Electronic Journal of Probability

Vol. 12 (2007), Paper no. 14, pages 420–446.

Journal URL
http://www.math.washington.edu/~ejpecp/

Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct

Amaury Lambert
Unit of Mathematical Evolutionary Biology
UMR 7625 Lab of Ecology and Evolution
École Normale Supérieure
46 rue d'Ulm
F-75230 Paris Cedex 05, France
E-mail: amaury.lambert@ens.fr

UMR 7625 Laboratoire d'Écologie
Université Pierre et Marie Curie-Paris 6
7 quai Saint Bernard
F-75252 Paris Cedex 05, France
URL: http://ecologie.snv.jussieu.fr/amaury

Abstract

We consider continuous-state branching (CB) processes which become extinct (i.e., hit 0) with positive probability. We characterize all the quasi-stationary distributions (QSD) for the CB-process as a stochastically monotone family indexed by a real number. We prove that the minimal element of this family is the so-called Yaglom quasi-stationary distribution, that is, the limit of one-dimensional marginals conditioned on being nonzero.

Next, we consider the branching process conditioned on not being extinct in the distant future, or Q-process, defined by means of Doob h-transforms. We show that the Q-process is distributed as the initial CB-process with independent immigration, and that under the $L \log L$ condition, it has a limiting law which is the size-biased Yaglom distribution (of the

CB-process).
More generally, we prove that for a wide class of nonnegative Markov processes absorbed at 0 with probability 1, the Yaglom distribution is always stochastically dominated by the stationary probability of the Q-process, assuming that both exist.
Finally, in the diffusion case and in the stable case, the Q-process solves a SDE with a drift term that can be seen as the instantaneous immigration.

Key words: Continuous-state branching process ; Lévy process ; Quasi-stationary distribution ; theorem ; h-transform ; Q-process ; Immigration ; Size-biased distribution ; Stochastic differential equations.

AMS 2000 Subject Classification: Primary 60J80; Secondary: 60K05, 60F05, 60H10, 60G18.

Submitted to EJP on July 28 2006, final version accepted April 5 2007.

1 Introduction

We study continuous-state branching processes (CB), which are the continuous (in time and space) analogue of Bienaymé–Galton–Watson processes (see (4; 9; 16; 19)). In particular, a CB-process Z is a strong Markov process with nonnegative values, and 0 as absorbing state. It is characterized by the Laplace exponent ψ of a Lévy process with nonnegative jumps. CB-processes can serve as models for population dynamics. In that setting, we will set $\rho := \psi'(0^+)$ and call $-\rho$ the *Malthusian parameter*, since one has $\mathbb{E}(Z_t) = \mathbb{E}(Z_0)\exp(-\rho t)$. Here special attention is given to such dynamics running over large amounts of time, in the cases when the CB-process hits 0 with probability 1, that is, those subcritical ($\rho > 0$) or critical ($\rho = 0$) CB-processes such that $\int^\infty 1/\psi$ converges.

First, we study *quasi-stationary distributions*, that is, those probability measures ν on $(0, \infty)$ satisfying for any Borel set A

$$\mathbb{P}_\nu(Z_t \in A \mid Z_t > 0) = \nu(A) \qquad t \geq 0.$$

We characterize all quasi-stationary distributions of the CB-process in the subcritical case, which form a *stochastically decreasing* family (ν_γ) of probabilities indexed by $\gamma \in (0, \rho]$. The probability ν_ρ is the so-called *Yaglom distribution*, in the sense that

$$\lim_{t\to\infty} \mathbb{P}_x(Z_t \in A \mid Z_t > 0) = \nu_\rho(A) \qquad x > 0.$$

As far as processes are concerned, conditioning the population to be still extant at some fixed time t yields time-inhomogeneous kernels. We therefore aim at defining the branching process conditioned on being never extinct. Of course when Z is absorbed with probability less than 1, this conditioning can be made in the usual sense, and the game is over. But when Z hits 0 with probability 1, one can condition it to non-extinction in the sense of h-transforms (martingale changes of measure). More precisely, for any $\mathcal{F}_t$-measurable Θ

$$\lim_{s\to\infty} \mathbb{P}_x(\Theta \mid Z_{t+s} > 0) = \mathbb{E}_x\left(\frac{Z_t}{x}\, e^{\rho t}, \Theta\right).$$

The process thus conditioned to be never extinct is denoted $Z^\uparrow$ and called Q-process as in the discrete setting (see (2)). We also show that it is distributed as a CB-process with immigration (CBI, see (21)). Under the $L\log L$ condition, the Q-process converges in distribution and we are able to characterize its limiting law as the *size-biased* Yaglom distribution aforementioned. In particular, the stationary probability of the Q-process dominates stochastically the quasi-stationary limit of the initial process. In a side result, we prove that this holds for all nonnegative Markov processes Y absorbed at 0 with probability 1, for which the mappings $x \mapsto \mathbb{P}_x(Y_t > 0)$ are nondecreasing (for every t).

The critical case is degenerate, since then the CB-process has no QSD, but has a Q-process, which is transient. However, if $\sigma := \psi''(0^+) < \infty$, there is a relationship of the size-biasing type between the limiting distributions of the *rescaled* processes. First, Z_t/t conditioned on being nonzero converges in distribution as $t \to \infty$ to an exponential variable with parameter $2/\sigma$. Second, $Z_t^\uparrow/t$ converges in distribution as $t \to \infty$ to the *size-biased* distribution of the same exponential variable.

In order to understand better how the immigration occurs (see also (25)), we study the diffusion case (ψ is a quadratic polynomial) and the stable case (ψ is a power function with power $\alpha \in (1, 2]$). In the diffusion case, there are $\sigma > 0$ and $r \leq 0$ such that

$$dZ_t = rZ_t dt + \sqrt{\sigma Z_t} dB_t,$$

where B is the standard Brownian motion, and then the Q-process $Z^{\uparrow}$ solves

$$dZ_t^{\uparrow} = rZ_t^{\uparrow} dt + \sqrt{\sigma Z_t^{\uparrow}} dB_t + \sigma dt.$$

In the stable case, we prove that the CB-process Z solves the following stochastic differential equation

$$dZ_t = Z_{t-}^{1/\alpha} dX_t,$$

where X is a spectrally positive Lévy process with Laplace exponent ψ, and that the Q-process solves

$$dZ_t^{\uparrow} = (Z_{t-}^{\uparrow})^{1/\alpha} dX_t + d\sigma_t,$$

where σ is a subordinator with Laplace exponent ψ' independent of X, which may then be seen as the instantaneous immigration.

For related results on h-transforms of branching processes, see (1; 32). For further applications of the Yaglom distribution, see (26). An alternative view on the various conditionings in terms of the (planar) Bienaymé–Galton–Watson tree itself can be found in (15; 30) and is briefly discussed in Subsection 2.1. Also, see (39) for a study of Yaglom-type results for the *Jirina process* (branching process in *discrete* time and continuous-state space). More generally, a census of the works on quasi-stationary distributions is regularly updated on the website of P.K. Pollett (36).

Finally, we want to point out that (part of) the results concerning the Q-process had been written in the author's PhD thesis (24) in 1999, but had remained unpublished. Nonempty intersection with these results, as well as other conditionings, were proved by analytical methods and published independently by Zeng-Hu Li in 2000 (29). Because the present work takes a probabilistic approach, puts the emphasis on sample-paths, and provides a self-contained display on quasi-stationary distributions and Q-processes as well as new results (in particular, results relating these), we believe it is of independent interest. We indicate in relevant places where our results intersect.

Actually, we have to add that the CB-process conditioned to be never extinct first appeared in complete generality in (37). This seminal paper initiated a series of papers on conditioned superprocesses, all of which focussed on continuous branching mechanism (see e.g. (10; 11; 13; 31; 40) and the references therein). In that case, the mass of the superprocess is a critical CB-process called the Feller diffusion, or squared Bessel process of dimension 0. It is a diffusion Z satisfying the SDE $dZ_t = 2\sqrt{Z_t} dB_t$. Properties of the (conditioned) Feller diffusion are studied in particular in (13; 14; 28; 35). These references are quoted more precisely in Subsection 5.2. In the superprocess setting, other interesting conditionings can be made, and specifically, a vast literature is dedicated to conditioning on *local survival* at a given site (see for example (41) and the references therein).

Let us also mention a paper of Tony Pakes (33), in which some of the problems we tackle here were considered for CB-processes which cannot hit 0, conditioning Z_t by the events $\{Z_{t+s} > \epsilon\}$ or $\{T(\epsilon) > t+s\}$, where $T(\epsilon)$ is the last hitting time of ϵ.

The next section recalls some classical results in the Bienaymé–Galton–Watson case and reviews results concerning Lévy processes and CB-processes. At the end of these preliminaries, a useful lemma is stated and proved. In the third section, we treat the question of quasi-stationary distributions for the CB-process and in the fourth one, we introduce the Q-process and study its main properties. The last section is dedicated to some comments and results on links between the Q-process and the Yaglom distribution, as well as further results in the diffusion as well as stable cases.

2 Preliminaries

In the first subsection, we remind the reader of classical definitions and results in the discrete case. This will ease understanding the sequel, thanks to the numerous similarities between the discrete and continuous cases.

