Adv. Appl. Prob. **57,** 1–18 (2025)
Printed in England

SUPPLEMENTARY MATERIAL: THE NUMBER OF INDIVIDUALS ALIVE IN A BRANCHING PROCESS GIVEN ONLY TIMES OF DEATHS

FRANK BALL,* ** *AND*
PETER NEAL,* *** *University of Nottingham*

In the supplementary material, the section counter continues from the main paper with the first section being Section 8. References to sections, theorems, equations or similar, with a section number less than 8 refer to statements or results in the main paper.

In Section 8 we consider two special cases of the phase-type lifetime distribution, the hyper-exponential distribution (mixture of $J > 1$ exponential distributions) and the Erlang distribution (sum of $J > 1$ *i.i.d.* exponential distributions). We study time-homogeneous branching processes with all deaths detected (for all $t \in \mathbb{R}$, $\mathbf{d}_t = \mathbf{1}$) and show that the Erlang lifetime distribution satisfies (3.16), which by Corollary 3.1, gives the number of individuals alive immediately after the k^{th} death as a mixture of negative binomial distributions. In Section 9, we discuss the case where $L \equiv 1$, a constant lifetime which arises as the limit as $J \to \infty$ of $L \sim \text{Gamma}(J, J)$. In Section 10 we derive the distribution of the number of individuals alive at the first detected death given we know the time of birth of the initial individual in Lemma 4. We comment on how Theorem 3.1 can be adapted to the scenario where the time of birth of the initial individual is known with Lemma 4 replacing Lemma 5.2. In Section 11, we provide two more numerical examples by applying the approximation given in Section 3.5.2 to simulated data including an SIR epidemic with an Erlang distributed infectious period.

8. Special cases of phase-type lifetime distributions

Throughout this section we consider a time-homogeneous branching process with $\beta_t = \alpha\,(> 0)$ $(t \in \mathbb{R})$ and all deaths detected, $\mathbf{d}_t = \mathbf{1}$. This leads to simplification in probabilities related to exploration process defined in Section 4.2, which no longer depend on t. Using results from Section 4.3, we have that, for $t \in \mathbb{R}$, $\tau > 0$, $0 \le u \le \tau$ and $i, j = 1, 2, \ldots, J$,

$$p_{ij}(t, u, \tau) = P_{ij}(\tau - u) \qquad (i, j = 1, 2, \ldots, J)$$

Submitted 23 January 2024; accepted 1 October 2024.
* Postal address: School of Mathematical Sciences, University of Nottingham, NG7 2RD, United Kingdom.
** Email address: frank.ball@nottingham.ac.uk
*** Email address: peter.neal@nottingham.ac.uk

and

$$q_i(t, u, \tau) = 0 \qquad (i = 1, 2, \ldots, J).$$

For $j = 1, 2, \ldots, J$ and $u \geq 0$, the probability that an individual is of type j, u units after they are born, is $\sum_{i=1}^{J} \chi_i P_{ij}(u)$, for which we employ the shorthand notation $\chi_j(u)$. Then, for $t \in \mathbb{R}$ and $\tau > 0$, it follows from (4.14) that $\psi_j(t;\tau) = \bar{\psi}_j(\tau)$, where

$$\bar{\psi}_j(\tau) = \int_0^{\tau} \alpha \exp(-\alpha u) \chi_j(u)\, du, \tag{8.1}$$

and from (4.15) that $\zeta(t;\tau) = \bar{\zeta}(\tau)$, where

$$\bar{\zeta}(\tau) = \exp(-\alpha\tau). \tag{8.2}$$

Let $\bar{\psi}(\tau) = \sum_{j=1}^{J} \bar{\psi}_j(\tau)$ and note that $\psi(\bar{\tau}) = \int_0^{\tau} \alpha \exp(-\alpha u) \mathbb{P}(L > u)\, du$. As in Section 3.3, let $\pi_0 = 1 - \bar{\psi}(\infty)$ and for $j = 1, 2, \ldots, J$, let $\eta_j^0 = \bar{\psi}_j(\infty)/\bar{\psi}(\infty)$. Setting $t = 0$ in (3.1) yields, for $\tau > 0$, that

$$\begin{aligned}
\pi_\tau &= 1 - \bar{\psi}(\tau) - \exp(-\alpha\tau) \sum_{i=1}^{J} \sum_{j=1}^{J} \bar{\psi}(\infty) P_{ij}(\tau) \\
&= \pi_0 + \bar{\psi}(\infty) - \bar{\psi}(\tau) - \exp(-\alpha\tau) \sum_{i=1}^{J} \sum_{j=1}^{J} \bar{\psi}(\infty) P_{ij}(\tau).
\end{aligned}$$

Now,

$$\begin{aligned}
\exp(-\alpha\tau) \sum_{i=1}^{J} \sum_{j=1}^{J} \bar{\psi}(\infty) P_{ij}(\tau) &= \sum_{i=1}^{J} \sum_{j=1}^{J} \int_0^{\infty} \alpha \exp(-\alpha(u+\tau)) \chi_i(u) P_{ij}(\tau)\, du \\
&= \int_0^{\infty} \alpha \exp(-\alpha(u+\tau)) \mathbb{P}(L > u + \tau)\, du \\
&= \int_{\tau}^{\infty} \alpha \exp(-\alpha u) \mathbb{P}(L > u)\, du \\
&= \bar{\psi}(\infty) - \bar{\psi}(\tau),
\end{aligned}$$

so $\pi_\tau = \pi_0$ for all $\tau \geq 0$.

Setting $t = 0$ in (3.2) now yields, for $\tau > 0$ and $j = 1, 2, \dots, J$, that

$$\begin{aligned}(1-\pi_0)\eta_j^\tau &= \bar{\psi}_j(\tau) + \exp(-\alpha\tau)\sum_{i=1}^{J}\bar{\psi}_i(\infty)P_{ij}(\tau)\\ &= \int_0^\tau \alpha\exp(-\alpha u)\chi_j(u)\,du + \sum_{i=1}^{J}\int_0^\infty \alpha\exp(-\alpha(u+\tau))\chi_i(u)P_{ij}(\tau)\,du\\ &= \int_0^\tau \alpha\exp(-\alpha u)\chi_j(u)\,du + \int_0^\infty \alpha\exp(-\alpha(u+\tau))\chi_j(t+u)\,du\\ &= \int_0^\infty \alpha\exp(-\alpha u)\chi_j(u)\,du\\ &= \bar{\psi}_j(\infty),\end{aligned}$$

so $\eta_j^\tau = \eta_j^0$ for all $\tau \geq 0$.

Let

$$\bar{\pi} = \mathbb{E}[\exp(-\alpha L)], \tag{8.3}$$

the probability that an individual does not give birth during their lifetime. It is straightforward to show that $\bar{\pi} = 1 - \bar{\psi}(\infty)$, so for all $t \geq 0$, we have that $\pi_t = \bar{\pi}$. For $j = 1, 2, \dots, J$, let $\bar{\eta}_j = \eta_j^0$, which can be expressed as

$$\bar{\eta}_j = \frac{\bar{\psi}_j(\infty)}{\bar{\psi}(\infty)} = \frac{\bar{\psi}_j(\infty)}{1-\bar{\pi}}. \tag{8.4}$$

Let $f_L(u)$ denote the probability density function of L. Since the probability that an individual is of type j, u units after they are born, is $\chi_j(u)$, we have that $f_L(u) = \sum_{j=1}^J \chi_j(u)\gamma_j$ $(u \geq 0)$ and from (8.4), (8.1) and (8.3) that

$$\begin{aligned}(1-\bar{\pi})\sum_{j=1}^{J}\bar{\eta}_j\gamma_j &= \sum_{j=1}^{J}\bar{\psi}_j(\infty)\gamma_j\\ &= \int_0^\infty \alpha\exp(-\alpha u)f_L(u)\,du\\ &= \alpha\mathbb{E}[\exp(-\alpha L)] = \alpha\bar{\pi}.\end{aligned} \tag{8.5}$$

For $\mathbf{a} > -(1-\bar{\pi})\bar{\boldsymbol{\eta}}$, let $\bar{\mathbf{W}}(\mathbf{a})$ denote the J-dimensional random variable with, for $\boldsymbol{\theta} \in [0,1]^J$, probability generating function (pgf)

$$\begin{aligned}\bar{\varphi}(\boldsymbol{\theta};\mathbf{a}) &= \mathbb{E}\left[\prod_{j=1}^{J}\theta_j^{\bar{W}_j(\mathbf{a})}\right]\\ &= \frac{1+\sum_{j=1}^J a_j\theta_j}{1+\sum_{j=1}^J a_j} \times \frac{\bar{\pi}}{1-(1-\bar{\pi})\sum_{j=1}^J \bar{\eta}_j\theta_j}.\end{aligned}$$

Thus $\bar{\mathbf{W}}(\mathbf{a})$ is the time-homogeneous version of $\mathbf{W}(t, \mathbf{a})$ defined in (3.4).

