SPLITTING TREES STOPPED WHEN THE FIRST CLOCK RINGS AND VERVAAT'S TRANSFORMATION

AMAURY LAMBERT,* *Université Pierre et Marie Curie-Paris 6*

PIETER TRAPMAN,** *Stockholm University*

Abstract

We consider a branching population where individuals have independent and identically distributed (i.i.d.) life lengths (not necessarily exponential) and constant birth rates. We let N_t denote the population size at time t. We further assume that all individuals, at their birth times, are equipped with independent exponential clocks with parameter δ. We are interested in the genealogical tree stopped at the first time T when one of these clocks rings. This question has applications in epidemiology, population genetics, ecology, and queueing theory. We show that, conditional on $\{T < \infty\}$, the joint law of $(N_T, T, X^{(T)})$, where $X^{(T)}$ is the jumping contour process of the tree truncated at time T, is equal to that of $(M, -I_M, Y'_M)$ conditional on $\{M \neq 0\}$. Here $M + 1$ is the number of visits of 0, before some single, independent exponential clock e with parameter δ rings, by some specified Lévy process Y without negative jumps reflected below its supremum; I_M is the infimum of the path Y_M, which in turn is defined as Y killed at its last visit of 0 before e; and Y'_M is the Vervaat transform of Y_M. This identity yields an explanation for the geometric distribution of N_T (see Kitaev (1993) and Trapman and Bootsma (2009)) and has numerous other applications. In particular, conditional on $\{N_T = n\}$, and also on $\{N_T = n, \ T < a\}$, the ages and residual lifetimes of the n alive individuals at time T are i.i.d. and independent of n. We provide explicit formulae for this distribution and give a more general application to outbreaks of antibiotic-resistant bacteria in the hospital.

Keywords: Branching process; splitting tree; Crump–Mode–Jagers process; contour process; Lévy process; scale function; resolvent; age and residual lifetime; undershoot and overshoot; Vervaat's transformation; sampling; detection; epidemiology; processor sharing

2010 Mathematics Subject Classification: Primary 60J80
Secondary 92D10; 92D25; 92D30; 92D40;
60J85; 60G17; 60G51; 60G55;
60K15; 60K25

1. Introduction

We consider a population of particles behaving independently from one another, where each particle gives birth at constant rate $b > 0$ during its lifetime (inter-birth durations are independent and identically distributed (i.i.d.) exponential random variables with parameter b), and where lifetime durations are i.i.d. on $(0, +\infty]$ (some particles may have infinite lifetimes) with probability distribution μ (not necessarily exponential).
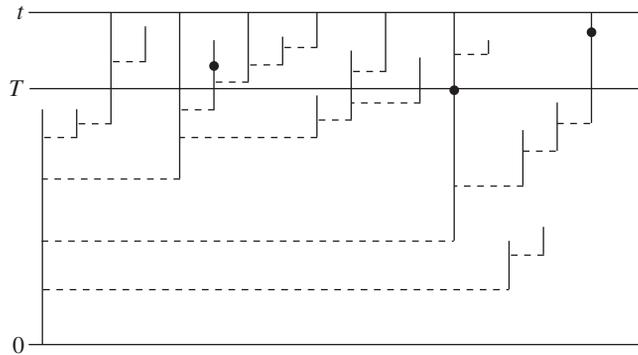
FIGURE 1: A realisation of a splitting tree with individual exponential clocks. Time flows from the bottom to the top; the horizontal dashed lines show filiation. The filled circles show individual ringing clocks. The time T when the first clock rings is indicated. Here $N_T = 7$.

The genealogical trees that we consider here are usually called *splitting trees* [6]. We define the *lifespan measure* as the measure on $(0, +\infty]$ with total mass b simply defined as $\pi := b\mu$.

The process $(N_t; t \geq 0)$, giving the number of extant particles at time t, belongs to a wide class of branching processes called *Crump–Mode–Jagers processes*. Actually, the processes we consider are homogeneous (constant birth rate) and binary (one birth at a time) but are more general than classical (simple) birth–death processes [7] in that the lifetime durations may follow a general distribution.

In addition, we assume that each particle is independently equipped with a random exponential clock with parameter $\delta > 0$. We are interested in the first time T when one of these clocks rings, called the *detection time*. See Figure 1 for a realisation of a splitting tree with individual clocks. Note that, on the extinction event, T can be infinite (no clock rings) with positive probability.

This question has applications in population genetics and ecology [4], [5], [12], [14] (T is then the first time when a new mutant or a new species arises), queueing theory [8], [10], [15] (because N is a time-changed processor-sharing queue, and then in the new timescale, T is a single, independent exponential clock), and epidemiology [3], [16] (T is then the first detection time of the epidemic). In this last setting, ages of individuals in the population at T are the times since infection of infectives in the detected outbreak, and in the final section we see how more easily available data, such as the length of stay in the hospital up to time T, can be incorporated.

Our main result is to characterise, on the event $\{T < \infty\}$, the joint law of $(N_T, T, X^{(T)})$, where $X^{(T)}$ is the jumping contour process of the tree truncated at time T, in terms of the Vervaat transform of the path of the (reflected) Lévy process X with jump measure π and slope -1. In particular, we recover the known fact [10], [16] that, conditional on $\{T < \infty\}$, N_T is geometrically distributed, and we characterise the joint law of T and N_T in terms of (joint) Laplace transforms of some hitting times of X. As a further example, restricting the main identity to the undershoots and overshoots of X whenever it crosses 0, we get the following application. Conditional on $\{N_T = n\}$, and also on $\{N_T = n, T < a\}$, the ages and residual lifetimes of the n alive individuals at time T are i.i.d. and independent of n, and follow the bivariate law of (und, ove) (respectively (und$_a$, ove$_a$)). Here the pair (und, ove) (respectively (und$_a$, ove$_a$)) is the undershoot and overshoot of the jump across 0 of X, at its first hitting time τ_0^+ of $(0, +\infty]$, conditional on τ_0^+ being smaller than some independent exponential time with

parameter δ, respectively, conditional on the same event intersected with $\inf_{0 \le s \le \tau_0^+} X_s > -a$). In the epidemics model, these statements are extended by taking into account, in addition to the age and residual lifetime (of individual infection at time T), the length of stay in the hospital up to infection time. In all cases, explicit formulae are also provided for these laws.

2. Splitting trees and Lévy processes

We assume that splitting trees are started with one unique progenitor born at time 0. We denote by $\mathbb{P}$ their law, and the subscript s in $\mathbb{P}_s$ means conditioning on the lifetime of the progenitor being s. Of course, if $\mathbb{P}$ bears no subscript, this means that the lifetime of the progenitor follows the usual distribution μ.

Lambert [13] considered, for $t > 0$, the so-called jumping chronological contour process (JCCP), here denoted by $X^{(t)}$, of the splitting tree truncated up to height (time) t, which starts at $s \wedge t$ (here and in what follows, $x \wedge y$ denotes the minimum of x and y), where s is the time of death of the progenitor, visits all existence times (smaller than t) of all individuals exactly once, and terminates at 0 (see Figure 2).

Lambert showed in [13, Theorem 4.3] that the JCCP is a Markov process; more specifically, the JCCP has the same law as the compound Poisson process X with jump measure π, compensated at rate -1, reflected below t, and killed upon hitting 0.

We denote the law of X by P, to make a distinction with the law $\mathbb{P}$ of the CMJ process. As seen previously, we record the lifetime duration, say s, of the progenitor, by writing P_s for its conditional law on $X_0 = s$. Using notation analogous to that defined for $\mathbb{P}$, it will be implicit in the absence of a subscript that X_0 under P has probability distribution μ.
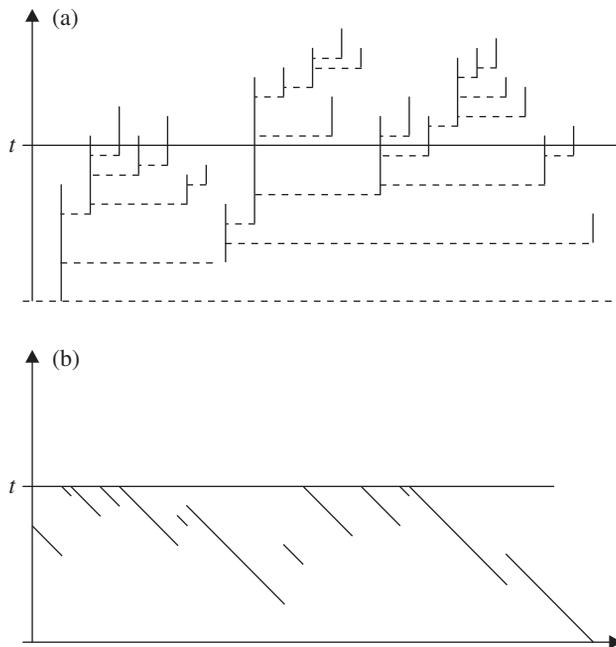
FIGURE 2: (a) A splitting tree. (b) The jumping chronological contour process associated with the same splitting tree after truncation at time t. Here $N_t = 9$.

