OPTIMISATION-BASED REPRESENTATIONS FOR BRANCHING PROCESSES

DAVID P. DRIVER, MICHAEL R. TEHRANCHI
UNIVERSITY OF CAMBRIDGE

Abstract. It is shown that a certain functional of a branching process has representations in terms of both a maximisation problem and a minimisation problem. A consequence of these representation is that upper and lower bounds on the functional can be found easily, yielding a non-asymptotic Trotter product formula. As an application, the speed of the right-most particle of a branching Lévy process is calculated.

1. Introduction

Consider a branching process $\{X_t^i : i \in I_t,\ t \geq 0\}$ constructed as follows. Initially, there is one particle sitting at a point x_0 in a Polish space $\mathcal{X}$. The position of the particle then evolves according to the law of a given right-continuous strong Markov process X started from $X_0 = x_0$. At time $T > 0$, the initial particle is killed and replaced with N particles, where both T and N are random. Each of these new particles then move and branch as independent copies of the initial particle, except that each new particle now starts from the final position X_T of the initial particle. We assume that the conditional law of the first branching time T given the Markov process X is

$$\mathbb{P}(T > t | X) = e^{-\int_0^t \lambda(X_s)ds} \text{ for all } t \geq 0,$$

for a given non-negative measurable function λ. We also assume that the conditional distribution of the number of offspring N given X and T only depends on X_T, the location of the initial particle at the time of branching. Letting I_t be the collection of particles alive at time t, a construction of such a branching process $\{X_t^i : i \in I_t,\ t \geq 0\}$ can be found in the paper of Ikeda–Nagasawa–Watanabe [11]. In what follows, we will let X_t denote the position at time t of the initial particle if it were allowed to continue living after the branching event.

We will assume that the branching rate $\lambda(x)$ and the mean number of new offspring $\mathbb{E}(N|X_T = x)$ per branching event are bounded functions of $x \in \mathcal{X}$. This is a sufficient condition that the branching process does not explode in finite time, so that $\mathbb{P}(|I_t| < \infty) = 1$ for all $t \geq 0$. See, for instance, the book of Athreya & Ney [1, Theorem III.2.1]

Our main result is the following:

Theorem 1.1. *Let $\mathcal{F}$ be the filtrations generated by X. Let $\mathcal{Z}$ be the set of bounded adapted processes, let $\mathcal{Z}^\circ$ be the set of bounded anticipative processes, and let $\mathcal{M}$ the set of non-negative martingales.*

Date: April 22, 2020.
Keywords and phrases: branching process, FKPP equation, stochastic control, front propagation .
Mathematics Subject Classification 2010: 60J80, 93E20, 35B40 .

Given a measurable function $f : \mathcal{X} \to [0,1]$ and $t \geq 0$, let

$$u = \mathbb{E}\left[\prod_{i \in I_t} f(X_t^i)\right]$$

$$M = \max_{Z \in \mathcal{Z}} \mathbb{E}\left[e^{\int_0^t Z_s ds} f(X_t) - \int_0^t e^{\int_0^s Z_r dr} h(X_s, Z_s + \lambda(X_s)) ds\right]$$

$$m = \min_{\zeta \in \mathcal{M}} \mathbb{E}\left[\text{ess sup}_{z \in \mathcal{Z}^\circ} \left\{ e^{\int_0^t z_s ds} f(X_t) - \int_0^t e^{\int_0^s z_r dr} \left(z_s + [h(X_s, z_s + \lambda(X_s)) - z_s] \frac{\zeta_t}{\zeta_s} \right) ds \right\}\right]$$

where the function h is defined by

$$h(x,z) = \max_{0 \leq \eta \leq 1} \{\eta z - \lambda(x)\mathbb{E}(\eta^N | X_T = x)\} \text{ for all } (x,z) \in \mathcal{X} \times \mathbb{R}.$$

where we set $\zeta_t/\zeta_s = 1$ on the event $\{\zeta_s = 0\}$. Then

$$u = M = m.$$

Remark 1.2. We are using the convention that all real processes have measurable sample paths, so that the pathwise integrals appearing in the statement of Theorem 1.1 are well-defined.

Remark 1.3. The proof will show that it is possible to replace the function h appearing in the statement of Theorem 1.1 with a function $\tilde{h}$ so long as $\tilde{h}(x,z) \geq h(x,z)$ for all (x,z) and $\tilde{h}(x,z) = h(x,z)$ when $\lambda(x)\mathbb{P}(N = 1|X_T = x) \leq z \leq \lambda(x)\mathbb{E}(N|X_T = x)$. For instance, we may take

$$\tilde{h}(x,z) = \max_{\eta \geq 0}\{\eta z - \lambda(x)\mathbb{E}(\eta^N | X_T = x)\}$$

The full proof of this result appears in Section 2. To put the above optimisation-based representations into context, we jump ahead a bit. The rough idea behind the equality $u = M$ appearing in Theorem 1.1 is that the value function of the stochastic optimal control problem defining M should satisfy the Bellman equation of the problem. However, we have chosen the data of the control problem in such a way that the associated Bellman equation is, essentially, the S-equation (in the terminology of Ikeda–Nagasawa–Watanabe [12, equation (4)]) of the branching process. Although we have not found this done explicitly in other papers, we acknowledge that this may not be surprising in this respect.

In contrast, the dual minimisation problem defining m is not in the form of a standard stochastic control problem, and so the usual dynamic programming arguments do not apply. In particular, there is no conventional Bellman equation to this minimisation problem. Our formulation of the dual problem is inspired by the pathwise stochastic control approach of Rogers [14]. The general formulation is a bit cumbersome, involving both a maximisation over anticipative processes and a minimisation over martingales. However, in the special case of dyadic branching, when the number of offspring is the constant $N = 2$, the pathwise maximisation problem can be solved explicitly, yielding the following corollary:

Corollary 1.4. *With the notation of Theorem 1.1, suppose $N = 2$ almost surely. Then*

$$u = 1 - \max_{\zeta \in \mathcal{M}} \mathbb{E}\left[\frac{(1 - f(X_t))\zeta_t e^{\int_0^t \lambda(X_s) ds}}{\zeta_t + (1 - f(X_t)) \int_0^t \zeta_s \lambda(X_s) e^{\int_s^t \lambda(X_r) dr} ds}\right]$$

A proof of this fact will be given in section 2. We find it somewhat surprising (maybe even mysterious) that the maximisation problem appearing in Corollary 1.4 is related to a dyadic branching process.