2.1 Classical results in the discrete case

Consider $(Z_n, n \geq 0)$ Bienaymé–Galton–Watson (BGW) process with offspring distribution $(\nu(k), k \geq 0)$ and associated probability generating function f, shorter called a DB(f). We call m the mean of ν. Assume that $m = f'(1) \leq 1$ (critical or subcritical case) and that $\nu(0)\nu(1) \neq 0$. It is well-known that

$$T := \inf\{n \geq 0 : Z_n = 0\}$$

is then a.s. finite. First, we briefly review the results on the distribution of Z_n conditional on $\{Z_n \neq 0\}$. Early work of Kolmogorov (22) on the expectation of Z_n conditional on $\{Z_n \neq 0\}$ culminated in so-called Yaglom's theorem, whose assumptions were refined in (18) and (38). Namely, in the subcritical case, there is a probability $(\alpha_k, k \geq 1)$, called the *Yaglom distribution*, such that

$$\sum_{k\geq 1} \alpha_k \mathbb{P}_k(Z_1 = j) = m\alpha_j \qquad j \geq 1, \tag{1}$$

and if g is its probability generating function, then

$$\lim_{n\to\infty} \mathbb{E}_x(s^{Z_n} \mid Z_n \neq 0) = g(s) \qquad s \in [0,1].$$

Also, provided that $\sum_{k\geq 1}(k \log k)\nu(k) < \infty$, one has $g'(1) < \infty$ and

$$\lim_{n\to\infty} \mathbb{E}_x(Z_n \mid Z_n \neq 0) = g'(1).$$

Second, we state without proof the results concerning the conditioning of Z on being non extinct in the distant future. They can be found in (2, pp. 56–59), although it seems that the *Q-process* first appeared in (17).

- The conditional probabilities $\mathbb{P}(\cdot \mid T \geq k)$ converge as $k \to \infty$ in the sense of finite-dimensional distributions to an honest probability measure $\mathbb{P}^{\uparrow}$. The probability $\mathbb{P}^{\uparrow}$ defines a new homogeneous Markov chain, denoted by $Z^{\uparrow}$ and called Q-process. The Q-process lives in the positive integers and its n-fold transition function is given by

$$\mathbb{P}(Z_n^{\uparrow} = j \mid Z_0^{\uparrow} = i) = P_{ij}(n)\frac{j}{i}m^{-n} \qquad i, j \geq 1, \tag{2}$$

  where $P_{ij}(n)$ denotes that of the initial BGW process.

- The Q-process has the following properties
  (i) if $m = 1$, then it is transient.
  (ii) if $m < 1$, then it is positive-recurrent iff

$$\sum_{k\geq 1}(k \log k)\nu(k) < \infty.$$

  (iii) In the positive-recurrent case, the Q-process has stationary measure $(k\alpha_k/g'(1), k \geq 1)$, where α is the Yaglom distribution described previously.

Next consider a BGW tree and add independently of the tree at each generation n a random number Y_n of particles, where the Y_i's are i.i.d. Give to these immigrating particles independent BGW descendant trees with the same offspring distribution. Then the width process (i.e., the process of generation sizes) of the modified tree is a Markov chain called a discrete-branching process with *immigration*. If f and g stand for the probability generating functions of resp. the offspring distribution and Y_1, we denote this Markov chain by DBI(f, g). It is then straightforward that

$$\mathbb{E}_i(s^{Z_1}) = \sum_{j\geq 0}\mathbb{P}(Z_1 = j \mid Z_0 = i)s^j = g(s)f(s)^i, \qquad i \geq 0.$$

Then recalling (2), and differentiating the last equality w.r.t. s when there is no immigration ($g \equiv 1$) yields

$$\sum_{j\geq 0}\mathbb{P}(Z_1^{\uparrow} = j \mid Z_0^{\uparrow} = i)s^j = \sum_{j\geq 0}\mathbb{P}(Z_1 = j \mid Z_0 = i)\frac{j}{i}m^{-1}s^j = \frac{sf'(s)}{m}f(s)^{i-1}.$$

The foregoing equality shows that $(Z_n^{\uparrow} - 1, n \geq 0)$ is a DBI(f, f'/m) and thus provides a useful recursive construction for a Q-process tree, called *size-biased tree* :

- at each generation, a particle is marked
- give to *unmarked* particles independent (sub)critical BGW descendant trees with offspring distribution ν
- give to each *marked* particle k children with probability $\mu(k)$, where μ is the *size-biased* distribution of ν, that is,

$$\mu(k) = \frac{k\nu(k)}{m} \qquad k \geq 1,$$

  and *mark* one of these children uniformly at random.

By construction, the width process of the tree obtained *after removing the marked particles* is a DBI(f, f'/m). This proves that the Q-process has the same law as the width process of the initial size-biased tree, which contains one infinite branch and one only, that of the marked particles. This infinite branch is usually called the *spine* of the tree. In continuous time models (especially in the superprocess literature), it is sometimes called the *immortal particle* (12).

Actually, this *spine decomposition* can be proved a little bit more precisely, by conditioning the tree itself to never become extinct, instead of the width process. The general idea is as follows. First, condition the tree upon having descendance at generation k (large). The probability that two or more individuals from generation n have descendants at generation k vanishes with k, so that, with very high probability, all individuals from generation k have the same one ancestor at generation n. The spine is the set of such individuals as n varies. The size-biasing comes from the fact that each such individual is uniformly distributed among its siblings. This description was made precise in (15; 30).

As a conclusion, remember that the immigrating mechanism in the Q-process is obtained from the branching mechanism by differentiation because of this spine decomposition with *size-biased* offspring. We will see that this relation carries over to the continuous setting. See (8; 25) for works focusing more specifically on the spine decomposition of continuous trees.

2.2 Continuous-state branching processes and Lévy processes

A continuous-state branching process, or CB-process, is a strong Markov process $(Z_t; t \geq 0)$ with values in $[0, \infty]$, 0 and ∞ being absorbing states. It is characterized by its branching mechanism function ψ and enjoys the following branching property. The sum of two independent CB(ψ) starting respectively from x and y, is a CB(ψ) starting from $x + y$. CB-processes can be seen as the analogue of BGW processes in continuous time and continuous state-space. Their branching mechanism function ψ is specified (4; 27) by the Lévy–Khinchin formula

$$\psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_0^\infty (e^{-\lambda r} - 1 + \lambda r)\Lambda(dr) \qquad \lambda \geq 0,$$

where $\beta \geq 0$ denotes the Gaussian coefficient, and Λ is a positive measure on $(0, \infty)$ such that $\int_0^\infty (r^2 \wedge r)\Lambda(dr) < \infty$.

Then ψ is also the Laplace exponent of a Lévy process X with no negative jumps. Throughout this paper, we will denote by P_y the law of X started at $y \in \mathbb{R}$, and by $\mathbb{P}_x$ that of the CB-process Z started at $x \geq 0$. Then for all $\lambda, t \geq 0$,

$$E_y(\exp -\lambda X_t) = E_0(\exp -\lambda(X_t + y)) = \exp(-\lambda y + t\psi(\lambda)),$$

and

$$\mathbb{E}_x(\exp -\lambda Z_t) = \exp(-x u_t(\lambda)),$$

where $t \mapsto u_t(\lambda)$ is the unique nonnegative solution of the integral equation (see e.g. (9))

$$v(t) + \int_0^t \psi(v(s))ds = \lambda \qquad \lambda, t \geq 0. \tag{3}$$

For general results about CB-processes that we state and do not prove, such as the last one, we refer the reader to (4; 16; 27). Set

$$\rho := \psi'(0+).$$

A CB-process is resp. called supercritical, critical or subcritical as $\rho < 0$, $= 0$, or > 0. Since we aim at conditioning on non-extinction (in two different ways), we have to assume that Z is a critical or a subcritical CB-process ($\rho \geq 0$), otherwise the conditioning would drive the paths to ∞. However, in the supercritical case, it is well-known that by conditioning Z to *extinction*, one recovers a subcritical CB-process. Indeed, if $\eta = \sup\{\lambda > 0 : \psi(\lambda) \leq 0\}$ is nonzero, then $\rho < 0$ (supercritical case), and the CB(ψ) conditioned on its extinction is a CB($\psi^\sharp$), where $\psi^\sharp(\lambda) = \psi(\lambda + \eta)$. As a consequence, every following statement concerning subcritical processes conditioned to survive up until generation, say n, can be applied to superprocesses conditioned to survive up until generation n, but to *eventually die out*.

Now we assume that X is not a subordinator, so the following equivalence holds (16)

$$\int^{\infty} \frac{d\lambda}{\psi(\lambda)} \text{ diverges } \Leftrightarrow T = \infty \quad \text{a.s.},$$

where T is the extinction time of the CB-process

$$T := \inf\{t \geq 0 : Z_t = 0\}.$$

Hence we always make the assumption that

$$\int^{\infty} \frac{d\lambda}{\psi(\lambda)} < \infty.$$

This implies in particular that the sample-paths of X and Z have infinite variation a.s.. The last two assumptions together imply that $T < \infty$ a.s.. Next put

$$\phi(t) := \int_t^{\infty} \frac{d\lambda}{\psi(\lambda)} \qquad t > 0.$$

The mapping $\phi : (0, \infty) \to (0, \infty)$ is bijective, and we write φ for its inverse mapping. From (3), it is straightforward to get

$$\int_{u_t(\lambda)}^{\lambda} \frac{dv}{\psi(v)} = t \qquad \lambda, t > 0,$$

so that

$$u_t(\lambda) = \varphi(t + \phi(\lambda)) \qquad \lambda, t > 0. \tag{4}$$

Note that the branching property implies $u_{t+s} = u_t \circ u_s$. Then check that, since $\phi(\infty) = 0$, one has $u_t(\infty) = \varphi(t)$, and for every $x, t > 0$,

$$\mathbb{P}_x(Z_t > 0) = \mathbb{P}_x(T > t) = 1 - \exp(-x\varphi(t)).$$

To avoid confusion, we define the hitting time of 0 by X as T_0

$$T_0 := \inf\{t \geq 0 : X_t = 0\}.$$

There is actually a sample-path relationship between the branching process Z and the Lévy process X stopped at T_0, called *Lamperti's transform* (27). Specifically, introduce

$$C_t = \int_0^t Z_s \, ds \qquad t \geq 0.$$

If $(\gamma_t, t \geq 0)$ denotes the right inverse of C

$$\gamma_t = \inf\{s \geq 0 : C_s > t\} \wedge T,$$

then for $x > 0$, the process $Z \circ \gamma$ under $\mathbb{P}_x$ has the same law as the Lévy process X started at x and stopped at T_0.

