume k is large enough so that $m\vee l<2^k$). Then

$$\begin{aligned}
|X^k(t)-X^k(s)|&\leq\left|X^k(t)-X^k\left(\frac{m}{2^k}\right)\right|\\
&\quad+\left|X^k\left(\frac{m}{2^k}\right)-X^k\left(\frac{l}{2^k}\right)\right|+\left|X^k\left(\frac{l}{2^k}\right)-X^k(s)\right|\\
&\leq\left|X^k\left(\frac{m+1}{2^k}\right)-X^k\left(\frac{m}{2^k}\right)\right|+\left|X^k\left(\frac{m}{2^k}\right)-X^k\left(\frac{l}{2^k}\right)\right|\\
&\quad+\left|X^k\left(\frac{l+1}{2^k}\right)-X^k\left(\frac{l}{2^k}\right)\right|.
\end{aligned}\tag{2.15}$$

Now for any $p,q\in\{0,\dots,2^k\}$ with $q\geq p$ we have

$$\begin{aligned}
&\int_{X^k(p/2^k)+\varepsilon}^{\infty} f^k\left(\frac{q}{2^k},y\right)dy\\
&\qquad=\int_{X^k(p/2^k)}^{\infty} f^k\left(\frac{p}{2^k},y\right)dy+\int_{X^k(p/2^k)}^{\infty}\left[f^k\left(\frac{q}{2^k},y\right)-f^k\left(\frac{p}{2^k},y\right)\right]dy\\
&\qquad\quad-\int_{X^k(p/2^k)}^{X^k(p/2^k)+\varepsilon} f^k\left(\frac{q}{2^k},y\right)dy.
\end{aligned}\tag{2.16}$$

The first term on the right-hand side equals 1. The second term corresponds to the amount of mass accumulated by f^k to the right of $X^k(\frac{p}{2^k})$ on the time interval $(\frac{p}{2^k},\frac{q}{2^k}]$. Using (FB1) it is not hard to see that this is bounded by the same quantity with f^k replaced by $\widehat{f}$, so using (2.5) we get the bound

$$\int_{X^k(p/2^k)}^{\infty}\left[\widehat{f}\left(\frac{q}{2^k},y\right)-\widehat{f}\left(\frac{p}{2^k},y\right)\right]dy\leq e^{q/2^k}-e^{p/2^k}\leq e\frac{q-p}{2^k}.$$

On the other hand, using the fact that $X^k(t)$ is increasing, it is clear that $f^k(\frac{q}{2^k}, x) \geq f^k(0, x) = f_0(x)$ for $x \geq X^k(\frac{q}{2^k})$. Therefore the last term on (2.16) is greater than or equal to

$$\int_{X^k(p/2^k)}^{X^k(p/2^k)+\varepsilon} f_0(y)\,dy.$$

Now $X^k(\frac{p}{2^k})$ is nonnegative and bounded by M by the preceding arguments, so the last integral is at least

$$L = \inf_{x \in [0,M]} \int_x^{x+\varepsilon} f_0(y)\,dy > 0,$$

where we used the fact that f_0 is strictly positive on the positive half-line. Putting the last two bounds together with (2.16) we get

$$\int_{X^k(p/2^k)+\varepsilon}^{\infty} f^k\left(\frac{q}{2^k}, y\right) dy \leq 1 + e\frac{q-p}{2^k} - L.$$

Now $|t - s| < \delta$ implies that $\frac{m-l}{2^k} \leq \delta + \frac{1}{2^k}$, and thus we deduce that

$$\int_{X^k(p/2^k)+\varepsilon}^{\infty} f^k\left(\frac{q}{2^k}, y\right) dy < 1$$

for small enough δ and large enough k and for $(p, q) \in \{(l, l+1), (m, m+1), (l, m)\}$. The preceding means that if δ is small enough and K is large enough then $|X^k(\frac{q}{2^k}) - X^k(\frac{p}{2^k})| < \varepsilon$ for $k \geq K$ and for these three pairs (p, q). Using (2.15) we obtain

$$\sup_{k \geq K} |X^k(t) - X^k(s)| < \varepsilon,$$

if $|t - s| < \delta$ and δ is small enough. Since the functions X^k are all uniformly continuous (on $[0, 1]$), it is clear that, by choosing δ even smaller if necessary, the same will hold also for $k = 1, \dots, K - 1$. This finishes the proof of the equicontinuity.

The last thing we need to show is that our sequence has a unique limit point. Consider two convergent subsequences $X^{n_k} \to \gamma_1$ and $X^{m_k} \to \gamma_2$. Let $t = \frac{i}{2^l}$ be any dyadic number in $[0, 1]$ and assume that k is large enough so that $n_k \wedge m_k \geq l$. Recall from the proof of Lemma 2.5 that $X^k(t)$ is nonincreasing in k for each fixed $t \in [0, 1]$. Since $F^{n_k}(t, x) = 1$ for all $x \leq X^{n_k}(t)$ we deduce that

$$F^{n_k}(t, x) = 1 \qquad \text{for all } x \leq \gamma_1(t). \tag{2.17}$$

Now given any k there is a k' such that $n_{k'} \geq m_k$, so by Lemma 2.5 we get

$$1 = F^{n_{k'}}(t, x) \leq F^{m_k}(t, x) \leq 1 \qquad \text{for all } x \leq \gamma_1(t).$$

This means that $X^{m_k}(t) \geq \gamma_1(t)$ for all large enough k, and taking $k \to \infty$ we deduce that $\gamma_2(t) \geq \gamma_1(t)$. By symmetry we get $\gamma_1(t) \geq \gamma_2(t)$. This gives $\gamma_1(t) = \gamma_2(t)$ for all dyadic $t \in [0, 1]$, and now the uniqueness follows from the continuity of γ_1 and γ_2. $\square$

2.3. *Properties of F and proof of the theorem.* To finish the proof of the Theorem 1 we need to show that F has a density which satisfies (FB) and the rest of the requirements of the theorem and then extend the convergence to the measure-valued process ν_t^N. The first step in doing that will be to derive an equation satisfied by F.

Suppose that $(g(t,x),\gamma(t))$ solves (FB) and let $G(t,x)=\int_x^\infty g(t,y)\,dy$. Then it is not difficult to check, repeating the arguments leading to (2.12), that $(G(t,x),\gamma(t))$ must solve the following free boundary problem (FB′):

$$\text{(FB1}'\text{)}\qquad \frac{\partial G}{\partial t}(t,x)=\int_{-\infty}^{\infty} G(t,y)\rho(x-y)\,dy \qquad \forall x>\gamma(t),$$

$$\text{(FB2}'\text{)}\qquad G(t,x)=1 \qquad \forall x\le\gamma(t)$$

with initial condition $G(0,x)=\int_x^\infty f_0(y)\,dy$. Moreover, if $(G(t,x),\gamma(t))$ solves (FB′) and $G(t,\cdot)$ is absolutely continuous for all t, then $(g(t,x),\gamma(t))$, where $g(t,\cdot)$ is the density of $G(t,\cdot)$, must solve (FB).