Let us be a little more specific about the JCCP. Recall that this process visits all existence times of all individuals of the truncated tree. For any individual of the tree, we denote by α its birth time and by ω its death time. When the visit of an individual v with lifespan $(\alpha(v), \omega(v))$ begins, the value of the JCCP is $\omega(v)$. The JCCP then visits all the existence times of v's lifespan at constant speed -1. If v has no child then this visit lasts exactly as long as the lifespan of v; if v has at least one child then the visit is interrupted each time a birth time of one of v's daughters, say w, is encountered (youngest child first since the visit started at the death level). At this point, the JCCP jumps from $\alpha(w)$ to $\omega(w) \wedge t$ and starts the visit of the existence times of w. Since a truncated tree has finite length, the visit of v has to terminate: it does so at the chronological level $\alpha(v)$ and continues the exploration of the existence times of v's mother, at the height (time) where it had been interrupted. This procedure then goes on recursively and terminates as soon as 0 is encountered (the birth time of the progenitor). See Figure 2 for an example.

Note that the genealogy of a splitting tree truncated at time t can be coded by associating each individual with a word of integers, such that $\varnothing$ is the root, 1 is the last daughter of the root born before time t, 2 is the penultimate daughter of the root,…, 11 is the last daughter of 1 (born before time t), and so on. Then the order in which individuals are first visited by the contour process is the lexicographical order associated with this (so-called Ulam–Harris–Neveu) labelling. Roughly speaking, if u and v are two distinct finite words of integers, and h_u and h_v are the first integers in the words u and v, respectively, coming immediately after their longest common prefix, then u comes first in the lexicographical order if and only if $h_u < h_v$. Here we assume that if $h \neq \varnothing$ then $h > \varnothing$.

Since the JCCP is Markovian (as seen earlier, it is a reflected, killed Lévy process), its excursions between consecutive visits of points at height t are i.i.d. excursions of X away from $(t, +\infty]$. Observe in particular that the number of visits of t by X is exactly the number N_t of individuals alive at time t. Therefore, it is easy to see that N_t has a shifted geometric distribution with parameters specified as follows. Let τ_A denote the first hitting time of the set A by X. We also use the shorthand notation

$$\tau_x := \tau_{\{x\}} \quad \text{and} \quad \tau_x^+ := \tau_{(x,+\infty]},$$

that is, τ_x is the first hitting time of x and τ_x^+ is the first hitting time of the open interval $(x, +\infty]$. Then, conditional on the initial progenitor having lived for s units of time, we have

$$\mathbb{P}_s(N_t = 0) = \mathrm{P}_s(\tau_0 < \tau_t^+), \tag{2.1}$$

and, recursively applying the strong Markov property,

$$\mathbb{P}_s(N_t = n \mid N_t \neq 0) = \mathrm{P}_t(\tau_t^+ < \tau_0)^{n-1} \, \mathrm{P}_t(\tau_0 < \tau_t^+). \tag{2.2}$$

Note that the subscript s in the last display is useless. Furthermore, because of the truncation of X at time t, this formula holds even when X drifts to $+\infty$. Also note that the spatial homogeneity of Lévy processes implies that $\mathrm{P}_t(\tau_0 < \tau_t^+) = \mathrm{P}_0(\tau_{-t} < \tau_0^+)$.

In addition, exact formulae can be deduced for (2.1) and (2.2) from the fact that the JCCP is a Lévy process with no negative jumps, using scale functions of the Lévy process X. This part is developed in Section 5.

Later, we see that the population size is not only (conditionally) geometric at fixed times, but also at the first detection time T, using the same decomposition of the contour process into excursions away from $(T, +\infty]$. This decomposition is given in Subsection 3.1.

Subsections 3.2 and 3.3 are devoted to path decompositions providing equalities in law for the whole contour process of the tree stopped at T, involving in particular Vervaat's transformation; see Theorems 3.1, 3.2, and especially 3.3 for the result stated in the abstract. In Section 4 we focus on the joint distribution of T and N_T, as well as the ages and residual lifetimes of the N_T alive individuals at time T. In Subsections 4.2 and 4.3 we provide explicit formulae (up to scale functions of the Lévy process X) for these distributions. The reader interested in applications is directed to the last statement of Section 4, Proposition 4.2. Finally, in Section 5 we extend these results to the example of a pathogen outbreak in the hospital modelled by a Crump–Mode–Jagers process with constant transmission rate b and i.i.d. (infection) lifetimes, but also taking into account the length of stay in the hospital up to infection.

In the rest of the paper we use the following notation: E is any expectation operator, A and B are any events, and Z is any (positive or integrable) random variable $\mathrm{E}(Z, A, B) := \mathrm{E}(Z \, \mathbf{1}_{\{A \cap B\}})$.

2.1. Intuition for the geometric distribution

In this section we show how to gain insight from the equivalence of the splitting tree and the corresponding contour process, as visualised in Figure 2, and, in particular, intuitively explain why the number of individuals at the first detection time is geometrically distributed [10], [16]. This intuition also gives the main ideas behind the rigorous proofs below.

We consider the event that $N_t = n$, $n > 0$, and no detection has occurred up to time t, i.e. the event $\{N_t = n, \ T > t\}$. This event occurs if the following events successively occur.

1. The Lévy process X following the contour of the tree truncated below time t starts with a typical jump (distributed according to μ) and hits the interval $(t, \infty]$ before it hits 0 again, and during this time no clock rings. The probability of this event is $\mathrm{E}(e^{-\delta \tau_t^+}, \tau_t^+ < \tau_0)$.

2. The process X started at t makes an excursion ending in the interval $(t, \infty]$ without hitting 0, and no clock rings during this excursion. Since the contour process of the tree truncated below t is started again at t, independently from the past, $n - 1$ such events occur successively, and each of them occurs independently with probability $\mathrm{E}_t(e^{-\delta \tau_t^+}, \tau_t^+ < \tau_0)$.

3. X starts at t and reaches 0 before hitting the interval $(t, \infty]$, and during this time no clock rings. This happens with probability $\mathrm{E}_t(e^{-\delta \tau_0}, \tau_t^+ > \tau_0)$.

The next step is to 'glue' the path described in event 3 to the path described in event 1, forming one excursion with infimum equal to 0 within which no clock rings. This is basically the inverse of Vervaat's transformation, see Figure 3, where the inverse is constructed in such a way that the infimum of the whole process is performed during this newly created (first) excursion.

Since X jumps at rate b and has slope -1 (and also is translation invariant), multiplying the probability of this concatenated path by $b \, dt$ gives the probability of an excursion of X away from $[0, +\infty)$ without a clock ringing and with infimum in $(-t - dt, -t)$. Then $b\mathbb{P}(N_t = n, T > t) \, dt$ is the probability that X makes n excursions away from $[0, +\infty)$ without making a clock ring, and that the infimum of the whole path is reached in the first excursion and belongs
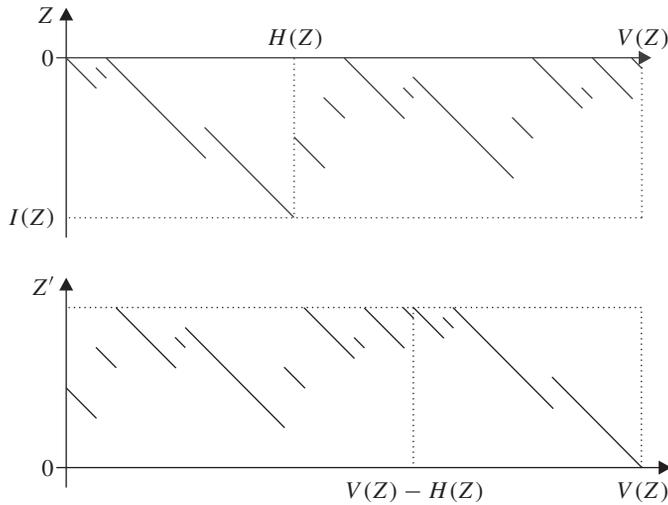
FIGURE 3: Vervaat's transformation. *Top*: a path Z with finite lifetime $V(Z)$ performing its infimum $I(Z)$ at time $H(Z)-$. *Bottom*: Vervaat's transform Z' of Z obtained by shifting Z by $-I(Z)$ and performing a circular time change starting at time $H(Z)$.