In the same direction, we stress that the law $P^\uparrow$ of X conditioned to stay positive is already well-known (6). It is submarkovian in the subcritical case ($\rho > 0$), but in the critical case ($\rho = 0$), it is defined thanks to the following absolute continuity relationship

$$P_x^\uparrow(\Theta) = E_x(\frac{X_t}{x}, \Theta, t < T_0), \qquad t \geq 0, \Theta \in \mathcal{F}_t.$$

Moreover, 0 is then an entrance boundary for $X^\uparrow$, that is, the measures $P_x^\uparrow$ converge weakly as $x \downarrow 0$ to a probability measure denoted by $P_0^\uparrow$.

We also have to say a word on CB-processes with immigration (21; 34). Recall from the previous subsection that in discrete branching processes with immigration (DBI), the total number of immigrants up until generation n is $\sum_{k=0}^n Y_k$, which is a *renewal process*. In the continuous setting, this role is played by a subordinator, which is characterized by its Laplace exponent, denoted by χ. Then the analogue of the DBI is denoted CBI(ψ, χ), and is a strong Markov process characterized by its Laplace transform

$$\mathbb{E}_x(\exp -\lambda Z_t) = \exp[-xu_t(\lambda) - \int_0^t \chi(u_s(\lambda))ds] \qquad \lambda \geq 0,$$

where $u_t(\lambda)$ is given by (4).

Finally, we state a technical lemma with a self-contained proof (beforehand, one part of this proof had to be found in (24) and the other one in (26)). The first convergence stated is also mentioned in (29, Theorem 3.1).

Lemma 2.1. *Assume $\rho \geq 0$ and let $G(\lambda) := \exp(-\rho\phi(\lambda))$. Then for any positive λ*

$$\lim_{t\to\infty} \frac{u_t(\lambda)}{\varphi(t)} = G(\lambda),$$

and for any nonnegative s

$$\lim_{t\to\infty} \frac{\varphi(t+s)}{\varphi(t)} = e^{-\rho s}.$$

When $\rho > 0$, the following identities are equivalent
(i) $G'(0+) < \infty$
(ii) $\int^\infty r \log r \, \Lambda(dr) < +\infty$
(iii)

$$\int_0 \left(\frac{1}{\rho\lambda} - \frac{1}{\psi(\lambda)}\right) d\lambda < \infty$$

(iv) There is a positive constant c such that $\varphi(t) \sim c\exp(-\rho t)$, as $t \to \infty$.
In that case, the constant c is implicitly defined by

$$\phi(c) = \int_0^c \left(\frac{1}{\rho\lambda} - \frac{1}{\psi(\lambda)}\right) d\lambda, \tag{5}$$

and then we have $G'(0+) = c^{-1}$.

Proof. The convergence of the ratios $\varphi(t+s)/\varphi(t)$ stems from the observation that

$$\int_{\varphi(t+s)}^{\varphi(s)} \frac{d\lambda}{\psi(\lambda)} = (\phi(\varphi(t+s)) - \phi(\varphi(s))) = t. \tag{6}$$

Indeed, for any $\varepsilon > 0$, there is a S such that for any $s > S$, $\lambda < \varphi(s)$ implies that $f(\lambda) := \psi(\lambda) - \rho\lambda < \varepsilon\lambda$. Next, since ψ is convex, f is nonnegative and for any $s > S$ and $t > 0$,

$$\begin{aligned} 0 \le \log \frac{\varphi(s)}{\varphi(t+s)} - \rho t &= \int_{\varphi(t+s)}^{\varphi(s)} \frac{d\lambda}{\lambda} - \int_{\varphi(t+s)}^{\varphi(s)} \frac{\rho d\lambda}{\psi(\lambda)} \\ &= \int_{\varphi(t+s)}^{\varphi(s)} \frac{f(\lambda)d\lambda}{\lambda\psi(\lambda)} \\ &\le \varepsilon t, \end{aligned}$$

which yields the result. From (4), we thus get

$$\lim_{t\to\infty} \frac{u_t(\lambda)}{\varphi(t)} = \lim_{t\to\infty} \frac{\varphi(t+\phi(\lambda))}{\varphi(t)} = e^{-\rho\phi(\lambda)}. \tag{7}$$

Next assume that $\rho > 0$. The proof of (ii)$\Leftrightarrow$(iii) was done in (16). From a similar calculation as that of (6), we get

$$\int_{u_t(\theta)}^{\theta} \left(\frac{1}{\rho\lambda} - \frac{1}{\psi(\lambda)}\right) d\lambda = \rho^{-1} \log\left(\frac{\theta}{u_t(\theta)\exp(\rho t)}\right),$$

hence (iv)$\Leftrightarrow$(iii). Furthermore, the previous equation implies

$$u_t(\theta) \sim \theta e^{-\rho t} \exp\left[-\rho \int_0^{\theta} \left(\frac{1}{\rho\lambda} - \frac{1}{\psi(\lambda)}\right) d\lambda\right] \qquad \text{as } t \to \infty.$$

The definition of the constant c comes from the necessary agreement between the last display and (7). This constant is well defined because the left-hand side and right-hand side of (5), as functions of c, are resp. bijective increasing and bijective decreasing from $(0,\infty)$ to $(0,\infty)$. Next observe that

$$G'(x) \sim \frac{1}{x} e^{-\rho\phi(x)} \qquad \text{as } x \to 0+,$$

and that for any $0 < x < \theta$,

$$\frac{1}{x} e^{-\rho\phi(x)} = \frac{1}{\theta} e^{-\rho\phi(\theta)} \exp\left[\int_x^{\theta} \left(\frac{1}{\lambda} - \frac{\rho}{\psi(\lambda)}\right) d\lambda\right],$$

which yields (i)$\Leftrightarrow$(iii). When (iii) holds, letting $x \to 0+$ and $\theta = c$ in the previous display, one gets

$$G'(0+) = \frac{1}{c} e^{-\rho\phi(c)} \exp\left[\int_0^{c} \left(\frac{1}{\lambda} - \frac{\rho}{\psi(\lambda)}\right) d\lambda\right] = \frac{1}{c},$$

and the proof is complete. $\square$

3 Quasi-stationary distributions

Roughly speaking, a quasi-stationary distribution (QSD) is a subinvariant distribution for a killed or transient Markov process. In the branching process setting, a QSD ν is a probability on $(0,\infty)$ satisfying

$$\mathbb{P}_\nu(Z_t \in A \mid T > t) = \nu(A).$$

Then by application of the simple Markov property,

$$\mathbb{P}_\nu(T > t+s) = \mathbb{P}_\nu(T > s)\mathbb{P}_\nu(T > t),$$

so that the extinction time T under $\mathbb{P}_\nu$ has an exponential distribution with parameter, say, γ. Then γ can be seen as the rate of mass decay of $(0,\infty)$ under $\mathbb{P}_\nu$. It is a natural question to characterize all the quasi-stationary probabilities associated to a given rate of mass decay γ.

Theorem 3.1. *Assume $\rho > 0$ (subcritical case). For any $\gamma \in (0,\rho]$ there is a unique QSD ν_γ associated to the rate of mass decay γ. It is characterized by its Laplace transform*

$$\int_{(0,\infty)} \nu_\gamma(dr)\, e^{-\lambda r} = 1 - e^{-\gamma\phi(\lambda)} \qquad \lambda \geq 0.$$

There is no QSD associated to $\gamma > \rho$.

In addition, the minimal QSD ν_ρ is the so-called Yaglom distribution, in the sense that for any starting point $x \geq 0$, and any Borel set A

$$\lim_{t\to\infty} \mathbb{P}_x(Z_t \in A \mid T > t) = \nu_\rho(A).$$

Actually, the last conditional convergence stated is due to Z.-H. Li, see (29, Theorem 4.3), where it is also generalized to conditionings of the type $\{T > t + r\}$ instead of $\{T > t\}$.

From now on, we will denote by Υ the r.v. with distribution ν_ρ. Since the Laplace transform of Υ is $1 - G$, Υ is integrable iff $\int^\infty r \log r\, \Lambda(dr) < \infty$, and then

$$\mathbb{E}(\Upsilon) = c^{-1},$$

where c is defined in Lemma 2.1.

