INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 9309–9331 PII: S0305-4470(02)37688-1

Population death sequences and Cox processes driven by interacting Feller diffusions

Gang Wei[1], **Peter Clifford**[2] **and Jianfeng Feng**[3]

[1] Department of Mathematics, Baptist University, Hong Kong, People's Republic of China
[2] Department of Statistics, 1 South Parks Road, Oxford OX1 3TG, UK
[3] COGS, Sussex University, Brighton BN1 9QH, UK

Received 6 June 2002
Published 22 October 2002

Abstract

We carry out a complete study on the relationship between Cox processes driven by interacting Feller diffusions and death sequences of immigration–emigration linked population networks. It is first proved that the Cox process driven by a Feller diffusion is equivalent to the death sequence of a birth and death process. The conclusion is then generalized to the case of Cox processes driven by interacting Feller diffusions and death sequences of interacting populations.

PACS numbers: 02.50.Ga, 87.18.Su

1. Introduction

Let $Y(t)$ be the Markovian population process (see [2, 7, 16, 18] or section 2 for an exact definition) with constants ν, λ, μ as the rates of immigration, birth and death. For the process $Y(t)$, assume that we are able to observe the total number of deaths $D(t)$ (death sequence) in $[0, t]$ for $t > 0$ (see figure 1).

On the other hand, we take into account a Cox process $C(t)$ (see figure 2 and section 2 for an exact definition) with a rate defined by

$$\mathrm{d}Z(t) = p\sqrt{Z(t)}\,\mathrm{d}B_t + (qZ(t) + r)\,\mathrm{d}t$$

where B_t is the standard Brownian motion, p, q, r are appropriate constants, i.e. $Z(t)$ is the Feller diffusion. Without a noise term, we see that the stable state of $Z(t)$ is $-r/q$. Hence $Z(t)$ is the process which describes the fluctuations around $-r/q$. To ensure that $\sqrt{Z(t)}$ is properly defined, it is required that $-r/q > 0$.

Establishing a relationship between a discrete process, say $Y(t)$, and a continuous process, say $Z(t)$, is a long-term endeavour in probability theory. Usually, it is easy to deal with a continuous process due to the powerful tools developed in stochastic analysis, but the discrete process might give us a more clearly defined physical meaning. Certainly, the two processes $Y(t)$ and $Z(t)$ are substantially different: the former is a discrete process and will by no means converge to a continuous process, comparing figure 1 with figure 2. As a consequence, one

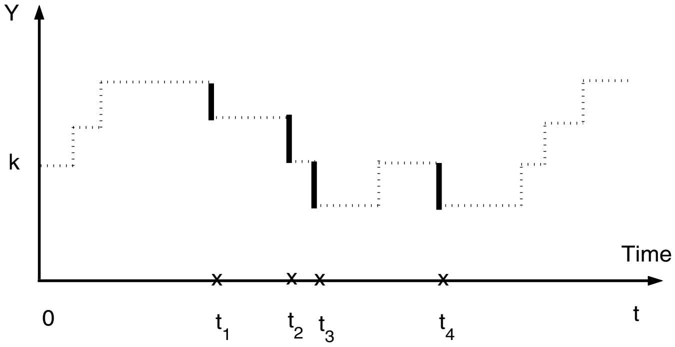

Figure 1. Schematic plot of the process $Y(t)$ starting with $Y(0) = k$. Thick, vertical lines indicate the death sequence of $Y(t)$. The observable events are the death sequence, as indicated by × on the time axis. In the figure we have $D(t) = 4$.

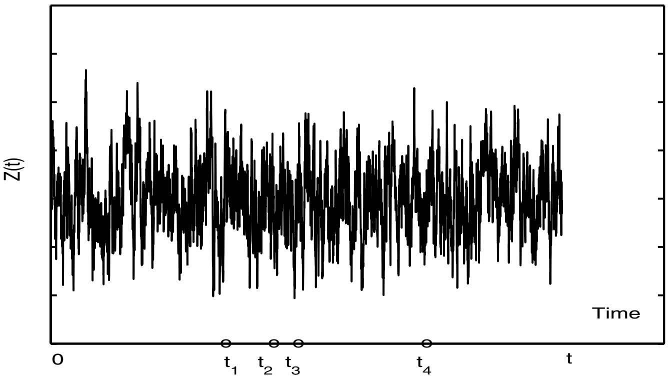

Figure 2. A Feller process $Z(t)$ starting with $Z(0) = 1$ and $p = 0.1, q = -1, r = 1$(line). The observable events (Cox process) are the process with intensity $Z(t)$, as indicated by o on the time axis. In the figure we have $C(t) = 4$.

might thus conclude that the two point processes $D(t)$ and $C(t)$ are not relevant at all. However, we show that by appropriately choosing the parameters (p, q, r) in the Feller process, the two processes $D(t)$ and $C(t)$ are equivalent.

Our motivation to establish the relationship between $D(t)$ and $C(t)$ is from neuroscience. Traditionally and as a first-order approximation, it is accepted in neuroscience that each neuron sends and receives signals taking the form of a Poisson process [8, 9, 23]. However, due to the large noise presented in the brain, it is natural to use the Cox process (see, for example, [1, 15]) rather than the Poisson process to fit experimental data. Recently, the approach has been found to be quite promising [1, 15]. One of the bottlenecks to applying the theory of Cox processes to experimental data lies in the fact that it is much harder to estimate parameters in the Cox process, in comparison with the Poisson process where only one parameter (the firing intensity) is needed to estimate. It is found that directly calculating the maximum likelihood function of the Cox process is very time consuming [5, 6, 11–14]. On the other hand, computationally it is much more efficient to calculate the maximum likelihood function of $D(t)$, albeit its form is still complex [4]. According to the theory developed in this paper, we assert that the aforementioned estimations based upon $D(t)$ and $C(t)$ are identical. We will report the estimation of parameters in the Cox process using experimental data in future publications [4] and in the present paper we concentrate on the establishment of the equivalent relationship between the two processes $C(t)$ and $D(t)$.

One might naturally ask whether our conclusions are true for more general situations. Consider a sequence of coupled birth and death processes: the death process of a current population is actually to emigrate from the population and immigrate to the next population. Are there corresponding Cox processes driven by Feller diffusions so that the final outcomes (the observable variables) of two classes of processes are equivalent? We prove that it is the case and the corresponding Cox processes are exactly given.

In summary, we prove the equivalent relationship of two classes of point processes which are generated by seemingly unrelated processes. The equivalent relationship enables us to gain some further insights into the statistics of the processes and to develop novel algorithms to estimate various model parameters. For the application of our theory to other areas, we refer the reader to [19–22].

The paper is organized as follows. In section 2, we first prove the equivalence between the death sequence of $Y(t)$ and a Cox process driven by a Feller diffusion. The main tool we employ is the probability generating functional (see the next section for an exact definition). In section 3, the interacting birth and death networks are considered and the equivalence between the outcomes of birth and death networks and Cox processes driven by interacting Feller processes are established. In appendices A and B, we show how to calculate various functions of the processes using the probability generating functional. The application of current theory to experimental data from neuroscience will be reported in a separate publication [4].