Finally, for $i = 1, 2, \ldots, J$ and $t \geq 0$, let

$$\bar{C}_i(t) = \frac{(1-\bar{\pi}) \sum_{j=1}^J \gamma_j \left[\bar{\psi}_j(t)\bar{\eta}_i - \bar{\psi}_i(t)\bar{\eta}_j\right]}{\sum_{j=1}^J \gamma_j [(1-\bar{\pi})\bar{\eta}_j - \bar{\psi}_j(t)]} \tag{8.6}$$

$$= \frac{\bar{D}_i(t)}{\bar{E}(t)}, \tag{8.7}$$

say, where $\bar{E}(t) > 0$. Note that (8.6) is the time-homogeneous version of (3.19) and $\bar{C}_i(t) = c_i(0, t; \mathbf{0})$ is defined in (3.8). Let $\bar{C}(t) = \sum_{i=1}^J \bar{C}_i(t)$ with $\bar{C}(t) = \bar{D}(t)/\bar{E}(t)$, where

$$\bar{D}(t) = \sum_{i=1}^J \bar{D}_i(t) = (1-\bar{\pi}) \sum_{i=1}^J \sum_{j=1}^J \gamma_j \left[\bar{\psi}_j(t)\bar{\eta}_i - \bar{\psi}_i(t)\bar{\eta}_j\right]. \tag{8.8}$$

8.1. Mixture of Exponential Distributions

For a mixture of J exponentials with the j^{th} mixture component being $\text{Exp}(\gamma_j)$, we have that $\chi_j(u) = \chi_j \exp(-\gamma_j u)$ giving

$$\bar{\pi} = \sum_{j=1}^J \frac{\chi_j \gamma_j}{\alpha + \gamma_j}, \tag{8.9}$$

and

$$(1-\bar{\pi})\bar{\eta}_j = \bar{\psi}_j(\infty) = \int_0^\infty \alpha \exp(-\alpha u) \chi_j \exp(-\gamma_j u)\, du = \frac{\chi_j \alpha}{\alpha + \gamma_j}. \tag{8.10}$$

To ease the presentation, let $\xi_i = \frac{\chi_i}{\alpha + \gamma_i}$ $(i = 1, 2, \ldots, J)$. Then $1 - \bar{\pi} = \alpha \sum_{j=1}^J \xi_j$, since $\sum_{j=1}^J \eta_j = 1$. It is trivial to show that, for all $i = 1, 2, \ldots, J$ and $\tau \geq 0$,

$$\bar{\psi}_i(\tau) = \alpha \xi_i \left[1 - \exp(-\{\alpha + \gamma_i\}\tau)\right] = (1-\bar{\pi})\bar{\eta}_i \left[1 - \exp(-\{\alpha + \gamma_i\}\tau)\right]. \tag{8.11}$$

From (8.10), (8.11) and (8.7), we have that

$$\bar{E}(t) = \alpha \exp(-\alpha t) \sum_{j=1}^J \gamma_j \xi_j \exp(-\gamma_j t)$$

$$= \alpha \exp(-\alpha t) g(t), \qquad \text{say.} \tag{8.12}$$

Similarly using (8.10), (8.11) and (8.8),

$$\bar{D}(t) = \alpha \sum_{i=1}^J \sum_{j=1}^J \gamma_j \left[\xi_i \bar{\psi}_j(t) - \xi_j \bar{\psi}_i(t)\right]$$

$$= \alpha^2 \exp(-\alpha t) \sum_{i=1}^J \sum_{j=1}^J \gamma_j \xi_i \xi_j [\exp(-\gamma_i t) - \exp(-\gamma_j t)]. \tag{8.13}$$

The summands are zero when $i = j$, so

$$\begin{aligned}\bar{D}(t) &= \alpha^2 \exp(-\alpha\tau) \sum_{i=1}^{J-1} \sum_{j=i+1}^{J} \xi_i \xi_j \{\gamma_j [\exp(-\gamma_i t) - \exp(-\gamma_j t)] + \gamma_i [\exp(-\gamma_j t) - \exp(-\gamma_i t)]\}, \\ &= \alpha^2 \exp(-\alpha\tau) \sum_{i=1}^{J-1} \sum_{j=i+1}^{J} \xi_i \xi_j (\gamma_j - \gamma_i)[\exp(-\gamma_i t) - \exp(-\gamma_j t)].\end{aligned}$$

Thus $\bar{D}(t) = \alpha^2 \exp(-\alpha t) f(t)$, where

$$f(t) = \sum_{i=1}^{J-1} \sum_{j=i}^{J} (\gamma_j - \gamma_i) \xi_i \xi_j \left[\exp(-\gamma_i t) - \exp(-\gamma_j t)\right]. \tag{8.14}$$

The term when $i = j$ has been included in the double summation to ease the subsequent algebra.

For all $t > 0$, if $\gamma_j \neq \gamma_i$,

$$(\gamma_j - \gamma_i) \left[\exp(-\gamma_i \tau) - \exp(-\gamma_j \tau)\right] > 0,$$

so it follows from (8.14) that, for $t > 0$, $\bar{D}(t) > 0$, and hence, $\bar{C}(t) > 0$. Since the numbers of individuals of each type alive immediately following the first death are distributed according to $\bar{\mathbf{W}}(\mathbf{0})$, it follows from Lemma 5.4, (5.29), that

$$\{\mathbf{X}_2 | T_2 = \tau\} \stackrel{D}{=} \bar{\mathbf{W}}(\bar{\mathbf{C}}(\tau)) + \bar{\mathbf{W}}(\mathbf{0}), \tag{8.15}$$

where $\bar{\mathbf{C}}(\tau) = (\bar{C}_1(\tau), \bar{C}_2(\tau), \ldots, \bar{C}_J(\tau))$ and $\bar{\mathbf{W}}(\bar{\mathbf{C}}(\tau))$ and $\bar{\mathbf{W}}(\mathbf{0})$ are independent. Therefore, it follows from Lemma 3.1 that the size of the population immediately after the second death,

$$\{X_2^* | T_2 = \tau\} \stackrel{D}{=} \mathrm{NegBin}(2, \bar{\pi}) + \mathrm{Bin}\left(1, \frac{\bar{C}(\tau)}{1 + \bar{C}(\tau)}\right),$$

where the two random variables on the right-hand side are independent. In Lemma 1 below we show that $\bar{C}(\tau)$ is increasing in τ, and consequently that, X_2^* is stochastically increasing in τ.

Lemma 1. *For $J \geq 2$ and $\tau \geq 0$,*

$$\bar{C}'(\tau) = \frac{\alpha}{g(\tau)^2} \left(\sum_{i=1}^{J} \gamma_i \xi_i\right) \left(\sum_{k=1}^{J-1} \sum_{m=k+1}^{J} \xi_k \xi_m (\gamma_k - \gamma_m)^2 \exp(-[\gamma_k + \gamma_m]\tau)\right) > 0. \tag{8.16}$$

Proof. Let $h(\tau) = g(\tau)f'(\tau) - f(\tau)g'(\tau)$. Then

$$h(\tau) = \left(\sum_{m=1}^{J} \gamma_m \xi_m \exp(-\gamma_m \tau)\right) \left(\sum_{i=1}^{J-1} \sum_{j=i}^{J} (\gamma_j - \gamma_i)\xi_i \xi_j \left[\gamma_j \exp(-\gamma_j \tau) - \gamma_i \exp(-\gamma_i \tau)\right]\right) + \left(\sum_{i=1}^{J-1} \sum_{j=i}^{J} (\gamma_j - \gamma_i)\xi_i \xi_j \left[\exp(-\gamma_i \tau) - \exp(-\gamma_j \tau)\right]\right) \left(\sum_{m=1}^{J} \gamma_m^2 \xi_m \exp(-\gamma_m \tau)\right), \tag{8.17}$$

so $h(\tau)$ admits the form

$$h(\tau) = \sum_{k=1}^{J} \sum_{m=k}^{J} \alpha_{km} \exp(-(\gamma_k + \gamma_m)\tau). \tag{8.18}$$

Using (8.17), for $k = 1, 2, \ldots, J$,

$$\begin{aligned} \alpha_{kk} = \gamma_k \xi_k & \left[-\sum_{j=k}^{J} (\gamma_j - \gamma_k)\xi_k \xi_j \gamma_k + \sum_{i=1}^{k} (\gamma_k - \gamma_i)\xi_i \xi_k \gamma_k \right] \\ & + \gamma_k^2 \xi_k \left[-\sum_{i=1}^{k} (\gamma_k - \gamma_i)\xi_i \xi_k + \sum_{j=k}^{J} (\gamma_j - \gamma_k)\xi_k \xi_j \right] \\ = 0. & \end{aligned} \tag{8.19}$$