Proof. There are multiple ways of proving the specific form of the QSD. The most straightforward way is the following

$$1 - e^{-\gamma t} = \mathbb{P}_{\nu_\gamma}(T < t) = \int_{(0,\infty)} \nu_\gamma(dr)\, e^{-r\varphi(t)},$$

so that, writing $t = \phi(\lambda)$, one gets

$$1 - e^{-\gamma\phi(\lambda)} = \int_{(0,\infty)} \nu_\gamma(dr)\, e^{-\lambda r} \qquad \lambda \geq 0.$$

Another way of getting this consists in proving that $\nu_\gamma Q = -\gamma Q + \gamma\delta_0$, where Q is the infinitesimal generator of the Feller process Z and δ_0 is the Dirac measure at 0. Taking Laplace transforms then leads to the differential equation

$$\gamma(1-\chi_\gamma(\lambda)) = -\psi(\lambda)\chi'_\gamma(\lambda) \qquad \lambda \geq 0,$$

where χ_γ stands for the Laplace transform of ν_γ. Solving this equation with the boundary condition $\chi(0) = 1$ yields the same result as given above.

Next recall that $\phi(\lambda) = \int_\lambda^\infty du/\psi(u)$, so that $\phi'(\lambda) \sim -1/\rho\lambda$ and $\phi(\lambda) \sim -\rho^{-1}\log(\lambda)$, as $\lambda \downarrow 0$. This entails

$$\int_{(0,\infty)} r\nu_\gamma(dr)\, e^{-\lambda r} \sim C(\lambda)\lambda^{\gamma/\rho-1} \qquad \text{as } \lambda \downarrow 0,$$

where C is slowly varying at 0^+, which would yield a contradiction if $\gamma > \rho$.

Before proving that $1 - G^{\gamma/\rho}$ is indeed a Laplace transform, we display the Yaglom distribution of Z. Observe that

$$\mathbb{E}_x(1-e^{-\lambda Z_t} \mid T > t) = \frac{\mathbb{E}_x(1-e^{-\lambda Z_t})}{\mathbb{P}_x(T > t)} = \frac{1-e^{-xu_t(\lambda)}}{1-e^{-x\varphi(t)}}$$

so that, by Lemma 2.1,

$$\lim_{t\to\infty} \mathbb{E}_x(e^{-\lambda Z_t} \mid T > t) = 1 - G(\lambda) \qquad \lambda > 0.$$

Since $G(0+) = 0$, this proves indeed that $1 - G$ is the Laplace transform of some probability measure ν_ρ on $(0,\infty)$. It just remains to show that when $\gamma \in (0,\rho)$, $1 - G^{\gamma/\rho}$ is indeed the Laplace transform of some probability measure ν_γ on $(0,\infty)$. Actually this stems from the following lemma applied to G and $\alpha = \gamma/\rho$. $\square$

Lemma 3.2. *If $1 - g$ is the Laplace transform of some probability measure on $(0,\infty)$, then so is $1 - g^\alpha$, $\alpha \in (0,1)$.*

Proof. It suffices to prove that, similarly as g, the n-th derivative of g^α has constant sign, equal to that of $(-1)^{n+1}$. By induction on n, it is elementary to prove that this n-th derivative is the sum of n functions $f_{n,k}$, $k = 1,\dots,n$, where

$$f_{n,k}(\lambda) = \alpha(\alpha-1)\cdots(\alpha-k+1)g^{\alpha-k}(\lambda) \sum_{\beta_1+\cdots+\beta_k=n} c_{n,k} \prod_{j=1}^k g^{(\beta_j)}(\lambda),$$

where the sum is taken over all k-tuples of positive integers summing to n, the coefficients $c_{n,k}$ are nonnegative, and $g^{(\beta)}$ denotes the β-th derivative of g. Because $g^{(\beta)}$ has the sign of $(-1)^{\beta+1}$, and the β_j's sum to n, the sum has constant sign equal to that of $(-1)^{n+k}$. Then since $\alpha < 1$, the sign of $f_{n,k}$ is that of $(-1)^{k-1}(-1)^{n+k} = (-1)^{n+1}$. $\square$

It is not difficult to get a similar result as the last theorem in the critical case. Assume $\rho = 0$ and $\sigma := \psi''(0+) < +\infty$ (Z has second-order moments). Variations on the arguments of the proof of Lemma 2.1 then show that $\varphi(t) \sim 2/\sigma t$ as $t \to \infty$, and

$$\lim_{t\to\infty} u_t(\lambda/t)/\varphi(t) = \frac{1}{1 + 2/\sigma\lambda} \qquad \lambda > 0.$$

Since

$$\mathbb{E}_x(1 - e^{-\lambda Z_t/t} \mid T > t) = \frac{1 - e^{-xu_t(\lambda/t)}}{1 - e^{-x\varphi(t)}} \qquad \lambda > 0,$$

the following statement is proved, which displays the usual 'universal' exponential limiting distribution of the conditioned, rescaled critical process.

Theorem 3.3. *Assume $\rho = 0$ and $\sigma := \psi''(0+) < +\infty$. Then*

$$\lim_{t\to\infty} \mathbb{P}_x(Z_t/t > z \mid T > t) = \exp(-2z/\sigma) \qquad z \geq 0.$$

Again, this conditional convergence is due to Z.-H. Li, see (29, Theorem 5.2), where it is generalized to conditionings of the type $\{T > t + r\}$ instead of $\{T > t\}$.

4 The Q-process

4.1 Existence

Recall that $\rho = \psi'(0^+) \geq 0$ is the negative of the Malthusian parameter of the CB-process. The next theorem states the existence in some special sense of the branching process conditioned to be never extinct, or Q-process.

Theorem 4.1. *Let $x > 0$.*

(i) The conditional laws $\mathbb{P}_x(\cdot \mid T > t)$ converge as $t \to \infty$ to a limit denoted by $\mathbb{P}_x^\uparrow$, in the sense that for any $t \geq 0$ and $\Theta \in \mathcal{F}_t$,

$$\lim_{s\to\infty} \mathbb{P}_x(\Theta \mid T > s) = \mathbb{P}_x^\uparrow(\Theta).$$

(ii) The probability measures $\mathbb{P}^\uparrow$ can be expressed as h-transforms of $\mathbb{P}$ based on the $(\mathbb{P}, (\mathcal{F}_t))$-martingale

$$D_t = Z_t e^{\rho t},$$

that is

$$d\mathbb{P}^\uparrow_{x|\mathcal{F}_t} = \frac{D_t}{x} \cdot d\mathbb{P}_{x|\mathcal{F}_t}$$

(iii) The process $Z^\uparrow$ which has law $\mathbb{P}_x^\uparrow$ is a CBI(ψ, χ) started at x, where χ is (the Laplace transform of a subordinator) defined by

$$\chi(\lambda) = \psi'(\lambda) - \psi'(0^+), \qquad \lambda \geq 0.$$

We point out that this result is originally due to S. Roelly and A. Rouault (37).

Proof. Recall that $\mathbb{P}_x(T < t) = \exp(-x\varphi(t))$, and, from Lemma 2.1, that

$$\lim_{s\to\infty} \frac{\varphi(s)}{\varphi(t+s)} = e^{\rho t}, \qquad t \geq 0.$$

(i) Now let $x \geq 0, s, t > 0$, and $\Theta \in \mathcal{F}_t$.
As a consequence of the foregoing convergence,

$$\lim_{s\to\infty} \frac{1-\exp(-Z_t\varphi(s))}{1-\exp(-x\varphi(t+s))} = \frac{Z_t}{x} e^{\rho t} \qquad \text{a.s.}$$

Moreover,

$$0 \leq \frac{1-\exp(-Z_t\varphi(s))}{1-\exp(-x\varphi(t+s))} \leq \frac{Z_t\varphi(s)}{1-\exp(-x\varphi(t+s))} \leq 2\frac{Z_t}{x}e^{\rho t},$$

for any s greater than some bound chosen independently of $Z_t(\omega)$. Hence by dominated convergence,

$$\begin{aligned}
\lim_{s\to\infty} \mathbb{P}_x(\Theta \mid T > t+s) &= \lim_{s\to\infty} \mathbb{E}_x\left(\frac{\mathbb{P}_{Z_t}(T > s)}{\mathbb{P}_x(T > t+s)}, \Theta, T > t\right) \\
&= \lim_{s\to\infty} \mathbb{E}_x\left(\frac{1-\exp(-Z_t\varphi(s))}{1-\exp(-x\varphi(t+s))}, \Theta, T > t\right) \\
&= \mathbb{E}_x\left(\frac{Z_t}{x}e^{\rho t}, \Theta\right).
\end{aligned}$$

(ii) It is well-known that

$$\mathbb{E}_x(Z_t) = xe^{-\rho t},$$

and the fact that D is a martingale follows from the simple Markov property.
(iii) Let us compute the Laplace transform of $Z_t^{\uparrow}$

$$\begin{aligned}
\mathbb{E}_x^{\uparrow}(e^{-\lambda Z_t}) &= \mathbb{E}_x\left(e^{-\lambda Z_t}\frac{Z_t}{x}e^{\rho t}\right) \\
&= -\frac{\partial}{\partial\lambda}(\mathbb{E}_x(e^{-\lambda Z_t}))\frac{e^{\rho t}}{x} \\
&= \frac{\partial}{\partial\lambda}(u_t(\lambda))e^{-xu_t(\lambda)}e^{\rho t}.
\end{aligned}$$

Now it is easy to prove thanks to (3), that

$$\frac{\partial}{\partial\lambda}(u_t(\lambda)) = \exp\left[-\int_0^t \psi'(u_s(\lambda))ds\right], \qquad \lambda, t \geq 0,$$

and then

$$\mathbb{E}_x^{\uparrow}(e^{-\lambda Z_t}) = \exp(-xu_t(\lambda))\exp\left(-\int_0^t \chi(u_s(\lambda))ds\right),$$

where

$$\chi(\lambda) = \psi'(\lambda) - \rho.$$

To check that χ is the Laplace transform of a subordinator, differentiate the Lévy–Khinchin formula for ψ, as

$$\chi(\lambda) = 2\beta\lambda + \int_0^\infty (1-e^{-\lambda r})r\Lambda(dr) \qquad \lambda \geq 0,$$

and the proof is complete. $\square$

4.2 Properties

We investigate the asymptotic properties of the Q-process $Z^\uparrow$ defined in the previous subsection. Since it is a CBI-process, criteria for convergence in distribution can readily be found in (34), but we will not need them. We remind the reader that $\mathbb{P}^\uparrow$ denotes the law of the Q-process, whereas $P^\uparrow$ is that of the Lévy process conditioned to stay positive. Also recall that the Yaglom r.v. Υ displayed in Theorem 3.1 is integrable as soon as $\int^\infty r\log r\,\Lambda(dr)<\infty$.