PROPOSITION 2.8. *$F(t,x)$ is differentiable in t for all $x>\gamma(t)$ and it satisfies (FB′).*

PROOF. We already proved [see (2.17)] that $(F(t,x),\gamma(t))$ satisfies (FB2′). For $x>\gamma(t)$ and by the definition of $F^k(t,x)$ [which implies that $F^k(t,x)$ is differentiable inside each dyadic subinterval] we may write

$$\begin{aligned}
F^k(t,x)&=F^k(0,x)+\sum_{m=1}^{n_k(t)}\biggl[F^k\Bigl(\frac{m}{2^k},x\Bigr)-F^k\Bigl(\frac{m-1}{2^k},x\Bigr)\biggr]\\
&\qquad+[F^k(t,x)-F^k(n_k(t),x)],\\
&=F^k(0,x)+\sum_{m=1}^{n_k(t)}\biggl[\int_{((m-1)/2^k,m/2^k)}\frac{\partial F^k}{\partial s}(s,x)\,ds\biggr]\\
&\qquad+\int_{(n_k(t)/2^k,t)}\frac{\partial F^k}{\partial s}(s,x)\,ds\\
&\qquad+\sum_{m=1}^{n_k(t)}\biggl[F^k\Bigl(\frac{m}{2^k},x\Bigr)-F^k\Bigl(\Bigl(\frac{m}{2^k}\Bigr)-,x\Bigr)\biggr]\\
&\qquad+[F^k(t,x)-F^k(t-,x)],
\end{aligned}$$

where $n_k(t)=\frac{\lfloor 2^k t\rfloor}{2^k}$. Recalling that $X^k(t)\downarrow\gamma(t)$, we can take k large enough so that $\gamma(t)\le X^k(t)<x$. Since $X^k(s)$ is increasing in s we deduce that $X^k(\frac{m}{2^k})<x$

for $m = 1, \dots, n_k(t)$, and therefore all the terms in the last line above are 0. On the other hand, observe that F^k must solve (FB1′) on each dyadic subinterval, which can be checked repeating again the calculations in (2.12). Therefore,

$$
\begin{aligned}
F^k(t,x) &= F^k(0,x) + \sum_{m=1}^{n_k(t)} \int_{((m-1)/2^k, m/2^k)} \int_{-\infty}^{\infty} F^k(s,y)\rho(x-y)\,dy\,ds \\
&\quad + \int_{(n_k(t)/2^k, t)} \int_{-\infty}^{\infty} F^k(s,y)\rho(x-y)\,dy\,ds \\
&= F^k(0,x) + \int_0^t \int_{-\infty}^{\infty} F(s,y)\rho(x-y)\,dy\,ds \\
&\quad + \sum_{m=1}^{n_k(t)} \int_{((m-1)/2^k, m/2^k)} \int_{-\infty}^{\infty} [F^k(s,y) - F(s,y)]\rho(x-y)\,dy\,ds \\
&\quad + \int_{((n_k(t))/2^k, t)} \int_{-\infty}^{\infty} [F^k(s,y) - F(s,y)]\rho(x-y)\,dy\,ds.
\end{aligned} \tag{2.18}
$$

Now for fixed y, $F^k(\cdot, y)$ is a decreasing sequence converging to $F(\cdot, y)$, so Dini's theorem implies that

$$
\Delta_k(y) = \sup_{t \in [0,1]} |F^k(t,y) - F(t,y)| \xrightarrow[k\to\infty]{} 0.
$$

The sum of the terms on the last two lines of (2.18) is bounded by

$$
\begin{aligned}
&\sum_{m=1}^{n_k(t)} \int_{((m-1)/2^k, m/2^k)} \int_{-\infty}^{\infty} \Delta_k(y)\rho(x-y)\,dy\,ds \\
&\quad + \int_{(n_k(t)/2^k, t)} \int_{-\infty}^{\infty} \Delta_k(y)\rho(x-y)\,dy\,ds \\
&\quad \le t \int_{-\infty}^{\infty} \Delta_k(y)\rho(x-y)\,dy,
\end{aligned}
$$

and this last integral goes to 0 as $k \to \infty$ by the dominated convergence theorem, because using (2.5) we get $\Delta_k(y) \le \sup_{t\in[0,1]} F^k(t,y) \le \int_y^\infty \widehat{f}(1,z)\,dz \le e$. Using this and taking $k \to \infty$ in (2.18) we get

$$
F(t,x) = F(0,x) + \int_0^t \int_{-\infty}^{\infty} F(s,y)\rho(x-y)\,dy\,ds. \tag{2.19}
$$

To finish the proof it is enough to show that the mapping $s \mapsto \int_{-\infty}^{\infty} F(s,y)\rho(x-y)\,dy$ is continuous, since if that is the case then we can differentiate (2.19) and deduce (FB1′). This actually follows easily from (2.19):

$$\left|\int_{-\infty}^{\infty} F(s+h,y)\rho(x-y)\,dy - \int_{-\infty}^{\infty} F(s,y)\rho(x-y)\,dy\right|$$

$$(2.20)\qquad = \int_{-\infty}^{\infty}\int_{s}^{s+h}\int_{-\infty}^{\infty} F(r,z)\rho(y-z)\rho(x-y)\,dz\,dr\,dy$$

$$\leq \int_{s}^{s+h}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(y-z)\rho(x-y)\,dz\,dy\,dr = h. \qquad \square$$

Let ν_t be the probability measure defined by $\nu_t([x,\infty)) = F(t,x)$. Since F satisfies (FB′) we have that for every $b > a > \gamma(t)$,

$$\frac{d}{dt}\nu_t([a,b]) = \int_{-\infty}^{\infty} \nu_t([a-y,b-y])\rho(y)\,dy$$

and thus by standard measure theory arguments we deduce that for every bounded and measurable φ with support contained in $(\gamma(t),\infty)$,

$$(2.21)\qquad \frac{d}{dt}\int_{-\infty}^{\infty}\varphi(y)\nu_t(dy) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x+y)\rho(y)\nu_t(dx)\,dy$$

[cf. (2.4)]. Now if $A \subseteq (\gamma(t),\infty)$ has zero Lebesgue measure and the support of φ is contained in A, then the right-hand side above is 0 and we deduce that $\nu_t(A)$ is constant. Since $\nu_0(A) = \int_A f_0(x)\,dx = 0$, we have proved that ν_t is absolutely continuous with respect to the Lebesgue measure. We will denote its density by $f(t,\cdot)$, and we obviously have $F(t,x) = \int_x^\infty f(t,y)\,dy$.

At this point we are ready to finish the proof of Theorem 1 by showing that F satisfies the desired properties and then using the convergence of the tail distributions $F^N(t,x)$ to obtain the convergence in distribution of the process ν_t^N in $D([0,1],\mathcal{P})$.

PROOF OF THEOREM 1. By Proposition 2.6 we know that $F^N(t,x) \to F(t,x)$ almost surely as $N \to \infty$ and that F can be written in terms of the integral of f. Now, by (2.19),

$$\int_x^\infty f(t,y)\,dy = \int_x^\infty f_0(y)\,dy + \int_0^t\int_{-\infty}^{\infty}\int_y^\infty f(s,z)\rho(x-y)\,dz\,dy\,ds$$

for $x > \gamma(t)$, so for any $h > 0$ we have

$$(2.22)\qquad \begin{aligned} &\frac{1}{h}\left[\int_{x+h}^\infty f(t,y)\,dy - \int_x^\infty f(t,y)\,dy\right] \\ &\quad = -\frac{1}{h}\int_x^{x+h} f_0(y)\,dy - \int_0^t\int_{-\infty}^{\infty}\frac{1}{h}\int_{x-u}^{x+h-u} f(s,z)\,dz\,\rho(u)\,du\,ds. \end{aligned}$$

The first term on the right-hand side goes to $f_0(x)$ as $h \to 0$ [recall that $f_0(x)$ is continuous for $x > 0$]. The second term goes to $\int_0^t\int_{-\infty}^\infty f(s,x-u)\rho(u)\,du\,ds$

by the dominated convergence theorem. On the other hand, the left-hand side of (2.22) goes to $\frac{\partial F}{\partial x}(t,x)$, which equals $-f(t,x)$ for almost every $x>\gamma(t)$. We deduce that

$$f(t,x)=f_0(x)+\int_0^t\int_{-\infty}^{\infty}f(s,y)\rho(x-y)\,dy\,ds \tag{2.23}$$

for almost every x.