to $(-t - \mathrm{d}t, -t)$. This yields

$$
b\mathbb{P}(N_t = n, \ T > t)\,\mathrm{d}t
$$
$$
= \mathrm{E}_0\left(\mathrm{e}^{-\delta\tau_0^+}, \ -\inf_{0 < s < \tau_0^+} X_s \in \mathrm{d}t\right)(\mathrm{E}_0(\mathrm{e}^{-\delta\tau_0^+}, \ \tau_0^+ < \tau_{-t}))^{n-1}
$$
$$
= \mathrm{E}_0\left(\mathrm{e}^{-\delta\tau_0^+}, \ -\inf_{0 < s < \tau_0^+} X_s \in \mathrm{d}t\right)\left(\mathrm{E}_0\left(\mathrm{e}^{-\delta\tau_0^+}, \ -\inf_{0 < s < \tau_0^+} X_s < t\right)\right)^{n-1}
$$
$$
= \frac{1}{n}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{E}_0\left(\mathrm{e}^{-\delta\tau_0^+}, \ -\inf_{0 < s < \tau_0^+} X_s < t\right)\right)^n \mathrm{d}t.
$$

Now observe that

$$
\mathbb{P}(N_t = n, \ T \in \mathrm{d}t) = \delta n\mathbb{P}(N_t = n, \ T > t)\,\mathrm{d}t = \frac{\delta}{b}\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{E}_0(\mathrm{e}^{-\delta\tau_0^+}, \ \tau_0^+ < \tau_{-t}))^n \,\mathrm{d}t.
$$

Integrating over t now gives

$$
\mathbb{P}(N_T = n, \ T < \infty) = \frac{\delta}{b}(\mathrm{E}_0(\mathrm{e}^{-\delta\tau_0^+}))^n.
$$

A little elaboration on this argument also reveals that the distributions of the ages and residual lifetimes at time T should be i.i.d. and independent of N_T. Precise proofs are given in the sections below.

3. Main general results

3.1. Decomposition of the splitting tree at the first detection time

In this subsection we call any càdlàg right continuous with left limits [7, p. 346]) path ϵ with lifetime $V(\epsilon) \in [0, +\infty]$, *an excursion*. We use the notation $\mathcal{E}$ for the space of excursions endowed with Skorokhod's topology and the associated Borel σ-field.

For any time $t > 0$, we set ρ_t as the first time $X^{(t)}$ exits the interval $(0, t)$ and we let w_0^t denote the finite path of the JCCP $X^{(t)}$ killed upon exiting $(0, t)$, that is, $w_0^t := (X_s^{(t)}; s \leq \rho_t)$. If the life length of the progenitor, say L, exceeds t, then w_0^t is reduced to the one-point process that maps 0 to L. Furthermore, on the event $N_t = n \geq 1$, for $i = 1, \ldots, n$, we define σ_i to be the ith hitting time of t by $X^{(t)}$ and we let w_i^t denote the path of the JCCP $X^{(t)}$ between times σ_i and σ_{i+1}, with the conventions that $\sigma_0 = 0$ and $\sigma_{n+1} = \tau_0$, that is,

$$w_i^t(s) := X_{s+\sigma_i}^{(t)}, \qquad s < \sigma_{i+1} - \sigma_i.$$

For $i = 0, \ldots, n$, we denote by $\ell_i := V(w_i^t) = \sigma_{i+1} - \sigma_i$ the lifetime of the excursion w_i^t (so that $w_n^t(\ell_n) = 0$), and, for $i = 0, \ldots, n-1$, we record the size of the jump made by the contour process before reflection by setting $w_i^t(\ell_i)$ equal to the time of death of the individual alive at time t and visited at time σ_{i+1}.

In particular, when t is fixed, we know from [13] that, conditional on $\{N_t = n\}$,

- the excursion w_0^t follows the law of X started at a jump distributed as μ, killed at τ_t^+, and conditioned on $\tau_t^+ < \tau_0$;

- the excursions w_i^t, $i = 1, \ldots, n-1$, are i.i.d. and follow the law of X started at t, killed at τ_t^+, and conditioned on $\tau_t^+ < \tau_0$;

- the excursion w_n^t follows the law of X started at t, killed at τ_0, and conditioned on $\tau_0 < \tau_t^+$.

Now recall that all individuals are equipped with an independent exponential clock with parameter δ, and that the time when the first of these clocks rings is denoted by T and called the detection time.

Proposition 3.1. *Let $p := \mathbb{P}(T < \infty)$ be the probability that at least one clock rings before extinction of the population. Then*

$$p = \mathrm{E}(1 - \mathrm{e}^{-\delta\tau_0}).$$

More specifically,

$$\mathbb{P}(N_t = 0, \ T > t) = \mathbb{P}(N_t = 0, \ T = \infty) = \mathrm{E}(\mathrm{e}^{-\delta\tau_0}, \ \tau_0 < \tau_t^+)$$

and, for any $n \geq 1$,

$$\mathbb{P}(N_t = n, \ T > t) = \mathrm{E}(\mathrm{e}^{-\delta\tau_t^+}, \ \tau_t^+ < \tau_0)(\mathrm{E}_t(\mathrm{e}^{-\delta\tau_t^+}, \ \tau_t^+ < \tau_0))^{n-1} \, \mathrm{E}_t(\mathrm{e}^{-\delta\tau_0}, \ \tau_0 < \tau_t^+).$$

Furthermore, let $n \geq 1$ and $G', G, F_0, F_1, \ldots, F_{n-1}$ be nonnegative, measurable functions on $\mathcal{E}$. Then

$$\mathbb{E}\left(G'(w_t^0)G(w_t^n) \prod_{i=1}^{n-1} F_i(w_t^i), \ N_T = n, \ T \in \mathrm{d}t \right)$$

$$= \delta n \, \mathrm{d}t \, \mathrm{E}(G'(X_s; s \leq \tau_t^+)\mathrm{e}^{-\delta\tau_t^+}, \ \tau_t^+ < \tau_0)$$

$$\times \left\{ \prod_{i=1}^{n-1} \mathrm{E}_t(F_i(X_s; s \leq \tau_t^+)\mathrm{e}^{-\delta\tau_t^+}, \ \tau_t^+ < \tau_0) \right\} \mathrm{E}_t(G(X_s; s \leq \tau_0)\mathrm{e}^{-\delta\tau_0}, \ \tau_0 < \tau_t^+).$$

Proof. Throughout the proof, e denotes an independent, exponential random variable with parameter δ. Since the JCCP has slope -1 and jump sizes equal to life lengths, the lifetime τ_0 of the contour process is exactly the sum of the lifespans of all individuals in the population. As a consequence,

$$p = \mathrm{P}(\tau_0 > e) = \mathrm{E}(1 - \mathrm{e}^{-\delta \tau_0}).$$

Applying this property to the truncated contour process $X^{(t)}$ and using the path decomposition preceding the statement of the proposition, we obtain

$$\mathbb{P}(N_t = n, T > t) = \mathbb{P}(N_t = n, \tau_0(X^{(t)}) < e)$$
$$= \mathbb{E}(\mathrm{e}^{-\delta \tau_0(X^{(t)})}, N_t = n)$$
$$= \mathrm{P}(\tau_t^+ < \tau_0)\,\mathrm{E}(\mathrm{e}^{-\delta V(w_0^t)})(\mathrm{P}_t(\tau_t^+ < \tau_0)\,\mathrm{E}(\mathrm{e}^{-\delta V(w_1^t)}))^{n-1}$$
$$\times \mathrm{P}_t(\tau_0 < \tau_t^+)\,\mathrm{E}(\mathrm{e}^{-\delta V(w_n^t)}),$$

which yields the desired expression. Note that here we have used the fact that w_i^t is defined conditional on $\tau_t^+ > \tau_0$ for $i = 0, 1, \ldots, n-1$ and w_n^t is defined conditional on $\tau_0 > \tau_t^+$.

Observing that the detection rate equals δn conditional on $\{N_t = n\}$, we finally obtain

$$\mathbb{E}\left(G'(w_t^0)G(w_t^n) \prod_{i=1}^{n-1} F_i(w_t^i),\ N_T = n,\ T \in \mathrm{d}t \right)$$
$$= \delta n\, \mathrm{d}t\, \mathbb{E}\left(G'(w_t^0)G(w_t^n) \prod_{i=1}^{n-1} F_i(w_t^i),\ N_t = n,\ T > t \right),$$

and the desired equality follows by the same method as previously.

Remark 3.1. Since the knowledge of the contour of the genealogical tree yields that of the tree itself, the previous proposition characterises the law of the splitting tree stopped at the first detection time (noting that, conditional on $N_T = n$, the marked individual is of course uniform among all n alive individuals).

3.2. Rephrasing with i.i.d. excursions

In this subsection ϵ denotes an excursion *distributed as* $\mathbf{X}$ *started at* 0 *and killed upon hitting* $(0, +\infty]$. Recall that ϵ takes only negative values, except at 0 ($\epsilon(0) = 0$) and at V, since $\epsilon(V) > 0$ on the event $\{V < \infty\}$ (on the complementary event, ϵ drifts to $-\infty$).