2. Equivalence between a single death sequence and a Cox process

$Y(t)$ is defined as a Markov chain taking values in $\{0, 1, 2, \ldots\}$ with the infinitesimal transition probabilities

$$P(Y(t+h) = n+m|Y(t) = n) = \begin{cases} \lambda_n h + o(h) & \text{if} \quad m = 1 \\ \mu_n h + o(h) & \text{if} \quad m = -1 \\ o(h) & \text{if} \quad |m| > 1 \end{cases}$$

where $h > 0$, $\lambda_n = n\lambda + \nu$ and $\mu_n = n\mu$.

A Cox process (or doubly stochastic Poisson process) $C(t)$ is a nonhomogeneous Poisson process with the rate function itself being a random process. In our considerations, the rate function is given by the Feller diffusion $Z(t)$.

The basic tool we employ to tackle the problem is the probability generating functional (pgf). Let $\mathcal{G}$ be the set of all non-negative and measurable functions bounded by 1, and $1 - \xi$ has finite support, i.e. $\xi(t) = 1$ when t is large enough.

Definition 1. *The* pgf *of a one-dimensional counting process* $N(t)$ *is*

$$G(\xi) = \mathbf{E}\left(\prod_i \xi(t_i)\right) \tag{2.1}$$

$$= \mathbf{E}\left\{\exp\left[\int_0^\infty \log \xi(u)\, \mathrm{d}N(u)\right]\right\} \tag{2.2}$$

where $\{t_i, i = 1, 2, \ldots\}$ *are the times when the points are counted (see, for example, figures 1 and 2) and* $\xi \in \mathcal{G}$.

If two point processes $X(t)$ and $Y(t)$ have identical pgfs, then the two processes are identical in distribution (see appendix B for details).

For concreteness of notation, we sometimes denote $\mathrm{IBD}(\nu, \lambda, \mu) = Y(t)$. For a birth and death process $Y(t)$, we sometimes denote $D(Y(t))$ is its death sequence; for a Feller diffusion $Z(t)$, $C(Z(t))$ is its corresponding Poisson process with the rate $Z(t)$.

2.1. The death sequence of IBD (ν, λ, μ)

By definition, given $Y(0) = k$, the pgf associated with $D(t) = D(Y(t))$ is

$$G_k(\xi) = \mathbf{E}\left\{\exp\left[\int_0^\infty \log\xi(u)\,\mathrm{d}D(u)\right]\middle|\, Y(0) = k\right\}.$$

Lemma 1. *The pgf of the death sequence* $D(t)$ *for the* IBD (ν, λ, μ) *process with initial population size* $Y(0) = k$ *is given by*

$$G_k(\xi) = f^k(0)\exp\left\{\nu\int_0^\infty (f(t) - 1)\,\mathrm{d}t\right\} \tag{2.3}$$

where f *satisfies (the Ricatti equation)*

$$f'(t) - (\lambda + \mu)f(t) + \lambda f^2(t) = -\mu\xi(t) \qquad t \geqslant 0 \tag{2.4}$$

with the boundary condition $f(t) = 1$ *for* $t \geqslant \limsup\{u : 1 - \xi(u) > 0\}$.

Proof. Let $X(t)$ be the population size in a simple birth and death process $\mathrm{BD}(\lambda, \mu)$ and let $D(t)$ be the total number of deaths in $[0, t]$.

We define the following conditional pgf,

$$f(t) \overset{\mathrm{def}}{=} \mathbf{E}\left\{\exp\left[\int_0^\infty \log\xi(t+u)\,\mathrm{d}D(u)\right]\middle|\, X(0) = 1\right\} \tag{2.5}$$

where $\xi \in \mathcal{G}$. When there is no risk of ambiguity, for convenience, we always write $f(t)$ for $f(t, \xi)$. Equivalently, by a simple time shift, this is the pgf of the death sequence generated by a population process starting with a single individual at time *t*.

The time *S* of the first event, be it birth or death, is a random variable with density

$$g(s) = (\lambda + \mu)\exp[-(\lambda + \mu)s] \qquad 0 \leqslant s < \infty. \tag{2.6}$$

The process $X(t)$ is strongly Markovian and the random time *S* is a stopping time. If the first event is a death, the process dies out; if the first event is a birth, the process splits into two independent processes, with each of them being stochastically equivalent to the original one except that the time origin is *S*.

The probability ratio of whether the birth or the death occurs first is $\lambda : \mu$. Hence the functional $f(t)$, as a function of *t*, satisfies the integral equation

$$f(t) = \int_0^\infty \mathrm{e}^{-(\lambda+\mu)s}\{\lambda f^2(t+s) + \mu\xi(t+s)\}\,\mathrm{d}s. \tag{2.7}$$

We rewrite this as

$$f(t) = \int_t^\infty \mathrm{e}^{-(\lambda+\mu)s}\,\mathrm{e}^{(\lambda+\mu)t}[\lambda f^2(s) + \mu\xi(s)]\,\mathrm{d}s$$

with the appropriate boundary condition. Hence, differentiating both sides of the equation above with respect to *t* and rearranging terms, we have

$$f'(t) - (\lambda + \mu)f(t) + \lambda f^2(t) = -\mu\xi(t). \tag{2.8}$$

Now we turn to consider $Y(t)$, the IBD(ν, λ, μ) process. Since the sequence of deaths for this process is the superposition of the death sequences generated by new immigrants and their descendants, the pgf starting with population size zero can be written as

$$G_0(\xi) = \mathbf{E}\left(\prod_j f(T_j)|Y(0) = 0\right) \tag{2.9}$$

where T_j is the time of the jth immigration.

Let $D_1(t)$ be the number of immigrations in $(0, t)$, that is a Poisson process with rate ν, then the well-known pgf of Poisson process leads to

$$G_0(\xi) = \mathbf{E}\left\{\exp\left[\int_0^\infty f(t)\,\mathrm{d}D_1(t)\right]\middle|Y(0) = 0\right\} \tag{2.10}$$

$$= \exp\left(\int_0^\infty \nu(f(t) - 1)\,\mathrm{d}t\right). \tag{2.11}$$

Combining (2.5) and (2.11) we have

$$G_k(\xi) = \mathbf{E}\left\{\exp\left[\int_0^\infty \log\xi(u)\,\mathrm{d}D(u)\right]\middle|Y(0) = k\right\} \tag{2.12}$$

$$= f^k(0)\exp\left[\nu\int_0^\infty (f(t) - 1)\,\mathrm{d}t\right] \tag{2.13}$$

which completes the proof. □

2.2. *Cox process driven by Feller diffusion*

We consider the Poisson process with the Feller diffusion as its intensity. The intensity process satisfies the stochastic differential equation

$$\mathrm{d}Z(t) = p\sqrt{Z(t)}\,\mathrm{d}B_t + (qZ_t + r)\,\mathrm{d}t \tag{2.14}$$

where $p^2 = 2\lambda\mu$, $q = \lambda - \mu$, and $r = \mu\nu$. Let $C(t)$ be the number of events in $[0, t]$ and let

$$H(z, \xi) = \mathbf{E}\left\{\exp\left[\int_0^\infty \log\xi(t)\,\mathrm{d}C(t)\right]\middle|Z(0) = z\right\}$$

denote the pgf of the point process, starting at $Z(0) = z$.