Now if $x_n\to x$, then by (2.23)

$$\begin{aligned}|f(t,x_n)-f(t,x)|&\le|f_0(x_n)-f_0(x)|\\&\quad+\int_0^t\int_{-\infty}^{\infty}f(s,y)|\rho(x_n-y)-\rho(x-y)|\,dy\,ds.\end{aligned}$$

The first term on the right-hand side goes to 0 as $n\to\infty$ because f_0 is continuous, while the second term goes to 0 by the continuity and boundedness of ρ and the dominated convergence theorem. We deduce that $f(t,x)$ is continuous in x and thus, in particular, (2.23) holds for every x. Hence $(f(t,x),\gamma(t))$ satisfies (FB). Moreover, since the above convergence can be achieved uniformly for t in compact intervals, it is easy to see that, if $f(t,x)$ is continuous in t for $x\neq\gamma(t)$ (as we will show next), then it is actually jointly continuous in t and x outside the curve $\{(t,\gamma(t)):t\ge 0\}$. The fact that $f(t,x)$ is differentiable (and thus continuous) in t for $x\neq\gamma(t)$ follows easily from (2.23) by repeating the arguments in (2.20). γ is strictly increasing because, according to the evolution defined by (FB), $f(t,x)$ always increases when $x>\gamma(t)$ [alternatively, differentiate (FB2) with respect to t to find that $\gamma'(t)>0$]. We also have that $f(t,x)>0$ for $x>\gamma(t)$ thanks to the facts that $f_0(x)>0$ for $x>0$ and that γ is increasing.

The only thing left to show before turning to the proof of the convergence in distribution of ν_t^N in $D([0,1],\mathcal{P})$ is that $(f(t,x),\gamma(t))$ is the unique solution of (FB). To do this, it is enough to show that if $(h(t,x),\sigma(t))$ is any given solution and $H(t,x)=\int_x^\infty h(t,y)\,dy$ then $H(t,x)=F(t,x)$. Indeed, if that is the case then the above arguments imply that $h(t,x)$ is jointly continuous in t and x outside the curve $\{(t,\gamma(t)):t\ge 0\}$, and thus $h(t,x)=f(t,x)$, while (FB2) and (FB3) imply that $\sigma(t)=\gamma(t)$.

The idea of the proof will be to compare $h(t,x)$ and $f^k(t,x)$ by adapting the coupling introduced in the proof of Lemma 2.4 to the deterministic system, so we will sketch the main ideas and leave the details to the reader. We will write

$$f^k(t,x)=g^k(t,x)+d^k(t,x)+b^k(t,x),$$

which is to be interpreted in a manner analogous to (2.6). We construct these three functions inductively on each dyadic subinterval of $[0,1]$. We start with $g^k(0,x)=f_0(x)$ and $d^k(0,x)=b^k(0,x)=0$ for all x. Next, for $t\in(0,\frac{1}{2^k})$ we let the three functions evolve according to the following system of integro-differential equations:

$$
\begin{aligned}
&\frac{\partial g^k}{\partial t}(t,x) = \mathbf{1}_{x>\sigma(t)} \int_{-\infty}^{\infty} g^k(t,y)\rho(x-y)\,dy,\\
&g^k(t,x) = 0 \qquad \text{for } x \le \sigma(t),\\
&\frac{\partial d^k}{\partial t}(t,x) = \mathbf{1}_{x\le\sigma(t)} \int_{-\infty}^{\infty} g^k(t,y)\rho(x-y)\,dy,\\
&\frac{\partial b^k}{\partial t}(t,x) = \int_{-\infty}^{\infty} [d^k(t,y)+b^k(t,y)]\rho(x-y)\,dy.
\end{aligned}
\tag{2.24}
$$

In words, the "good mass" $g^k(t,x)$ is constrained to be to the right of $\sigma(t)$ and evolves by the analog of branching (the convolution term on the right-hand side of the first equation), which is also constrained to the half-line $(\sigma(t),\infty)$; the "dangerous mass" $d^k(t,x)$ evolves by acquiring the mass due to the branching of the good mass to the left of $\sigma(t)$; and the "bad mass" $b^k(t,x)$ arises from acquiring the mass due to the branching of both the dangerous mass and the bad mass.

Adding the first, third, and fourth equations above we see clearly that $f^k(t,x)$ satisfies (2.1) on this interval as required. At time $\frac{1}{2^k}$ we need to shave off the extra mass in f^k. Observe that $g^k((\frac{1}{2^k})-,x) = h((\frac{1}{2^k})-,x)$ for all x, so $X_1^k \ge \sigma(\frac{1}{2^k})$. Thus all the mass to the left of $\sigma(\frac{1}{2^k})$ needs to be erased, so we put $d^k(\frac{1}{2^k},x)=0$ for all x. The rest of the mass to be erased from f^k will come from both g^k and b^k. This leaves us with

$$
g^k\left(\frac{1}{2^k},x\right) \le h\left(\frac{1}{2^k},x\right), \qquad d^k\left(\frac{1}{2^k},x\right) = 0 \quad \text{and} \quad b^k\left(\frac{1}{2^k},x\right) \ge 0.
$$

We continue the construction inductively. Assume that the above holds at time $\frac{m}{2^k}$ for some $1 \le m < 2^k$. On the interval $(\frac{m}{2^k},\frac{m+1}{2^k})$ we let g^k, d^k and b^k evolve according to (2.24). At time $\frac{m+1}{2^k}$ we need to shave off the extra mass in f^k, and as in the proof of Lemma 2.4 it is not hard to see that after doing that we may rebalance d^k and b^k in such a way that $d^k(\frac{m+1}{2^k},x)=0$ for all x and the total mass of $b^k(\frac{m+1}{2^k},x)$ is at most the total mass of $b^k((\frac{m+1}{2^k})-,x)$. Observe that this construction preserves the inequality $g^k(t,x)\le h(t,x)$ thanks to the observation following the proof of Proposition 2.1.

We continue in a similar way to the proof of Lemma 2.4. Fix a dyadic interval $[\frac{m}{2^k},\frac{m+1}{2^k})$ and observe that, for t on this interval, if we let

$$
\overline{d}_t^k = \int_{-\infty}^{\infty} d^k(t,x)\,dx \quad \text{and} \quad \overline{b}_t = \int_{-\infty}^{\infty} b^k(t,x)\,dx
$$

be the total dangerous and bad masses, respectively, then these quantities satisfy the differential inequalities $\frac{d}{dt}\overline{d}_t^k \le 1$ and $\frac{d}{dt}\overline{b}_t \le \overline{d}_t^k + \overline{b}_t$, whence we deduce that

$\overline{b}_t$ satisfies

$$\frac{d}{dt}\overline{b}_t \leq t - \frac{m}{2^k} + \overline{b}_t.$$

Thus the same argument we used to obtain (2.11) gives

$$\sup_{m=0,\dots,2^k} \overline{b}^k_{m/2^k} = \frac{e-1}{e^{2^{-k}}-1} 2^{-2k} \xrightarrow[k\to\infty]{} 0.$$

Now let $t = \frac{l}{2^k}$ be any dyadic number in $[0,1]$ and assume that $k \geq l$. Using the fact that $d^k(t,x) = 0$ for all $x \in \mathbb{R}$ we can write

$$\begin{aligned} 0 &= \int_x^\infty \big(h(t,y) - f^k(t,y)\big)\,dy \\ &= \int_x^\infty \big(h(t,y) - g^k(t,y)\big)\,dy - \int_x^\infty b^k(t,y)\,dy \end{aligned} \tag{2.25}$$

so $\int_x^\infty (h(t,y) - g^k(t,y))\,dy \leq \overline{b}^k_t \to 0$. Thus, since $h(t,y) \geq g^k(t,y)$, we deduce that $\int_x^\infty g^k(t,y)\,dy \to \int_x^\infty h(t,y)\,dy$ as $k \to \infty$ uniformly in x, which in turn implies by (2.25) that

$$\lim_{k\to\infty} F^k(t,x) = \lim_{k\to\infty} \int_x^\infty f^k(t,y)\,dy = \int_x^\infty h(t,y)\,dy = H(t,x)$$

uniformly in x and for every dyadic number $t \in [0,1]$. Since $(H(t,x), \sigma(t))$ solves (FB′), H is continuous in t and we deduce that $H(t,x) = F(t,x)$ for all $t \in [0,1]$.