Theorem 4.2. *(i) (Lamperti's transform) If $\rho=0$, then*

$$\lim_{t\to\infty} Z_t^\uparrow = +\infty \qquad a.s.$$

Moreover, set

$$C_t=\int_0^t Z_s^\uparrow\,ds \qquad t\geq 0,$$

and let γ be its inverse. Then for $x>0$, the process $Z^\uparrow\circ\gamma$ under $\mathbb{P}_x$ has law $P_x^\uparrow$. In addition, if $\sigma:=\psi''(0+)<+\infty$, then

$$\lim_{t\to\infty}\mathbb{P}_x(Z_t^\uparrow/t>z)=\left(\frac{2}{\sigma}\right)^2\int_z^\infty u\exp(-2u/\sigma)\,du \qquad z\geq 0.$$

(ii) If $\rho>0$, the following dichotomy holds.
(a) If $\int^\infty r\log r\,\Lambda(dr)=\infty$, then

$$\lim_{t\to\infty} Z_t^\uparrow \overset{P}{=} +\infty.$$

(b) If $\int^\infty r\log r\,\Lambda(dr)<\infty$, then $Z_t^\uparrow$ converges in distribution as $t\to\infty$ to a positive r.v. $Z_\infty^\uparrow$ which has the distribution of the size-biased Yaglom distribution

$$\mathbb{P}(Z_\infty^\uparrow\in dr)=\frac{r\mathbb{P}(\Upsilon\in dr)}{\mathbb{E}(\Upsilon)} \qquad r>0.$$

We point out that the conditional convergence in distribution in (i) is due to Z.-H. Li (29, Theorem 5.1).

Proof. We start with the proof of (ii). From the proof of Theorem 4.1, we have

$$\mathbb{E}_x^\uparrow(e^{-\lambda Z_t})=\frac{\partial}{\partial\lambda}(u_t(\lambda))e^{-xu_t(\lambda)}e^{\rho t}.$$

Then since $\int_{u_t(\lambda)}^\lambda \frac{d\theta}{\psi(\theta)}=t$, we get

$$\frac{\partial}{\partial\lambda}(u_t(\lambda))=\frac{\psi(u_t(\lambda))}{\psi(\lambda)},$$

and

$$\mathbb{E}_x^\uparrow(e^{-\lambda Z_t})=\exp(-xu_t(\lambda))\exp(\rho t)\frac{\psi(u_t(\lambda))}{\psi(\lambda)}. \tag{8}$$

Now by convexity of ψ,

$$0 \le \int_{u_t(\lambda)}^{\lambda} \left(\frac{1}{\rho\theta} - \frac{1}{\psi(\theta)}\right) d\theta = \rho^{-1} \log\left(\frac{\lambda}{u_t(\lambda)\exp(\rho t)}\right).$$

Then from Lemma 2.1

$$\int^{\infty} r \log r \, \Lambda(dr) = \infty \Leftrightarrow \int_0 \left(\frac{1}{\rho\theta} - \frac{1}{\psi(\theta)}\right) d\theta = \infty.$$

In this case, for any $\lambda > 0$,

$$\lim_{t\to\infty} \mathbb{E}_x^{\uparrow}(e^{-\lambda Z_t}) = \lim_{t\to\infty} \frac{\psi(u_t(\lambda))}{\psi(\lambda)} \exp(-xu_t(\lambda)) \exp(\rho t) = 0.$$

In the opposite case,

$$\begin{aligned}
\lim_{t\to\infty} \mathbb{E}_x^{\uparrow}(e^{-\lambda Z_t}) &= \lim_{t\to\infty} u_t(\lambda) \frac{\rho}{\psi(\lambda)} \exp(-xu_t(\lambda)) \exp(\rho t) \\
&= \lim_{t\to\infty} \frac{\rho\lambda}{\psi(\lambda)} \exp\left[-\int_{u_t(\lambda)}^{\lambda} \left(\frac{1}{\theta} - \frac{\rho}{\psi(\theta)}\right) d\theta\right] \\
&= \frac{\rho\lambda}{\psi(\lambda)} \exp\left[-\int_0^{\lambda} \left(\frac{1}{\theta} - \frac{\rho}{\psi(\theta)}\right) d\theta\right].
\end{aligned}$$

Now from Lemma 2.1 again, recall that the Yaglom r.v. Υ has Laplace transform $1 - G$, where $G(\lambda) = \exp(-\rho\phi(\lambda))$, and $c^{-1} = \mathbb{E}(\Upsilon)$ is implicitly defined by

$$\phi(c) = \int_0^c \left(\frac{1}{\rho\lambda} - \frac{1}{\psi(\lambda)}\right) d\lambda.$$

This implies that

$$\exp\left[-\int_0^{\lambda} \left(\frac{1}{\theta} - \frac{\rho}{\psi(\theta)}\right) d\theta\right] = \exp[-\log(\lambda/c) - \rho\phi(\lambda)] = cG(\lambda)/\lambda,$$

so that

$$\lim_{t\to\infty} \mathbb{E}_x^{\uparrow}(e^{-\lambda Z_t}) = c\, G'(\lambda),$$

which completes the proof of (ii).

(i) For any nonnegative measurable functional F, $x > 0$, $t \ge 0$, since γ_t is a stopping time for the natural filtration of Z,

$$\begin{aligned}
\mathbb{E}_x^{\uparrow}(F(Z_{\gamma_s}, s \le t)) &= \mathbb{E}_x(F(Z_{\gamma_s}, s \le t) x^{-1} Z_{\gamma_t}, \gamma_t < T) \\
&= \mathbb{E}_x(F(X_s, s \le t) x^{-1} X_t, t < T_0) \\
&= E_x^{\uparrow}(F(X_s, s \le t)),
\end{aligned}$$

which proves the second part of the statement. As in (ii), it is still true that $Z_t^{\uparrow}$ converges in probability to $+\infty$ as $t \to \infty$. Hence $Z^{\uparrow}$ is a.s. not bounded and an application of the Markov property entails that $\lim_{t\to\infty} C_t = \infty$ $\mathbb{P}^{\uparrow}$-a.s. We conclude recalling that $\lim_{t\to\infty} X_t = \infty$ $P^{\uparrow}$-a.s. Finally, the limiting law of $Z_t^{\uparrow}/t$ can be obtained using (8) and the calculations preceding Theorem 3.3. $\square$

5 More about...

5.1 The quasi-stationary distribution and the Q-process

In Theorem 4.2, we proved that, in the subcritical case ($\rho > 0$) when $\int^{\infty} r \log r \, \Lambda(dr) < \infty$, the stationary probability of the Q-process (exists and) is the size-biased distribution of the Yaglom random variable Υ. Let us make three points about this result.

First, we would like to give a probabilistic interpretation of this result in terms of random trees. Consider a cell population, where each cell contains a certain (integer) number of parasites which proliferate *independently* and leave to *each* of the two daughter cells a random number of offspring, with the same law ξ. Assume that $m := \mathbb{E}\xi \in (0, 1)$. Then the number of parasites contained in a random line of descent of the cell population is a *subcritical* BGW-process and the overall population of parasites is a *supercritical* BGW-process. Then it is proved in (3) that on non-extinction of parasites,

- the fraction of *infected* cells of generation n containing k parasites converges in probability (as $n \to \infty$) to the k-th mass of the *Yaglom distribution* Υ of the subcritical BGW-process
- the fraction of *infected* cells of generation $n + p$ whose ancestor at generation n contained k parasites converges in probability (as p, and then $n \to \infty$) to the k-th mass of the *size-biased distribution* of Υ.

These results are the exact analogue of the relation between the Yaglom distribution and the stationary probability of the Q-process, but can be explained more easily in the present setting: because descendances of parasites *separate into disjoint lines of descent* with high probability, a *uniform pick* in the set $\partial\mathbb{T}^\star$ of *infinite lines of infected cells*, roughly amounts to a *size-biased pick* at generation n. Indeed, if there are two infected cells at generation n, the first one containing 1 parasite and the second one containing k parasites, then the probability that a uniform line in $\partial\mathbb{T}^\star$ descends from a parasite (this makes sense because of separation) in the second cell is k times greater than its complementary (size-biasing of Υ).