A consequence of the connection between branching processes and the various optimisation problems is that lower and upper bounds of certain functionals of the branching process can be derived immediately, simply by evaluating the objective functions of the optimisation problems at feasible controls. In particular, this technique can be used in principle to derive asymptotic estimates on the behaviour of the branching process.

As an illustrative application, we consider the branching process where X is a real-valued Lévy process, and where the rate of branching λ is a positive constant and the distribution of the number of offspring N is independent of the position of the particles. Let K be the cumulant generating function of the underlying Lévy process, defined by

$$\mathbb{E}_x[e^{\theta X_t}] = e^{\theta x + tK(\theta)} \text{ for all } t \geq 0.$$

Suppose that K is finite in a neighbourhood of $\theta = 0$. Recall that by the Lévy–Khintchine formula we have

$$K(\theta) = b\theta + \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}\setminus\{0\}} [e^{\theta y} - 1 - \theta y \mathbb{1}_{\{|y|\leq 1\}}]\nu(dy)$$

for some constants b, σ and measure ν, where we are supposing that $\int (e^{\theta y} \wedge y^2)\nu(dy) < \infty$ for all θ in some neighbourhood of $\theta = 0$.

Theorem 1.5. *Let $\mu = \mathbb{E}(N) - 1$ be the mean net number of new particles created at a branching event and suppose $\mu > 0$. Conditional on the event $\{I_t \neq \emptyset$ for all $t \geq 0\}$ that the branching Lévy process never becomes extinct, we have*

$$\frac{1}{t} \max_{i \in I_t} X_t^i \to \inf_{\theta > 0} \frac{K(\theta) + \lambda\mu}{\theta} \text{ in probability}$$

Remark 1.6. The condition $\mu > 0$ is necessary and sufficient for supercriticality of the branching process, that is $\mathbb{P}(I_t \neq \emptyset$ for all $t \geq 0) > 0$. See, for instance, the book of Athreya & Ney [1, Theorem III.4.1].

Remark 1.7. Consider the case where the Lévy process X is a standard Brownian motion, so that $K(\theta) = \frac{1}{2}\theta^2$. Then Theorem 1.5 says that, conditional on the branching process not becoming extinct, the speed of right-most particle is

$$\inf_{\theta > 0} \left(\frac{\theta}{2} + \frac{\lambda\mu}{\theta} \right) = \sqrt{2\lambda\mu}.$$

Remark 1.8. Versions of Theorem 1.5 are known, see for instance Biggins [3, Corollary 2], but our precise formulation seems new and requires fewer assumptions. More importantly, our proof will be rather different, using estimates derived from Theorem 1.1, rather than renewal theory.

The remainder of the paper is structured as follows. Section 2 contains the proof of Theorem 1.1. The key ingredient is a more general representation result given by Theorem 2.1. Section 3 gives the main take-away implications of Theorem 2.1: easy to apply bounds on the solution to certain reaction-diffusion-type equations. Section 4 contains the proof of Theorem 1.5 which finds the speed of the right-most particle of a branching Lévy process.

2. Proof of Theorem 1.1

In this section, we prove Theorem 1.1. We first prove a more general result. As in the introduction, let X be a right-continuous strong Markov process valued in a Polish space $\mathcal{X}$. As in Theorem 1.1, we let $\mathcal{Z}$, $\mathcal{Z}^\circ$ and $\mathcal{M}$ be the set of bounded adapted processes, bounded anticipative processes and non-negative martingales, respectively.

Theorem 2.1. *Let $\phi : \mathcal{X} \times \mathbb{R} \to \mathbb{R}$ be measurable and such that $\phi(x, \cdot)$ is concave and differentiable with a derivative bounded uniformly in $x \in \mathcal{X}$. Suppose the bounded function $v : \mathbb{R}_+ \times \mathcal{X} \to \mathbb{R}$ satisfies the integral equation*

$$v(t,x) = \mathbb{E}_x\left[v(0,X_t) + \int_0^t \phi(X_s, v(t-s,X_s))ds\right]$$

for all (t,x). Then

$$v(t,x) = \min_{Z\in\mathcal{Z}} \mathbb{E}_x\left[e^{\int_0^t Z_s ds} v(0,X_t) - \int_0^t e^{\int_0^s Z_r dr}\psi(X_s, Z_s)ds\right]$$

where

$$\psi(x,z) = \inf_{\eta\in\mathbb{R}}\{\eta z - \phi(x,\eta)\} \text{ for all } (x,z) \in \mathcal{X}\times\mathbb{R}.$$

For fixed (t,x), a minimiser is given by the adapted control

$$Z^*_s = \frac{\partial\phi}{\partial v}(X_s, v(t-s,X_s)).$$

If $v(t,x) \geq 0$ for all (t,x) and $\phi(x,0) = 0$ for all x, then

$$v(t,x) = \max_{\zeta\in\mathcal{M}} \mathbb{E}_x\left[\text{ess inf}_{z\in\mathcal{Z}^\circ}\left\{e^{\int_0^t z_s ds} v(0,X_t) - \int_0^t e^{\int_0^s z_r dr}\psi(X_s, Z_s)ds\right\}\right]$$

For fixed (t,x), a maximiser is given by the non-negative martingale

$$\zeta^*_s = v(t-s,X_s)e^{\int_0^s \theta(X_r, v(t-r,X_r))dr}$$

where $\theta(x,\eta) = \phi(x,\eta)/\eta$ for all $x \in \mathcal{X}, \eta > 0$. For the martingale ζ^*, the essential infimum is attained for the control $z^* = Z^*$.

Remark 2.2. Formally, the differential form of the integral equation appearing Theorem 2.1 is

$$\frac{\partial v}{\partial t} = \mathcal{L}v + \phi(x,v)$$

where $\mathcal{L}$ is the infinitesimal generator of the Markov process X. The hypothesis can be reworded to say that v is a mild solution of the above differential equation.

We also note in passing that when X is a diffusion process in finite-dimensional Euclidean space, so that $\mathcal{L}$ is a second order differential operator, the semi-linear partial differential equation is of the reaction-diffusion type.

Under our assumption that $\phi(x,\cdot)$ is uniformly Lipschitz, one can show by a standard Picard iteration argument that given a bounded initial condition $v(0,\cdot)$ the integral equation has a unique solution v bounded on any bounded time intervals $[0,t]$. See for instance the paper of Cabré & Roquejoffre [5, Section 2.3]. In Theorem 2.1, we take this for granted and simply assume that the solution v exists.