To show that the sequence of processes ν^N_t converges in distribution in $D([0,1],\mathcal{P})$ to the deterministic process ν_t in this space defined by having its densities evolve according to (FB), it is enough to prove that this sequence is tight. In fact, if $\nu^{N_k}_t$ is any convergent subsequence, then its limit ν_t is completely defined by its tail distribution at each time t, which we know must be $F(t,\cdot)$. To show that ν^N_t is tight it is enough, by Theorem 2.1 of Roelly-Coppoletta (1986), to show that for any continuous and bounded function φ on $\mathbb{R}$ the sequence of real-valued processes $\langle \nu^N_t, \varphi\rangle = \int_{-\infty}^\infty \varphi(x)\nu^N_t(dx)$ is tight in $D([0,T],\mathbb{R})$. Fix one such function φ. By Aldous' criterion [which we take from Theorem 2.2.2 in Joffe and Métivier (1986) and the corollary that preceeds it in page 34], we need to prove that the following two conditions hold:

(i) For every rational $t \in [0,T]$ and every $\varepsilon > 0$, there is an $L > 0$ such that

$$\sup_{N>0} \mathbb{P}(|\langle \nu^N_t, \varphi\rangle| > L) \leq \varepsilon.$$

(ii) If $\mathfrak{T}^N_T$ is the collection of stopping times with respect to the natural filtration associated to $\langle \nu^N_t, \varphi\rangle$ that are almost surely bounded by T, then for every $\varepsilon > 0$

$$\lim_{r\to 0} \limsup_{N\to\infty} \sup_{\substack{s<r \\ \tau\in\mathfrak{T}^N_T}} \mathbb{P}\big(\big|\langle \nu^N_{(\tau+s)\wedge T}, \varphi\rangle - \langle \nu^N_\tau, \varphi\rangle\big| > \varepsilon\big) = 0.$$

The first condition holds trivially in our case by taking $L > \|\varphi\|_\infty$. To get the second one fix $N > 0$, $\varepsilon > 0$, $0 < s < r$ and $\tau \in \mathfrak{T}_T^N$ and let K be the number of branchings in η_t^N on the interval $[\tau, (\tau + s) \wedge T]$. Observing that

$$\left|\langle \nu^N_{(\tau+s)\wedge T}, \varphi\rangle - \langle \nu^N_\tau, \varphi\rangle\right| \le \frac{2\|\varphi\|_\infty}{N} K$$

and $\mathbb{E}(K) \le Ns < Nr$, we deduce by Markov's inequality that

$$\begin{aligned}\mathbb{P}\left(\left|\langle \nu^N_{(\tau+s)\wedge T}, \varphi\rangle - \langle \nu^N_\tau, \varphi\rangle\right| > \varepsilon\right) &\le \mathbb{P}\left(\frac{2\|\varphi\|_\infty}{N} K > \varepsilon\right)\\ &\le \frac{2\|\varphi\|_\infty \mathbb{E}(K)}{\varepsilon N} < \frac{2\|\varphi\|_\infty r}{\varepsilon}\end{aligned}$$

and (ii) follows. □

3. Proof of the results for the finite system. Now we turn to the properties of the finite system. Recall the explicit construction of η_t^N we gave before the proof of Lemma 2.4: given an i.i.d. family $(U_i^N)_{i\ge 1}$ with uniform distribution on $\{1, \dots, N\}$, an i.i.d. family $(R_i)_{i\ge 1}$ with distribution ρ and the jump times $(T_i^N)_{i\ge 1}$ of a Poisson process with rate N, we construct η_t^N by letting $\eta^N_{T_i^N}(U_i^N)$ branch at time T_i^N, using R_i for the displacement, erasing the leftmost particle, and then relabeling the particles to keep the ordering.

This construction allows us to give a monotone coupling for two copies of the process. As in the proof of Proposition 2.2, for $\mu, \nu \in \mathcal{M}$ we will say that $\mu \preceq \nu$ whenever $\mu([x, \infty)) \le \nu([x, \infty))$ for all $x \in \mathbb{R}$. Observe that if $\mu = \sum_{i=1}^{N_1} \delta_{x_i}$ with $x_1 \ge \cdots \ge x_{N_1}$ and $\nu = \sum_{i=1}^{N_2} \delta_{y_i}$ with $y_1 \ge \cdots \ge y_{N_2}$ then $\mu \preceq \nu$ if and only if $N_1 \le N_2$ and $x_i \le y_i$ for $i = 1, \dots, N_1$. It is easy to check that if $\eta_t^{N_1}$ and $\xi_t^{N_2}$ are two copies of our process (note that we allow them to have different total number of particles) with $\eta_0^{N_1} \preceq \xi_0^{N_2}$ (i.e., in the sense that $\sum_{i=1}^{N_1} \delta_{\eta_0^{N_1}(i)} \preceq \sum_{i=1}^{N_2} \delta_{\xi_0^{N_2}(i)}$), then if we use the same branching times and displacements and the same uniform variables for the particles in $\eta_t^{N_1}$ and the leftmost N_1 particles in $\eta_t^{N_2}$, then we have $\eta_t^{N_1} \preceq \xi_t^{N_2}$ for all $t \ge 0$.

For most of the proof of Theorem 2 it will more convenient to work with the discrete time version of our process η_n^N which is defined as follows: at each time step, choose one particle uniformly at random, and then branch that particle and remove the leftmost particle among the $N + 1$. The variables U_i^N and R_i can be used to decide which particle to branch and where to send its offspring at each time step.

PROOF OF THEOREM 2(a). We will first prove that each of the two limits exists with probability 1 and in L^1 and that the limits are nonrandom. We will do this for the discrete time process and leave to the reader the (easy) extension to the continuous time case. We borrow the proof from that of Proposition 2 in Bérard and Gouéré (2010). Since it is simple we include it for convenience. We observe that, as in the cited proof, by translation invariance and the monotonicity property of the coupling discussed above, it is enough to prove the result when all particles start at the origin.

The result is a consequence of the subadditive ergodic theorem. To see why, suppose we run the process up to time k, restart it with all N particles at $\max \eta_k^N$, and then run it for an extra l units of time. Then the resulting configuration will dominate the configuration that we would get by running the process for time $k+l$. To apply the subadditive ergodic theorem we will need to make this precise by introducing an appropriate coupling.