Second, we want to point out that this relationship could have been proved with less knowledge on the Q-process than required the proof of the theorem. Specifically, we want to show that

$$\lim_{t\to\infty} \lim_{s\to\infty} \mathbb{P}_y(Z_t \in dx \mid Z_{t+s} > 0) = \frac{x\mathbb{P}(\Upsilon \in dx)}{\mathbb{E}(\Upsilon)} \qquad y > 0.$$

By the simple Markov property, get

$$\mathbb{P}_y(Z_t \in dx \mid Z_{t+s} > 0) = \mathbb{P}_y(Z_t \in dx)\frac{\mathbb{P}_x(Z_s > 0)}{\mathbb{P}_y(Z_{t+s} > 0)},$$

and then observe, thanks to Lemma 2.1, that

$$\lim_{s\to\infty} \frac{\mathbb{P}_x(Z_s > 0)}{\mathbb{P}_y(Z_{t+s} > 0)} = \frac{x}{y}e^{\rho t},$$

so that

$$\lim_{s\to\infty} \mathbb{P}_y(Z_t \in dx \mid Z_{t+s} > 0) = x\mathbb{P}(Z_t \in dx \mid Z_t > 0)\frac{\mathbb{P}_y(Z_t > 0)}{y\exp(-\rho t)},$$

which, by another application of Lemma 2.1, converges as $t \to \infty$ to $cx\mathbb{P}(\Upsilon \in dx)$.

Third, we would like to give a more general viewpoint on this relationship. Consider any Markov process Y with values in $[0, \infty)$ which is absorbed at 0 with probability 1, and assume it has a Yaglom distribution Υ, defined as for the CB-process. A general question is to know whether there is any kind of relationship between the asymptotic distribution of Y_t conditioned on not yet being absorbed (the Yaglom distribution) and the asymptotic distribution of Y_t conditioned on not being absorbed in the distant future (the stationary probability of the Q-process). Intuitively, the second conditioning is more stringent than the first one, and should thus charge more heavily the paths that stay away from 0, than the first conditioning. One would think that this elementary observation should translate mathematically into a stochastic domination (of the first distribution by the second one). This is indeed the case when Y is a CB-process, since $Z_\infty^\uparrow$ is the size-biased Υ, that is, $Z_\infty^\uparrow$ is absolutely continuous w.r.t. Υ with increasing Radon–Nikodym derivative (the identity), which yields the domination.

Theorem 5.1. *If for any $t > 0$ the mapping $x \mapsto \mathbb{P}_x(Y_t > 0)$ is nondecreasing, then for any starting point $x > 0$, and for any $t, s > 0$,*

$$\mathbb{P}_x(Y_t > a \mid Y_{t+s} > 0) \geq \mathbb{P}_x(Y_t > a \mid Y_t > 0) \qquad a > 0.$$

Then, if there exists a Q-process $Y^\uparrow$, by letting $s \to \infty$,

$$\mathbb{P}_x(Y_t^\uparrow > a) \geq \mathbb{P}_x(Y_t > a \mid Y_t > 0) \qquad a > 0.$$

In addition, if the Q-process converges in distribution to a r.v. $Y_\infty^\uparrow$, and if the Yaglom r.v. Υ exists, then by letting $t \to \infty$,

$$Y_\infty^\uparrow \overset{stoch}{\geq} \Upsilon.$$

Remark. By a standard coupling argument, the monotonicity condition for the probabilities $x \mapsto \mathbb{P}_x(Y_t > 0)$ is satisfied for any strong Markov process with no negative jumps.

Proof. It takes a standard application of Bayes' theorem to get that for any $a, x, s, t > 0$,

$$\begin{array}{ccc} \mathbb{P}_x(Y_t > a \mid Y_{t+s} > 0) & \geq & \mathbb{P}_x(Y_t > a \mid Y_t > 0) \\ & \Updownarrow & \\ \mathbb{P}_x(Y_{t+s} > 0 \mid Y_t > a) & \geq & \mathbb{P}_x(Y_{t+s} > 0 \mid Y_t > 0). \end{array}$$

Next for any $m \geq 0$, set X_m the r.v. defined as

$$\mathbb{P}(X_m \in dr) = \mathbb{P}_x(Y_t \in dr \mid Y_t > m) \qquad r > 0.$$

For any $m \leq m'$, and $u \geq 0$, check that

$$\mathbb{P}(X_{m'} > u) \geq \mathbb{P}(X_m > u),$$

which means $X_{m'} \overset{stoch}{\geq} X_m$, so in particular $X_a \overset{stoch}{\geq} X_0$. Finally, observe that

$$\mathbb{P}_x(Y_{t+s} > 0 \mid Y_t > a) \geq \mathbb{P}_x(Y_{t+s} > 0 \mid Y_t > 0) \Leftrightarrow \mathbb{E}(f(X_a)) \geq \mathbb{E}(f(X_0)),$$

where $f(x) := \mathbb{P}_x(Y_s > 0)$. Since f is nondecreasing, the proof is complete. $\square$

5.2 Diffusions

In this subsection, we focus on the case when the CB-process is a diffusion.

First observe that if Z has continuous paths, then by Lamperti's time-change, it is also the case of the associated Lévy process, so that the branching mechanism must be of the form $\psi(\lambda) = \sigma\lambda^2/2 - r\lambda$. Using again Lamperti's time-change $C_t := \int_0^t Z_s ds$ and its right-inverse γ, $X := Z \circ \gamma$ is a (killed) Lévy process with continuous paths, namely $X_t = \sqrt{\sigma}\beta_t + rt$, where β is a standard Brownian motion. As a consequence, $X(C_t) - rC_t$ is a local martingale with increasing process σC_t, or equivalently, $Z_t - r\int_0^t Z_s ds$ is a local martingale with increasing process $\sigma \int_0^t Z_s ds$. This entails

$$dZ_t = rZ_t dt + \sqrt{\sigma Z_t} dB_t \qquad t > 0,$$

where B is a standard Brownian motion. Such diffusions are generally called *Feller diffusions* (eventhough this term is sometimes exclusive to the case $r = 0$), and when $r = 0$ and $\sigma = 4$, *squared Bessel processes with dimension 0.* Feller diffusions were introduced in (14), where they were proved to be limits of rescaled BGW processes. Such a convergence was also studied in (28) for *critical* BGW processes *conditioned to be never extinct*, and the limit was found to be a *squared Bessel process with dimension 4*. But as a special case of Theorem 4.1, this process is also known (35, Theorem 1) to be the critical Feller diffusion (squared Bessel process with dimension 0) conditioned to be never extinct, suggesting that *rescalings and conditionings commute.* Also, let us mention that specializations to the critical Feller diffusion of Theorems 3.3 (conditional convergence of Z_t/t to the exponential) and 4.2 (convergence of $Z_t^\uparrow/t$ to the size-biased exponential) were proved in (13, Lemma 2.1). Generalized Feller diffusions modelling population dynamics with density-dependence (e.g. competitive interactions) are shown in (5) to have QSD's and a Q-process. More interestingly, it is proved in (5) that if competition intensity increases rapidly enough with population size, then the diffusion 'comes down from infinity' and has a *unique* QSD.

By elementary calculus, check that if $r = 0$, then for any $t > 0$,

$$\phi(t) = \varphi(t) = 2/\sigma t,$$

whereas if $r \neq 0$,

$$\phi(t) = -r^{-1}\log(1 - 2r/\sigma t) \quad \text{and} \quad \varphi(t) = (2r/\sigma)e^{rt}/(e^{rt} - 1).$$

Note that $\rho = \psi'(0^+) = -r$.

The quasi-stationary distributions. Here we assume that $r < 0$ (subcritical case), so that $\rho = -r > 0$. Then from Theorem 3.1, for any $\gamma \in (0, \rho]$, the Laplace transform of the QSD ν_γ is

$$\int_0^\infty \nu_\gamma(dr)\, e^{-\lambda r} = 1 - \left(\frac{\lambda}{\lambda + 2\rho/\sigma}\right)^{\gamma/\rho}$$

In particular, whenever $\gamma < \rho$, ν_γ has *infinite* expectation, and it takes only elementary calculations to check that for any $\gamma < \rho$, ν_γ has a density f_γ given by

$$f_\gamma(t) = \frac{2\rho/\sigma}{\Gamma(1 - \gamma/\rho)\Gamma(\gamma/\rho)} \int_0^1 ds\ s^{\gamma/\rho}(1 - s)^{-\gamma/\rho}\ e^{-2\rho t s/\sigma} \qquad t > 0.$$

This can also be expressed as

$$\nu_\gamma((t,\infty)) = \mathbb{E}(\exp(-2\rho t X/\sigma)) \qquad t>0,$$

where X is a random variable with law $\mathrm{Beta}(\gamma/\rho, 1-\gamma/\rho)$.

Finally, for $\gamma=\rho$, the Laplace transform is easier to invert and provides the Yaglom distribution. The Yaglom r.v. Υ with distribution ν_ρ is an exponential variable with parameter $2\rho/\sigma$

$$\mathbb{P}(\Upsilon \in dx) = (2\rho/\sigma)e^{-2\rho x/\sigma} \qquad x\geq 0.$$

The Q-process. Here we assume that $r\leq 0$. From Theorem 4.1, the Q-process is a CBI-process with branching mechanism ψ and immigration $\chi=\psi'-\rho$.