Proof. Fix (t,x) and let

$$M_s = V_s + \int_0^s \Phi_r dr$$

where $V_s = v(t-s, X_s)$ and $\Phi_s = \phi(X_s, V_s)$. Note that $(M_s)_{0\le s\le t}$ is a martingale.

The key observation is that

$$\begin{aligned} e^{\int_0^t Z_s ds} v(0,X_t) - \int_0^t e^{\int_0^s Z_r dr}\Psi_s ds =& M_t + \int_0^t (M_t - M_s) Z_s e^{\int_0^s Z_r dr} ds \\ &+ \int_0^t (V_s Z_s - \Phi_s - \Psi_s) e^{\int_0^s Z_r dr} ds \end{aligned}$$

where $\Psi_s = \psi(X_s, Z_s)$. Note that the two path-wise Lebesgue integrals on the right-hand side are well-defined, though the second one might take the value $-\infty$. Indeed, the integrand in the first integral is Lebesgue integrable almost surely, since by the assumed boundedness of Z there is a constant $c > 0$ such that

$$\mathbb{E}_x\left(\int_0^t |(M_t - M_s) Z_s e^{\int_0^s Z_r dr}| ds \right) \le c\, \mathbb{E}_x(|M_t|) < \infty$$

and

$$\mathbb{E}_x\left(\int_0^t (M_t - M_s) Z_s e^{\int_0^s Z_r dr} ds \right) = 0$$

by Fubini's theorem and the tower property of conditional expectation. The integrand in the second integral is non-positive by the Fenchel–Young inequality:

$$\phi(x,v) + \psi(x,z) \le vz.$$

with equality if

$$z = \frac{\partial \phi}{\partial v}(x,v).$$

Hence

$$\mathbb{E}\left(e^{\int_0^t Z_s ds} v(0,X_t) - \int_0^t e^{\int_0^s Z_r dr} \Psi_s \right) \ge \mathbb{E}(M_t) = v(t,x)$$

with equality if $Z = Z^*$. Note Z^* is bounded, and hence feasible, by the assumption of uniform boundedness of $\partial\phi/\partial v$. This proves that $v(t,x)$ is the value of the minimisation problem.

Now consider the max-min problem. Fix a non-negative martingale ζ and note that by Fubini's theorem and iterating expectations we have

$$\begin{aligned} v(t,x) &= \mathbb{E}\left(e^{\int_0^t Z^*_s ds} v(0,X_t) - \int_0^t e^{\int_0^s Z^*_r dr} \psi(X_s, Z^*_s) ds \right) \\ &= \mathbb{E}\left(e^{\int_0^t Z^*_s ds} v(0,X_t) - \int_0^t e^{\int_0^s Z^*_r dr} \psi(X_s, Z^*_s) \frac{\zeta_t}{\zeta_s} ds \right) \\ &\ge \mathbb{E}\left(\text{ess inf}_z \left\{ e^{\int_0^t z_s ds} v(0,X_t) - \int_0^t e^{\int_0^s z_r dr} \psi(X_s, z_s) \frac{\zeta_t}{\zeta_s} ds \right\} \right). \end{aligned}$$

Since ζ is arbitrary, computing the supremum of the right-hand side yields the lower bound.

It remains to show that there is no duality gap. We now assume $\phi(x, 0) = 0$ for all x. Under the uniform Lipschitz assumption, the function θ is bounded. Let $\Theta_s = \theta(X_s, v(t-s, X_s))$ and

$$\zeta_s^* = V_s e^{\int_0^s \Theta_r dr}$$

Note that by Fubini's theorem

$$\zeta_s^* = M_s + \int_0^s (M_s - M_r)\Theta_r e^{\int_0^r \Theta_q dq} dr$$

so ζ^* is a non-negative bounded martingale. Similar to the key observation above, we have for any anticipative process z that

$$\begin{aligned} e^{\int_0^t z_s ds} v(0, X_t) - \int_0^t e^{\int_0^s z_r dr} \Psi_s \frac{\zeta_t^*}{\zeta_s^*} ds =& \zeta_t^* + \int_0^t (V_s z_s - \Psi_s - \Phi_s) e^{\int_0^s z_r dr} \frac{\zeta_t^*}{\zeta_s^*} ds \\ &\geq \zeta_t^* \end{aligned}$$

where here $\Psi_s = \psi(X_s, z_s)$. Note there is equality when $z_s = Z_s^*$ for all s. This shows

$$\begin{aligned} \mathbb{E}\left[\min_z \left\{ e^{\int_0^t z_s ds} v(0, X_t) - \int_0^t e^{\int_0^s z_r dr} \psi(X_s, z_s) \frac{\zeta_t^*}{\zeta_s^*} ds \right\}\right] &= \mathbb{E}[\zeta_t^*] \\ &= v(t, x) \end{aligned}$$

□

To prove Theorem 1.1 we need one more ingredient, the so-called S-equation, due to Skorokhod [15, equation (4)]. We provide a proof for completeness.

Theorem 2.3. *Let $\{X_t^i : i \in I_t, t \geq 0\}$ be the branching process described in the introduction. Fix a measurable $f : \mathcal{X} \to [0, 1]$ and for all $(t, x) \in \mathbb{R}_+ \times \mathcal{X}$, let*

$$u(t, x) = \mathbb{E}_x \left[\prod_{i \in I_t} f(X_t^i) \right].$$

Then

$$u(t, x) = \mathbb{E}_x \left[f(X_t) + \int_0^t g(X_s, u(t-s, X_s)) ds \right]$$

where

$$g(x, \eta) = \lambda(x) \left(\mathbb{E}[\eta^N | X_T = x] - \eta \right) \text{ for all } (x, \eta) \in \mathcal{X} \times [0, 1].$$

Proof. Letting

$$G(x, \eta) = \mathbb{E}[\eta^N | X_T = x] \text{ for all } (x, \eta) \in \mathcal{X} \times [0, 1]$$

be the conditional probability generating function of the offspring distribution, we have

$$\begin{aligned} \mathbb{E}_x \left[\mathbb{1}_{\{t \geq T\}} \prod_{i \in I_t} f(X_t^i) \right] &= \mathbb{E}_x \left[\mathbb{1}_{\{t \geq T\}} u(t-T, X_T)^N \right] \\ &= \mathbb{E}_x \left[\mathbb{1}_{\{t \geq T\}} G(X_T, u(t-T, X_T)) \right] \\ &= \mathbb{E}_x \left[\int_0^t e^{-\int_0^s \lambda(X_r) dr} \lambda(X_s) G(X_s, u(t-s, X_s)) ds \right]. \end{aligned}$$