Consider the variables U_i^N and R_i used to construct η_n^N. For each $k \geq 0$ let $(\eta_{k,n}^N)_{n\geq 0}$ be a copy of our process, started at η_0, constructed as follows: if $\eta_{k,n}^N$ is given then we let $\eta_{k,n+1}^N$ be specified by adding a particle at $\eta_{k,n}^N(U_{n+k}^N) + R_{n+k}$ and then removing the leftmost particle. That is, the index n in $\eta_{k,n}^N$ corresponds to time while the index k indicates that the kth copy of the process $(\eta_{k,n}^N)_{n\geq 0}$ uses the random the variables U_i^N and R_i starting from the kth one. With this definition and in view of the preceding paragraph it is not difficult to see that for any $k, l \geq 0$,

$$\max \eta_{0,k+l}^N \leq \max \eta_{0,k}^N + \max \eta_{k,l}^N.$$

Moreover, for any given $d \geq 1$ the family $(\eta_{dm,d}^N)_{m\geq 1}$ is i.i.d., because to compute $\eta_{dm,d}^N$ we only need to use the variables U_i^N and R_i for $i = dm, \dots, dm+d-1$. Also observe that the distribution of $(\eta_{k,n}^N)_{n\geq 0}$ does not depend on k. Now for $0 \leq k \leq n$ define $\xi_{k,n}^N = \eta_{k,n-k}^N$. Then using the above facts we see that

$$\max \xi_{0,k+l}^N \leq \max \xi_{0,k}^N + \max \xi_{k,k+l}^N,$$

the family $(\max \xi_{dm,d(m+1)}^N)_{m\geq 1}$ is i.i.d. for any $d \geq 1$ and the distribution of the sequence $(\max \xi_{k,n+k}^N)_{n\geq 0}$ does not depend on k. It is not hard to check that $\max \xi_{k,n}^N$ satisfies the rest of the hypotheses of the subadditive ergodic theorem [see Theorem 6.6.1 in Durrett (2004)] and thus $\lim_{n\to\infty} \max \xi_{0,n}^N/n$ exists almost surely and in L^1, and moreover the limit is nonrandom. Since $(\xi_{0,n}^N)_{n\geq 0}$ has the same distribution as $(\eta_n^N)_{n\geq 0}$, the same holds for $\lim_{n\to\infty} \max \eta_n^N/n$.

The above proof can be straightforwardly adapted to obtain the existence of the limit for $\min \eta_t^N/t$. To show that the two limits are equal it is enough to prove that $(\max \eta_t^N - \min \eta_t^N)/t \to 0$ in probability as $t \to \infty$. Observe that if we follow the genealogy of the particle at $\max \eta_t^N$ back in time and go back in time N generations then we will necessarily reach a particle that is not in η_t^N (at time t), and that is thus to the left of $\min \eta_t^N$. If we call X_t the position of this particle and let N_t be the number of branchings in the system up to time t, then clearly $\max \eta_t^N - X_t \leq$

$\sum_{i=N_t-N}^{N_t}|R_i|$. Thus for any $\varepsilon > 0$ we have that

$$\begin{aligned}\mathbb{P}\left(\left|\frac{\max\eta_t^N - \min\eta_t^N}{t}\right| > \varepsilon\right) \\ &\leq \mathbb{P}\big(|R_i| > \sqrt{t} \text{ for some } N_t - N \leq i \leq N_t\big) + \mathbf{1}_{N\sqrt{t}/t>\varepsilon} \\ &= 1 - \mathbb{P}\big(|R_1| \leq \sqrt{t}\big)^N + \mathbf{1}_{N/\sqrt{t}>\varepsilon} \xrightarrow[t\to\infty]{} 0.\end{aligned}$$

The monotone coupling introduced above allows to deduce that a_N is nondecreasing. On the other hand, in the case $N = 1$ we have that $\eta_t^1(1)$ is simply a random walk jumping at rate 1 whose jump distribution is that of $R_1 \vee 0$. Therefore $\mathbb{E}(\eta_t^1(1)) = bt$ with $b = \int_0^\infty x\rho(x)\,dx > 0$, and thus $a_N \geq a_1 = b > 0$ for all $N \geq 1$. $\square$

To prove parts (b) and (c) of Theorem 2 we will work with the discrete time version of the shifted process: $\Delta_n^N = (\Delta_n^N(1), \dots, \Delta_n^N(N))$ with

$$\Delta_n^N(j) = \eta_n^N(j) - \eta_n^N(N).$$

PROPOSITION 3.1. *Δ_n^N is a positive recurrent Harris chain.*

PROOF. Following Athreya and Ney (1978), in order to show that Δ_n^N is Harris recurrent we need to show that there is a set $A \subseteq \mathcal{X}_N$ such that:

(i) $\mathbb{P}^\xi(\tau_A < \infty) = 1$ for all $\xi \in \mathcal{X}_N$, where $\tau_A = \inf\{n \geq 0 : \Delta_n^N \in A\}$.
(ii) There exists a probability measure q on A, a $\lambda > 0$ and a $k \in \mathbb{N}$ so that $\mathbb{P}^\xi(\Delta_k^N \in B) \geq \lambda q(B)$ for all $\xi \in A$ and all $B \subseteq A$.

To achieve this, choose some $L > 0$ so that $\delta = \rho((0, L)) > 0$ and let

$$A = \{\xi \in \mathcal{X}_N : \xi(i) - \xi(i+1) \in (0, L) \text{ for } i = 1, \dots, N-1 \text{ and } \xi(N) = 0\}.$$

Then for any initial condition $\Delta_0^N \in \mathcal{X}_N$ we can get to A in $N - 1$ steps via the following path: at time 1 we choose to branch the rightmost particle [the one at $\Delta_0^N(1)$, which happens with probability N^{-1}] and send the newborn particle to a location $x_1 \in (\Delta_0^N(1), \Delta_0^N(1) + L)$ (which happens with probability at least δ). Next we branch the particle at x_1 and send the newborn particle to a location $x_2 \in (x_1, x_1 + L)$ (which happens with probability at least δ/N). If we continue this for $N - 1$ steps we will end up with a configuration in A, and thus

$$\mathbb{P}^\xi(\Delta_{N-1}^N \in A) \geq \left(\frac{\delta}{N}\right)^{N-1}. \tag{3.1}$$

The bound is independent of the initial condition ξ, so by the Borel–Cantelli lemma it follows that (i) holds. Moreover, if $B \subseteq A$ is of the form $B = \{\xi \in$

$\mathcal{X}_N : \xi(i) - \xi(i+1) \in B_i \subseteq (0, L)$ for $i = 1, \dots, N-1$ and $\xi(N) = 0\}$, then the preceding argument implies that

$$\mathbb{P}^{\xi}(\Delta^N_{N-1} \in B) \geq \frac{\rho(B_1)\cdots\rho(B_{N-1})}{N^{N-1}},$$

so by taking $\lambda = N^{-N+1}$ and q to be the normalized Lebesgue measure on the first $N-1$ coordinates of the configurations in A, we deduce that (ii) also holds.