For $1 \leq k < m \leq J$,

$$\begin{aligned} \alpha_{km} = \gamma_k \xi_k & \left[-\sum_{j=m}^{J} (\gamma_j - \gamma_m)\xi_m \xi_j \gamma_m + \sum_{i=1}^{m} (\gamma_m - \gamma_i)\xi_i \xi_m \gamma_m \right] \\ & + \gamma_m \xi_m \left[-\sum_{j=k}^{J} (\gamma_j - \gamma_k)\xi_k \xi_j \gamma_k + \sum_{i=1}^{k} (\gamma_k - \gamma_i)\xi_i \xi_k \gamma_k \right] \\ & + \gamma_k^2 \xi_k \left[\sum_{j=m}^{J} (\gamma_j - \gamma_m)\xi_m \xi_j - \sum_{i=1}^{m} (\gamma_m - \gamma_i)\xi_i \xi_m \right] \\ & + \gamma_m^2 \xi_m \left[\sum_{j=k}^{J} (\gamma_j - \gamma_k)\xi_k \xi_j - \sum_{i=1}^{k} (\gamma_k - \gamma_i)\xi_i \xi_k \right]. \end{aligned} \tag{8.20}$$

Thus α_{km} takes the form

$$\alpha_{km} = \sum_{i=1}^{J} \beta_i^{km} \xi_i \xi_k \xi_m. \tag{8.21}$$

For $i \leq k < m$, only the 2nd, 4th, 6th and 8th sums in (8.20) give contributions to β_i^{km}, so

$$\begin{aligned}\beta_i^{km} &= \gamma_k(\gamma_m - \gamma_i)\gamma_m + \gamma_m(\gamma_k - \gamma_i)\gamma_k - \gamma_k^2(\gamma_m - \gamma_i) - \gamma_m^2(\gamma_k - \gamma_i) \\ &= \gamma_i(\gamma_k - \gamma_m)^2.\end{aligned}$$

Similar calculations for the case $k < i < m$, when the 2nd, 3rd, 6th and 7th sums in (8.20) give contributions to β_i^{km}, and the case $k < m \leq i$, when the 1st, 3rd, 5th and 7th sums in (8.20) give contributions to β_i^{km}, show that in both of these cases $\beta_i^{km} = \gamma_i(\gamma_k - \gamma_m)^2$ also. Recalling (8.19), it then follows using (8.18) and (8.21) that

$$h(\tau) = \left(\sum_{k=1}^{J}\sum_{l=k}^{J} \xi_k \xi_m (\gamma_k - \gamma_m)^2 \exp(-[\gamma_k + \gamma_m]\tau)\right)\left(\sum_{i=1}^{J} \gamma_i \xi_i\right),$$

and (8.16) follows as $\bar{C}'(\tau) = \alpha \frac{h(\tau)}{g(\tau)^2}$. □

8.2. Erlang Distribution

In general, if $L = \sum_{j=1}^{J} E_j$, where the $E_j \sim \text{Exp}(\tilde{\mu}_j)$s are independent, then individuals start in type 1 and transition through the types in order before leaving type J via death. For $j = 1, 2, \ldots, J$, and $u \geq 0$, the probability that an individual born at time 0 is of type j at time u is

$$\chi_j(u) = \mathbb{P}\left(\sum_{i=1}^{j} E_i > u\right) - \mathbb{P}\left(\sum_{i=1}^{j-1} E_i > u\right) \tag{8.22}$$

with $\sum_{i=1}^{j-1} E_i = 0$ if $j = 1$. Further, $\chi_1(= \chi_1(0)) = 1$, $\chi_i(= \chi_i(0)) = 0$ $(i = 2, 3, \ldots, J)$, $\gamma_i = 0$ $(i = 1, 2, \ldots, J-1)$ and $\gamma_J = \tilde{\mu}_J$. Since for a random variable Z with support on $[0, \infty)$,

$$\int_0^\infty \alpha \exp(-\alpha z)\mathbb{P}(Z > z)\, dz = 1 - \mathbb{E}\left[\exp(-\alpha Z)\right],$$

it is straightforward to show, using (8.1) and (8.4), that

$$\begin{aligned}(1 - \bar{\pi})\bar{\eta}_j &= \int_0^\infty \alpha \exp(-\alpha u)\chi_j(u)\, du \\ &= \int_0^\infty \alpha \exp(-\alpha u)\left\{\mathbb{P}\left(\sum_{i=1}^{j} E_i > u\right) - \mathbb{P}\left(\sum_{i=1}^{j-1} E_i > u\right)\right\} du \\ &= \left\{1 - \prod_{i=1}^{j} \frac{\tilde{\mu}_i}{\alpha + \tilde{\mu}_i}\right\} - \left\{1 - \prod_{i=1}^{j-1} \frac{\tilde{\mu}_i}{\alpha + \tilde{\mu}_i}\right\} \\ &= \frac{\alpha}{\alpha + \tilde{\mu}_j} \prod_{i=1}^{j-1} \frac{\tilde{\mu}_i}{\alpha + \tilde{\mu}_i}.\end{aligned}$$

For the Erlang distribution, where each $\tilde{\mu}_j = \mu$ so successive jumps to the next type form a Poisson process having rate μ, we have that

$$\chi_j(u) = \frac{(u\mu)^{j-1}}{(j-1)!}\exp(-u\mu) \qquad (j = 1, 2, \dots, J).$$

Therefore, recalling (8.1),

$$\bar{\psi}_1(\tau) = \frac{\alpha}{\alpha+\mu}\left[1 - \exp(-\{\alpha+\mu\}\tau)\right], \tag{8.23}$$

and for $j > 1$, using integration by parts,

$$\begin{aligned}\bar{\psi}_j(\tau) &= \alpha\mu^{j-1}\left[-\frac{u^{j-1}}{(j-1)!} \times \frac{\exp(-\{\alpha+\mu\}u)}{\alpha+\mu}\right]_0^\tau \\ &\quad + \frac{\alpha\mu^{j-1}}{\alpha+\mu}\int_0^\tau \frac{u^{j-2}}{(j-2)!}\exp(-\{\alpha+\mu\}u)\,du \\ &= -\frac{\alpha}{\alpha+\mu}\left[\frac{(\mu\tau)^{j-1}}{(j-1)!}\exp(-\{\alpha+\mu\}\tau)\right] + \frac{\mu}{\alpha+\mu}\bar{\psi}_{j-1}(\tau). \end{aligned} \tag{8.24}$$

Noting that the solution of the difference equation

$$x_i = a_i + \theta x_{i-1} \qquad (i = 2, 3, \dots)$$

is

$$x_i = \theta^{i-1}x_1 + \sum_{j=1}^{i-1} a_j \theta^{i-1-j} \qquad (i = 2, 3, \dots),$$

it follows from (8.23), (8.24) and a little algebra, for $i = 1, 2, \dots, J$, that

$$\bar{\psi}_i(\tau) = \frac{\alpha}{\alpha+\mu}\left[\left(\frac{\mu}{\mu+\alpha}\right)^{i-1} - \sum_{k=0}^{i-1}\frac{(\mu\tau)^k}{k!}\left(\frac{\mu}{\mu+\alpha}\right)^{i-1-k} \mathrm{e}^{-(\alpha+\mu)\tau}\right]. \tag{8.25}$$

Thus

$$\bar{\psi}_i(\infty) = \frac{\alpha}{\alpha+\mu}\left(\frac{\mu}{\mu+\alpha}\right)^{i-1} \qquad (i = 1, 2, \dots, J),$$

$\bar{\psi}(\infty) = \sum_{i=1}^J \bar{\psi}_i(\infty) = 1 - \left(\frac{\mu}{\mu+\alpha}\right)^J$ and

$$\bar{\pi} = 1 - \bar{\psi}(\infty) = \left(\frac{\mu}{\mu+\alpha}\right)^J. \tag{8.26}$$

Recalling $\sum_{j=1}^J \bar{\eta}_j = 1$, the definitions of $\bar{C}_i(\tau)$, $\bar{C}(\tau)$, $\bar{D}_i(\tau)$, $\bar{D}(\tau)$ and $\bar{E}(\tau)$ given in (8.6)-(8.8), and using (8.5), we have that

$$\begin{aligned}\bar{E}(\tau) &= \sum_{j=1}^J \gamma_j\left[(1-\bar{\pi})\bar{\eta}_j - \bar{\psi}_j(\tau)\right] = \alpha\bar{\pi} - \mu\bar{\psi}_J(\tau) \\ &= \frac{\mu\alpha}{\mu+\alpha}\mathrm{e}^{-(\alpha+\mu)\tau}\sum_{k=0}^{J-1}\frac{(\mu\tau)^k}{k!}\left(\frac{\mu}{\mu+\alpha}\right)^{J-1-k} \end{aligned} \tag{8.27}$$

and

$$\bar{D}_i(\tau) = (1-\bar{\pi}) \sum_{j=1}^{J} \gamma_j \left[\bar{\eta}_i \bar{\psi}_j(\tau) - \bar{\eta}_j \bar{\psi}_i(\tau)\right]$$
$$= \mu(1-\bar{\pi}) \left\{\bar{\eta}_i \bar{\psi}_J(\tau) - \bar{\eta}_J \bar{\psi}_i(\tau)\right\}. \tag{8.28}$$