Hence

$$u(t,x) = \mathbb{E}_x\left[\mathbb{1}_{\{t<T\}} f(X_t) + \mathbb{1}_{\{t\geq T\}}\prod_{i\in I_t} f(X_t^i)\right]$$
$$= \mathbb{E}_x\left[e^{-\int_0^t \lambda(X_s)ds} f(X_t) + \int_0^t e^{-\int_0^s \lambda(X_r)dr}\lambda(X_s)G(X_s, u(t-s,X_s))ds\right].$$

Fixing (t,x), the process

$$M_s = e^{-\int_0^s \lambda_r dr} U_s + \int_0^s e^{-\int_0^r \lambda_v dv}\lambda_r G(X_s, U_s)dr$$

is a martingale, where $\lambda_s = \lambda(X_s)$ and $U_s = u(t-s, X_s)$. Then by Fubini's theorem we have

$$f(X_t) + \int_0^t g(X_s, U_s))ds = M_t + \int_0^t \lambda_s e^{\int_0^s \lambda_r dr}(M_t - M_s)ds.$$

By assumption, the function λ is bounded and hence the pathwise integral on the right-hand side is integrable, with mean zero. Since $\mathbb{E}_x(M_t) = M_0 = u(t,x)$ we are done. □

Remark 2.4. McKean [13] noted that that the solution of the FKPP equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + u^2 - u,$$

named after Fisher [7] and Kolmogorov–Petrovskii–Piskunov [10], can be represented in terms of a branching Brownian motion with unit $\lambda = 1$ branching rate and binary $N = 2$ offspring distribution. This observation is an important special case of Theorem 2.3. In this context, it is often called the McKean representation of the solution of the FKPP equation; see the lecture notes of Berestycki [2, Section 2.3] for instance.

Proof of Theorem 1.1. Let $v(t,x) = 1 - u(t,x)$ where $u(t,x)$ is defined in Theorem 2.3. Hence, the function v satisfies the hypothesis of Theorem 2.1 with

$$\phi(x,\eta) = -g(x, 1-\eta),$$

where the function g is defined in Theorem 2.3. Note that $\phi(x,0) = -g(x,1) = 0$. The optimisation representations for v yield the desired optimisation representations for u after some manipulation. □

Finally, we consider the case where the number of offspring is constant $N = 2$.

Proof of Corollary 1.4. We appeal to Remark 1.3 now, and replace the function h appearing in Theorem 1.1 with the function

$$h(x,z) = \max_{\eta\in\mathbb{R}}\{\eta z - \lambda(x)\eta^2\} = \frac{z^2}{4\lambda(x)}.$$

This yields

$$m = 1 - \max_{\zeta}\mathbb{E}\left[\text{ess sup}_z\left\{e^{\int_0^t z_s ds}(1-f_t) + \zeta_t\int_0^t e^{\int_0^s z_r dr}\frac{(z_s-\lambda_s)^2}{4\lambda_s\zeta_s}ds\right\}\right]$$

where we use the notation $f_t = f(X_t)$ and $\lambda_s = \lambda(X_s)$.

Letting $w_s = \frac{1}{2}(z_s - \lambda_s)$ we see

$$\int_0^t \frac{e^{\int_0^s \lambda_r dr}}{\lambda_s \zeta_s} \left(w_s e^{\int_0^s w_r dr} \right)^2 ds \geq \frac{(e^{\int_0^t w_s ds} - 1)^2}{\int_0^t e^{-\int_0^s \lambda_r dr} \lambda_s \zeta_s ds}$$

by the Cauchy–Schwarz inequality, with equality if $w_s e^{\int_0^s w_r dr} = -e^{-\int_0^s \lambda_r dr} \lambda_s \zeta_s$. Also letting $W_t = e^{\int_0^t w_s ds}$ we have by completing the square that

$$(1 - f_t) e^{\int_0^t \lambda_s ds} W_t^2 + \frac{\lambda_t (W_t - 1)^2}{\int_0^t e^{-\int_0^s \lambda_r dr} \lambda_s \zeta_s ds} \geq \frac{\zeta_t (1 - f_t) e^{\int_0^t \lambda_s ds}}{\zeta_t + (1 - f_t) \int_0^t e^{\int_s^t \lambda_r dr} \lambda_s \zeta_s ds}$$

with equality if $W_t = \frac{\zeta_t}{\zeta_t + (1-f_t) \int_0^t e^{\int_s^t \lambda_s} \lambda_s \zeta_s ds}$.

Finally, note that both equality conditions are satisfied for the martingale $\zeta_s^* = v_s e^{\int_0^s \lambda_r (1 - v_r) dr}$ and the control $w_s^* = -\lambda_s v_s$, where $v_s = v(t - s, X_s)$ and

$$v(t, x) = 1 - \mathbb{E}_x \left[\prod_{i \in I_t} f(X_t^i) \right].$$

□

3. Bounding solutions

In this section, we explore a simple consequence of Theorem 2.1. We now assume that the non-linearity ϕ appearing in the integral equation is such that there is concave, differentiable and Lipschitz function $\hat{\phi} : \mathbb{R} \to \mathbb{R}$ such that $\phi(x, \eta) = \hat{\phi}(\eta)$ for all $(x, \eta) \in \mathcal{X} \times \mathbb{R}$. In order to avoid overburdening the notation, we will drop that ˆ and simply write this function as ϕ.

We also introduce the following notation. We let $\mathbf{V}_t$ be the operator that sends the bounded measurable function $v_0 : \mathcal{X} \to \mathbb{R}$ to $v(t, \cdot)$, where $v : \mathbb{R}_+ \times \mathcal{X} \to \mathbb{R}$ is the unique bounded solution to the integral equation

$$v(t, x) = \mathbb{E}_x \left[v_0(X_t) + \int_0^t \phi(v(t - s, X_s)) ds \right] \text{ for all } (t, x) \in \mathbb{R}_+ \times \mathcal{X},$$

so that $v(t, x) = \mathbf{V}_t(v_0)(x)$. We let $\mathbf{P}_t$ be the transition operator of the Markov process, such that

$$\mathbf{P}_t(f)(x) = \mathbb{E}_x[f(X_t)] \text{ for all } (t, x) \in \mathbb{R}_+ \times \mathcal{X},$$

for all bounded measurable f.