To check that Δ^N_n is actually positive recurrent it is enough to check that $\sup_{\xi \in \mathcal{X}_N} \mathbb{E}^{\xi}(\tau_A) < \infty$. This follows from (3.1) and the strong Markov property by writing

$$\begin{aligned}
\mathbb{E}^{\xi}(\tau_A) &= \sum_{n \geq 1} \mathbb{P}^{\xi}(\tau_A \geq n) \leq N - 1 + \sum_{i \geq 1} \sum_{n=iN}^{(i+1)N-1} \mathbb{P}^{\xi}(\tau_A \geq n) \\
&\leq N - 1 + (N-1) \sum_{i \geq 1} \left[1 - \left(\frac{\delta}{N}\right)^{N-1}\right]^i = (N-1)\left(\frac{N}{\delta}\right)^{N-1} < \infty
\end{aligned}$$

for any $\xi \in \mathcal{X}_N$. $\square$

PROOF OF THEOREM 2(b). The result now follows from Proposition 3.1. The fact that Δ^N_n is positive recurrent implies that the invariant measure whose existence is assured by the Harris recurrence is finite. The absolute continuity of μ_N is a direct consequence of Theorem 2(c), which we prove below, together with the fact that if the initial condition for Δ^N_t is absolutely continuous, then so is the distribution of the process at all times. $\square$

PROOF OF THEOREM 2(c). Let $A \subseteq \mathcal{X}_N$ and $k = N - 1$ be the objects which we found satisfy (i) and (ii) in the proof of Proposition 3.1. It is enough to prove the result along the k subsequences of the form $(\Delta^N_{km+j})_{m \geq 0}$ with $0 \leq j < k$. Moreover, using the Markov property at time j we see that it is enough to prove the result along the subsequence $(\Delta^N_{km})_{m \geq 0}$, which is an aperiodic recurrent Harris chain. The result for this subsequence follows from Theorem 4.1(ii) in Athreya and Ney (1978) as long as we have that $\sup_{\xi} \mathbb{P}^{\xi}(\tau_A > t) < 1$ for some $t > 0$, where τ_A is the hitting time of A. This follows easily from the estimate in (3.1) (which is uniform in ξ). $\square$

4. Proof of Theorem 4. Recall that in this part we are assuming that ρ (and hence its tail distribution R) has exponential decay and, consequently, that the moment generating function of ρ, $\phi(\theta) = \int_{-\infty}^{\infty} e^{\theta x} \rho(x)\, dx$, is finite for $\theta \in (-\Theta, \Theta)$ [see (1.2) and (1.7)]. Before getting started with the proof of Theorem 4 we need to prove the claim we implicitly made in (1.6).

LEMMA 4.1.

$$\min_{\lambda\in(0,\Theta)} \frac{\phi(\lambda)}{\lambda} = a,$$

where a is the asymptotic speed defined in Theorem 3. Moreover, letting $\lambda^* \in (0,\Theta)$ be the number such that $\phi(\lambda^*)/\lambda^* = a$, we have that $\phi'(\lambda^*) = a$, $\phi(\lambda)/\lambda$ is strictly convex on $(0,\Theta)$ and the sign of $\phi'(\lambda) - \phi(\lambda)/\lambda$ equals that of $\lambda - \lambda^*$.

PROOF. Define $c(\lambda) = \phi(\lambda)/\lambda$ for $\lambda \in (0,\lambda^*)$. A little calculus shows that $c(\lambda)$ is strictly convex:

$$c''(\lambda) = \frac{\phi''(\lambda)}{\lambda} - \frac{2\phi'(\lambda)}{\lambda^2} + \frac{2\phi(\lambda)}{\lambda^3} = \frac{1}{\lambda^3}\int_{-\infty}^{\infty} [(\lambda x - 1)^2 + 1] e^{\lambda x}\rho(x)\,dx > 0.$$

It is clear that $c(\lambda) \to \infty$ as $\lambda \to 0$. On the other hand, (1.4) implies $c(\lambda) \to \infty$ as $\lambda \to \Theta-$ as well. Thus the minimum of c is attained at some $\lambda^* \in (0,\Theta)$, and we have $c'(\lambda^*) = 0$, or $\phi'(\lambda^*) = \phi(\lambda^*)/\lambda^*$, which will give the second claim in the lemma once we show that $c(\lambda^*) = a$.

Recalling the characterization of a given after (1.5), we need to show that

$$\sup_{\theta>0} [\theta c(\lambda^*) - \phi(\theta)] = 0.$$

This is easy: using the definition of c we get

$$\sup_{\theta>0} [\theta c(\lambda^*) - \phi(\theta)] \geq \lambda^* c(\lambda^*) - \phi(\lambda^*) = 0,$$

while for all $\theta > 0$

$$\theta c(\lambda^*) - \phi(\theta) \leq \theta c(\theta) - \phi(\theta) = 0.$$

Finally, to get the last claim in the lemma recall that $c'(\lambda^*) = 0$ and c is convex, so

$$\phi'(\lambda) - \frac{\phi(\lambda)}{\lambda} = \lambda c'(\lambda)$$

is negative for $\lambda \in (0,\lambda^*)$ and positive for $\lambda \in (\lambda^*,\Theta)$. $\square$

Recall [see (1.8)] that

$$k(x) = \frac{\lambda}{\phi(\lambda)} e^{\lambda x} R(x).$$

As we explained in Section 1.3 the proof of Theorem 4 will depend on looking for positive solutions to the equation (1.9). We will actually consider a slightly more general equation:

$$U(x) = \int_0^{\infty} U(y)k(x-y)\,dy \qquad \forall x \geq 0, \tag{4.1}$$

where we look for a nondecreasing solution U, continuous except at the origin, with $U(x) = 0$ for all $x < 0$. In Spitzer's terminology, a solution U with these properties is a P^*-solution of (4.1). When $\lim_{x\to\infty} U(x) = 1$ we call U a P-solution, and think of it as the distribution function of a nonnegative random variable. The following summarizes the two results of Spitzer that we will need.

THEOREM 4.2 [Theorems 2 and 4 in Spitzer (1957)].

(a) *If* $\int_{-\infty}^{\infty} xk(x)\,dx \le 0$ *then there is a unique* (*up to multiplicative constant*) P^*-*solution of* (4.1).
(b) *If* $\int_{-\infty}^{\infty} xk(x)\,dx < 0$ *then there is a unique* P-*solution of* (4.1) *which can be obtained as the limit* $U(x) = \lim_{n\to\infty} U_n(x)$ *of the iterative procedure defined by* $U_{n+1}(x) = \int_0^\infty U_n(y)k(x-y)\,dy$ *starting with an arbitrary continuous* U_0 *corresponding to the distribution function of a nonnegative random variable.*
(c) *If* $\int_{-\infty}^{\infty} xk(x)\,dx \ge 0$ *then* (4.1) *has no* P-*solution.*

Repeating the arguments we used to show that $F(t,x)$ had a density [see (2.21)] we see that if U is a P^*-solution of (4.1) then there is a nonnegative function u such that $U(x) = \int_0^x u(y)\,dy$. Again repeating previous arguments (see the first part of the proof of Theorem 1), we deduce that u satisfies (1.9), while obviously $u(x) = 0$ for $x \le 0$. u is continuous except possibly at the origin by the dominated convergence theorem thanks to the fact that k is continuous. Multiplying $u(x)$ by $e^{\lambda x}$ will allow us to obtain a solution for (TW) with the desired properties.

PROOF OF THEOREM 4. Let $c \ge a$ and take $\lambda \in (0, \lambda^*]$ such that $\phi(\lambda)/\lambda = c$ as above. The uniqueness of the solutions of (TW) in this case follows from Theorem 4.2(a). In fact, if w_1 and w_2 are two solutions of (TW) then $u_i(x) = e^{\lambda x} w_i(x)$ solves (1.9) for $i = 1, 2$, and thus the functions $U_i(x) = \int_0^x u_i(y)\,dy$ are P^*-solutions of (4.1), and they are continuous because the u_i are locally integrable. Hence $U_1(x) = AU_2(x)$ for all $x \in \mathbb{R}$ and some $A > 0$, which implies that $w_1(x) = Aw_2(x)$ for all $x \in \mathbb{R}$. Integrating this relation we get $A = 1$ and uniqueness follows.