Noting from (8.4) that $(1-\bar{\pi})\bar{\eta}_j = \bar{\psi}_j(\infty)$ $(j = 1, 2, \ldots, J)$, it follows from (8.28) that

$$\bar{D}(\tau) = \sum_{j=1}^{J} \bar{D}_j(\tau) = \mu[\bar{\psi}(\infty)\bar{\psi}_J(\tau) - \bar{\psi}_J(\infty)\bar{\psi}(\tau)]. \tag{8.29}$$

Now $\bar{\psi}(\tau) = \sum_{i=1}^{J} \bar{\psi}_i(\tau)$, so, using (8.25),

$$\bar{\psi}(\tau) = \sum_{i=1}^{J} \frac{\alpha}{\alpha+\mu} \left[\left(\frac{\mu}{\mu+\alpha}\right)^{i-1} - \sum_{k=0}^{i-1} \frac{(\mu\tau)^k}{k!} \left(\frac{\mu}{\mu+\alpha}\right)^{i-1-k} \mathrm{e}^{-(\alpha+\mu)\tau}\right]$$
$$= 1 - \left(\frac{\mu}{\mu+\alpha}\right)^{J} - \frac{\alpha}{\alpha+\mu}\mathrm{e}^{-(\alpha+\mu)\tau} \sum_{k=0}^{J-1} \frac{(\mu\tau)^k}{k!} \sum_{i=k+1}^{J} \left(\frac{\mu}{\mu+\alpha}\right)^{i-1-k}$$
$$= 1 - \left(\frac{\mu}{\mu+\alpha}\right)^{J} - \mathrm{e}^{-(\alpha+\mu)\tau} \sum_{k=0}^{J-1} \frac{(\mu\tau)^k}{k!} \left[1 - \left(\frac{\mu}{\mu+\alpha}\right)^{J-k}\right],$$

whence

$$\bar{\psi}_J(\infty)\bar{\psi}(\tau)$$
$$= \frac{\alpha}{\alpha+\mu} \left(\frac{\mu}{\mu+\alpha}\right)^{J-1} \left\{1 - \left(\frac{\mu}{\mu+\alpha}\right)^{J} - \mathrm{e}^{-(\alpha+\mu)\tau} \sum_{k=0}^{J-1} \frac{(\mu\tau)^k}{k!} \left[1 - \left(\frac{\mu}{\mu+\alpha}\right)^{J-k}\right]\right\}.$$

Further,

$$\bar{\psi}(\infty)\bar{\psi}_J(\tau)$$
$$= \left[1 - \left(\frac{\mu}{\mu+\alpha}\right)^{J}\right] \left(\frac{\alpha}{\alpha+\mu}\right) \left[\left(\frac{\mu}{\mu+\alpha}\right)^{J-1} - \sum_{k=0}^{J-1} \frac{(\mu\tau)^k}{k!} \left(\frac{\mu}{\mu+\alpha}\right)^{J-1-k} \mathrm{e}^{-(\alpha+\mu)\tau}\right],$$

so using (8.29),

$$\bar{D}(\tau) = -\frac{\mu\alpha}{\mu+\alpha}\mathrm{e}^{-(\alpha+\mu)\tau} \sum_{k=0}^{J-1} \left[\left(\frac{\mu}{\mu+\alpha}\right)^{J-1-k} - \left(\frac{\mu}{\mu+\alpha}\right)^{J-1}\right] \frac{(\mu\tau)^k}{k!}. \tag{8.30}$$

Hence, $\bar{D}(\tau) < 0$, and consequently, $\bar{C}(\tau) < 0$. Therefore it follows from (8.15) and Lemma 3.1 that

$$X_2^*|T_2 = \tau \sim \begin{cases} \text{NegBin}(2, \bar{\pi}) & \text{with probability } \frac{1+\bar{C}(\tau)-\bar{\pi}}{(1+\bar{C}(\tau))(1-\bar{\pi})} \\ \text{Geom}(\bar{\pi}) & \text{with probability } \frac{-\bar{C}(\tau)\bar{\pi}}{(1+\bar{C}(\tau))(1-\bar{\pi})}. \end{cases} \tag{8.31}$$

Equations (8.27) and (8.30) yield

$$\bar{C}(\tau) = \frac{\bar{D}(\tau)}{\bar{E}(\tau)} = \frac{\displaystyle\sum_{k=0}^{J-1} \frac{(\mu\tau)^k}{k!}}{\displaystyle\sum_{k=0}^{J-1} \frac{((\mu+\alpha)\tau)^k}{k!}} - 1. \tag{8.32}$$

It is shown in Lemma 2 below that, for $J \geq 2$, $\bar{C}(\tau)$ is strictly decreasing in τ on $[0,\infty)$, and hence, that X_2^* is stochastically decreasing in τ.

Lemma 2. *For $J \geq 2$, $\bar{C}(\tau)$ is strictly decreasing in τ on $[0,\infty)$.*

Proof. Let $n = J-1$, $\rho = \frac{\mu+\alpha}{\mu}$ and define the function $\tilde{C} : [0,\infty) \to [0,1)$ by

$$\tilde{C}(t) = \frac{\displaystyle\sum_{k=0}^{n} \frac{t^k}{k!}}{\displaystyle\sum_{k=0}^{n} \frac{(\rho t)^k}{k!}}.$$

Then, by (8.32),

$$\bar{C}(\tau) = \tilde{C}(\mu\tau) - 1.$$

Let $\tilde{f}(t) = \sum_{k=0}^n \frac{t^k}{k!}$, $\tilde{g}(t) = \sum_{k=0}^n \frac{(\rho t)^k}{k!}$ and $\tilde{h}(t) = \tilde{g}(t)\tilde{f}'(t) - \tilde{f}(t)\tilde{g}'(t)$. Then

$$\tilde{h}(t) = \sum_{i=0}^{2n-1} \tilde{\alpha}_i t^i,$$

where

$$\tilde{\alpha}_i = \sum_{k=\max(0,i-n+1)}^{\min(i,n)} \frac{\rho^k}{k!(i-k)!} - \sum_{k=\max(0,i-n+1)}^{\min(i,n)} \frac{\rho^{i-k+1}}{k!(i-k)!}. \tag{8.33}$$

For $i = 0, 1, \ldots, n-1$,

$$\begin{aligned}\tilde{\alpha}_i &= \sum_{k=0}^{i} \frac{\rho^k}{k!(i-k)!} - \sum_{k=0}^{i} \frac{\rho^{i-k+1}}{k!(i-k)!} \\ &= \sum_{k=0}^{i} \frac{\rho^k}{k!(i-k)!} - \sum_{k=0}^{i} \frac{\rho^{k+1}}{k!(i-k)!} \\ &< 0,\end{aligned}$$

as $\rho > 1$.

For $i = n, n+1, \dots, 2n-1$, substituting $l = k - (i+1-n)$ in (8.33) yields

$$\begin{aligned}
\tilde{\alpha}_i &= \sum_{l=0}^{2n-1-i} \frac{\rho^{l+i+1-n}}{(l+i+1-n)!(n-l-1)!} - \sum_{l=0}^{2n-1-i} \frac{\rho^{n-l}}{(l+i+1-n)!(n-l-1)!} \\
&= \rho^{i+1-n} \left[\sum_{l=0}^{2n-1-i} \frac{\rho^{l}}{(l+i+1-n)!(n-l-1)!} - \sum_{l=0}^{2n-1-i} \frac{\rho^{2n-1-i-l}}{(l+i+1-n)!(n-l-1)!} \right] \\
&= \rho^{i+1-n} \left[\sum_{l=0}^{2n-1-i} \frac{\rho^{l}}{(l+i+1-n)!(n-l-1)!} - \sum_{l=0}^{2n-1-i} \frac{\rho^{l}}{(l+i-n)!(n-l)!} \right] \\
&= \frac{\rho^{i+1-n}}{i!} \sum_{l=0}^{2n-1-i} \rho^{l} \left[\binom{i}{n-l-1} - \binom{i}{n-l} \right].
\end{aligned}$$

Substituting $k = 2n-1-i-l$ in the final sum above yields

$$\tilde{\alpha}_i = \frac{\rho^{i+1-n}}{i!} \sum_{k=0}^{2n-1-i} \rho^{2n-1-i-k} \left[\binom{i}{n-k} - \binom{i}{n-k-1} \right],$$

so

$$\tilde{\alpha}_i = \frac{1}{2} \frac{\rho^{i+1-n}}{i!} \sum_{l=0}^{2n-1-i} \left(\rho^{l} - \rho^{2n-1-i-l} \right) \left[\binom{i}{n-l-1} - \binom{i}{n-l} \right]. \qquad (8.34)$$

It is easily checked that

$$\binom{i}{n-l-1} - \binom{i}{n-l} \begin{cases} > 0 & \text{if } l < n - \frac{i+1}{2}, \\ = 0 & \text{if } l = n - \frac{i+1}{2}, \\ < 0 & \text{if } l > n - \frac{i+1}{2}, \end{cases}$$

and, as $\rho > 1$,

$$\rho^{l} - \rho^{2n-1-i-l} \begin{cases} > 0 & \text{if } l > n - \frac{i+1}{2}, \\ = 0 & \text{if } l = n - \frac{i+1}{2}, \\ < 0 & \text{if } l < n - \frac{i+1}{2}. \end{cases}$$

Hence it follows from (8.34) that $\tilde{\alpha}_i < 0$ $(i = n, n+1, \dots, 2n-2)$ and $\tilde{\alpha}_{2n-1} = 0$. Thus $\tilde{C}'(t) < 0$ for all $t > 0$ and the lemma follows. □

Suppose now that $L \sim \text{Gamma}(J, J)$, so $\mu = J$, and consider the limit as $J \to \infty$, in which case L converges in distribution to a unit mass at one. This provides a link between phase-type distributions and the analysis of a constant lifetime distribution presented in Section 9.