Generally speaking a CBI(ψ, χ) has infinitesimal generator A whose action on the exponential functions $x\mapsto e_\lambda(x) := \exp(-\lambda x)$ is given by

$$Ae_\lambda(x) = (x\psi(\lambda)-\chi(\lambda))e_\lambda(x) \qquad x\geq 0.$$

This last expression stems directly from the definition of generator and the Laplace transforms of CBI-processes at fixed time given in the Preliminaries. In the present case, $\psi(\lambda)=\sigma\lambda^2/2-r\lambda$ and $\chi(\lambda)=\sigma\lambda$, so that

$$Ae_\lambda(x) = (x\frac{\sigma\lambda^2}{2} - xr\lambda - \sigma\lambda)e_\lambda(x) \qquad x\geq 0,$$

which yields, for any twice differentiable function f,

$$Af(z) = \frac{\sigma}{2}xf''(x) + rxf'(x) + \sigma f'(x) \qquad x\geq 0,$$

which can equivalently be read as

$$dZ_t^\uparrow = rZ_t^\uparrow dt + \sqrt{\sigma Z_t^\uparrow}dB_t + \sigma dt,$$

where $Z^\uparrow$ stands for the Q-process. Note that the immigration can readily be seen in the additional deterministic term σdt.

Now if $r<0$, according to Theorem 4.2, the Q-process converges in distribution to the r.v. $Z_\infty^\uparrow$ which is the size-biased Υ. But Υ is an exponential r.v. with parameter $2\rho/\sigma$, so that

$$\mathbb{P}(Z_\infty^\uparrow \in dx) = (2\rho/\sigma)^2\ xe^{-2\rho x/\sigma} \qquad x\geq 0,$$

or equivalently,

$$Z_\infty^\uparrow \overset{(d)}{=} \Upsilon_1+\Upsilon_2,$$

where Υ_1 and Υ_2 are *two independent copies of the Yaglom r.v.* Υ.

5.3 Stable processes

In this subsection, we consider the case when X is a spectrally positive α-stable process $1 < \alpha \leq 2$, that is a Lévy process with Laplace exponent ψ proportional to $\lambda \mapsto \lambda^\alpha$. In particular, $\rho = 0$ (critical case). Note that the associated Q-process was mentioned in (10, Corollary 1.5).

We show that the associated Q-process is the solution of a certain stochastic differential equation (SDE), which enlightens the immigration mechanism.

Theorem 5.2. *The branching process with branching mechanism ψ is the unique solution in law to the following SDE*

$$dZ_t = Z_{t-}^{1/\alpha} dX_t, \tag{9}$$

where X is a spectrally positive α-stable Lévy process with Laplace exponent ψ. Moreover the branching process conditioned to be never extinct is solution to

$$dZ_t = Z_{t-}^{1/\alpha} dX_t + d\sigma_t, \tag{10}$$

where σ is an $(\alpha - 1)$-stable subordinator with Laplace exponent ψ', independent of X.

Remark 1. The comparison between (9) and (10) allows to see the jumps of the subordinator σ as some instantaneous immigration added to the initial CB(ψ) in order to obtain the Q-process, which is a CBI(ψ, ψ').

Remark 2. More generally, CB-processes and CBI-processes can be shown (Theorem 5.1 and equation (5.3) in (7)) to be strong solutions to two different classes of stochastic equations, both driven by Brownian motions and Poisson random measures (provided that the intensity of the Poisson measure corresponding to the immigration mechanism has finite first-order moment –which is not the case here). As in (9) and (10), the second class of equation differs from the first one by an independent additional term (a deterministic drift and a pure jump part) that can readily be identified as the immigration term.

Proof. The uniqueness in law of (9) follows from (43, Theorem 1). Whether or not uniqueness holds for (10) remains an open question.

We turn to the law $\mathbb{P}$. By Lamperti's time-change,

$$Z_t = X_{C_t}, \qquad t \geq 0,$$

where X is a spectrally positive Lévy process with Laplace exponent ψ, and C is the non-decreasing time-change $C_t = \int_0^t Z_s ds$. Now by Theorem 4.1 in (20), there is a copy X' of X such that

$$X_{C_t} = Z_t = X_0 + \int_0^t Z_{s-}^{1/\alpha} dX'_s, \qquad t \leq T,$$

which entails (9).

We now show the result concerning $\mathbb{P}^\uparrow$. First recall from Theorem 4.2 that Lamperti's time-change still holds between Q-processes and processes conditioned to stay positive. We will thus find a SDE satisfied by X under $P^\uparrow$ and then conclude by time-change for $\mathbb{P}^\uparrow$.

By marking the jumps of X under $P^\uparrow$, we split the point process of jumps into two independent point processes. We will then identify their laws and deduce the SDE satisfied by X.

Let (A_t, U_t) be a Poisson point process in $(0,\infty)\times(0,1)$ with characteristic measure $\Lambda\otimes\lambda$, where λ stands for Lebesgue measure on $(0,1)$. The process $(A_t, t\geq 0)$ is the point process of jumps of some process X with law $\mathbb{P}$ defined by

$$X_t = \lim_{\varepsilon\downarrow 0}\left(\sum_{s\leq t} A_s \mathbf{1}_{\{A_s>\varepsilon\}} - t\int_{[\varepsilon,\infty)} r\Lambda(dr)\right).$$

Define two nonnegative r.v.'s Δ_t and δ_t by

$$(\Delta_t,\delta_t) = \begin{cases} (0, X_{t-}^{1/(\alpha-1)}A_t) & \text{if } X_{t-}>0 \text{ and } U_t<\frac{A_t}{X_t} \\ (A_t, 0) & \text{otherwise.}\end{cases}$$

In particular, Δ and δ never jump simultaneously and $\Delta X_t = A_t = \Delta_t + X_{t-}^{-1/(\alpha-1)}\delta_t$, $t\geq 0$. For any nonnegative predictable F, nonnegative bivariate f vanishing on the diagonal, and $x\geq 0$, we compute by optional projection the following expectation, after change of probability measure

$$E_x^\uparrow(\sum_{s\leq t} F_s f(\Delta_s,\delta_s)) = E_x^\uparrow\left[\sum_{s\leq t} F_s(f(0, A_s X_{s-}^{1/(\alpha-1)})\mathbf{1}_{U_s<A_s/X_s} + f(A_s,0)\mathbf{1}_{U_s\geq A_s/X_s})\right]$$

$$\begin{aligned}
&= \mathbb{E}_x\int_0^t ds F_s \mathbf{1}_{\{t<T_0\}}\int_0^\infty \Lambda(dz)\frac{X_s+z}{x}\left[\frac{z}{X_s+z}f(0,zX_s^{1/(\alpha-1)}) + \frac{X_s}{X_s+z}f(z,0)\right]\\
&= E_x^\uparrow\int_0^t ds F_s\int_0^\infty \Lambda(dz)\left[\frac{z}{X_s}f(0,zX_s^{1/(\alpha-1)}) + f(z,0)\right].
\end{aligned}$$

But the Lévy measure $\Lambda(dz)$ is proportional to $z^{-(\alpha+1)}dz$, hence putting $r = zX_s^{1/(\alpha-1)}$, $\int_0^\infty \Lambda(dz) z X_s^{-1} f(0, zX_s^{1/(\alpha-1)}) = \int_0^\infty \Lambda(dr) r f(0,r)$, and the last displayed quantity equals

$$E_x^\uparrow\int_0^t ds F_s\int_0^\infty \Lambda(dz)\left[zf(0,z)+f(z,0)\right].$$

Therefore under $P^\uparrow$, $\Delta X_t = \Delta_t + X_{t-}^{-1/(\alpha-1)}\delta_t$, where Δ and δ are two independent Poisson point processes with characteristic measures $\Lambda(dz)$ and $z\Lambda(dz)$, respectively. Moreover, for any positive ε and t,

$$\begin{aligned}
&\sum_{s\leq t}\Delta X_s \mathbf{1}_{\{\Delta X_s>\varepsilon\}} - t\int_{[\varepsilon,\infty)} r\Lambda(dr)\\
&= \left(\sum_{s\leq t}\Delta_s\mathbf{1}_{\{\Delta_s>\varepsilon\}} - t\int_{[\varepsilon,\infty)} r\Lambda(dr)\right) + \sum_{s\leq t}\frac{\delta_s}{X_{s-}^{1/(\alpha-1)}}\mathbf{1}_{\{\delta_s>\varepsilon\}}.
\end{aligned}$$

It is known that the first term in the r.h.s. converges a.s. as $\varepsilon\downarrow 0$ to the value at time t, X'_t of an α-stable Lévy process. By absolute continuity between $P^\uparrow$ and $\mathbb{P}$, the same holds for the quantity in the l.h.s. And the last quantity converges to

$$\sum_{s\leq t}\frac{\delta_s}{X_{s-}^{1/(\alpha-1)}} = \int_0^t \frac{d\sigma_s}{X_{s-}^{1/(\alpha-1)}},$$

where σ is an $(\alpha-1)$-stable subordinator independent from X defined by $\sigma_t = \sum_{s\leq t} \delta_s$. Indeed, the point process of jumps $(\delta_t, t \geq 0)$ of σ has compensation measure $\mu(dz) = z\Lambda(dz)$, and since then σ has Laplace exponent ψ'. We conclude that under $P^\uparrow$, the canonical process X satisfies

$$dX_t = dX'_t + \frac{1}{(X_{t-})^{1/(\alpha-1)}} d\sigma_t, \qquad t \geq 0,$$

where X' has law $\mathbb{P}$.