To show existence we start by integrating by parts to obtain

$$\int_{-\infty}^{\infty} xk(x)\,dx = \frac{\lambda}{\phi(\lambda)} \int_{-\infty}^{\infty} xe^{\lambda x} R(x)\,dx = \frac{1}{\phi(\lambda)} \int_{-\infty}^{\infty} \left(x - \frac{1}{\lambda}\right) e^{\lambda x} \rho(x)\,dx.$$

Thus the sign of $\int_{-\infty}^{\infty} xk(x)\,dx$ is the same as that of $\phi'(\lambda) - \phi(\lambda)/\lambda$. This last quantity is strictly negative for $\lambda \in (0, \lambda^*)$ and vanishes for $\lambda = \lambda^*$ by Lemma 4.1.

If $c > a$, then $\lambda < \lambda^*$ and thus Theorem 4.2(b) provides us with a P-solution of (4.1) to which, by the discussion preceding this proof, we can associate a continuous (except at the origin) function u satisfying (1.9) and corresponding to the density of a nonnegative random variable. Now let $A = \int_0^\infty e^{-\lambda x} u(x)\,dx$ which is clearly finite (actually $A < 1$). Define $w(x) = A^{-1} e^{-\lambda x} u(x)$. Then w is the density of a nonnegative random variable and it is easy to check that it satisfies (TW1) with $c = \phi(\lambda)/\lambda$. The random variable w obviously satisfies (TW2), (TW3) and (TW4) as well, it is continuous except possibly at the origin because so is u and, by definition, $\int_0^\infty e^{\lambda x} w(x)\,dx = A^{-1} < \infty$. The last thing left to show in this case is that w is differentiable (except at the origin), but this follows easily from writing, for $x \ge 0$,

$$w(x) = \frac{1}{c}\int_0^\infty w(y)\int_{x-y}^\infty \rho(z)\,dz\,dy = \frac{1}{c}\int_x^\infty \int_0^\infty w(y)\rho(z-y)\,dy\,dz,$$

and using the fact that the integrand above is continuous. This establishes (a) in the case $c > a$.

The equality $\int_0^\infty e^{\lambda x} w(x)\,dx = A^{-1} < \infty$ obtained above gives the first claim in (b). To prove the second claim in (b) we may obviously assume that $\widetilde{\lambda} \in (\lambda, \lambda^*)$. Let $W(x) = \int_x^\infty w(y)\,dy$. It is not hard to check that W satisfies

$$W(x) = \frac{1}{c}\int_{-\infty}^\infty W(y)R(x-y)\,dy \qquad \forall x \ge 0.$$

Then if $A = \sup_{x\ge 0} e^{\widetilde{\lambda} x} W(x) < \infty$ we have that

$$\begin{aligned} W(x)e^{\widetilde{\lambda}x} &= \frac{1}{c}\int_{-\infty}^\infty W(y)e^{\widetilde{\lambda}y}e^{\widetilde{\lambda}(x-y)}R(x-y)\,dy \\ &\le \frac{A}{c}\int_{-\infty}^\infty e^{\widetilde{\lambda}(x-y)}R(x-y)\,dy = \frac{A}{c}\frac{\phi(\widetilde{\lambda})}{\widetilde{\lambda}}. \end{aligned}$$

Taking supremum in $x \ge 0$ and recalling that $c = \phi(\lambda)/\lambda$ the above says that

$$A \le \frac{\phi(\widetilde{\lambda})}{\widetilde{\lambda}}\frac{\lambda}{\phi(\lambda)}A,$$

and since $A > 0$ this says that $\phi(\widetilde{\lambda})/\widetilde{\lambda} \ge \phi(\lambda)/\lambda$. But Lemma 4.1 implies exactly the opposite for $\lambda < \widetilde{\lambda} < \lambda^*$. This is a contradiction, and thus $A = \infty$, which finishes the proof of (b).

The case $c = a$ is similar so we will skip some details. Now we have $\lambda = \lambda^*$ and thus $\int_{-\infty}^\infty xk(x)\,dx = 0$. Theorem 4.2(a) provides us now with a P^*-solution of (TW) to which corresponds a function u which is the density of a measure supported on $[0,\infty)$ and which satisfies (1.9). Let $A = \int_0^\infty e^{-\lambda^* x}u(x)\,dx$. We need this quantity to be finite, so that $w(x) = A^{-1}e^{-\lambda^* x}u(x)$ is a continuous probability density. This follows from Theorem 6.2 of Engibaryan (1996), which assures that $U(x) = \int_0^x u(y)\,dy \le Cx$ for some $C > 0$, and integration by parts:

$$\begin{aligned} A &= \int_0^\infty e^{-\lambda^* x}u(x)\,dx = \lim_{x\to\infty} e^{-\lambda^* x}U(x) - U(0) + \lambda^*\int_0^\infty e^{-\lambda^* x}U(x)\,dx \\ &\le C\lambda^*\int_0^\infty e^{-\lambda^* x}x\,dx < \infty. \end{aligned}$$

It is easy again to verify that w satisfies (TW), and its differentiability follows from the same reasons as above. Hence we have established (a) for the case $c = a$. Clearly

$$\int_0^x e^{\lambda^* y} w(y)\,dy = A^{-1} \int_0^x u(y)\,dy = O(x), \tag{4.2}$$

which is the second claim in (c). The first and third claims in (c) follow from two other consequences of the cited result in Engibaryan (1996), namely that $U(x) \to \infty$ as $x \to \infty$ and that U is subadditive. For the third claim use the first of these properties of U and the first equality in (4.2), while for the first one integrate by parts to get

$$\begin{aligned} e^{\lambda^* x} \int_x^\infty w(y)\,dy &= \int_x^\infty e^{-\lambda^*(y-x)} u(y)\,dy = \int_0^\infty e^{-\lambda^* z} u(x+z)\,dz \\ &= \lim_{z\to\infty} e^{-\lambda^* z} U(x+z) - U(x) + \int_0^\infty \lambda^* e^{-\lambda^* z} U(x+z)\,dz \\ &\le -U(x) + \int_0^\infty \lambda^* e^{-\lambda^* z} [U(x) + U(z)]\,dz \\ &\le C \int_0^\infty \lambda^* e^{-\lambda^* z} z\,dz = \frac{C}{\lambda^*}. \end{aligned}$$

We are only left showing (d), that is, that there are no solutions of (TW) when $c < a$. We start by observing that if w were a solution then the above arguments would imply that w is differentiable on the positive axis, and thus if we set $f_0(x) = w(x)$ in (FB) we get $\gamma(t) = ct$. Therefore, to show the nonexistence of solutions for $c < a$ it is enough to show that, given any $\varepsilon > 0$ and any f_0 supported on $[0, \infty)$, there is a $T > 0$ such that the solution $(f(t,x), \gamma(t))$ of (FB) satisfies

$$\gamma(T) > (a - \varepsilon)T. \tag{4.3}$$

Recall the definition of the process $\widehat{\nu}_t$ in the proof of Proposition 2.1, which corresponded to the deterministic measure valued limit of the branching random walk $\widehat{\nu}_t^N$, and observe that we can run this process started with any initial measure (not necessarily an absolutely continuous one). Moreover, (2.4) still holds in this case by Theorem 5.3 of Fournier and Méléard (2004). Consider a copy of $\widehat{\nu}_t$ started with a unit mass at 0 and let $\widehat{F}(t,x) = \widehat{\nu}_t([x,\infty))$. For $T > 0$ we define $\chi_T > 0$ to be such that $\widehat{F}(T, \chi_T) = 1$.