Lemma 3. *Under the above limit,* $\bar{\pi} \to \exp(-\alpha)$ *and* $\bar{C}(\tau) \to -r(\tau)$*, where*

$$r(\tau) = \begin{cases} 1 - \exp(-\alpha\tau) & \textit{if } \tau \le 1, \\ 1 - \exp(-\alpha) & \textit{if } \tau > 1. \end{cases}$$

Proof. It follows from (8.26) that

$$\bar{\pi} = \left(\frac{J}{J+\alpha}\right)^J \to \exp(-\alpha),$$

as required.

Let $\bar{F}(\tau) = \bar{C}(\tau) + 1$. Using (8.32),

$$\bar{F}(\tau) = \frac{\displaystyle\sum_{k=0}^{J-1} \frac{(J\tau)^k}{k!}}{\displaystyle\sum_{k=0}^{J-1} \frac{((J+\alpha)\tau)^k}{k!}} = \frac{\exp(J\tau)\mathbb{P}(X \leq J-1)}{\exp((J+\alpha)\tau)\mathbb{P}(Y \leq J-1)},$$

where $X \sim \text{Po}(J\tau)$ and $Y \sim \text{Po}((J+\alpha)\tau)$. Let $W \sim \text{Gamma}(J,1)$. Then $\mathbb{P}(X \leq J-1) = \mathbb{P}(W > J\tau)$ and $\mathbb{P}(Y \leq J-1) = \mathbb{P}(W > (J+\alpha)\tau)$, so

$$\bar{F}(\tau) = \exp(-\alpha\tau) f_J(\tau) \tag{8.35}$$

where

$$f_J(\tau) = \frac{g_J(J\tau)}{g_J(J+\alpha)\tau)},$$

with

$$g_J(x) = \int_x^\infty u^{J-1} \mathrm{e}^{-u} \,\mathrm{d}u.$$

Thus

$$\log f_J(\tau) = h_J(J\tau) - h_J((J+\alpha)\tau),$$

where $h_J(x) = \log g_J(x)$.

Now

$$h'_J(x) = \frac{-x^{J-1}\mathrm{e}^{-x}}{\int_x^\infty u^{J-1}\mathrm{e}^{-u}\,\mathrm{d}u} = -\frac{1}{\int_0^\infty \left(1+\frac{v}{x}\right)^{J-1} \mathrm{e}^{-v}\,\mathrm{d}v},$$

so, by the mean value theorem,

$$\log f_J(\tau) = \frac{\alpha\tau}{\int_0^\infty \left(1+\frac{v}{x}\right)^{J-1} \mathrm{e}^{-v}\,\mathrm{d}v},$$

for some $x \in (J\tau, (J+\alpha)\tau)$. Further,

$$\lim_{J\to\infty} \left(1+\frac{v}{J\tau}\right)^{J-1} = \lim_{J\to\infty} \left(1+\frac{v}{(J+\alpha)\tau}\right)^{J-1} = \mathrm{e}^{\frac{v}{\tau}}, \tag{8.36}$$

so since $J\tau < x < (J+\alpha)\tau$,

$$\begin{aligned}\lim_{J\to\infty} \log f_J(\tau) &= \frac{\alpha\tau}{\int_0^\infty \mathrm{e}^{\frac{v}{\tau}}\mathrm{e}^{-v}\,\mathrm{d}v} \\ &= \begin{cases} 0 & \text{if } \tau \leq 1, \\ \alpha(\tau-1) & \text{if } \tau > 1. \end{cases}\end{aligned}$$

(The two sequences in (8.36) are increasing, so the monotone convergence theorem can be used to justify the above limit.) The second part of the lemma now follows using (8.35), since $\bar{C}(\tau) = \bar{D}(\tau)/\bar{E}(\tau) = \bar{F}(\tau) - 1$. □

In the limit as $J \to \infty$, we focus on $0 < \tau \leq 1$ since for $\tau > 1$ the probability that the second death occurs time τ after the first death tends to 0 as $J \to \infty$. Using (8.15), it follows from Lemma 3, that, for $0 < \tau < 1$,

$$\begin{aligned}\mathbb{E}\left[s^{X_2^*} \middle| T_2 = \tau \right] &= \frac{1 + s\bar{C}(\tau)}{1 + \bar{C}(\tau)} \left(\frac{\bar{\pi}}{1 - (1 - \bar{\pi})s} \right)^2 \\ &\to \frac{1 - (1 - \exp(-\alpha\tau))s}{\exp(-\alpha\tau)} \left(\frac{\bar{\pi}}{1 - (1 - \bar{\pi})s} \right)^2, \end{aligned} \tag{8.37}$$

as $J \to \infty$. Applying Lemma 3.1 to (8.37), yields after straightforward algebraic manipulation, that in the limit as $J \to \infty$, for $0 < \tau < 1$,

$$\{X_2 | T_2 = \tau\} \sim \begin{cases} \text{NegBin}(2, \bar{\pi}) & \text{with probability } 1 - h(\tau) \\ \text{Geom}(\bar{\pi}) & \text{with probability } h(\tau) \end{cases}, \tag{8.38}$$

where

$$h(\tau) = \frac{\exp(\alpha\tau) - 1}{\exp(\alpha) - 1}. \tag{8.39}$$

Note that $h(1) = 1$, so in the limit as $J \to \infty$, $\{X_2 | T_2 = 1\} \sim \text{Geom}(\exp(-\alpha))$. This has a simple explanation, since if all lifetimes are equal to one and the second death occurs one time unit after the first death, then the initial individual had only one child, who was born as the initial individual dies. Therefore the population just after the second death comprises solely of the descendants of the second individual at its death, which follows a Geom(exp(−α)) distribution.

9. Constant Lifetime distribution

In this section, we explore further $L \equiv 1$ and as in Section 8 we assume that all deaths are detected. However, we allow a time-inhomogeneous birth rate. That is, an individual born at time t is alive on the interval $[t, t + 1)$ and during this time gives birth at the points of a time-inhomogeneous Poisson point process with rate β_u at time u, so if there are x individuals alive in the population at time u the infinitesimal birth rate is $x\beta_u$. Given that the first death is at time 0, the initial individual starts their lifetime at time $t = -1$ and we require β_u to be defined for $u \geq -1$, but unlike the general phase-type model given in Section 2 we do not require the birth rate to be constant before the first (detected) death.

For $s \geq -1$ and $0 < \tau \leq 1$, let $Z(s, \tau)$ denote the number of offspring alive at time $s + 1$ given there is a single individual alive at time s who dies at time $s + \tau$. It is straightforward using the exploration process outlined in Section 4.2, with minor modifications, to show that $Z(s, \tau)$ is a zero-modified Geometric random variable with probability mass function,

$$\mathbb{P}(Z(s,\tau)=0) = \exp\left(-\int_s^{s+\tau} \beta_u \, du\right) \tag{9.1}$$

$$\mathbb{P}(Z(s,\tau)=k) = \left[1-\exp\left(-\int_s^{s+\tau} \beta_u \, du\right)\right] [1-\tilde{\pi}_{s+1}]^{k-1} \tilde{\pi}_{s+1}, \qquad (k=1,2,\ldots). \tag{9.2}$$

where for $t \geq 0$, $\tilde{\pi}_t = \exp(-\int_{t-1}^t \beta_u \, du)$ is the probability that an individual alive on the interval $(t-1,t]$ has no offspring. The main observations in deriving (9.1) and (9.2) are that (9.1) is the probability that the initial individual has no offspring and since $L \equiv 1$ that any individual born in $(t-1,t]$ will be alive at time t. For $0 < \tau \leq 1$,

$$\exp\left(-\int_s^{s+\tau} \beta_u \, du\right) \geq \tilde{\pi}_{s+1}, \tag{9.3}$$

and it follows using Lemma 3.1 that

$$Z(s,\tau) \stackrel{D}{=} \begin{cases} \text{Geom}(\tilde{\pi}_{s+1}) & \text{with probability } \frac{1-\exp\left(-\int_s^{s+\tau} \beta_u \, du\right)}{1-\tilde{\pi}_{s+1}}, \\ 0 & \text{with probability } \frac{\exp\left(-\int_s^{s+\tau} \beta_u \, du\right)-\tilde{\pi}_{s+1}}{1-\tilde{\pi}_{s+1}}. \end{cases} \tag{9.4}$$