Finally, we let

$$R_t(r_0) = \mathbf{V}_t(r_0 \mathbb{1}) \text{ for all } (t, r_0) \in \mathbb{R}_+ \times \mathbb{R}$$

where $\mathbb{1}(x) = 1$ for all $x \in \mathcal{X}$. That is to say, if $r : \mathbb{R}_+ \to \mathbb{R}$ solves the ordinary differential equation

$$\dot{r} = \phi(r), \quad r(0) = r_0$$

then $R_t(r_0) = r(t)$. Now let $\mathbf{R}_t$ be the operator defined by

$$\mathbf{R}_t(f)(x) = R_t(f(x)).$$

The main result of this section is the following non-asymptotic form of the Trotter product formula:

Corollary 3.1. *Fix all bounded measurable f and integers $n \geq 1$ we have*

$$(\mathbf{P}_{t/n} \circ \mathbf{R}_{t/n})^n(f)(x) \leq \mathbf{V}_t(f)(x) \leq (\mathbf{R}_{t/n} \circ \mathbf{P}_{t/n})^n(f)(x)$$

for all $(t, x) \in \mathbb{R}_+ \times \mathcal{X}$.

Remark 3.2. Our Corollary 3.1 is very much in the spirit of a result of Cliff, Goldstein & Wacker [6, Theorem 18], though our method of proof is rather different to theirs.

Remark 3.3. Recall that a solution v to the integral equation can be interpreted as the mild solution to the reaction diffusion-type equation

$$\frac{\partial v}{\partial t} = \mathcal{L}v + \phi(v)$$

where $\mathcal{L}$ is the generator of the Markov process X. The 'diffusion' term corresponds to the Markov (linear) semigroup $(\mathbf{P}_t)_{t\geq 0}$ generated by $\mathcal{L}$, while the 'reaction' term corresponds to the non-linear semigroup $(\mathbf{R}_t)_{t\geq 0}$ generated by the concave (state-independent) function ϕ. Finally, $(\mathbf{V}_t)_{t\geq 0}$ is the non-linear 'reaction-diffusion' semigroup generated by the sum $\mathcal{L} + \phi$. An interesting reformulation of Corollary 3.1 is

$$(e^{t\mathcal{L}/n} e^{t\phi/n})^n \leq e^{t(\mathcal{L}+\phi)} \leq (e^{t\phi/n} e^{t\mathcal{L}/n})^n$$

Proof. The key ingredient of the proof is that by Theorem 2.1 we have

$$R_t(r_0) = \min_z \left\{ e^{\int_0^t z_s ds} r_0 - \int_0^t e^{\int_0^s z_r dr} \psi(z_s) ds \right\}$$

where

$$\psi(z) = \inf_\eta \{ z\eta - \phi(\eta) \}.$$

and the minimum is over deterministic bounded measurable functions $z : [0, t] \to \mathbb{R}$.

We first consider the case $n = 1$. For the upper bound, note that by Theorem 2.1 we have

$$\begin{aligned}
\mathbf{V}_t(f)(x) &\leq \inf_z \mathbb{E}_x \left[e^{\int_0^t z_s ds} f(X_t) - \int_0^t e^{\int_0^s z_r dr} \psi(z_s) ds \right] \\
&= \min_z \left\{ e^{\int_0^t z_s ds} \mathbb{E}_x[f(X_t)] - \int_0^t e^{\int_0^s z_r dr} \psi(z_s) ds \right\} \\
&= \mathbf{R}_t \circ \mathbf{P}_t(f)(x).
\end{aligned}$$

Similarly, letting $(Z_s^*)_{0\leq s\leq t}$ be the maximiser of the minimisation in Theorem 2.1, we have

$$\begin{aligned}
\mathbf{V}_t(f)(x) &= \mathbb{E}_x \left[e^{\int_0^t Z_s^* ds} f(X_t) - \int_0^t e^{\int_0^s Z_r^* dr} \psi(Z_s^*) ds \right] \\
&\geq \mathbb{E}_x \left[\min_z \left\{ e^{\int_0^t z_s ds} f(X_t) - \int_0^t e^{\int_0^s z_r dr} \psi(z_s) ds \right\} \right] \\
&= \mathbb{E}_x[R_t(f(X_t))] \\
&= \mathbf{P}_t \circ \mathbf{R}_t(f)(x).
\end{aligned}$$

Now, note that each of the operators $\mathbf{P}_t$, $\mathbf{R}_t$ and $\mathbf{V}_t$ are increasing. In particular, we have

$$\begin{aligned}\mathbf{V}_{s+t}(f)(x) &= \mathbf{V}_s \circ \mathbf{V}_t(f)(x) \\ &\geq \mathbf{V}_s \circ \mathbf{P}_t \circ \mathbf{R}_t(f)(x) \\ &\geq \mathbf{P}_s \circ \mathbf{R}_s \circ \mathbf{P}_t \circ \mathbf{R}_t(f)(x).\end{aligned}$$

The same argument works for the upper bound. Induction completes the proof. □

Remark 3.4. Alternatively, in the case where $\phi(0) = 0$, we could insert the martingale $\zeta_s = 1$ into the objective of the max-min problem in Theorem 2.1 to obtain the lower bound.

Remark 3.5. From the proof of Theorem 1.1, we say that an interesting case is when $\phi(\eta) = \lambda(1-\eta-G(1-\eta))$ where $\lambda > 0$ is constant and G is the probability generating function of a non-negative integer-valued random variable N. This corresponds to the case of a branching process with a constant branching rate λ and the distribution of the number of particles N produced at a branching event is independent of the event's location. In this case, we have the formula

$$R_t(r_0) = 1 - \mathbb{E}[(1-r_0)^{|I_t|}]$$

where I_t is the set of particles alive at times $t \geq 0$.

4. An application to a branching Lévy process

In this section, we prove Theorem 1.5. Recall that here the branching process $\{X_t^i : i \in I_t, t \geq 0\}$ is constructed from a real-valued Lévy process X starting from $X_0 = x_0$. Given the result we are trying to prove, there is no loss assuming $x_0 = 0$. Recall also that the branching rate is a positive constant λ and the distribution of the number of particles N produced at a branching event is independent of the position of the particles. Recall also that the cumulant generating function K of X is assumed finite in a neighbourhood of the origin. In what follows, we will let $\hat{X}$ be the Lévy process with the transition distribution of $-X$. Note that the function K plays the role of the Laplace exponent of $\hat{X}$:

$$\mathbb{E}_x[e^{-\theta \hat{X}_t}] = e^{\theta x + tK(\theta)}$$

for all $t \geq 0$, where here the subscript x denotes conditioning on the event $\{\hat{X}_0 = x\}$.