Lemma 2. *The Poisson process with the Feller diffusion $Z(t)$ as its intensity has pgf*

$$H(z, \xi) = \exp\left\{(f(0) - 1)z/\mu + \nu\int (f(t) - 1)\,\mathrm{d}t\right\} \tag{2.15}$$

where z is the initial value of $Z(t)$ and f satisfies

$$f'(t) - (\lambda + \mu)f(t) + \lambda f^2(t) = -\mu\xi(t) \tag{2.16}$$

with the boundary condition $f(t) \to 1$ as $t \to \infty$.

Proof. Since it is well known that the pgf of the Poisson$(Z(t))$ is

$$H(z, \xi) = \mathbf{E}\left[\mathbf{E}\left[\exp\left\{\int_0^\infty \log\xi(t)\,\mathrm{d}C(t)\right\}\middle|Z(t), 0 \leqslant t < \infty\right]\middle|Z(0) = z\right]$$

$$= \mathbf{E}\left[\exp\left\{\int_0^\infty (1 - \xi(u)Z(u))\,\mathrm{d}u\right\}\middle|Z(0) = z\right]$$

we define

$$A(t) = \int_0^t (1 - \xi(u))Z(u)\,\mathrm{d}u \tag{2.17}$$

so that the pgf associated with $C(t)$ is given by

$$\mathbf{E}\left\{\exp\left[-\int_0^\infty (1 - \xi(u))Z(u)\,\mathrm{d}u\right]\middle| Z(0) = z\right\} = \lim_{t\to\infty} \mathbf{E}(\mathrm{e}^{-A(t)}|Z(0) = z). \tag{2.18}$$

Let

$$M(t) = \exp\{-A(t) + a(t)Z(t) + b(t)\}. \tag{2.19}$$

We intend to find functions $a(t)$ and $b(t)$ so that $M(t)$ is a martingale. Writing

$$g(x, t) = \exp\{-A(t) + a(t)x + b(t)\} \tag{2.20}$$

we have

$$\begin{cases} \dfrac{\partial g}{\partial x} = ga(t) \\ \dfrac{\partial g}{\partial t} = g[-A'(t) + a'(t)x + b'(t)] \\ \dfrac{\partial^2 g}{\partial x^2} = ga^2(t). \end{cases} \tag{2.21}$$

Hence, using Ito's formula, the infinitesimal mean of $M(t)$ has the following form,

$$\begin{aligned} \mu_M(t) &= g\left[a(t)\mu_{Z(t)} + (-A'(t) + a'(t)Z(t) + b'(t)) + \tfrac{1}{2}a^2(t)\sigma^2_{Z(t)}\right] \\ &= g\left[a(t)(qZ(t) + r) - A'(t) + b'(t) + a'(t)Z(t) + \tfrac{1}{2}a^2(t)\sigma^2_{Z(t)}\right] \\ &= g\left[Z(t)\left(a(t)q + a'(t) + \tfrac{1}{2}p^2a^2(t) - 1 + \xi(t)\right) + b'(t) + a(t)r\right] \end{aligned}$$

where $\mu_{Z(t)}$ and $\sigma_{Z(t)}$ are the drift and diffusion terms of $Z(t)$ respectively.

We know that $\mu_M(t) = 0$ is a sufficient condition to ensure that $M(t)$ is a martingale. This implies that $a(t)$ and $b(t)$ should satisfy the following differential equations:

$$\begin{cases} a'(t) + qa(t) + \tfrac{1}{2}p^2a^2(t) - 1 = \xi(t) \\ b'(t) + a(t)r = 0. \end{cases} \tag{2.22}$$

Now define $a(t)$ by

$$a(t) = -\frac{1}{\mu}(1 - f(t)) \tag{2.23}$$

so that

$$a'(t) = \frac{1}{\mu}f'(t) \qquad a^2(t) = \frac{1}{\mu^2}(1 - 2f(t) + f^2(t)).$$

Putting all related terms into the left-hand side of (2.22) we get

$$a'(t) + qa(t) + \frac{1}{2}p^2a^2(t) - 1 = \frac{1}{\mu}[f'(t) - (\lambda + \mu)f(t) + \lambda f^2(t)]. \tag{2.24}$$

Therefore equation (2.22) is identical to equation (2.8) or (2.16). Furthermore

$$M(t) = \exp\left\{-A(t) + \frac{1}{\mu}(f(t) - 1)Z(t) - \nu\int_0^t (f(u) - 1)\,\mathrm{d}u\right\} \tag{2.25}$$

is an L^2-martingale since $0 \leqslant M(t) \leqslant 1$. Hence with the basic property of a martingale, condition on $Z(0) = z$, we have

$$\begin{aligned}\lim_{t\to\infty} \mathbf{E}M(t) &= M(0)\\ &= \exp\left\{\frac{z}{\mu}(f(0)-1)\right\}.\end{aligned}$$

Also from the definition

$$\lim_{t\to\infty} \mathbf{E}M(t) = \lim_{t\to\infty} \mathbf{E}(\mathrm{e}^{-A(t)}) \exp\left[-\nu\int_0^\infty (f(t)-1)\,\mathrm{d}t\right]. \tag{2.26}$$

Thus, together with equation (2.18), the pgf of the $C(t)$ condition on $Z(0) = z$ is

$$H(z,\xi) = \lim_{t\to\infty} \mathbf{E}(\mathrm{e}^{-A(t)}|Z(0)=z) \tag{2.27}$$

$$= \exp\left\{-\frac{z}{\mu}(1-f(0)) - \nu\int_0^\infty (1-f(t))\,\mathrm{d}t\right\}. \tag{2.28}$$

which completes the proof. □

2.3. Equivalence between two processes $D(t)$ and $C(t)$

We establish equivalent relationships between two processes $D(t)$ and $C(t)$ with different initial distributions. From now on we always assume that $\mu > \lambda$. The first equivalent relationship is about the processes with stationary initial distributions.