For $t \geq 0$, let $\mathcal{S}_t$ denote the death times of all individuals who die before time t. Let $\mathcal{R}_t = \mathcal{S}_t \backslash \mathcal{S}_{t-1}$, the set of death times in the interval $(t-1,t]$. Then the cardinality of $\mathcal{R}_t$ is $R_t = Y^*(t-1)$, the number of individuals alive at time $t-1$ and

$$Y^*(t)|\mathcal{S}_t \stackrel{D}{=} Y^*(t)|\mathcal{R}_t,$$

since no individuals born before time $t-1$ are alive at time t. Write $\mathcal{R}_t = \{(t-1+\rho_1),(t-1+\rho_2),\ldots,(t-1+\rho_{R_t})\}$, so $0 < \rho_i \leq 1$ is the remaining lifetime at time $t-1$ of the i^{th} to die in the interval $(t-1,t]$. Therefore if $\tilde{Z}_1, \tilde{Z}_2, \ldots, \tilde{Z}_{R_t}$ are independent random variables with $\tilde{Z}_i \stackrel{D}{=} Z(t-1,\rho_i)$, we have that

$$Y^*(t)|\mathcal{R}_t \stackrel{D}{=} \sum_{i=1}^{R_t} \tilde{Z}_i. \tag{9.5}$$

It follows straightforwardly from (9.4) and (9.5) that $Y^*(t)|\mathcal{R}_t$ is a mixture of $\{\text{NegBin}(k,\tilde{\pi}_t); k = 0,1,\ldots,R_t\}$.

We note that $Z(s,1) \sim \text{Geom}(\tilde{\pi}_{s+1})$, so for $X_k^*|\mathbf{T}_{2:k} = \mathbf{t}_{2:k} \stackrel{D}{=} Y^*(s_k)|\mathcal{R}_{s_k}$ is a mixture of $\{\text{NegBin}(k,\tilde{\pi}_t); k=1,2,\ldots,R_{s_k}\}$. In particular, for $k=2$ with $0 < t_2 (= s_2) \leq 1$, we have that

$$X_2^*|T_2 = t_2 \stackrel{D}{=} \begin{cases} \text{NegBin}(2,\tilde{\pi}_{t_2}) & \text{with probability } \frac{1-\exp\left(-\int_{t_2-1}^0 \beta_u \, du\right)}{1-\tilde{\pi}_{t_2}} \\ \text{Geom}(\tilde{\pi}_{t_2}) & \text{with probability } \frac{\exp\left(-\int_{t_2-1}^0 \beta_u \, du\right)-\tilde{\pi}_{t_2}}{1-\tilde{\pi}_{t_2}} \end{cases} \tag{9.6}$$

In the time-homogeneous case where $\beta_u = \alpha$ ($u \in \mathbb{R}$), we have that $\tilde{\pi}_t = \bar{\pi} = \exp(-\alpha)$ ($t \in \mathbb{R}$) and (9.6) becomes

$$X_2^*|T_2 = t_2 \sim \begin{cases} \text{NegBin}(2, \bar{\pi}) & \text{with probability } \frac{1-\exp(-(1-t_2)\alpha)}{1-\exp(-\alpha)} = 1 - h(t_2) \\ \text{Geom}(\bar{\pi}) & \text{with probability } \frac{\exp(-(1-t_2)\alpha)-\exp(-\alpha)}{1-\exp(-\alpha)} = h(t_2), \end{cases}$$

where $h(\cdot)$ is defined in (8.39), in agreement with (8.38).

10. Initial individual's birth time known

In this section we consider the case where the birth time of the initial individual, S_0, is known and is equal to $-t_0$, say, for some $t_0 > 0$. We derive the distribution of $\mathbf{X}_1$ in this case without any restrictions on the birth-rate, β_t, or detection probabilities, $\mathbf{d}_t$, $(-t_0 \leq t \leq 0)$ prior to the first detected death.

Let $\breve{\pi}_0 = 1 - \psi(-t_0; t_0)$ and for $j = 1, 2, \ldots, J$, let

$$\breve{\eta}_j^0 = \frac{\psi_j(-t_0; t_0)}{1 - \breve{\pi}_0} = \frac{\psi_j(-t_0; t_0)}{\psi(-t_0; t_0)}. \tag{10.1}$$

For $\mathbf{a} > -(1 - \breve{\pi}_0)\breve{\boldsymbol{\eta}}^0$, let $\breve{\mathbf{W}}(\mathbf{a})$ denote a J-dimensional random variable with, for $\boldsymbol{\theta} \in [0, 1]^J$, probability generating function (pgf)

$$\begin{aligned} \breve{\varphi}(\boldsymbol{\theta}; \mathbf{a}) &= \mathbb{E}\left[\prod_{j=1}^{J} \theta_j^{\breve{W}_j(\mathbf{a})}\right] \\ &= \frac{1 + \sum_{j=1}^J a_j \theta_j}{1 + \sum_{j=1}^J a_j} \times \frac{\breve{\pi}_0}{1 - (1 - \breve{\pi}_0)\sum_{j=1}^J \breve{\eta}_j^t \theta_j}. \end{aligned} \tag{10.2}$$

Note the similarity between the pgf of $\breve{\mathbf{W}}(\mathbf{a})$ and the pgf of $\mathbf{W}(t, \mathbf{a})$ defined in (3.4).

Lemma 4. *Suppose that the initial individual is born at time $S_0 = -t_0$ for some $t_0 > 0$. For $j = 1, 2, \ldots, J$, let*

$$\breve{c}_j = \frac{\sum_{i=1}^J \chi_i \sum_{l=1}^J d_{0,l}\gamma_l \left[\psi_l(-t_0; t_0) p_{ij}(-t_0, 0, t_0) - \psi_j(-t_0; t_0) p_{il}(-t_0, 0, t_0)\right]}{\sum_{i=1}^J \chi_i \sum_{l=1}^J d_{0,l}\gamma_l p_{il}(-t_0, 0, t_0)} \tag{10.3}$$

Then

$$\mathbf{X}_1 \sim \breve{\mathbf{W}}(\mathbf{0}) + \breve{\mathbf{W}}(\breve{\mathbf{c}}), \tag{10.4}$$

where $\breve{\mathbf{W}}(\mathbf{0})$ and $\breve{\mathbf{W}}(\breve{\mathbf{c}})$ are independent random variables.

Proof. The proof is similar to the proof of Lemma 5.2. We show that

$$\mathbb{E}\left[\prod_{j=1}^J \theta_j^{X_1^j} \middle| S_0 = -t_0\right] = \frac{1 + \sum_{j=1}^J \breve{c}_j \theta_j}{1 + \sum_{j=1}^J \breve{c}_j} \times \left[\frac{\breve{\pi}_0}{1 - (1 - \breve{\pi}_0)\sum_{j=1}^J \breve{\eta}_j^t \theta_j}\right]^2, \tag{10.5}$$

from which the lemma follows immediately.

For $i = 1, 2, \ldots, J$, the initial individual is type i at birth with probability χ_i. Therefore, for any $\boldsymbol{\theta} \in [0,1]^J$, we have that

$$\mathbb{E}\left[\prod_{j=1}^{J} \theta_j^{X_1^j} \middle| S_0 = -t_0\right] = \frac{\sum_{i=1}^J \chi_i H_D(\boldsymbol{\theta}; -t_0, t_0; \mathbf{e}_i)}{\sum_{i=1}^J \chi_i H_D(\mathbf{1}; -t_0, t_0; \mathbf{e}_i)}. \tag{10.6}$$

(*c.f.* (5.7)). Using Corollary 5.1 along with (5.2), we have that

$$H_D(\boldsymbol{\theta}; -t_0, t_0; \mathbf{e}_i) = \sum_{l=1}^{J} d_{0,l}\gamma_l \frac{\zeta(-t_0; t_0)}{\left[1 - \sum_{j=1}^J \psi_j(-t_0; t_0)\theta_j\right]^2} \times \left[p_{il}(-t_0, 0, t_0)\left(1 - \sum_{j=1}^{J} \psi_j(-t_0; t_0)\theta_j\right) + \psi_l(-t_0; t_0)\sum_{j=1}^{J} p_{ij}(-t_0; t_0)\theta_j\right]. \tag{10.7}$$

By summing (10.7) over i and simplifying, we have that

$$\sum_{i=1}^{J} \chi_i H_D(\boldsymbol{\theta}; -t_0, t_0; \mathbf{e}_i) = \frac{\zeta(-t_0; t_0)\sum_{i=1}^J \chi_i \sum_{l=1}^J d_{l,0}\gamma_l p_{il}(-t_0, 0, t_0)}{\left[1 - \sum_{j=1}^J \psi_j(-t_0; t_0)\theta_j\right]^2}\left[1 + \sum_{j=1}^{J} \check{c}_j\theta_j\right]. \tag{10.8}$$

By setting $\boldsymbol{\theta} = \mathbf{1}$ in (10.8) and substituting into (10.6), we have that (10.5) and the lemma is proved.