The key step of our proof of Theorem 1.5 is the following proposition:

Proposition 4.1. *Let $v : \mathbb{R}_+ \times \mathbb{R} \to [0,1]$ solve the integral equation*

$$v(t,x) = \mathbb{P}_x(\hat{X}_t < 0) + \mathbb{E}_x \int_0^t \phi(v(t-s, \hat{X}_s))ds$$

where $\phi : [0,1] \to \mathbb{R}$ is concave and differentiable with $\phi(0) = 0 = \phi(\beta)$ where $0 < \beta \leq 1$, and $\phi'(0) = \gamma > 0$. Set

$$q = \inf_{\theta > 0} \frac{K(\theta) + \gamma}{\theta}$$

Then we have

$$v(t, rt) \to \begin{cases} \beta & \text{if } r < q \\ 0 & \text{if } r > q \end{cases}$$

Before we prove Proposition 4.1, we show how it can be used to find the asymptotic speed of the right-most particle:

Proof of Theorem 1.5. Let $u(t,\cdot)$ be the distribution function of $M_t = \sup_{i\in I_t} X_t^i$, with $M_t = -\infty$ when I_t is empty. Note that by the translational invariance of the transition distribution of a Lévy process

$$\begin{aligned} u(t,x) &= \mathbb{P}(M_t \le x) \\ &= \mathbb{P}_x(\min_{i\in I_t} \hat{X}_t^i \ge 0) \\ &= \mathbb{E}_x\left[\prod_{i\in I_t} \mathbb{1}_{\{\hat{X}_t^i \ge 0\}}\right] \end{aligned}$$

According to Theorem 2.3 applied to the branching process $\{\hat{X}_t^i : i \in I_t,\ t \ge 0\}$ we have

$$u(t,x) = \mathbb{P}_x(\hat{X}_t \ge 0) + \mathbb{E}_x \int_0^t g(u(t-s,\hat{X}_s))ds$$

where

$$g(\eta) = \lambda(\mathbb{E}[\eta^N] - \eta) \text{ for } 0 \le \eta \le 1.$$

Note that g is convex with $g(1) = 0$ and $g(0) = \mathbb{P}(N = 0) \ge 0$. By the assumption that $\mathbb{E}[N] > 1$, we have $g'(1) > 0$ and hence there exists a smaller root $0 \le \alpha < 1$ such that $g(\alpha) = 0$.

Note that $v = 1 - u$ satisfies the conditions of Proposition 4.1 with $\beta = 1 - \alpha$. Now recall that $\alpha = \mathbb{P}(E)$ where $E = \{I_t = \emptyset \text{ for some } t > 0\}$ is the event that population eventually becomes extinct. See the book of Athreya & Ney [1, Theorem III.4.1]. Hence, we have shown

$$\mathbb{P}\left(M_t \le rt\right) \to \begin{cases} \mathbb{P}(E) & \text{if } r < q \\ 1 & \text{if } r > q \end{cases}$$

Noting that $\mathbb{P}\left(\big\{M_t \le rt\big\} \cap E\right) \to \mathbb{P}(E)$ since

$$\begin{aligned} \mathbb{P}(E) &\ge \mathbb{P}\left(\big\{M_t \le rt\big\} \cap E\right) \\ &\ge \mathbb{P}\left(\big\{M_t \le rt\big\} \cap \{I_t = \emptyset\}\right) \\ &= \mathbb{P}\left(I_t = \emptyset\right) \\ &\to \mathbb{P}(E). \end{aligned}$$

the conclusion follows since

$$\begin{aligned} \mathbb{P}\left(M_t \le rt \;\middle|\; E^c\right) &= \frac{1}{\mathbb{P}(E^c)}[\mathbb{P}\left(M_t \le rt\right) - \mathbb{P}\left(\{M_t \le rt\} \cap E\right)] \\ &\to \begin{cases} 0 & \text{if } r < q \\ 1 & \text{if } r > q. \end{cases} \end{aligned}$$

This shows that for any $\varepsilon > 0$ we have

$$\mathbb{P}(|\tfrac{1}{t}M_t - q| > \varepsilon \mid E^c) \to 0$$

as desired. $\square$

The rest of this section contains the proof of Proposition 4.1. The case where $\hat{X}$ is degenerate, in the sense that $\hat{X}_t = x + bt$ for a constant b is immediate. Therefore, we will assume without loss that $\hat{X}$ is non-degenerate, so that $K''(0) = \mathrm{Var}(\hat{X}_1) > 0$.

Of the two bounds, the upper bound is easier to obtain. Using the $n = 1$ case of Theorem 3.1, we have

$$v(t, x) \le R_t(\mathbb{P}_x(X_t < 0)).$$

By the concavity of ϕ we have

$$\phi(v) \le \gamma v$$

and hence by Grönwall's inequality

$$R_t(r_0) \le r_0 e^{\gamma t}.$$

Now by Markov's inequality we have

$$\mathbb{P}_x(\hat{X}_t < 0) \le e^{-x\theta + tK(\theta)}$$

for any $\theta > 0$ and $t \ge 0$. Putting this together, we have shown

$$v(t, rt) \le e^{t(K(\theta)+\gamma - r\theta)}.$$

If $r > \frac{1}{\theta}(\Lambda(\theta) + \gamma)$ then the right-hand side vanishes as $t \to \infty$, as claimed.

For the lower bound, we will introduce some more notation. Let

$$F_t(y) = \mathbb{P}_0(\hat{X}_t \le y)$$

be the conditional distribution function of the random variable $\hat{X}_t$ given $\hat{X}_0 = 0$. Note that by spacial homogeneity of the Lévy process, we have

$$\mathbb{P}_x(\hat{X}_t \le y) = F_t(y - x).$$

Let F_t^{-1} be the quantile function, defined as

$$F_t^{-1}(p) = \inf\{x : F_t(x) \ge p\},$$

so that $F_t(x) \ge p \Leftrightarrow x \ge F_t^{-1}(p)$.

The key estimates are the following:

Lemma 4.2. *For all $0 < b < \beta$, $n \ge 1$, $t > 0$ and $x \in \mathbb{R}$ we have*

$$v(t, x) \ge bF_\delta \left(-x - (n-1)F_\delta^{-1} \left(\frac{Q_\delta^{-1}(b)}{b} \right) \right)$$

where $\delta = t/n$.

Remark 4.3. It is interesting to note that Lemma 4.2 actually holds with no assumption on law of the Lévy process. In particular, it holds for processes, such as stable processes, for which the Laplace exponent $K(\theta)$ is infinite for all $\theta \ne 0$.

Proof of Lemma 4.2. We fix δ and use induction on n. We first consider the $n = 1$ case.