Theorem 1. *In equilibrium, the Cox process $C(t)$ driven by Feller diffusion $Z(t)$ (defined in equation (2.14)) and the death time process $D(t)$ of the* IBD (ν,λ,μ) *process are stochastically equivalent, i.e. they are identical in distribution.*

Proof. In equilibrium $Z(0)$ has a gamma distribution with parameters (δ,β) so that

$$\mathbf{E}(\mathrm{e}^{uZ(0)}) = (1-\mu\beta u)^{-\delta} \tag{2.29}$$

where $\delta = \nu/\lambda$, and $\beta = \lambda/(\mu-\lambda)$. Taking expectations on both sides of equation (2.15) we have the pgf of the Cox process in equilibrium,

$$H(\xi) = \left(1+\mu\beta\frac{1}{\mu}(1-f(0))\right)^{-\delta} \exp\left[-\nu\int_0^\infty (1-f(t))\,\mathrm{d}t\right] \tag{2.30}$$

$$= \left(\frac{\mu-\lambda}{\mu-\lambda f(0)}\right)^{\delta} \exp\left[-\nu\int_0^\infty (1-f(t))\,\mathrm{d}t\right]. \tag{2.31}$$

Similarly, when the IBD process is in equilibrium $Y(0)$ has a NegBin(c,δ) distribution with

$$\mathbf{E}u^{Y(0)} = [(1-c)/(1-cu)]^{\delta}$$

where $c = \lambda/\mu$, $\delta = \nu/\lambda$. Again taking expectations in both sides of equation (2.3) we have

$$\begin{aligned}G(\xi) &= \sum_{k=0}^{\infty} \pi_k G_k(\xi)\\ &= \left[\frac{1-c}{1-cf(0)}\right]^{\delta} \exp\left[-\nu\int_0^\infty (1-f(t)\,\mathrm{d}t\right]\\ &= \left[\frac{\mu-\lambda}{\mu-\lambda f(0)}\right]^{\delta} \exp\left[-\nu\int_0^\infty (1-f(t))\,\mathrm{d}t\right].\end{aligned} \tag{2.32}$$

Hence the two pgfs agree with each other. From the results in appendix B, we conclude that $D(t)$ and $C(t)$ are identical in distribution. ☐

Next we consider two more special cases: $Z(0) = 0$ and $Z(0) = z_0 > 0$, a constant.

Corollary 1. *The death sequence $D(t)$ of the* IBD (ν, λ, μ) *process with initial population $Y(0) = 0$ is stochastically equivalent to the Cox process driven by Feller diffusion (2.14) with initial intensity $Z(0) = 0$. The corresponding pgf is*

$$G(\xi) = \exp\left[-\nu \int_0^\infty (1 - f(t))\, \mathrm{d}t\right]. \tag{2.33}$$

Proof. This directly follows from equations (2.15) and (2.3). ☐

Corollary 2. *The Cox process $C(t)$ driven by Feller diffusion (2.14) conditioned on the initial intensity $Z(0) = z_0$ is stochastically equivalent to the death time process $D(t)$ of the* IBD (ν, λ, μ) *process with the initial population sampled from* $\mathrm{Poisson}(z_0/\mu)$. *The corresponding pgf is equation (2.28).*

Proof. When $Y(0)$ is sampled from $\mathrm{Poisson}(z_0/\mu)$,

$$\mathbf{E} f(0)^{Y(0)} = \exp\{-(1 - f(0)) z_0/\mu\} \tag{2.34}$$

which implies that the pgf of $D(t)$ with the random initial distribution is coincident with equation (2.28). ☐

3. Equivalence between a population death sequence and a Cox process driven by interacting Feller processes

In this section, we extend the results of the previous section to the point processes associated with multi-population systems and interacting Feller diffusions. We start by showing that the equivalent results in section 2 can be generalized to a nonhomogeneous birth and death process with an immigration rate depending on time.

3.1. IBD$(\nu(t), \lambda, \mu)$ process and generalized Feller diffusion

In section 2 we showed that the pgf of the death time sequence of the stable IBD(ν, λ, μ) process takes the form

$$G(\xi) = \left[\frac{\mu - \lambda}{\mu - \lambda f(0)}\right]^\delta \exp\left[-\int_0^\infty \nu(1 - f(t))\, \mathrm{d}t\right] \tag{3.1}$$

where f is defined in equation (2.4). We note that the proofs of the equivalence results in section 2 remain valid when we allow ν to depend on time as an a.s. continuous positive function. We then conclude as follows.

Theorem 2. *The sequence of death times $D(t)$ for the* IBD $(\nu(t), \lambda, \mu)$ *process is stochastically equivalent to the Cox process $C(t)$ driven by generalized Feller diffusion $Z(t)$ defined by*

$$\mathrm{d}Z(t) = p\sqrt{Z(t)}\, \mathrm{d}B(t) + (qZ(t) + r)\, \mathrm{d}t \tag{3.2}$$

with $p^2 = 2\lambda\mu$, $q = \lambda - \mu$, $r = \mu\nu(t)$ and the initial population size $Y(0)$ being distributed as $\mathrm{Poisson}(Z_0/\mu)$. *The corresponding pgf is*

$$H(\xi) = \exp\left\{-\frac{Z_0}{\mu}(1 - f(0)) - \int_0^\infty (1 - f(t))\nu(t)\, \mathrm{d}t\right\}. \tag{3.3}$$

In particular, the result is true when the population and diffusion processes both start from 0.

Proof. Here we will not repeat the proofs similar to those in section 2, instead we merely mention two crucial facts. The first is that the function f does not depend on $\nu(t)$. The second is that the exponential martingale becomes

$$M(t) = \exp\left[\int_0^t (1-\xi(u))Z(u)\,\mathrm{d}u + \frac{Z(t)}{\mu}(f(t)-1) - \int_0^t \nu(u)(f(u)-1)\,\mathrm{d}u\right]. \tag{3.4}$$

The other arguments remain valid. □

Note that the equivalence depends on the particular relationship between the initial states of the population process and the Feller diffusion. For brevity of the later arguments, we formally introduce the following definition to describe such a relationship.

Definition 2. *Suppose that $M_Z(u)$ is the moment generating function of a positive and continuous random variable Z and $m_Y(u)$ is the generating function of a discrete random Y. If*

$$M_Z(u) = m_Y(u+1) \qquad \textit{for} \quad -2 \leqslant u \leqslant 0. \tag{3.5}$$

then we say that Z and Y are compatible.

Lemma 3. *The random variables $Z(0)/\mu$ and $Y(0)$ are compatible if and only if $Y(0) \overset{\text{dist}}{=}$* Poisson$(Z(0)/\mu)$.

Proof. This is elementary. □

It is well known that the IBD(ν, λ, μ) process, when $\mu > \lambda$, has equilibrium distribution as negative binomial which is a gamma mixture of Poisson distributions. This gamma distribution is the equilibrium distribution of the corresponding Feller diffusion, divided by μ, which is taken as the intensity of the Cox process. Now we generalize this result to the non-stable situation with immigration rate as an a.s. continuous non-negative function of time t. The conclusion below is needed in the following developments.

Lemma 4. *Let $Y(t)$ be the population process* IBD $(\nu(t), \lambda, \mu)$ *and let $Z(t)$ be the generalized Feller diffusion characterized by the stochastic differential equation*

$$\mathrm{d}Z(t) = p\sqrt{Z(t)}\,\mathrm{d}B(t) + (qZ(t) + r)\,\mathrm{d}t \tag{3.6}$$

with $p^2 = 2\lambda\mu, \quad q = \lambda - \mu, \quad r = \mu\nu(t)$. If $Z(t)/\mu$ and $Y(t)$ are compatible at time $t = 0$ then they are compatible for all $t > 0$.