It is straightforward to modify Theorem 3.1 to obtain $\mathbf{X}_k|\mathbf{T}_{2:k}$ in the case where $S_0 = -t_0$ using Lemma 4 for $\mathbf{X}_1$. We set $\pi_0 = \check{\pi}_0$ and $\boldsymbol{\eta}^0 = \check{\boldsymbol{\eta}}^0$ and for $t > 0$, construct π_t and $\boldsymbol{\eta}^t$ using (3.1) and (3.2), respectively. Given the base step $\mathbf{X}_1 \stackrel{D}{=} \check{\mathbf{W}}(\mathbf{0}) + \check{\mathbf{W}}(\check{\mathbf{c}})$, the inductive step proceeds as in Theorem 3.1 (b), with $\mathbf{X}_{2:k}|\mathbf{T}_{2:k} = \mathbf{t}_{2:k}$ now being a mixture of $k!$ random variables each consisting of the sum of $k+1$ independent zero-modified geometric random variables.

11. Numerical Results

In this section we present two additional examples to demonstrate the approximation given in Section 3.5.2.

The first example is a simulated branching process with $L \sim \text{Gamma}(4,4)$ up until the 500^{th} death was observed. For $k = 1, 2, \ldots, 500$, let α_k and ε_k denote the birth rate and the probability of detecting the death of a type-4 individual, respectively, between the $(k-1)^{st}$ and k^{th} detected death. (Remember that since the lifetime is an Erlang distribution only type-4 individuals can die.) The birth rate and detection probability changed after every 100 detected deaths with $(\alpha_k, \varepsilon_k) = (\alpha_{200+k}, \varepsilon_{200+k}) = (\alpha_{400+k}, \varepsilon_{400+k}) = (2.0, 0.25)$ and $(\alpha_{100+k}, \varepsilon_{100+k}) = (\alpha_{300+k}, \varepsilon_{300+k}) = (0.5, 0.5)$ $(k =$

$1, 2, \ldots, 100$). Since $\mathbb{E}[L] = 1$ the branching process alternates after every 100 detected deaths between being super-critical ($R_0 = 2$) and sub-critical ($R_0 = 0.5$). In Figure 5, we plot the number of individuals alive, of each type and total number, immediately after a detected death against the number of detected deaths from a single realisation of the branching process, along with the median ($\hat{X}_k^j, j = 1, 2, 3;\ \hat{X}_k^*$) of the approximate distribution derived in Section 3.5.2. We also include the 5% and 95% quantiles of the approximate distribution, denoted l_k and u_k, with $[l_k, u_k]$ shaded for $k = 1, 2, \ldots, 500$. We observe that there is very different behaviour over time in the number of individuals of the four types (four stages of the Erlang distribution). For all four types the approximation captures the trajectories of the number of individuals alive of that type.

As in Section 7, we assess the performance of the approximate distribution based on 100 branching processes realisations using a P-P plot (Figure 6), of $\tilde{\mathbf{u}}_{\cdot,500,j}$, the ordered $\mathbf{u}_{\cdot,500,j}$ ($j = 1, 2, 3, 4$), where $u_{i,k,j}$ ($i = 1, 2, \ldots, 100; k = 1, 2, \ldots, 500; j = 1, 2, 3, 4$) are obtained using (7.1). The P-P plots demonstrate good performance of the approximate distribution for number of each type alive after the 500^{th} detected death. We also considered the above branching process with the detection probability changed in the sub-critical phases, in particular, we considered (a) $\varepsilon_{100+k} = \varepsilon_{300+k} = 0.25$, no change in the detection probability and (b) $\varepsilon_{100+k} = \varepsilon_{300+k} = 0.75$, a more signficant change in the detection probability. We noted better performance of the approximation distribution in (a) and worse performance of the approximation distribution in (b) supporting the notion that the approximation becomes worse as the detection probability changes more dramatically.

The second example is a simulated epidemic in a population of 2000 individuals with $L \sim \text{Gamma}(2, 2)$, infection rate $\beta = 1.5 (= R_0)$, detection probability $\varepsilon = 0.4$ (the probability an individual is detected on entering the removed state) and 1 initial infective in an otherwise susceptible population. The epidemic resulted in 480 detected removals (out of a total of 1239 removals) and we estimate the number of infectives of each type (each stage of the Erlang distribution) immediately after each removal. For $k = 1, 2, \ldots, 480$, let α_k denote the birth rate between the $(k-1)^{st}$ and k^{th} detected removal (death) of the approximating branching process. We set

$$\alpha_k = \beta \frac{1}{N} \left\{ N - \frac{k-1}{\varepsilon} - \mathbb{E}[\hat{X}_k^* | \mathbf{T}_{2:k-1} = \mathbf{t}_{2:k-1}] \right\} \tag{11.1}$$

where $(k-1)/\varepsilon$ and $\mathbb{E}[\hat{X}_{k-1}^* | \mathbf{T}_{2:k-1} = \mathbf{t}_{2:k-1}]$ are the estimated mean numbers of removed and infectives immediately after the $(k-1)^{st}$ detected removal. Note that if $\varepsilon = 1$, all removals are detected, the equation for α_k given by (11.1) reduces to the equation given in [2], Section 7. In Figure 7, we plot the number of infectives, of each type and in total, immediately after a detected removal against the number of detected removals, along with the approximate median ($\hat{X}_k^j, j = 1, 2; \hat{X}_k^*$) calculated using Section 3.5.2. We also include the 5% and 95%, l_k and u_k with $[l_k, u_k]$ shaded for $k = 1, 2, \ldots, 480$. We observe that the branching process approximation provides a good approximation to the trajectory of the total number, and the number of each type, of infectives over the entire course of the epidemic.

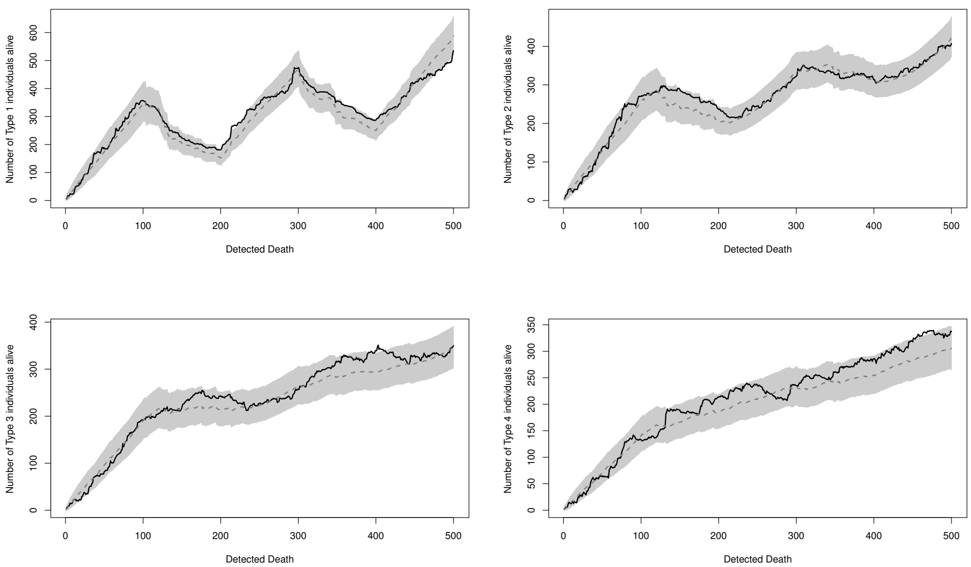

FIGURE 5: Number of individuals alive (solid line) and median of approximate distribution $\hat{X}_k^z|\mathbf{t}_{2:k}$ ($z = 1,2,3,4$) (dashed line) up to the 500^{th} detected death with $L \sim$ Gamma(4, 4), $(\alpha_k, \varepsilon_k) = (\alpha_{200+k}, \varepsilon_{200+k}) = (\alpha_{400+k}, \varepsilon_{400+k}) = (2.0, 0.25)$ and $(\alpha_{100+k}, \varepsilon_{100+k}) = (\alpha_{300+k}, \varepsilon_{300+k}) = (0.5, 0.5)$ ($k = 1, 2, \dots, 100$). The shaded area represents the probability mass between the 5% and 95% quantiles of $\hat{X}_k^z|\mathbf{t}_{2:k}$. Top left: Number of Type 1 individuals alive; Top right: Number of Type 2 individuals alive; Bottom left: Number of Type 3 individuals alive; Bottom right: Number of Type 4 individuals alive.

We again assess the performance of the approximate distribution based on 100 epidemic realisations with the population size and parameters given above. We restrict attention to epidemics which take-off and result in at least 100 detected removals with the number of detected removals ranging from 383 to 548. In Figure 8, we use P-P plots of the $\tilde{\mathbf{u}}_{\cdot,k,j}$ ($k = 100, 200, 300,\ j = 1, 2$). The plots show that using the branching process approximation with birth rate given by (11.1) provides a good approximation for the number of infectives of each type in the epidemic process.

References

[2] BALL, F. AND NEAL, P. (2023) The size of a Markovian SIR epidemic given only removal data. *Adv. Appl. Prob.* **55**, 895–926.