Since the points 0 and $\beta \le 1$ are fixed points of ϕ, we have $R_\delta(0) = 0$ and $R_\delta(1) \ge \beta$. In particular, we have

$$\begin{aligned} v(\delta, x) &\ge \mathbf{P}_\delta \circ \mathbf{R}_\delta \mathbb{1}_{(-\infty,0]}(x) \\ &\ge \beta \mathbb{P}_x(\hat{X}_\delta \le 0) \\ &= \beta F_\delta(-x) \end{aligned}$$

To do the inductive step, we will make use of the following observation: for any $0 < b < \beta$ and $k \in \mathbb{R}$ we have

$$R_\delta[bF_\delta(k)] \ge b\mathbb{1}_{\{F_\delta(k) \ge R_\delta^{-1}(b)/b\}}$$

since R_δ is increasing on $[0,\beta]$. Now suppose the claim is true for $n=m$, we have

$$\begin{aligned} v((m+1)\delta,x) &\geq \mathbf{P}_\delta \circ \mathbf{R}_\delta \left[bF_\delta\left(-\cdot -(m-1)F_\delta^{-1}\left(\frac{R_\delta^{-1}(b)}{b}\right)\right)\right](x) \\ &\geq b\ \mathbb{P}_x\left[F_\delta\left(-\hat{X}_\delta - (m-1)F_\delta^{-1}\left(\frac{R_\delta^{-1}(b)}{b}\right)\right) \geq \frac{R_\delta^{-1}(b)}{b}\right] \\ &= b\ F_\delta\left(-x - mF_\delta^{-1}\left(\frac{R_\delta^{-1}(b)}{b}\right)\right). \end{aligned}$$

□

Lemma 4.4. *For all $0 < c < \gamma = \phi'(0)$ and all $0 < b < \beta$, where β is the larger root of ϕ, there exists $\delta^* > 0$ such that $R_\delta^{-1}(b) \leq be^{-c\delta}$ for all $\delta \geq \delta^*$.*

Proof. Fix a $q^* \in (0,\beta)$, for instance $q^* = \beta/2$ and let

$$H(q) = \int_{q^*}^{q} \frac{ds}{\phi(s)}.$$

Note that the differential equation defining R can be solved as

$$R_\delta(r_0) = H^{-1}(H(r_0)+\delta)$$

for $0 < r_0 < \beta$, and hence

$$R_\delta^{-1}(r_0) = H^{-1}(H(r_0)-\delta).$$

In this notation, we must prove that

$$H(b) - \delta \leq H(be^{-c\delta})$$

or equivalently

$$\frac{1}{\delta}\int_0^\delta \frac{bce^{-cx}dx}{\phi(be^{-cx})} \leq 1$$

for δ large enough. To do this, note that the limit of the left-hand side as $\delta \to \infty$ is $c/\gamma < 1$ by l'Hôpital's rule. □

Lemma 4.5. *For all $r < q$ there exists a $c < \gamma$ and a $\delta^* > 0$ such that $F_\delta^{-1}(e^{-c\delta}) \leq -r\delta$ for all $\delta \geq \delta^*$.*

Proof. Note that $q > -\mathbb{E}_0(\hat{X}_1)$ with strict inequality since since $K(\theta) > -\theta\mathbb{E}_0(\hat{X}_1)$ by Jensen's inequality. Hence we need only consider r such that

$$-\mathbb{E}_0(\hat{X}_1) < r < q.$$

In particular, we may invoke Cramér large deviation principle to conclude that,

$$\log F_\delta(-r\delta) = -\hat{K}(r)\delta(1+o(1))$$

as $\delta \to \infty$, where the large deviation rate function $\hat{K}$ is the Legendre transform of K, defined by

$$\hat{K}(\eta) = \sup_\theta \left[\eta\theta - K(\theta)\right].$$

Hence, it is enough to show that

$$\hat{K}(r) < \gamma.$$

Now, since $r > K'(0) = -\mathbb{E}_0(\hat{X}_1)$, there exists an $\varepsilon > 0$ such that $r > K'(\varepsilon)$, since K' is continuous and increasing in a neighbourhood of $\theta = 0$. By the convexity of K we have the inequality

$$r\theta - K(\theta) \leq r\varepsilon - K(\varepsilon)$$

for $\theta < \varepsilon$ and hence

$$\begin{aligned}\hat{K}(r) &= \sup_{\theta \geq \varepsilon} [r\theta - K(\theta)] \\ &\leq -\varepsilon(q - r) + \sup_{\theta \geq \varepsilon} [q\theta - K(\theta)] .\end{aligned}$$

The conclusion follows since $q\theta - K(\theta) \leq \gamma$ for all $\theta > 0$ by the definition of q. □

Proof of Proposition 4.1. Fix $0 < b < \beta$ and $r < q$. Pick $\bar{r}$ such that $r < \bar{r} < q$. By Lemma 4.5 there exists a c and δ_1^* such that $F_\delta^{-1}(e^{-c\delta}) \leq -\bar{r}\delta$ for all $\delta \geq \delta_1^*$. By Lemma 4.4 there exists δ_2^* such that $R_\delta^{-1}(b) \leq be^{-c\delta}$ for all $\delta \geq \delta_2^*$.

Let $m = 1 + \mathbb{E}_0(\hat{X}_1)$. By the weak law of large numbers

$$F_\delta(m\delta) = \mathbb{P}_0(\hat{X}_\delta/\delta \leq m) \to 1.$$

So given $\varepsilon > 0$, there exists δ_3^* such that $F_\delta(m\delta) \geq 1 - \varepsilon$ for $\delta \geq \delta_3^*$.

Let $n \geq \frac{\bar{r}+m}{\bar{r}-r}$ and $t \geq n \max_i\{\delta_i^*\}$. Finally, let $\delta = t/n$, so $\delta \geq \max_i\{\delta_i^*\}$ and hence

$$\begin{aligned}v(t, rt) &= v(n\delta, rn\delta) \\ &\geq bF_\delta(-rn\delta - (n-1)F_\delta^{-1}\left(R_\delta^{-1}(b)/b\right)) \\ &\geq bF_\delta(-rn\delta - (n-1)F_\delta^{-1}(e^{-c\delta})) \\ &\geq bF_\delta(-rn\delta + (n-1)\bar{r}\delta) \\ &= bF_\delta([n(\bar{r} - r) - \bar{r}]\delta) \\ &\geq bF_\delta(m\delta) \\ &\geq b(1 - \varepsilon).\end{aligned}$$

Since $b < \beta$ and $\varepsilon > 0$ are arbitrary, the conclusion follows. □

Remark 4.6. It is possible to express the speed q of the travelling wave front in several ways. The above proof show that q can be rewritten as

$$q = \sup\{r : \hat{K}(r) < \gamma\},$$

where $\hat{K}$ is the Legendre transform of Λ. This formulation for the speed of the right-most particle appears in the paper of Biggins [3] or, more recently, in the paper of Groisman & Jonckheere [8].