Proof. Consider the population process. Let $p_i(t) = \text{prob}(Y(t) = i)$. The forward equations are

$$p_i'(t) = -[(\lambda+\mu)i + \nu]p_i(t) + [(i-1)\lambda + \nu]p_{i-1}(t) + \mu(i+1)p_{i+1}(t). \tag{3.7}$$

Let $m_{Y(t)} = G(u,t) = \sum_{i=0}^{\infty} p_i(t)u^i$, then we have

$$\begin{aligned}\frac{\partial G}{\partial u} &= \sum_{i=1}^{\infty} i p_i(t) u^{i-1} \\ \frac{\partial G}{\partial t} &= -(\lambda+\mu)u\frac{\partial G}{\partial u} - \nu G + \lambda u^2 \frac{\partial G}{\partial u} + \nu u G + \mu \frac{\partial G}{\partial u} \\ &= \nu(u-1)G + [\mu - (\lambda+\mu)u + \lambda u^2]\frac{\partial G}{\partial u}.\end{aligned} \tag{3.8}$$

Hence

$$\frac{\partial G}{\partial t} = [\lambda u - \mu][u-1]\frac{\partial G}{\partial u} + \nu(u-1)G \tag{3.9}$$

where $\nu = \nu(t)$ is the time-dependent immigration rate.

For the diffusion $Z(t)$, writing $X(t) = Z(t)/\mu$, we have

$$\mathrm{d}X(t) = a\sqrt{X(t)}\,\mathrm{d}B(t) + (bX(t) + c)\,\mathrm{d}t \tag{3.10}$$

where $a^2 = 2\lambda$, $b = q = \lambda - \mu$ and $c = \nu(t)$.

The corresponding forward equation is

$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial}{\partial y^2}[a^2 yp] - \frac{\partial}{\partial y}[(by+c)p]. \tag{3.11}$$

Consider the moment generating function

$$M_{X(t)} = H(t,u) = \int_0^\infty p(t,y)\,\mathrm{e}^{uy}\,\mathrm{d}y \tag{3.12}$$

we have

$$\begin{aligned}\frac{\partial H}{\partial t} &= \int_0^\infty \frac{\partial p}{\partial t}\,\mathrm{e}^{uy}\,\mathrm{d}y \\ &= \frac{1}{2}a^2\int_0^\infty y\,\mathrm{e}^{uy}\frac{\partial^2 p}{\partial y^2}\,\mathrm{d}y + \int_0^\infty (a^2 - by + c)\frac{\partial p}{\partial y}\,\mathrm{e}^{uy}\,\mathrm{d}y - bH.\end{aligned}$$

And since

$$\begin{aligned}\int_0^\infty y\,\mathrm{e}^{uy}\frac{\partial^2 p}{\partial y^2}\,\mathrm{d}y &= -\int_0^\infty \frac{\partial p}{\partial y}[1+uy]\,\mathrm{e}^{uy}\,\mathrm{d}y \\ &= \int_0^\infty \mathrm{e}^{uy}[u + u(1+uy)]p\,\mathrm{d}y \\ &= 2u\int_0^\infty p\,\mathrm{e}^{uy}\,\mathrm{d}y + u^2\int_0^\infty yp\,\mathrm{e}^{uy}\,\mathrm{d}y \\ &= 2uH + u^2\frac{\partial H}{\partial u}\end{aligned}$$

$$\begin{aligned}\int_0^\infty (a^2 - by + c)\frac{\partial p}{\partial y}\,\mathrm{e}^{uy}\,\mathrm{d}y &= (a^2 - by + c)\,\mathrm{e}^{uy}p|_{y=0}^{\infty} \\ &\quad - \int_0^\infty p\,\mathrm{e}^{uy}[-b + u(a^2 + c - by)]\,\mathrm{d}y \\ &= (b - (a^2 + c)u)H + bu\frac{\partial H}{\partial u}\end{aligned}$$

eventually

$$\frac{\partial H}{\partial t} = u[\lambda u + (\lambda - \mu)]\frac{\partial H}{\partial u} + \nu uH. \tag{3.13}$$

By comparing this equation with the forward equation for the generating function $m_{Y(t)}$, we conclude that if $H(t,u) = G(t,u+1)$ at time $t = 0$ then the relationship holds true for all $t > 0$. The lemma follows.

Recall that for the equivalence result to hold we require that $Y(0)$ and $Z(0)/\mu$ are compatible. The preceding lemma shows that if the processes are compatible at the beginning of some period of time they are compatible at the end. At this time, it is then possible to alter the parameter λ to another constant and proceed with another period of equivalence which also preserves compatibility, and so on. It follows that the parameter λ may also vary with time,

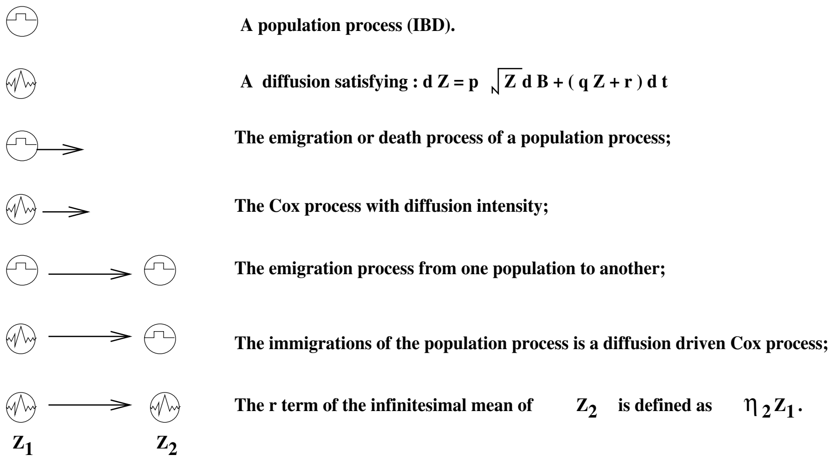

Figure 3. Key to symbols used in diagrams of stochastic systems.

without affecting the equivalence. This could be realized by the standard limit procedure of approximating the a.s. continuous function by step functions. Hence we have the following more general theorem. ☐

Lemma 5. *Lemma 4 remains true if the rates of immigration and birth are a.s. continuous positive functions of time t, which could be the sample paths of independent Markov processes.*

3.2. Linear chains of populations

Lemma 5 is concerned with inhomogeneous populations and enables us to investigate more general population systems. Let us start from a simple one.

For given positive numbers $\eta_1, \eta_2, \ldots, \eta_k$ and functions $\lambda_j(t) \geqslant 0$, $j = 1, \ldots, k$ with $\eta_j > \lambda_j(t)$, we define a sequence of birth and death processes[4] $Y_{1j}(t) = \mathrm{IBD}(I_{1j}(t), \lambda_j(t), \eta_j)$, $j = 1, \ldots, k$, where $I_{10}(t) = \nu(t)$, $I_{1j}(t) = \eta_{j-1} Y_{1(j-1)}$, for $j = 2, \ldots, k$. Hence the decrease epochs of $Y_{1(j-1)}(t)$ actually mark the emigration from $Y_{1(j-1)}(t)$ to $Y_{1j}(t)$. What we observe is the outcome (the death sequence) $D_1(t)$ from $Y_{1k}(t)$. Denote the defined hierarchical population as $\mathcal{S}_1 = \{Y_{1i}, i = 1, \ldots, k\}$.