Set $J(\epsilon) := \inf_s \epsilon(s)$. On the event $\{J \neq -\infty\}$ (which coincides almost surely with $\{V < \infty\}$), we denote by $h(\epsilon)$ the unique time h such that $\epsilon(h-) = J$. Also, we denote by $\epsilon^{\leftarrow}$ the pre-h process and by $\epsilon^{\rightarrow}$ the post-h process, i.e.

$$\epsilon^{\leftarrow}(s) := \epsilon(s), \qquad 0 \leq s < h(\epsilon),$$

with $\epsilon^{\leftarrow}(h(\epsilon)) = \epsilon^{\leftarrow}(h(\epsilon)-) = J(\epsilon)$ and

$$\epsilon^{\rightarrow}(s) := \epsilon(s + h(\epsilon)), \qquad 0 \leq s \leq V(\epsilon) - h(\epsilon).$$

Note that, with positive probability, $V(\epsilon) = h(\epsilon)$, so that $\epsilon^{\rightarrow}$ is then reduced to the one-point process that maps 0 to $\epsilon(V)$.

Let $n \geq 1$, and let $\epsilon_1, \ldots, \epsilon_n$ denote i.i.d. excursions distributed as ϵ. Set

$$I_n := \min_k J(\epsilon_k) \quad \text{and} \quad K_n := \arg\min_k J(\epsilon_k).$$

The next result is a consequence of the following two lemmas and Proposition 3.1.

Theorem 3.1. *Let G, G', $F_1, \ldots, F_n$ be nonnegative, measurable functions on $\mathcal{E}$. Then*

$$\mathbb{E}\left(G'(w_t^0)G(w_t^n) \prod_{i=1}^{n-1} F_i(w_t^i),\ N_T = n,\ T \in \mathrm{d}t \right)$$

$$= \frac{\delta}{b} \mathbb{E}\left(G(\epsilon_{K_n}^{\leftarrow} + t)G'(\epsilon_{K_n}^{\rightarrow} + t) \prod_{k \neq K_n} F_{k-K_n \bmod(n)}(\epsilon_k + t) \prod_{k=1}^{n} \mathrm{e}^{-\delta V(\epsilon_k)},\ -I_n \in \mathrm{d}t \right).$$

$$(3.1)$$

Remark 3.2. Note that the expression inside the expectation on the right-hand side of (3.1) has zero probability when one of the excursions has infinite lifetime, that is, when there is some k such that $V(\epsilon_k) = +\infty$.

Lemma 3.1. *Let G and G' be two nonnegative, measurable functions on $\mathcal{E}$. Then*

$$\mathrm{E}(G(\epsilon^{\leftarrow})G'(\epsilon^{\rightarrow}),\ -J \in \mathrm{d}t)$$
$$= b\,\mathrm{d}t\,\mathrm{E}_0(G(X_s; s \leq \tau_{-t}),\ \tau_{-t} < \tau_0^+)\,\mathrm{E}(G'(X_s - t; s \leq \tau_t^+),\ \tau_t^+ < \tau_0).$$

Proof. Applying the strong Markov property at τ_{-t} yields

$$\mathrm{E}(G(\epsilon^{\leftarrow})G'(\epsilon^{\rightarrow}),\ -J \in \mathrm{d}t) = \mathrm{E}(G(\epsilon^{\leftarrow})G'(\epsilon^{\rightarrow}),\ -J \in \mathrm{d}t,\ \tau_{-t} < \tau_0^+)$$
$$= \mathrm{E}_0(G(X_s; s \leq \tau_{-t}),\ \tau_{-t} < \tau_0^+)$$
$$\times \mathrm{d}t \int_{(0,+\infty]} \pi(\mathrm{d}y)\,\mathrm{E}_{y-t}(G'(X_s; s \leq \tau_0^+),\ \tau_0^+ < \tau_{-t}),$$

which yields the result.

Lemma 3.2. *Let $G, G', F_1, \ldots, F_n$ be nonnegative, measurable functions on $\mathcal{E}$. Then, for $j = 1, \ldots, n$,*

$$\mathrm{E}\left(G(\epsilon_{K_n}^{\leftarrow})G'(\epsilon_{K_n}^{\rightarrow}) \prod_{k \neq K_n} F_k(\epsilon_k),\ -I_n \in \mathrm{d}t,\ K_n = j \right)$$

$$= b\,\mathrm{d}t\,\mathrm{E}_0(G(X_s; s \leq \tau_{-t}),\ \tau_{-t} < \tau_0^+)\,\mathrm{E}(G'(X_s - t; s \leq \tau_t^+),\ \tau_t^+ < \tau_0)$$
$$\times \prod_{k \neq j} \mathrm{E}_0(F_k(X_s; s \leq \tau_0^+),\ \tau_0^+ < \tau_{-t}).$$

$$(3.2)$$

Proof. The expression on the left-hand side of (3.2) equals

$$\mathrm{E}(G(\epsilon_j^{\leftarrow})G'(\epsilon_j^{\rightarrow}),\ -J(\epsilon_j) \in \mathrm{d}t)\,\mathrm{E}\left(\prod_{k \neq j} F_k(\epsilon_k),\ -J(\epsilon_k) < t \text{ for all } k \neq j \right)$$

$$= \mathrm{E}(G(\epsilon_j^{\leftarrow})G'(\epsilon_j^{\rightarrow}),\ -J(\epsilon_j) \in \mathrm{d}t) \prod_{k \neq j} \mathrm{E}_0(F_k(X_s; s \leq \tau_0^+),\ \tau_0^+ < \tau_{-t}),$$

and the conclusion follows from the previous lemma.

3.3. Rephrasing with Vervaat's transformation

Forgetting about the terminal jump of each excursion (a piece of information that is actually useful in the next section), Theorem 3.1 can be expressed in a more elegant way.

For any càdlàg path Z with finite lifetime $V(Z)$ and law locally absolutely continuous with respect to X, we set $I(Z) := \inf Z$ and define $H(Z)$ as the unique time t such that $Z(t-) = I(Z)$. Finally, we let Z' denote Vervaat's transform of Z, defined as the path with lifetime $V(Z)$ such that $Z'(V(Z)) = 0$ and

$$Z'(s) = Z(s + H(Z) \bmod (V(Z))) - I(Z), \qquad 0 \le s < V(Z).$$

More specifically,

$$Z'(s) := \begin{cases} Z(s + H(Z)) - I(Z) & \text{if } 0 \le s < V(Z) - H(Z), \\ Z(s + H(Z) - V(Z)) - I(Z) & \text{if } V(Z) - H(Z) \le s < V(Z), \\ Z'(s) = 0 & \text{if } s = V(Z). \end{cases}$$

Note that Z' takes positive values, apart from its terminal value equal to 0 (and that Z' is left continuous at this point).

Now let Y_n denote the concatenation of the n i.i.d. excursions $(\epsilon_i)_{i=1,\dots,n}$ of the last subsection. In particular, Y_n is equally distributed as the Lévy process X reflected below its supremum and killed at its $(n+1)$th hitting time of 0. Then observe that $I_n = I(Y_n)$ and let Y'_n denote Vervaat's transformation of Y_n. We have the following corollary of Theorem 3.1.

Theorem 3.2. *For any $n \ge 1$ and $t > 0$, and any excursion ε,*

$$\mathbb{P}(N_T = n, \, T \in \mathrm{d}t, \, X^{(T)} \in \mathrm{d}\varepsilon) = \frac{\delta}{b} \mathrm{e}^{-\delta V(\varepsilon)} \, \mathrm{P}(-I_n \in \mathrm{d}t, \, Y'_n \in \mathrm{d}\varepsilon).$$

Proof. From Theorem 3.1 we obtain

$$\mathbb{E}\left(G'(w_t^0) G(w_t^n) \prod_{i=1}^{n-1} F_i(w_t^i), \, N_T = n, \, T \in \mathrm{d}t \right)$$

$$= \frac{\delta}{b} \mathbb{E}\left(G(\overset{\leftarrow}{\epsilon_{K_n}} - I_n) G'(\overset{\rightarrow}{\epsilon_{K_n}} - I_n) \prod_{k \ne K_n} F_{k-K_n \bmod(n)}(\epsilon_k - I_n) \mathrm{e}^{-\delta V(Y'_n)}, \, -I_n \in \mathrm{d}t \right),$$

which, by a monotone class theorem [9, p. 2], ensures that, for any nonnegative, measurable function F on $\mathscr{E}$,

$$\mathbb{E}(F(X^{(t)}), \, N_T = n, \, T \in \mathrm{d}t) = \frac{\delta}{b} \mathbb{E}(F(Y'_n) \mathrm{e}^{-\delta V(Y'_n)}, \, -I_n \in \mathrm{d}t),$$

completing the proof. $\qquad\blacksquare$

One can now push this path decomposition even further by starting from a path, say Y, of X reflected below its supremum as before, but not stopped at the $(n+1)$th hitting time of 0. Furthermore, use an independent, exponential random variable e with parameter δ. Note that Y is the mere concatenation of a sequence $(\epsilon_i)_{i \ge 1}$ of i.i.d. excursions distributed as ϵ (stopped at the first one with infinite lifetime). Then let M be the unique nonnegative integer such that e falls into the $(M+1)$th excursion of Y away from 0, i.e.

$$M := \max\{n \ge 0 : V(\epsilon_1) + \cdots + V(\epsilon_n) < e\},$$

with the usual convention that an empty sum is 0. As previously, define Y_M as the path Y killed at its $(M + 1)$th hitting time of 0, set $I_M := \inf Y_M$, and let Y'_M denote Vervaat's transform of Y_M.