Following an idea in the paper of Hiriart-Urruty & Martínez-Legaz [9], an inverse to the function $\hat{K}$ can be calculated as follow. First, define a new function K° by the formula

$$K^\circ(\theta) = \begin{cases} +\infty & \text{if } \theta \geq 0 \\ -\theta K(-1/\theta) & \text{if } \theta < 0. \end{cases}$$

Note that the function K° is convex, and indeed, it is related to the perspective function of the Laplace exponent K. Define its Legendre transform in the usual fashion

$$\hat{K}^\circ(\eta) = \sup_\theta [\eta\theta - K^\circ(\theta)] .$$

Then it can be shown that an inverse function to $\hat{K}$ is the function $-\hat{K}^\circ$. In particular, the speed q can be rewritten as

$$q = -\hat{K}^\circ(\gamma).$$

Simplifying the above formula recovers the formula in Proposition 4.1.

Remark 4.7. Consider the case of dyadic branching Brownian motion, where $N = 2$ and $K(\theta) = \frac{1}{2}\theta^2$. Letting $m(t)$ be the median, defined by

$$\mathbb{P}(\max_{i \in I_t} X_t^i \le m(t)) = 1/2$$

we have

$$-\sqrt{t}\Phi^{-1}\left(\frac{1}{e^t+1}\right) \ge m(t) \ge -\frac{\sqrt{t}}{\sqrt{n}}\Phi^{-1}\left(\frac{1}{2b}\right) - \frac{(n-1)\sqrt{t}}{\sqrt{n}}\Phi^{-1}\left(\frac{1}{e^{t/n}(1-b)+b}\right) \tag{4.1}$$

for all $1/2 < b < 1, n \ge 1$, where

$$\Phi(z) = \int_{-\infty}^{z} \frac{e^{-s^2/2}}{\sqrt{2\pi}} ds$$

is the standard normal distribution function. Indeed, the upper bound follows from the upper bound $1/2 \le v(t,m) = R_t[\mathbb{P}(X_t \le -m)]$ and the calculation $R_t(r_0) = \frac{r_0}{r_0 + e^{-t}(1-r_0)}$ in the case when $\phi(v) = v(1-v)$. The lower bound is implied by Lemma 4.2.

Using $\Phi^{-1}(\varepsilon) = -\sqrt{2\log(1/\varepsilon)}(1+o(1))$ as $\varepsilon \downarrow 0$ yields

$$m(t) = \sqrt{2}t + o(t).$$

On the other hand, a famous result of Bramson [4] says

$$m(t) = \sqrt{2}t - \frac{3}{2\sqrt{2}}\log t + O(1).$$

Since $\Phi^{-1}(\varepsilon) = -\sqrt{2\log(1/\varepsilon)} + O\left(\frac{\log\log(1/\varepsilon)}{\sqrt{\log(1/\varepsilon)}}\right)$ the upper bound in equation (4.1) actually recovers the correct order of magnitude of the second term of the expansion. It would be interesting to see if, by optimising over the free parameters b and n, it is possible to recover the $\log t$ term in the lower bound as well.

5. Acknowledgements

David Driver was supported, in part, by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/H023348/1 for the University of Cambridge Centre for Doctoral Training, the Cambridge Centre for Analysis. Mike Tehranchi thanks the Cambridge Endowment for Research in Finance for their support. Finally, both of us thank Nathanael Berestycki for his comments and suggestions.

References

[1] K.B. Athreya and P.E. Ney. *Branching Processes.* Die Grundlehren der mathematischen Wissenscraften in Einzeldarstellungen, Band 196. Springer-Verlag. (1972)

[2] J. Berestycki. *Topics on Branching Brownian motion.* `http://www.stats.ox.ac.uk/~berestyc/Articles/EBP182.pdf` (2014)

[3] J.D. Biggins. The growth and spread of the general branching random walk. *Annals of Applied Probability* **5**(4): 1008–1024. (1995)

[4] M. Bramson. Maximal displacement of branching Brownian motion. *Communications on Pure and Applied Mathematics* **31**(5): 531–581. (1978)
[5] X. Cabré and J.M. Roquejoffre. The influence of fractional diffusion in Fisher-KPP equations. *Communications in Mathematical Physics* **320**(3): 679–722. (2013)
[6] M. Cliff, J.A. Goldstein, M. Wacker. Positivity, Trotter products, and blow-up. *Positivity* **8**: 187–208. (2004)
[7] R.A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics* **7**(4): 355–369. (1937)
[8] P. Groisman, M. Jonckheere. Front propagation and quasi-stationary for one-dimensional Lévy processes. *Electronic Communications in Probability.* (2018)
[9] J.-B. Hiriart-Urruty and J.-E. Martínez-Legaz. New formulas for the Legendre–Fenchel transform. *Journal of Mathematical Analysis and Applications.* **288**: 544–555 (2003)
[10] A.N. Kolmogorov, I.G. Petrovskii and N.S. Piskunov. A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem. *Moscow University Bulletin of Mathematics* **1**(6): 1–26. (1937) In *Selected Works of A.N. Kolmogorov. Volume 1: Mathematics and Mechanics.* V.M. Tikhomirov editor, V.M. Volosov translator. Mathematics and Its Applications. Kluwer Academic Press. (1991)
[11] N. Ikeda, M. Nagasawa, S. Watanabe. A construction of branching Markov processes. *Proceedings of the Japan Academy* **42**(4): 380–384. (1966)
[12] N. Ikeda, M. Nagasawa, S. Watanabe. Branching Markov processes I. *Kyoto Journal of Mathematics* **8**(2): 233–278. (1968)
[13] H.P. McKean. Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. *Communications on Pure and Applied Mathematics* **28**(3): 323–331. (1975)
[14] L.C.G. Rogers. Pathwise stochastic optimal control. *SIAM Journal on Control and Optimization* **46**(3):1116–1132. (2007)
[15] A.V. Skorokhod. Branching diffusion processes. *Theory of Probability & Its Applications* **9**(3): 445–449. (1964)

Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, UK
E-mail address: d.driver.maths@gmail.com, m.tehranchi@statslab.cam.ac.uk