We define the system of interacting diffusions $\mathcal{S}_k = \{Z_1, \ldots, Z_k\}$ by

$$\mathrm{d}Z_i = p_i \sqrt{Z_i}\,\mathrm{d}B(t) + (q_i Z_i + r_i)\,\mathrm{d}t \qquad i = 1, \ldots, k \tag{3.14}$$

where $p_i^2 = 2\lambda_i(t)\eta_i$, $q_i = (\lambda_i(t) - \eta_i)$, $r_{ki} = \eta_i Z_{i-1}$. The process $Z_0(t)$ is introduced for notational convenience and equals $\nu(t)$ for all $t > 0$. Let $C_k(t)$ be the Poisson process with intensity $Z_k(t)$.

To carry out a complete comparison between the death sequence $D_1(t)$ and the Cox process $C_k(t)$, in fact we introduce an array of birth and death processes and Cox processes (see figures 3 and 4). The ith system $\mathcal{S}_i$ consists of $k - i + 1$ population processes $Y_{ii}, \ldots, Y_{ik}$, linked in a way similar to $\mathcal{S}_1$. More exactly the population Y_{ij} is an immigration, birth, emigration (death) process $\mathrm{IBD}(I_{ij}, \lambda_j, \eta_j)$ where $\eta_j > \lambda_j$, and $I_{ij} = \eta_{j-1} * Y_{i(j-1)}$ for $j = i + 1, \ldots, k$. The immigration process for Y_{ij} coincides with the emigration process from $Y_{i(j-1)}$ for $j = i + 1, \ldots, k$. Furthermore the immigration into Y_{ii} is a Poisson process with

---

[4] We use the convention that $Y_{i(j)} = Y_{ij}$.

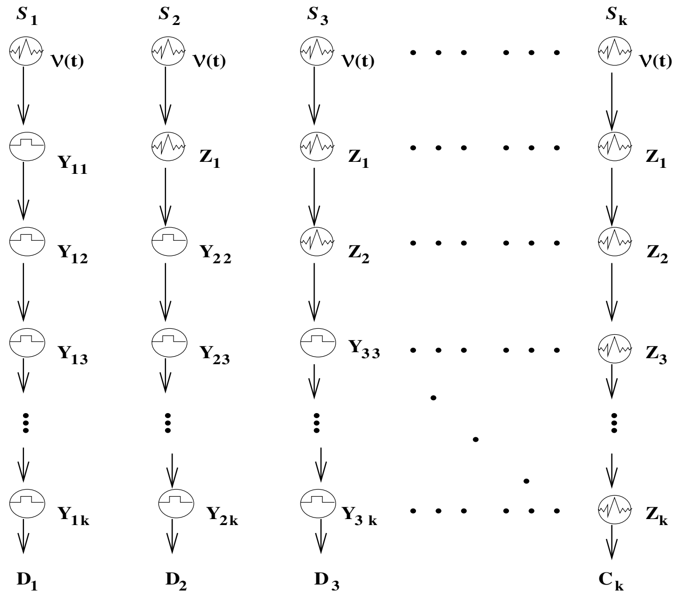

Figure 4. Linearly connected stochastic systems (see context for definition).

intensity Z_{i-1}. Finally, we denote by $D_i(t)$ the emigration process (death sequence) in the last population of the *i*th system.

For reasons of brevity, we introduce some symbols to represent the complicated systems we consider. The basic symbols are shown in figure 3.

Figure 4 illustrates the interacting diffusion processes and population processes defined above. Then we have the following theorem.

Theorem 3. *The point processes* $(D_1, D_2, \ldots, D_{k-1}, C_k)$ *constructed as in figure 4 are stochastically equivalent to each other, provided* $Z_j(0)/\eta_j$ *and* $Y_{ij}(0)$ *are compatible for all* $j = 1, \ldots, k, i = 1, \ldots, k$.

Proof. The proof follows from the theorems in the previous section by induction. □

3.3. Simultaneous death sequences and Cox processes

In this section, we consider system $\mathcal{S}_i, i = 1, \ldots, k$ as defined above but it allows both death and emigration at the intermediate steps (see figure 5). We first define Cox processes driven by

$$\mathrm{d}Z_i = p_i\sqrt{Z_i}\,\mathrm{d}B_i + (q_iZ_i + r_i)\,\mathrm{d}t \tag{3.15}$$

where

$$p_i^2 = 2\lambda_i(t)(\mu_i + \eta_i) \qquad q_i = \lambda_i(t) - (\mu_i + \eta_i) \tag{3.16}$$

$$r_1 = (\mu_1 + \eta_1)\nu(t) \tag{3.17}$$

$$r_i = (\mu_i + \eta_i)Z_{k,i-1}\eta_{i-1}/(\mu_{i-1} + \eta_{i-1}) \qquad i = 2, 3, \ldots, k. \tag{3.18}$$

For each diffusion, C_i is defined to be a Poisson process with rate $\mu_iZ_i/(\mu_i + \eta_i)$ for $i = 1, \ldots, k$.

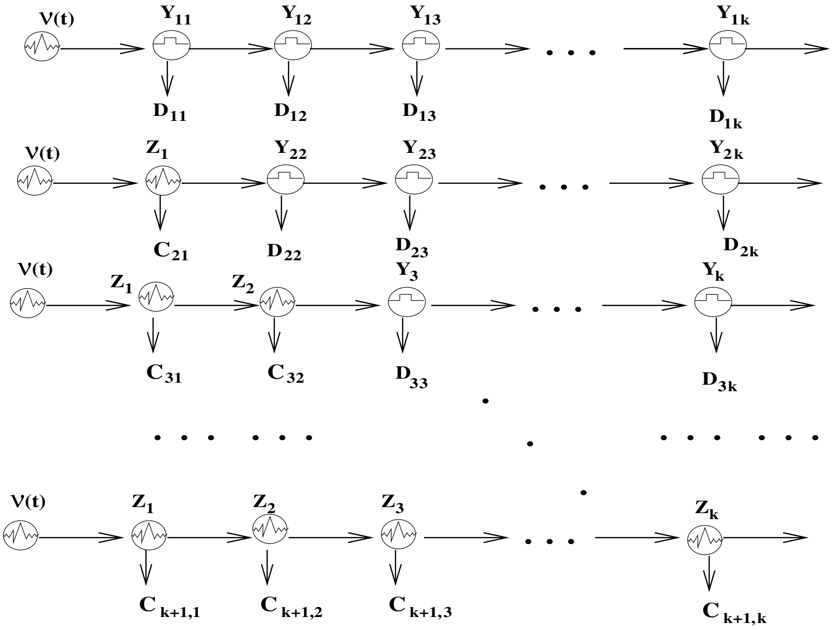

Figure 5. The mixed one-direction immigration linked systems.

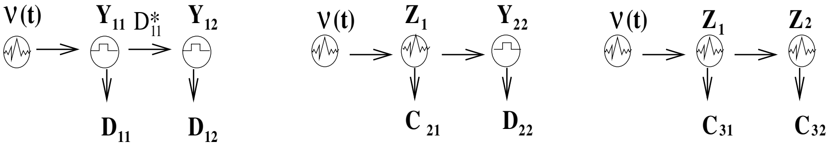

Figure 6. The one-direction system for $k = 2$.