Theorem 3.3. *For any $n \geq 1$ and $t > 0$, and any excursion ε,*

$$\mathbb{P}(N_T = n, \ T \in dt, \ X^{(t)} \in d\varepsilon) = \frac{\delta}{b \, E_0(1 - e^{-\delta \tau_0^+})} \, P(M = n, \ -I_M \in dt, \ Y'_M \in d\varepsilon).$$

Proof. Using the definition of M, we obtain

$$
\begin{aligned}
P(M = n, \ &-I_M \in dt, \ Y'_M \in d\varepsilon) \\
&= P(V(Y'_n) < e < V(Y'_n) + V(\epsilon_{n+1}), \ -I_n \in dt, \ Y'_n \in d\varepsilon) \\
&= e^{-\delta V(\varepsilon)} \, E(1 - e^{-\delta V(\epsilon_{n+1})}) \, P(-I_n \in dt, \ Y'_n \in d\varepsilon) \\
&= e^{-\delta V(\varepsilon)} \, E_0(1 - e^{-\delta \tau_0^+}) \, P(-I_n \in dt, \ Y'_n \in d\varepsilon),
\end{aligned}
$$

and an appeal to Theorem 3.2 yields the result, where we have used the fact that, on the event $Y'_n \in \varepsilon$, $V(Y'_n) = V(\varepsilon)$ is deterministic. □

4. Applications and explicit formulae

4.1. Some lower-dimensional marginals of interest

In this subsection we give the joint law of T and N_T, as well as the joint law of the ages and residual lifetimes $(A_1, R_1, \ldots, A_{N_T}, R_{N_T})$ of the N_T alive individuals at time T on the event $\{T < \infty\}$.

The next statement follows from Theorem 3.1 by taking all functionals equal to 1. Note that $E_0(e^{-\delta \tau_0^+}, \tau_0^+ < \tau_{-t})$ can be read as $E_0(e^{-\delta \tau_0^+}, -\inf_{0 \leq s \leq \tau_0^+} X_s < t)$, so it is differentiable with derivative equal to $E_0(e^{-\delta \tau_0^+}, -\inf_{0 \leq s \leq \tau_0^+} X_s \in dt)/dt$.

Corollary 4.1. *Let $n \geq 1$ and $y, t > 0$. The joint law of T and N_T is given by*

$$
\begin{aligned}
\mathbb{P}(N_T = n, \ &T \in dt) \\
&= \frac{\delta}{b} \, E\left(\prod_{k=1}^{n} e^{-\delta V(\epsilon_k)}, \ -I_n \in dt \right) \\
&= \frac{n\delta}{b} \, E_0\left(e^{-\delta \tau_0^+}, \ -\inf_{0 \leq s \leq \tau_0^+} X_s \in dt \right) (E_0(e^{-\delta \tau_0^+}, \ \tau_0^+ < \tau_{-t}))^{n-1}.
\end{aligned}
$$

Integrating the variable t over $(0, y)$ yields

$$\mathbb{P}(N_T = n, \ T < y) = \frac{\delta}{b} (E_0(e^{-\delta \tau_0^+}, \ \tau_0^+ < \tau_{-y}))^n,$$

and, finally, letting $y \to \infty$, we obtain

$$\mathbb{P}(N_T = n) = \frac{\delta}{b} (E_0(e^{-\delta \tau_0^+}))^n.$$

The next statement follows from Theorem 3.1 by reducing the functionals to functions of the bivariate random variable $(-\epsilon(V(\epsilon)-), \epsilon(V(\epsilon)))$, known as the undershoot and overshoot

of ϵ at its first upcrossing of the x-axis. Indeed, recall that the age A_i and the residual lifetime R_i of the ith individual in the population at time t in the order of the contour are seen directly on the JCCP as the undershoot and overshoot of w_i^t across t (which occurs at time $V(w_i^t)$), with terminal value equal to the date of death of this ith individual). We use the notation $\mathrm{und}(\epsilon) := -\epsilon(V(\epsilon)-)$ and $\mathrm{ove}(\epsilon) = \epsilon(V(\epsilon))$ for the undershoot and overshoot of ϵ.

Corollary 4.2. *The joint law of the ages and residual lifetimes* $(A_1, R_1, \ldots, A_{N_T}, R_{N_T})$ *of the* N_T *alive individuals at time* T *is given by*

$$\mathbb{P}(N_T = n, \ T \in \mathrm{d}t, \ A_i \in \mathrm{d}a_i, \ R_i \in \mathrm{d}r_i, \ i = 1, \ldots, n)$$

$$= \frac{\delta}{b} \mathrm{E}\Big(\prod_{k=1}^{n} \mathrm{e}^{-\delta V(\epsilon_k)}, \ -I_n \in \mathrm{d}t, \ \mathrm{und}(\epsilon_{i-1+K_n \bmod(n)}) \in \mathrm{d}a_i,$$

$$\mathrm{ove}(\epsilon_{i-1+K_n \bmod(n)}) \in \mathrm{d}r_i, \ i = 1, \ldots, n\Big)$$

$$= \frac{n\delta}{b} \mathrm{E}_0\Big(\mathrm{e}^{-\delta \tau_0^+}, \ -\inf_{0 \le s \le \tau_0^+} X_s \in \mathrm{d}t, \ -X_{\tau_0^+-} \in \mathrm{d}a_1, \ X_{\tau_0^+} \in \mathrm{d}r_1\Big)$$

$$\times \prod_{k=2}^{n} \mathrm{E}_0(\mathrm{e}^{-\delta \tau_0^+}, \ -X_{\tau_0^+-} \in \mathrm{d}a_k, \ X_{\tau_0^+} \in \mathrm{d}r_k, \ \tau_0^+ < \tau_{-t}).$$

Recall from Theorem 3.3 (or observe from the last statement) that A_1 and R_1 are the undershoot and overshoot of the excursion where the infimum is reached. In order to lose this information (which is certainly not available to the observers at the beginning of the epidemic), we reshuffle the labels of the individuals at T, on the event $\{T < \infty, N_T = n\}$, by drawing independently a uniform permutation ς on $\{1, \ldots, n\}$ and setting

$$(A_i', R_i') := (A_{\varsigma(i)}, R_{\varsigma(i)}), \qquad i = 1, \ldots, n.$$

The first equality in the next statement is a mere reformulation of Corollary 4.2 using the previous definition. The integration part comes from the same argument as that mentioned before Corollary 4.1, i.e. by writing the event $\{\tau_0^+ < \tau_{-t}\}$ in the form $\{-\inf_{0 \le s \le \tau_0^+} X_s < t\}$.

Corollary 4.3. *The joint law of the (reshuffled) ages and residual lifetimes* $(A_1', R_1', \ldots, A_{N_T}', R_{N_T}')$ *of the* N_T *alive individuals at time* T *is given by*

$$\mathbb{P}(N_T = n, \ T \in \mathrm{d}t, \ A_i' \in \mathrm{d}a_i, \ R_i' \in \mathrm{d}r_i, \ i = 1, \ldots, n)$$

$$= \frac{\delta}{b} \sum_{i=1}^{n} \mathrm{E}_0\Big(\mathrm{e}^{-\delta \tau_0^+}, \ -\inf_{0 \le s \le \tau_0^+} X_s \in \mathrm{d}t, \ -X_{\tau_0^+-} \in \mathrm{d}a_i, \ X_{\tau_0^+} \in \mathrm{d}r_i\Big)$$

$$\times \prod_{k \ne i} \mathrm{E}_0(\mathrm{e}^{-\delta \tau_0^+}, \ -X_{\tau_0^+-} \in \mathrm{d}a_k, \ X_{\tau_0^+} \in \mathrm{d}r_k, \ \tau_0^+ < \tau_{-t}).$$

Integrating the variable t *over* $(0, y)$ *yields*

$$\mathbb{P}(N_T = n, \ T < y, \ A_i' \in \mathrm{d}a_i, \ R_i' \in \mathrm{d}r_i, \ i = 1, \ldots, n)$$

$$= \frac{\delta}{b} \prod_{k=1}^{n} \mathrm{E}_0(\mathrm{e}^{-\delta \tau_0^+}, \ -X_{\tau_0^+-} \in \mathrm{d}a_k, \ X_{\tau_0^+} \in \mathrm{d}r_k, \ \tau_0^+ < \tau_{-y}),$$

and, finally, letting $y \to \infty$, we obtain

$$\mathbb{P}(N_T = n, \; A'_i \in \mathrm{d}a_i, \; R'_i \in \mathrm{d}r_i, \; i = 1, \dots, n)$$

$$= \frac{\delta}{b} \prod_{k=1}^{n} \mathrm{E}_0(\mathrm{e}^{-\delta \tau_0^+}, \; -X_{\tau_0^+ -} \in \mathrm{d}a_k, \; X_{\tau_0^+} \in \mathrm{d}r_k).$$

Remark 4.1. We observe that, conditional on $N_T = n$ and/or conditional on $N_T = n$ and $T < y$, the random pairs (A_i, R_i), $1 \leq i \leq n$, are i.i.d. and their common distribution does not depend on n.