Go back to Lamperti's time-change to find the SDE satisfied by Z under $\mathbb{P}^\uparrow$. Write $Z_t = Z_0 + X_{C_t}$, with the same notations as previously, Z a (α-stable) Q-process and X a (α-stable) Lévy process conditioned to stay positive. Once again thanks to (20),

$$\int_0^t \frac{d\sigma_s}{(X_{s-})^{1/(\alpha-1)}} = \sigma'(\int_0^t \frac{ds}{X_s}), \qquad t \geq 0,$$

where σ' is a copy of σ, and

$$X'_{C_t} = \int_0^t Z_{s-}^{1/\alpha} dX''_s, \qquad t \geq 0,$$

where X'' is a copy of X'. Therefore

$$\begin{aligned} dZ_t &= dX_{C_t} = dX'_{C_t} + d\sigma'_t \\ &= Z_{t-}^{1/\alpha} dX''_t + d\sigma'_t, \end{aligned}$$

and (10) is proved. It thus remains to show the independence between X'' and σ'. Observing that

$$\sigma'_t = Z_t - Z_0 - X'_{C_t}, \qquad t \geq 0,$$

and

$$X''_t = \int_0^t Z_{s-}^{-1/\alpha} dX'_{C_s}, \qquad t \geq 0,$$

we see that the jump processes of X'' and σ' are $(\mathcal{G}_t)$-Poisson processes, with $\mathcal{G}_t = \sigma(Z_s, X'_{C_s}; s \leq t)$. Moreover, X'' and σ' never jump simultaneously as by construction the same holds for X' and σ, and

$$[\Delta\sigma'_t > 0 \text{ and } \Delta X''_t > 0] \Leftrightarrow [\Delta\sigma_{C_t} > 0 \text{ and } \Delta X'_{C_t} > 0].$$

Hence the jump processes of σ' and X'' are independent. Now since neither σ' nor X'' has a Gaussian coefficient, they are independent. $\square$

Acknowledgments. Part of this work had been written in April 1999, as a small chapter of my PhD dissertation (24), and has been, later on, utilized in (23, chapter 10). My thanks go to Alain Rouault who insisted that it was worth having it published and to Jean Bertoin who was my PhD advisor at that time. I am also indebted to Patrick Cattiaux, Pierre Collet, Régis Ferrière, Servet Martínez and Sylvie Méléard, for the nice ideas they provided and the deep discussions we had together about quasi-stationary distributions, which brought up many improvements to the first draft. Finally, I am grateful to two anonymous referees who provided me with some important references that I had shamefully missed out.

References

[1] Athreya, K.B. (2000)
Change of measures for Markov chains and the $L \log L$ theorem for branching processes. *Bernoulli* **6** 323–338. MR1748724

[2] Athreya, K.B., Ney, P.E. (1972)
Branching processes. Springer-Verlag, New York. MR0373040

[3] Bansaye, V. (2007)
Proliferating parasites in dividing cells : KimmelŠs branching model revisited. Preprint arXiv:math/0701917v1.

[4] Bingham, N.H. (1976)
Continuous branching processes and spectral positivity. *Stoch. Proc. Appl.* **4** 217–242. MR0410961

[5] Cattiaux, P., Collet, P., Lambert, A., Martinez, S., Méléard, S., San Martin, J. (2007)
Quasi-stationary distributions and diffusion models in population dynamics. Preprint arXiv:math/0703781v1

[6] Chaumont, L. (1996)
Conditionings and path decompositions for Lévy processes. *Stoch. Proc. Appl.* **64** 39–54.

[7] Dawson, D.A., Li, Z.-H. (2006)
Skew convolution semigroups and affine Markov processes. *Ann. Probab.* **34** 1103–1142. MR2243880

[8] Duquesne, T. (2007)
Continuum random trees and branching processes with immigration. *Stoch. Proc. Appl.* In press.

[9] Duquesne, T., Le Gall, J.-F. (2002)
Random Trees, Lévy Processes and Spatial Branching Processes. Astérisque 281. MR1954248

[10] Etheridge, A., Williams, D.R.E. (2003)
A decomposition of the $(1+\beta)$-superprocess conditioned on survival. *Proc. Roy. Soc. Edinburgh A* **133** 829–847. MR2006204

[11] Evans, S.N. (1992)
The entrance space of a measure-valued Markov branching process conditioned on non-extinction. *Can. Math. Bull.* **35** 70–74. MR1157466

[12] Evans, S.N. (1993)
Two representations of a conditioned superprocess. *Proc. Roy. Soc. Edinburgh A* **123** 959–971. MR1249698

[13] Evans, S.N., Perkins, E. (1990)
Measure-valued Markov branching processes conditioned on nonextinction. *Israel J. Math.* **71** 329–337. MR1088825

[14] Feller, W. (1951)
Diffusion processes in genetics. *Proc. 2nd Berkeley Symp. Math. Statist. Prob.* 227–246. U. California Press, Berkeley and Los Angeles. MR0046022

[15] Geiger, J. (1999)
Elementary new proofs of classical limit theorems for Galton–Watson processes. *J. Appl. Prob.* **36** 301–309. MR1724856

[16] Grey, D.R. (1974)
Asymptotic behaviour of continuous-time, continuous state-space branching processes. *J. Appl. Prob.* **11** 669–677. MR0408016

[17] Harris, T.E. (1951)
Some mathematical models for branching processes. *Proc. 2nd Berkeley Symp. Math. Statist. Prob.* 305–328. U. California Press, Berkeley and Los Angeles. MR0045331

[18] Heathcote, C.R., Seneta, E., Vere-Jones, D. (1967)
A refinement of two theorems in the theory of branching processes. *Teor. Verojatnost. i Primenen.* **12** 341–346. MR0217889

[19] Jirina, M. (1958)
Stochastic branching processes with continuous state space. *Czech. Math. J.* **8** 292–312. MR0101554

[20] Kallenberg, O. (1992)
Some time-change representations of stable integrals, via predictable transformations of local martingales. *Stoch. Proc. Appl.* **40** 199–223. MR1158024

[21] Kawazu, K., Watanabe, S. (1971)
Branching processes with immigration and related limit theorems. *Teor. Verojanost i Primenen* **16** 34–51. MR0290475

[22] Kolmogorov, A.N. (1938)
Zur Lösung einer biologischen Aufgabe. *Comm. Math. Mech. Chebyshev Univ. Tomsk* **2** 1–6.

[23] Kyprianou, A.E. (2006)
Introductory Lectures on Fluctuations of Lévy Processes with Applications. Springer-Verlag, Berlin.

[24] Lambert, A. (2001)
Arbres, excursions et processus de Lévy complètement asymétriques. Thèse de doctorat de l'U. P. et M. Curie, Paris.

[25] Lambert, A. (2002)
The genealogy of continuous-state branching processes with immigration. *Probab. Theory Relat. Fields* **122** (1) 42–70. MR1883717

[26] Lambert, A. (2003)
Coalescence times for the branching process. *Adv. Appl. Prob.* **35** (4) 1071–1089. MR2014270

[27] Lamperti, J. (1967)
Continuous-state branching processes. *Bull. Amer. Math. Soc.* **73** 382–386. MR0208685

[28] Lamperti, J., Ney, P. (1968)
Conditioned branching processes and their limiting diffusions. *Theor. Probab. Appl.* **13** 128–139. MR0228073

[29] Li, Z.-H. (2000)
Asymptotic behavior of continuous time and state branching processes. *J. Aus. Math. Soc. Series A* **68** 68–84. MR1727226

[30] Lyons, R., Pemantle, R., Peres, Y. (1995)
Conceptual proofs of $L\log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23** 1125–1138. MR1349164

[31] Overbeck, L. (1993)
Conditioned super-Brownian motion. *Probab. Theory Relat. Fields* **96** 545–570. MR1234623

[32] Overbeck, L. (1994)
Martin boundaries of some branching processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **30** 181–195.

[33] Pakes, A.G. (1998)
Conditional limit theorems for continuous time and state branching processes. Technical report. To appear in *Records and Branching Processes* (M. Ahsanullah and G. Yanev Eds) NOVA Science Publishers Inc.

[34] Pinsky, M.A. (1972)
Limit theorems for continuous state branching processes with immigration. *Bull. Amer. Math. Soc.* **78** 242–244. MR0295450

[35] Pitman, J., Yor, M. (1982)
A decomposition of Bessel bridges. *Z. Wahrsch. verw. Gebiete* **59** 425–457. MR0656509

[36] Pollett, P.K.
Quasi-stationary distributions : a bibliography. Regularly updated, and available at http://www.maths.uq.edu.au/∼pkp/papers/qsds/qsds.pdf

[37] Roelly, S., Rouault, A. (1989)
Processus de Dawson–Watanabe conditionné par le futur lointain. *C. R. Acad. Sci. Paris Ser. I* **309** 867–872.

[38] Seneta, E., Vere-Jones, D. (1966)
On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Prob.* **3** 403–434. MR0207047

[39] Seneta, E., Vere-Jones, D. (1968)
On the asymptotic behaviour of subcritical branching processes with continuous state space. *Z. Wahrsch. verw. Gebiete* **10** 212-Ű225.

[40] Serlet, L. (1996)
The occupation measure of super-Brownian motion conditioned to nonextinction. *J. Theoret. Probab.* **9** 561–578. MR1400587

[41] Winter, A. (2002)
Multiple scale analysis of spatial branching processes under the Palm distribution. *Elec. J. Prob.* **7** Paper 13. MR1921742

[42] Yaglom, A.M. (1947)
Certain limit theorems of the theory of branching stochastic processes (in Russian). *Dokl. Akad. Nauk. SSSR* (n.s.) **56** 795–798. MR0022045

[43] Zanzotto, P.A. (2002)
On stochastic differential equations driven by a Cauchy process and other stable Lévy motions. *Ann. Probab.* **30** 802–825.