Let us now introduce the splitting between the emigration and death for a population process $Y_{ij}(t)$ or a Cox process $Z_{i,j+1}(t)$ to another population process $Y_{i,j+1}(t)$. We only introduce the emigration and death for Y_{11} and Y_{12} or for Z_{21} and Y_{22}. The general situation is similar. A decrease in Y_{11} is generated with a rate $(\eta_1 + \mu_1)Y_{11}$, and then taken as a death event in D_{11} or an emigration event in D^*_{11} with probability ratio $\mu_1 : \eta_1$. This is shown in figures 7.

For the death and emigration from a Cox process to a birth and death process, as shown in figure 7, C_{11} and C^*_{11} are independently conditioned on Z_1 and are taken as two conditional independent Poisson processes with intensities $\mu_1 Z_1/(\mu_1+\eta_1)$ and $\eta_1 Z_1/(\mu_1+\eta_1)$ respectively.

Theorem 4. *For the stochastic systems shown in figure 5, the point processes satisfy*

$$(D_{11}, D_{12}, D_{13}, \ldots, D_{1k}) \sim (C_{21}, D_{22}, D_{23}, \ldots, D_{2k}) \sim \cdots \sim (C_{k1}, C_{k2}, \ldots, C_{kk})$$

provided that the initial conditions are compatible, where $\sim$ *means equivalent in distribution.*

Proof. We need only prove the claim for the case $k = 2$. In the case $k = 2$ the systems have the form shown in figure 6.

It is straightforward to see that (D_{11}, D_{12}) and (C_{21}, D_{22}) are equivalent. From the superposition property of Poisson processes, C_{21} and C^*_{21} (see figure 7) are independently

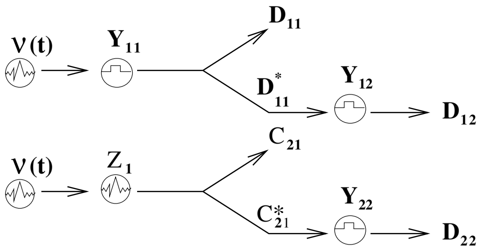

Figure 7. Upper panel: the detail of the mechanism between Y_{11} and Y_{12}. Bottom panel: the detail of the mechanism between Z_1 and Y_{22}.

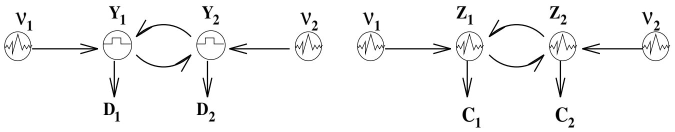

Figure 8. Bi-directional immigration–emigration systems.

conditioned on Z_{21}. Also C_{31}, C_{32} in the diffusion system are conditionally independent of the same parameters. Hence

$$(D_{11}, D_{12}) \sim (C_{21}, D_{22}) \sim (C_{31}, C_{32})$$

which completes our proof. □

We now complete our discussions on systems with one-direction interactions. It is certainly more interesting to take into account the case of mutual interactions, which is the case of the next subsection.

3.4. *Population networks and interacting Feller diffusions*

A very natural conjecture is that the theorems concerning equivalence in previous subsections should hold true for population systems where there are immigrations to and from populations as shown in figure 8.

In other words, for each $i = 1, 2$, the process Y_i is IBDE$(I_i, \lambda_i, \mu_i, \eta_i)$ with an immigration process consisting of new arrivals at rate ν_i together with emigrants from the other population process. For example, $I_1 = \nu_1 + Y_2\eta_2$. We are interested in the death sequences D_1 and D_2 of the two populations.

The diffusion system corresponds to stochastic differential equations

$$\mathrm{d}Z_1 = p_1\sqrt{Z_1}\,\mathrm{d}B_1 + (q_1 Z_1 + r_1)\,\mathrm{d}t \tag{3.19}$$

$$\mathrm{d}Z_2 = p_2\sqrt{Z_2}\,\mathrm{d}B_2 + (q_2 Z_2 + r_2)\,\mathrm{d}t \tag{3.20}$$

where

$$p_1^2 = 2\lambda_1(\mu_1 + \eta_1) \qquad q_1 = \lambda_1 - (\mu_1 + \eta_1)$$
$$r_1 = (\mu_1 + \eta_1)(\nu_1 + Z_2\eta_2/(\mu_2 + \eta_2))$$
$$p_2^2 = 2\lambda_2(\mu_2 + \eta_2) \qquad q_2 = \lambda_2 - (\mu_2 + \eta_2)$$
$$r_2 = (\mu_2 + \eta_2)(\nu_2 + Z_1\eta_1/(\mu_1 + \eta_1)).$$

We are interested in the coupled Cox processes C_1, C_2 given by Poisson processes with rates $Z_i(t)\mu_i/(\mu_i + \eta_i)$ for $i = 1, 2$. The interaction between Z_1 and Z_2 is linear, i.e.

$$\begin{aligned} dZ_1 &= p_1\sqrt{Z_1}\,dB_1 + [q_1 Z_1 + (\mu_1 + \eta_1)(\nu_1 + Z_2\eta_2/(\mu_2 + \eta_2))]\,dt \\ &= p_1\sqrt{Z_1}\,dB_1 + [-k_1 Z_1 + k_2 Z_2 + k_3]\,dt \end{aligned}$$

where $k_1 > 0, k_2 > 0$ are constants. Hence the interaction between Z_1 and Z_2 is scaled diffusive, the most common interaction and extensively studied in the literature.

Theorem 5. *For the systems described above, with compatible initial distributions, we have* $(D_1, D_2) \sim (C_1, C_2)$, *provided that* $\lambda_1(t) = \lambda_1$ *(constant) and* $\lambda_2(t) = \lambda_2$ *(constant).*

Proof. One approach to this problem is to define a bivariate probability generating functional for the output processes. Suppose that ξ and ζ are in $\mathcal{V}$ then the bivariate pgf of D_1, D_2 is given by

$$G(\xi, \zeta) = \mathbf{E}\left[\prod_{s_i \in D_1}\prod_{t_j \in D_2} \xi(s_i)\zeta(t_j)\right] = \exp\left[\int_0^\infty \log\xi(u)D_1(du) + \int_0^\infty \log\zeta(u)D_2(du)\right].$$

As in the single population case, we start by considering the conditional pgfs when there is no immigration from outside the system. We define

$$f_i(t) = \mathbf{E}\left\{\exp\left[\int_0^\infty \log\xi(t+s)D_1(ds) + \int_0^\infty \log\zeta(t+s)D_2(ds)\right]\middle| Y_i(0) = 1\right\} \quad \text{for} \quad i = 1, 2.$$