4.2. Completely asymmetric Lévy processes

In this subsection we seek to provide the reader with more explicit formulae regarding the quantities considered in Subsection 4.1, taking advantage of background knowledge on Lévy processes. Except for the proposition stated at the end of this subsection, all results stated here and the references to their original contributors can be found in [1] and [2].

Instead of the jump measure π of the Lévy process X with no negative jumps, it can be convenient to handle its Laplace exponent ψ defined as

$$\psi(a) := a - \int_{(0,+\infty]} \pi(\mathrm{d}x)(1 - \mathrm{e}^{-ax}), \qquad a \geq 0.$$

Recall that the real number $\pi(\{\infty\})$ can be positive, since particles may have infinite lifetimes. It is also the killing rate of X. The function ψ is differentiable and convex, and we denote by η its largest root. Then ψ is increasing on $[\eta, +\infty)$ and we denote by ϕ its inverse mapping on this set. Furthermore, the so-called two-sided exit problem (exit of an interval from the bottom or from the top by X) has a simple solution of the form

$$\mathrm{P}_s(\tau_0 < \tau_t^+) = \frac{W(t-s)}{W(t)}, \tag{4.1}$$

where the so-called scale function W is the nonnegative, nondecreasing, differentiable function such that $W(0) = 1$, characterised by its Laplace transform

$$\int_0^\infty \mathrm{d}x \, \mathrm{e}^{-ax} W(x) = \frac{1}{\psi(a)}, \qquad a > \eta.$$

Equation (4.1) gives the probability that X exits the interval $(0, t]$ from the bottom. The following formula gives the Laplace transform of the first exit time $\rho_t := \tau_0 \wedge \tau_t^+$ on this event. For any $q > 0$,

$$\mathrm{E}_s(\mathrm{e}^{-q\rho_t}, \; \tau_0 < \tau_t^+) = \frac{W^{(q)}(t-s)}{W^{(q)}(t)},$$

where the so-called q-scale function $W^{(q)}$ is the nonnegative, nondecreasing, differentiable function such that $W^{(q)}(0) = 1$, characterised by its Laplace transform

$$\int_0^\infty \mathrm{d}x \, \mathrm{e}^{-ax} W^{(q)}(x) = \frac{1}{\psi(a) - q}, \qquad a > \phi(q).$$

The q-resolvent of the process killed upon exiting $(0, t]$ is given by the formula ($s, y \in (0, t]$)

$$u_t^q(s, y) := \mathrm{E}_s\left(\int_0^{\rho_t} \mathrm{e}^{-qv} \mathbf{1}_{\{X_v \in \mathrm{d}y\}} \, \mathrm{d}v\right) \Big/ \mathrm{d}y = \frac{W^{(q)}(t-s)W^{(q)}(y)}{W^{(q)}(t)} - \mathbf{1}_{\{y>s\}} W^{(q)}(y-s). \tag{4.2}$$

We also need the q-resolvent of the process killed upon exiting $(-\infty, 0]$ ($s, y \geq 0$):

$$u^q(s, y) := \mathrm{E}_{-s}\left(\int_0^{\tau_0^+} \mathrm{e}^{-qv} \mathbf{1}_{\{-X_v \in \mathrm{d}y\}} \, \mathrm{d}v\right) \Big/ \mathrm{d}y = \mathrm{e}^{-\phi(q)y} W^{(q)}(s) - \mathbf{1}_{\{s > y\}} W^{(q)}(s - y).$$
(4.3)

Lastly, we have the following expression for the bivariate law of the undershoot and overshoot on the event that the process exits $(0, t]$ from the top

$$\mathrm{E}_s(\mathrm{e}^{-q\rho_t}, \tau_t^+ < \tau_0, X_{\rho_t-} \in \mathrm{d}y, X_{\rho_t} - X_{\rho_t-} \in \mathrm{d}z)$$
$$= u_t^q(s, y) \, \mathrm{d}y \pi(\mathrm{d}z), \qquad z + y > t, \; y \in (0, t);$$
(4.4)

an analogous result holds for the exit from $(-\infty, 0]$.

The next statement deals with the following quantities of interest in relation to Subsection 4.1. For any $t > 0$ and $q \geq 0$, set

$$G_q(t) := 1 - \mathrm{E}_0(\mathrm{e}^{-q\tau_0^+}, \tau_0^+ < \tau_{-t}).$$

In particular, as $t \to \infty$, $G_q(t)$ converges to $G_q(\infty) := \mathrm{E}_0(1 - \mathrm{e}^{-q\tau_0^+})$.

Proposition 4.1. *For any $q, r \geq 0$ and $0 < a < t$,*

$$\mathrm{E}_0(\mathrm{e}^{-q\tau_0^+}, -X_{\tau_0^+-} \in \mathrm{d}a, X_{\tau_0^+} \in \mathrm{d}r, \tau_0^+ < \tau_{-t}) = \frac{W^{(q)}(t - a)}{W^{(q)}(t)} \, \mathrm{d}a \, \mathrm{d}\pi(a + r)$$

and

$$\mathrm{E}_0(\mathrm{e}^{-q\tau_0^+}, -X_{\tau_0^+-} \in \mathrm{d}a, X_{\tau_0^+} \in \mathrm{d}r) = \mathrm{e}^{-\phi(q)a} \, \mathrm{d}a \, \mathrm{d}\pi(a + r).$$

For any $q, t \geq 0$,

$$G_q(t) = \frac{1 + q \int_0^t W^{(q)}(s) \, \mathrm{d}s}{W^{(q)}(t)}$$

and

$$\phi(q) = q + b \, \mathrm{E}(1 - \mathrm{e}^{-q\tau_0}) = \frac{q}{G_q(\infty)}.$$

Proof. The first two displays follow from evaluating (4.2) or (4.3), respectively, and (4.4) at $s = t$ or $s = 0$, respectively, using the spatial homogeneity of Lévy processes.

Writing $\bar{\pi}(x) := \pi((x, +\infty])$, $x > 0$, from the first display we obtain

$$\mathrm{E}_0(\mathrm{e}^{-q\tau_0^+}, -X_{\tau_0^+-} \in \mathrm{d}a, \tau_0^+ < \tau_{-t}) = \frac{W^{(q)}(t - a)}{W^{(q)}(t)} \, \mathrm{d}a \bar{\pi}(a),$$

so that

$$1 - G_q(t) = \frac{g_q(t)}{W^{(q)}(t)},$$

where

$$g_q(t) := \int_0^t \mathrm{d}a \, W^{(q)}(t - a)\bar{\pi}(a), \qquad t \geq 0.$$

The third display is proved as in Theorem 8.1(iii) of [11] (note that there the process X is mirrored in the horizontal axis).

The last two equalities are classical results in fluctuation theory of Lévy processes [1]. To be more specific, the first equality is the well-known fact that the inverse mapping of the Laplace exponent of a Lévy process without negative jumps is the Laplace exponent of its dual ladder time process. Since $q \mapsto G_q(\infty) = \mathrm{E}_0(1 - \mathrm{e}^{-q\tau_0^+})$ is the ladder time process of X, the Wiener–Hopf factorisation yields the second equality (which could also be proved in the same fashion as the third display, using the second display).

4.3. Summary statement with explicit formulae

The analytical results of Proposition 4.1 can be applied straightforwardly to rephrase the conceptual results of Subsection 4.1 at the preference of the reader. The next statement is one of the practical ways of doing this. It provides explicit formulae, up to the knowledge (or numerical computation) of δ-scale functions (occasionally via G_δ, but then use Proposition 4.1) and ϕ (which is fast to compute as the inverse mapping of ψ) for various marginals of interest of the splitting tree stopped when the first clock rings.