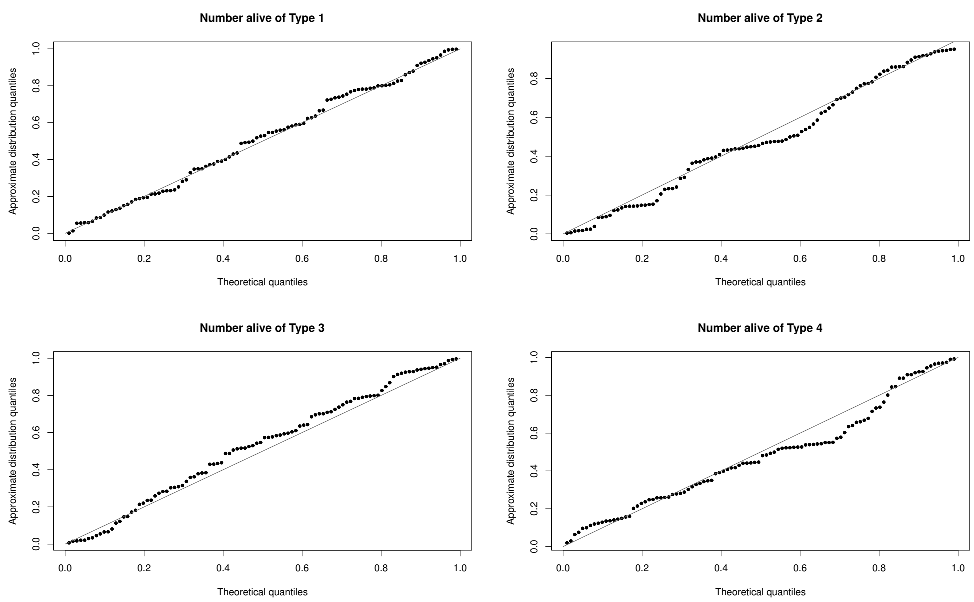

FIGURE 6: P-P plots based on 100 simulations of the ordered quantiles of $\tilde{\mathbf{u}}_{\cdot,500,j}$ ($j = 1, 2, 3, 4$), where $L \sim \text{Gamma}(4, 4)$, $(\alpha_k, \varepsilon_k) = (\alpha_{200+k}, \varepsilon_{200+k}) = (\alpha_{400+k}, \varepsilon_{400+k}) = (2.0, 0.25)$ and $(\alpha_{100+k}, \varepsilon_{100+k}) = (\alpha_{300+k}, \varepsilon_{300+k}) = (0.5, 0.5)$ ($k = 1, 2, \ldots, 100$). Top left: Number of Type 1 individuals alive; Top right: Number of Type 2 individuals alive; Bottom left: Number of Type 3 individuals alive; Bottom right: Number of Type 4 individuals alive.

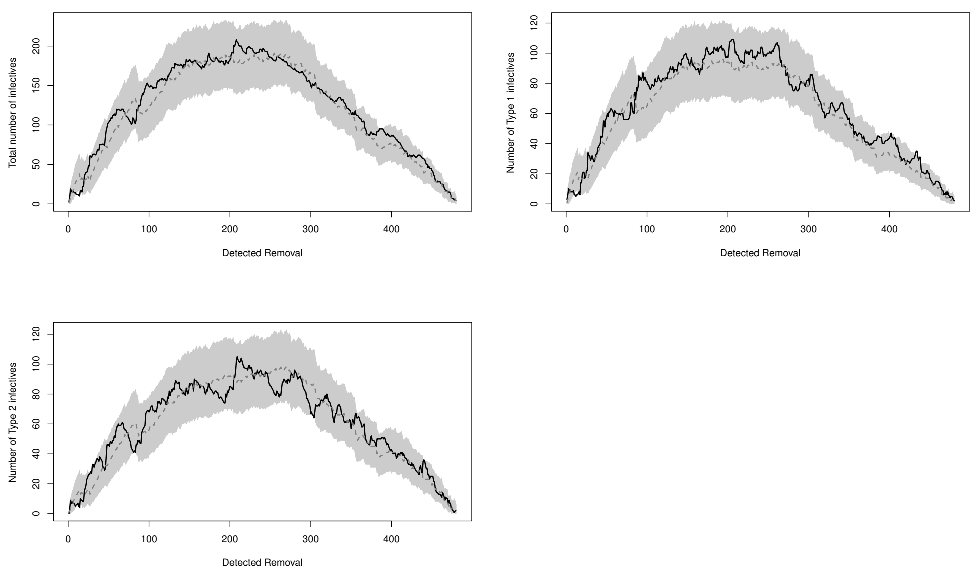

FIGURE 7: Number of infectives (solid line) and median $\hat{X}_k^z|\mathbf{t}_{2:k}$ $(z = *, 1, 2, 3)$ (dashed line) after each detected removal k in an epidemic in a population of size 2000 which infects 480 detected removals with infection rate $\beta = 1.5$, removal detection probability $\varepsilon = 0.4$ and $L \sim \text{Gamma}(2, 2)$. For the branching process approximation $\alpha_k = \beta\{N - (k-1)/\varepsilon - \mathbb{E}[\hat{X}_{k-1}^*|\mathbf{T}_{2:k-1} = \mathbf{t}_{2:k-1}]\}/N$. The shaded area represents the probability mass between the 5% and 95% quantiles of $\hat{X}_k^z|\mathbf{t}_{2:k}$. Top left: Total number of infectives; Top right: Number of Type 1 infectives; Bottom left: Number of Type 2 infectives

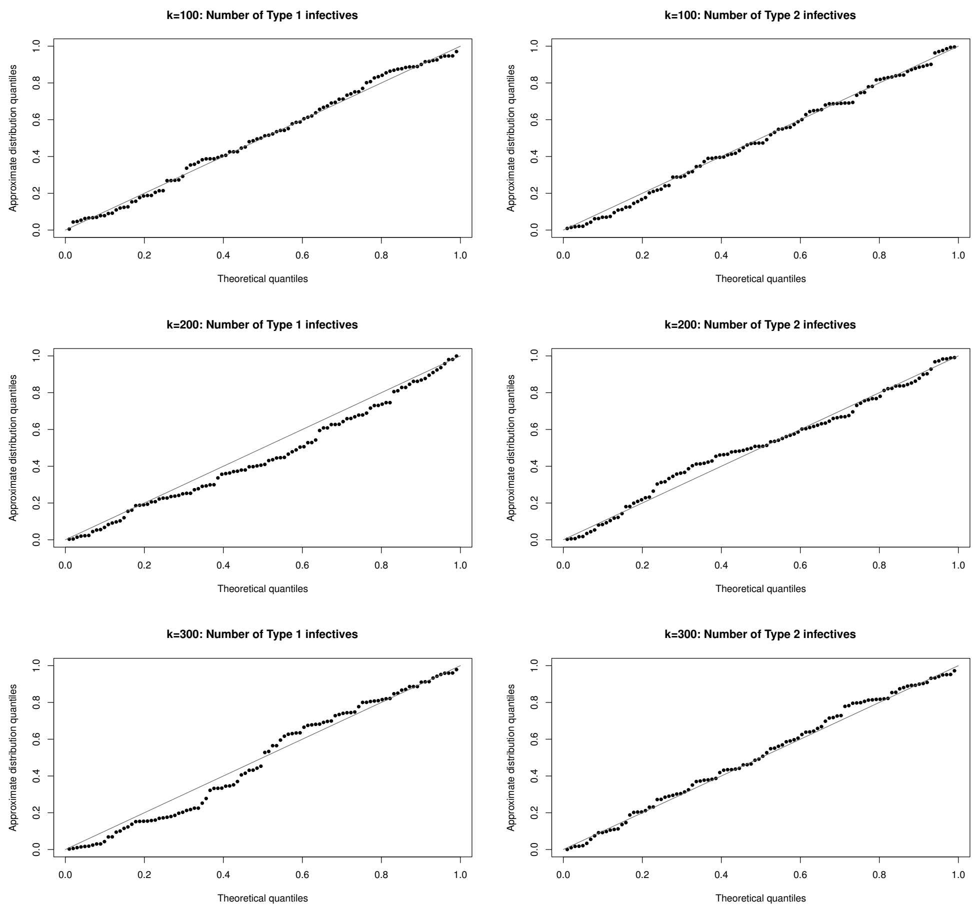

FIGURE 8: P-P plots based on 100 simulations of the ordered quantiles of $\tilde{\mathbf{u}}_{\cdot,k,j}$ for $k = 100, 200, 300$ and $j = 1, 2$, where $L \sim \text{Gamma}(2, 2),$, $\alpha_k = \beta\{N - (k-1)/\varepsilon - \mathbb{E}[\hat{X}^*_{k-1}|\mathbf{T}_{2:k-1} = \mathbf{t}_{2:k-1}]\}/N$. Top left: Number of Type 1 infectives $k = 100$; Top right: Number of Type 2 infectives $k = 100$; Middle left: Number of Type 1 infectives $k = 200$; Middle right: Number of Type 2 infectives $k = 200$; Bottom left: Number of Type 1 infectives $k = 300$; Bottom right: Number of Type 2 infectives $k = 300$.