Applying (2.4) to $\varphi(y) = \mathbf{1}_{y\in[x,\infty)}$, we see that $\widehat{F}(t,x)$ satisfies (FB1′) for all $x \in \mathbb{R}$. Now consider a copy of ν_t started at the product measure ν_0 defined by f_0. Then $(F(t,x), \gamma(t))$ satisfies (FB′) and, since γ is strictly increasing, it satisfies (FB1′) for $x = \gamma(T)$ and all $t \in [0, T)$. We deduce that

$$\frac{\partial}{\partial t}\big(F(t,\gamma(T)) - \widehat{F}(t,\gamma(T))\big) = \int_{-\infty}^\infty \big(F(t,y) - \widehat{F}(t,y)\big)\rho\big(\gamma(T) - y\big)\,dy$$

for all $t \in [0, T)$. Since $F(0, x) = \int_x^\infty f_0(y)\,dy$ and $\widehat{F}(0, x) = \mathbf{1}_{x \le 0}$, we have that $F(0, x) \ge \widehat{F}(0, x)$ for all x and thus the above equation implies that

$$F(t, \gamma(T)) \ge \widehat{F}(t, \gamma(T))$$

for all $t \in [0, T)$. Now $\widehat{F}(t, x)$ is clearly continuous in t, while

$$\begin{aligned} F(T, \gamma(T)) - F(t, \gamma(T)) &= F(t, \gamma(t)) - F(t, \gamma(T)) = \int_{\gamma(t)}^{\gamma(T)} f(t, y)\,dy \\ &\le \int_{\gamma(t)}^{\gamma(T)} \widehat{f}(t, y)\,dy \le \int_{\gamma(t)}^{\gamma(T)} \widehat{f}(T, y)\,dy \xrightarrow[t \to T-]{} 0 \end{aligned}$$

thanks to the continuity of γ. Therefore the inequality also holds for $t = T$, which gives $1 = F(T, \gamma(T)) \ge \widehat{F}(T, \gamma(T))$, whence

$$\chi_T \le \gamma(T).$$

To finish the proof of (4.3) we need to show that

$$\chi_T > (a - \varepsilon)T \tag{4.4}$$

for large enough T. Observe that $\widehat{\nu}_t$ (which we recall is started with $\widehat{\nu}_0 = \delta_0$), corresponds to the mean measure of the branching random walk ξ_t^1 (started with just one particle located at the origin). This can be made precise by writing down the formula for the generator of ξ_t^1 and applying it to functions of the form $\xi \mapsto \langle \xi, \varphi \rangle = \sum_{i=1}^{N(\xi)} \varphi(\xi(i))$, where $N(\xi)$ is the number of particles in the branching random walk configuration ξ, to deduce that after taking expectations the resulting equation is the same as (2.4). We leave the details to the reader, and instead only state that the above implies that

$$\widehat{\nu}_T([cT, \infty)) = \mathbb{E}(\xi_T^1([cT, \infty))) \tag{4.5}$$

for every c, where $\xi_T^1([cT, \infty))$ denotes the number of particles in the branching random walk to the right of cT at time T. On the other hand it is well known that

$$\mathbb{E}(\xi_T^1([cT, \infty))) = e^T \mathbb{P}(S_T \ge cT),$$

where S_t is defined as in (1.5) [see, e.g., the third equation in the proof of Proposition I.1.21 in Liggett (1999)], and thus (1.5) implies that

$$\lim_{T \to \infty} \frac{1}{T} \log(\mathbb{E}(\xi_T^1([cT, \infty)))) = \Lambda(c) + 1.$$

Since Λ is strictly decreasing on $[0, \infty)$ and $\Lambda(a) = -1$ we deduce that

$$\lim_{T \to \infty} \frac{1}{T} \log(\mathbb{E}(\xi_T^1([cT, \infty)))) > 0 \qquad \text{for all } 0 \le c < a.$$

This together with (4.5) implies that

$$\widehat{\nu}_T\big([(a - \varepsilon)T, \infty)\big) > 1$$

for large enough T. Therefore (4.4) holds and the proof is complete. $\square$

Acknowledgments. The authors would like to thank Nathanaël Berestycki and Lee Zhuo Zhao for pointing out a mistake in an earlier version of one of the proofs, and an anonymous referee for valuable comments and suggestions.

REFERENCES

ATHREYA, K. B. and NEY, P. (1978). A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* **245** 493–501. MR0511425

BÉRARD, J. and GOUÉRÉ, J.-B. (2010). Brunet–Derrida behavior of branching-selection particle systems on the line. *Comm. Math. Phys.* **298** 323–342. MR2669438

BIGGINS, J. D. (1977). Chernoff's theorem in the branching random walk. *J. Appl. Probab.* **14** 630–636. MR0464415

BRUNET, E. and DERRIDA, B. (1997). Shift in the velocity of a front due to a cutoff. *Phys. Rev. E* (3) **56** 2597–2604. MR1473413

BRUNET, É., DERRIDA, B., MUELLER, A. H. and MUNIER, S. (2007). Effect of selection on ancestry: An exactly soluble case and its phenomenological generalization. *Phys. Rev. E* (3) **76** 041104, 20. MR2365627

CHAYES, L. and SWINDLE, G. (1996). Hydrodynamic limits for one-dimensional particle systems with moving boundaries. *Ann. Probab.* **24** 559–598. MR1404521

DURRETT, R. (2004). *Probability: Theory and Examples*, 3rd ed. Duxbury Press, Belmont, CA.

DURRETT, R. and MAYBERRY, J. (2010). Evolution in predator-prey systems. *Stochastic Process. Appl.* **120** 1364–1392. MR2639750

ENGIBARYAN, N. B. (1996). Convolution equations containing singular probability distributions. *Izv. Ross. Akad. Nauk Ser. Mat.* **60** 21–48. MR1399417

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications. Vol. II*, 2nd ed. Wiley, New York. MR0270403

FOURNIER, N. and MÉLÉARD, S. (2004). A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.* **14** 1880–1919. MR2099656

GRAVNER, J. and QUASTEL, J. (2000). Internal DLA and the Stefan problem. *Ann. Probab.* **28** 1528–1562. MR1813833

JOFFE, A. and MÉTIVIER, M. (1986). Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. in Appl. Probab.* **18** 20–65. MR0827331

KREĬN, M. G. (1962). Integral equations on a half-line with kernel depending upon the difference of the arguments. *Trans. Amer. Math. Soc.* (2) **22** 163–288. Original (Russian): *Uspekhi Mat. Nauk.* **13** (1958) 3–120. MR0102721

LANDIM, C., OLLA, S. and VOLCHAN, S. B. (1998). Driven tracer particle in one-dimensional symmetric simple exclusion. *Comm. Math. Phys.* **192** 287–307. MR1617558

LIGGETT, T. M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **324**. Springer, Berlin. MR1717346

MEIRMANOV, A. M. (1992). *The Stefan Problem. de Gruyter Expositions in Mathematics* **3**. de Gruyter, Berlin. MR1154310

PALEY, R. E. A. C. and WIENER, N. (1987). *Fourier Transforms in the Complex Domain. American Mathematical Society Colloquium Publications* **19**. Amer. Math. Soc., Providence, RI. MR1451142

ROELLY-COPPOLETTA, S. (1986). A criterion of convergence of measure-valued processes: Application to measure branching processes. *Stochastics* **17** 43–65. MR0878553

SPITZER, F. (1957). The Wiener–Hopf equation whose kernel is a probability density. *Duke Math. J.* **24** 327–343. MR0090917

WIENER, N. and HOPF, E. (1931). *Üeber Eine Klasse Singulärer Integralgleichungen*. Sitzungber. Akad. Wiss, Berlin 696706.

DEPARTMENT OF MATHEMATICS
DUKE UNIVERSITY
BOX 90320
DURHAM, NORTH CAROLINA 27708-0320
USA
E-MAIL: rtd@math.duke.edu

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
40 ST. GEORGE ST.
TORONTO, ONTARIO M5S 2E4
CANADA
E-MAIL: dremenik@math.toronto.edu