Proposition 4.2. *Let $n \geq 1$ and $y, t > 0$. The joint law of T and N_T is given by*

$$\mathbb{P}(N_T = n, \ T \in \mathrm{d}t) = -\frac{n\delta}{b} G'_\delta(t)(1 - G_\delta(t))^{n-1} \, \mathrm{d}t.$$

As a consequence,

$$\mathbb{P}(N_T = n, \ T < y) = \frac{\delta}{b}(1 - G_\delta(y))^n,$$

with respective one-dimensional marginals

$$\mathbb{P}(N_T = n) = \frac{\delta}{b}(1 - G_\delta(\infty))^n = \frac{\delta}{b}\left(1 - \frac{\delta}{\phi(\delta)}\right)^n$$

$$and \quad \mathbb{P}(T < y) = \frac{\delta}{b}\frac{1 - G_\delta(y)}{G_\delta(y)}.$$

In particular,

$$\mathbb{P}(N_T = n \mid T = t) = n G_\delta^2(t)(1 - G_\delta(t))^{n-1}$$

$$and \quad \mathbb{P}(N_T = n \mid T < y) = G_\delta(y)(1 - G_\delta(y))^{n-1}.$$

Also, the probability p that $T < \infty$ equals

$$\frac{\delta}{b}\frac{1 - G_\delta(\infty)}{G_\delta(\infty)} = \frac{\phi(\delta) - \delta}{b}.$$

Conditional on $\{N_T = n, \ T < y\}$, $y \leq \infty$, the ages and residual lifetimes of the n alive individuals at time T are i.i.d., distributed as the random variable $(A(y), R(y))$ (independent of n). If $y < \infty$,

$$\mathbb{P}(A(y) \in \mathrm{d}a, \ R(y) \in \mathrm{d}r) = \frac{1}{1 - G_\delta(y)}\frac{W^{(\delta)}(y - a)}{W^{(\delta)}(y)} \, \mathrm{d}a \, \mathrm{d}\pi(a + r)$$

$$= \frac{W^{(\delta)}(y - a)}{W^{(\delta)}(y) - 1 - \delta \int_0^y W^{(\delta)}(s) \, \mathrm{d}s} \, \mathrm{d}a \, \mathrm{d}\pi(a + r).$$

If $y = \infty$,

$$\mathbb{P}(A(\infty) \in \mathrm{d}a, \ R(\infty) \in \mathrm{d}r) = \frac{1}{1 - G_\delta(\infty)}\mathrm{e}^{-\phi(\delta)a} \, \mathrm{d}a \, \mathrm{d}\pi(a + r)$$

$$= \frac{\phi(\delta)}{\phi(\delta) - \delta}\mathrm{e}^{-\phi(\delta)a} \, \mathrm{d}a \, \mathrm{d}\pi(a + r).$$

5. A (more general) model of epidemics

As in [3] and [16], we aim to model the spread of some antibiotic resistant bacteria, such as MRSA (methicillin-resistant *Staphylococcus aureus*) in a hospital. Once in a while, a patient is colonised by MRSA (presumably by introduction from outside) and this may cause an outbreak in the hospital. The first time T that the outbreak is detected occurs either when the first symptoms appear in a carrier or at the first positive medical exam of a carrier.

We make the following assumptions.

- Patients have i.i.d. lengths of stay in the hospital, all distributed as some positive random variable K with finite expectation.

- The outbreak starts with the infection of a randomly chosen patient.

- The length of stay is not influenced by whether or not an individual carries MRSA (neutrality, or exchangeability assumption).

- During an outbreak no further introductions from outside occur (no immigration).

- Carriers are infective from the first time they were infected till their departure from the hospital.

- While infective, patients independently transmit MRSA to other individuals at times of a Poisson process with parameter b (susceptible individuals are always assumed to be in excess, so that effects of the finite size of the hospital are ignored).

- As a consequence of the renewal theorem (assuming stationarity of the regenerative set of arrivals at the hospital), the length of stay of a patient conditional on infection is a size-biased version of K, and the time at which he/she is infected is independent, uniformly distributed during his/her stay.

- Each patient is classified as a carrier only after an independent, exponential time with parameter δ running from the beginning of his/her infection (time of screening or of developing symptoms). The first time T when a carrier is detected is called the *detection time*.

- At the detection time, all patients in the hospital are screened with a perfect test, so all carriers at T are immediately identified.

Remark 5.1. The second assumption above can be disputable, since MRSA is often introduced by a patient who already carries MRSA before entering the hospital (personal communication with Martin Bootsma). Changing this assumption on introduction of MRSA for a more realistic one would make the analysis harder, although possible, and obscure the illustrative character of the example provided in this section.

It is not possible to obtain useful data from patients who have already left the hospital at the moment of detection. Indeed, most carriers leaving the hospital will soon lose MRSA because—in the absence of antibiotic pressure—the antibiotic resistant strains will soon be outcompeted by antibiotic susceptible strains. Thus, our goal is to infer the parameters of the epidemics by using available medical data belonging to the detected carriers.

The model is thus a Crump–Mode–Jagers branching process where every birth event is interpreted as an infection, and individuals are endowed with i.i.d. bivariate random variables distributed as the pair (U, V), with V the lifetime (as an infective), i.e. the time between

infection and departure from the hospital, and U the time already spent in the hospital before infection. Individuals 'give birth' at a constant rate b during their (infective) lifetime (length V) to copies of themselves. Finally, the joint law of (U, V) is given by

$$\mathbb{E}(f(U, V)) = m^{-1} \int_{(0,\infty)} \mathbb{P}(K \in dz) \int_{(0,z)} dx f(x, z - x), \qquad (5.1)$$

where $m := \mathbb{E}(K)$ and f is any nonnegative Borel function.

At detection time T, all carriers $i = 1, \ldots, N_T$ are identified and we focus on the following medical data belonging to them:

- U_i is the time already spent in the hospital by carrier i upon his/her infection;

- A_i is the time elapsed between infection of carrier i and T ('age' of the infection);

- R_i is the remaining length of stay of carrier i in the hospital after T ('residual lifetime' of the infection);

- $V_i := A_i + R_i$ is the total infective lifetime of carrier i;

- $H_i := U_i + A_i$ is the time elapsed between entrance to the hospital of carrier i and time T.

Note that (U_i, V_i) is merely the typical pair (U, V) attached to carrier i, and that A_i and R_i have the interpretations given in the previous section. The quantities of empirical interest are the random variables H_i, which should be easy to obtain from the hospital administrations. Also, the distribution of K should be easy to estimate from hospital data.

It is not difficult to see that, with this extra information, Proposition 4.2 still holds with μ (and, hence, $\psi, \phi, W^{(\delta)}, \ldots$) defined thanks to (5.1) as

$$\mu(dx) := \mathbb{P}(V \in dx) = m^{-1}\mathbb{P}(K > x) \, dx. \qquad (5.2)$$

Corollary 5.1. *Conditional on $\{N_T = n\}$, the triples (U_i, A_i, R_i) of the n (randomly labelled) carriers at time T are i.i.d., distributed as the random variable (U, A, R) (independent of n), where*

$$\mathbb{E}(f(U, A, R)) = \frac{b}{m} \frac{\phi(\delta)}{\phi(\delta) - \delta} \int_{u=0}^{\infty} du \int_{a=0}^{\infty} da \int_{z=u+a}^{\infty} \mathbb{P}(K \in dz) e^{-\phi(\delta)a} f(u, a, z - u - a).$$

In particular, the times $H_i = U_i + A_i$ spent in the hospital up to time T are i.i.d., distributed as the random variable H:

$$\mathbb{P}(H \in dy) = \frac{b/m}{\phi(\delta) - \delta}\mathbb{P}(K > y)(1 - e^{-\phi(\delta)y}) \, dy. \qquad (5.3)$$

The only novelty here is that individuals are not only endowed with a lifetime V and an exponential clock, they are now endowed with a time U previously spent at the hospital. Now the epidemics model is constructed only from the Vs and independent Poisson processes of transmission. So, the time U_i associated to the ith detected individual depends on the epidemics only through V_i. Therefore, Proposition 4.2 gives the law of the (i.i.d.) pairs (A_i, R_i) and then, conditional on all the (A_i, R_i)s, the U_is are independent and the conditional distribution of U_i is the distribution of U conditional on $V = A_i + R_i$. As a consequence, the triples (U_i, A_i, R_i) are i.i.d., and their common law is given by

$$\mathbb{P}(U_i \in dx, A_i \in da, R_i \in dr) = \mathbb{P}(U \in dx \mid V = a + r)\mathbb{P}(A(\infty) \in da, R(\infty) \in dr),$$

then we use (5.1), (5.2), and the last line of Proposition 4.2 to obtain the first display of the corollary. The second display arises from an integration.

Remark 5.2. From the definition of $\phi(a)$ we deduce that

$$\delta = \phi(\delta) - b \int_0^\infty \mu(\mathrm{d}x)(1 - \mathrm{e}^{-\phi(\delta)x}) \quad \Longleftrightarrow \quad \frac{b}{\phi(\delta) - \delta} = \frac{1}{\int_0^\infty \mu(\mathrm{d}x)(1 - \mathrm{e}^{-\phi(\delta)x})},$$

and (5.3) might be rewritten as

$$\mathbb{P}(H \in \mathrm{d}y) = \frac{1/m}{\int_0^\infty \mu(\mathrm{d}x)(1 - \mathrm{e}^{-\phi(\delta)x})} \mathbb{P}(K > y)(1 - \mathrm{e}^{-\phi(\delta)y}) \, \mathrm{d}y.$$

Finally, using (5.2), we obtain

$$\mathbb{P}(H \in \mathrm{d}y) = \frac{\mathbb{P}(K > y)(1 - \mathrm{e}^{-\phi(\delta)y}) \, \mathrm{d}y}{\int_0^\infty \mathbb{P}(K > x)(1 - \mathrm{e}^{-\phi(\delta)x}) \, \mathrm{d}x}.$$

The right-hand side depends only on K (which might be estimated from independent hospital data) and $\phi(\delta)$.

Now assume that various outbreaks in various hospitals are observed at their detection times. If the sizes of outbreaks (all distributed as N_T) are the only observable statistics then, as stressed in [3] and [16], the fact that N_T is geometrically distributed allows only for the estimation of a single epidemiological parameter. Enlarging this information to, e.g. the times H_i spent in the hospital before T, we can hope to make finer inferences on the dynamical characteristics of these epidemics.

Assume that n outbreaks are observed of sizes $x_1, x_2, \ldots, x_n \in \mathbb{N}_{>0}$ and $s(n) := \sum_{i=1}^n x_i$ carriers are detected, which at the time of detection have been in the hospital for $y_1, y_2, \ldots, y_{s(n)} \in \mathbb{R}_+$ time units. We also assume that, since the distribution of K may be estimated from independent hospital data, its distribution is known exactly.

Using Remark 5.2, the likelihood of the observations $L(b, \delta; x_1, \ldots, x_n, y_1, \ldots, y_{s(n)})$ is given by

$$L(b, \delta) = \left(\prod_{i=1}^n \frac{\delta}{\phi(\delta)} \left(1 - \frac{\delta}{\phi(\delta)} \right)^{x_i - 1} \right) \prod_{j=1}^{s(n)} \mathbb{P}(H \in \mathrm{d}y_j)$$

$$= \left(\frac{\delta}{\phi(\delta)} \right)^n \left(1 - \frac{\delta}{\phi(\delta)} \right)^{s(n)-n} \prod_{j=1}^{s(n)} \frac{\mathbb{P}(K > y_j)(1 - \mathrm{e}^{-\phi(\delta)y_j}) \, \mathrm{d}y_j}{\int_0^\infty \mathbb{P}(K > x)(1 - \mathrm{e}^{-\phi(\delta)x}) \, \mathrm{d}x}. \qquad (5.4)$$

We write $L = L_1 L_2$, where

$$L_1(b, \delta) = \left(\frac{\delta}{\phi(\delta)} \right)^n \left(1 - \frac{\delta}{\phi(\delta)} \right)^{s(n)-n}$$

and

$$L_2(b, \delta) = \prod_{j=1}^{s(n)} \frac{\mathbb{P}(K > y_j)(1 - \mathrm{e}^{-\phi(\delta)y_j}) \, \mathrm{d}y_j}{\int_0^\infty \mathbb{P}(K > x)(1 - \mathrm{e}^{-\phi(\delta)x}) \, \mathrm{d}x}.$$

We observe that L_1 is the likelihood function for n realisations $x_1, x_2, \ldots, x_n$ of i.i.d. geometric random variables with parameter $g_1 := g_1(b, \delta) := \delta/\phi(\delta)$, while L_2 (assuming that the distribution of K is exactly known) depends only on $g_2 := g_2(b, \delta) := \phi(\delta)$. Observe that $\delta = g_1 g_2$ and (recalling that $m := \mathbb{E}(K)$),

$$b = \frac{\phi(\delta) - \delta}{\int_0^\infty \mu(\mathrm{d}x)(1 - \mathrm{e}^{-\phi(\delta)x})} = \frac{mg_2(1 - g_1)}{\int_0^\infty \mathbb{P}(K > x)(1 - \mathrm{e}^{-g_2 x})\,\mathrm{d}x}.$$

Furthermore, reparametrization of $L_1(b, \delta)$ in a function of g_1 and g_2 results in a function which is independent of g_2, while reparametrization of $L_2(b, \delta)$ in a function of g_1 and g_2 results in a function which is independent of g_1. It is straightforward to deduce that the maximum likelihood estimator (MLE) of g_1, say $\hat{g}_1$, is given by

$$\hat{g}_1 = \frac{n}{s(n)},$$

while, since L_1 does not depend on g_2, the MLE of g_2, say $\hat{g}_2$, is given by

$$\hat{g}_2 = \arg\max_{g_2} \prod_{j=1}^{s(n)} \frac{\mathbb{P}(K > y_j)(1 - \mathrm{e}^{-g_2 y_j})\,\mathrm{d}y_j}{\int_0^\infty \mathbb{P}(K > x)(1 - \mathrm{e}^{-g_2 x})\,\mathrm{d}x}.$$

Standard theory on maximum likelihood estimation reveals that the MLEs of b, say $\hat{b}$, and δ, say $\hat{\delta}$, are given by

$$\hat{b} = \frac{m\hat{g}_2(1 - \hat{g}_1)}{\int_0^\infty \mathbb{P}(K > x)(1 - \mathrm{e}^{-\hat{g}_2 x})\,\mathrm{d}x} \quad \text{and} \quad \hat{\delta} = \hat{g}_1 \hat{g}_2.$$

If H is exponentially distributed with parameter ν [3], [16] then $\hat{g}_1 = n/s(n)$, while

$$\hat{g}_2 = \arg\max_{g_2} \prod_{j=1}^{s(n)} \left(\frac{(\nu + g_2)\nu}{g_2}(1 - \mathrm{e}^{-g_2 y_j})\mathrm{e}^{-\nu y_j} \right)$$

$$= \arg\max_{g_2} \left(s(n) \log\left(1 + \frac{\nu}{g_2}\right) + \sum_{j=1}^{s(n)} \log(1 - \mathrm{e}^{-g_2 y_j}) \right),$$

and the MLE of b and δ are given by $\hat{b} = (1 - \hat{g}_1)(\nu + \hat{g}_2)$ and $\hat{\delta} = \hat{g}_1 \hat{g}_2$.

Note that it is possible to allow for differences in the distributions of lengths of stay (the random variable K) and infection rates (the parameter b) for different hospitals, while keeping the biologically governed rate of onset of symptoms (δ) the same for all hospitals. In that case we use the likelihood (5.4) with hospital specific parameters and observations for estimation.

Derivation of similar formulae for models relaxing too simplistic assumptions (see Remark 5.1), and applications to real hospital data, will be addressed in a future work.

Acknowledgements

References

[1] BERTOIN, J. (1996). *Lévy Processes*. Cambridge University Press.

[2] BERTOIN, J. (1997). Exponential decay and ergodicity of completely asymmetric Lévy processes in a finite interval. *Ann. Appl. Prob.* **7,** 156–169.

[3] BOOTSMA, M. C. J., WASSENBERG, M. W. M., TRAPMAN, P. AND BONTEN, M. J. M. (2011). The nosocomial transmission rate of animal-associated ST398 meticillin-resistant *Staphylococcus aureus*. *J. R. Soc. Interface* **8,** 578–584.

[4] CHAMPAGNAT, N. AND LAMBERT, A. (2011). Splitting trees with neutral Poissonian mutations II: largest and oldest families. Preprint. Available at http://arxiv.org/abs/1108.4812v1.

[5] CHAMPAGNAT, N. AND LAMBERT, A. (2012). Splitting trees with neutral Poissonian mutations I: small families. *Stoch. Process. Appl.* **122,** 1003–1033.

[6] GEIGER, J. AND KERSTING, G. (1997). Depth-first search of random trees, and Poisson point processes. In *Classical and Modern Branching Processes* (Minneapolis, 1994; IMA Math. Appl. Vol. **84**), Springer, New York.

[7] GRIMMETT, G. R. AND STIRZAKER, D. R. (1992). *Probability and Random Processes*, 2nd edn. Oxford University Press, New York.

[8] GRISHECHKIN, S. (1992). On a relationship between processor-sharing queues and Crump–Mode–Jagers branching processes. *Adv. Appl. Prob.* **24,** 653–698.

[9] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.

[10] KITAEV, M. Y. (1993). The M/G/1 processor-sharing model: transient behavior. *Queueing Systems* **14,** 239–273.

[11] KYPRIANOU, A. E. (2006). *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Springer, Berlin.

[12] LAMBERT, A. (2009). The allelic partition for coalescent point processes. *Markov Process. Relat. Fields* **15,** 359–386.

[13] LAMBERT, A. (2010). The contour of splitting trees is a Lévy process. *Ann. Prob.* **38,** 348–395.

[14] LAMBERT, A. (2011). Species abundance distributions in neutral models with immigration or mutation and general lifetimes. *J. Math. Biol.* **63,** 57–72.

[15] LAMBERT, A., SIMATOS, F. AND ZWART, B. (2012). Scaling limits via excursion theory: interplay between Crump–Mode–Jagers branching processes and processor-sharing queues. Preprint. Available at http://arxiv.org/abs/1102.5620v2.

[16] TRAPMAN, P. AND BOOTSMA, M. C. J. (2009). A useful relationship between epidemiology and queueing theory: the distribution of the number of infectives at the moment of the first detection. *Math. Biosci.* **219,** 15–22.