As before, f_1, f_2 satisfy Ricatti equations

$$0 = f_1' - (\mu_1 + \lambda_1 + \eta_1)f_1 + \lambda_1 f_1^2 + \mu_1\xi + \eta_1 f_2 \tag{3.21}$$

$$0 = f_2' - (\mu_2 + \lambda_2 + \eta_2)f_2 + \lambda_2 f_2^2 + \mu_2\zeta + \eta_2 f_1 \tag{3.22}$$

and the joint pgf of D_1, D_2 conditional on $Y_1(0) = y_1$ and $Y_2(0) = y_2$ can be written as

$$G(\xi, \zeta\,|\,y_1, y_2) = \prod_{i=1}^{2}\left\{f_i(0)^{y_i}\,\exp\left[\int_0^\infty \nu_i(t)(f_i(t) - 1)\,dt\right]\right\}.$$

For the diffusion system, we have the martingale $M(t)$ given by ($\xi_1 = \xi, \xi_2 = \zeta$)

$$\log M(t) = \sum_{i=1}^{2} -\frac{\mu_i}{\mu_i + \eta_i}\int_0^t (1 - \xi_i)Z_i\,du + \frac{Z_i}{\mu_i + \eta_i}(f_i - 1) - \int_0^t \nu_i(f_i - 1)\,du$$

so that by the arguments used for the one population case, we obtain the joint pgf of C_1, C_2 given $Z_1 = z_1$ and $Z_2 = z_2$ as

$$H(\xi, \zeta\,|\,z_1, z_2) = \prod_{i=1}^{2}\left\{\exp\left[\frac{z_i}{\mu_i + \eta_i}(f_i(0) - 1) + \int_0^\infty \nu_i(f_i(u) - 1)\,du\right]\right\}.$$

Using the usual compatibility conditions for the initial conditions G and H equality of the pgfs follows and the theorem is proved. □

Note that in the above proof we have used the assumption that the birth rates are constant, because only then do we have the neat form of the pfgs. But such an assumption is not necessary, and a more heuristic proof is given below without referring to the pgf explicitly. The proof covers the case with inhomogeneous birth rates.

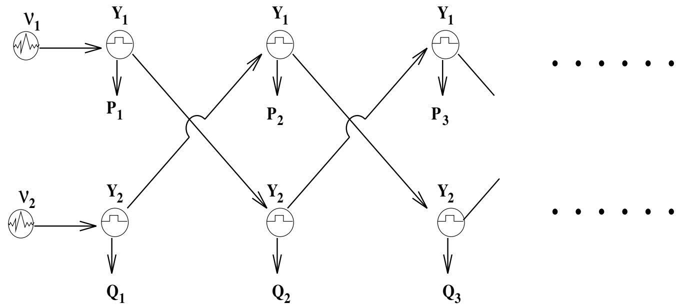

Figure 9. Mutual immigration systems.

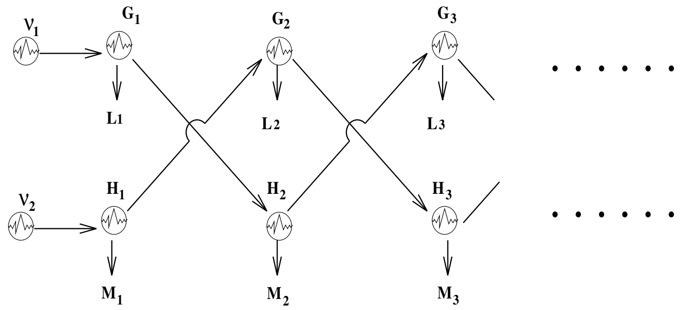

Figure 10. The diffusion system corresponding to split population systems.

A heuristic proof. Since each individual behaves independently, according to the possible path of each individual in the population, the population system can be split into the system shown in figure 9, where $D_1 = \sum_{i=1}^{\infty} P_i$ and $D_2 = \sum_{i=1}^{\infty} Q_i$.

Corresponding to this system we have a diffusion system as shown in figure 10.

From previous theorems, we know that

$$(P_1, Q_2, P_3, Q_4, \ldots) \sim (L_1, M_2, L_3, M_4, \ldots)$$

and they are independent of

$$(Q_1, P_2, Q_3, P_4, \ldots) \sim (M_1, L_2, M_3, L_4, \ldots).$$

Hence

$$D_1 = \sum_{i=1}^{\infty} P_i \sim \sum_{i=1}^{\infty} L_i = C_1' \qquad D_2 = \sum_{i=1}^{\infty} Q_i \sim \sum_{i=1}^{\infty} M_i = C_2' \qquad (D_1, D_2) \sim (C_1', C_2').$$

The $G_i, H_i, i = 1, 2, \ldots$ satisfy the stochastic differential equations

$$\begin{aligned}
\mathrm{d}G_1 &= p_1\sqrt{G_1}\,\mathrm{d}B_1 + [q_1 G_1 + (\mu_1 + \eta_1)\nu_1]\,\mathrm{d}t \\
\mathrm{d}G_2 &= p_1\sqrt{G_2}\,\mathrm{d}B_2 + \left[q_1 G_2 + (\mu_1 + \eta_1)H_1 \frac{\eta_2}{\mu_2 + \eta_2}\right]\mathrm{d}t \\
\mathrm{d}G_3 &= p_1\sqrt{G_3}\,\mathrm{d}B_3 + \left[q_1 G_3 + (\mu_1 + \eta_1)H_2 \frac{\eta_2}{\mu_2 + \eta_2}\right]\mathrm{d}t \\
\cdots &\quad \cdots \\
\mathrm{d}H_1 &= p_2\sqrt{H_1}\,\mathrm{d}B_1' + [q_2 H_1 + (\mu_2 + \eta_2)\nu_2]\,\mathrm{d}t
\end{aligned}$$

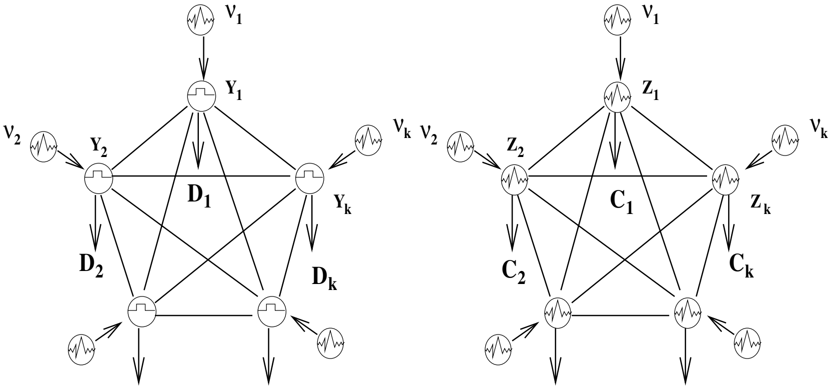

Figure 11. The multi-population net and the corresponding diffusion system.

$$\begin{aligned} \mathrm{d}H_2 &= p_2\sqrt{H_2}\,\mathrm{d}B_2' + \left[q_2H_2 + (\mu_2+\eta_2)G_1\frac{\eta_1}{\mu_1+\eta_1}\right]\mathrm{d}t \\ \mathrm{d}H_3 &= p_2\sqrt{H_3}\,\mathrm{d}B_3' + \left[q_2H_3 + (\mu_2+\eta_2)G_2\frac{\eta_1}{\mu_1+\eta_1}\right]\mathrm{d}t \\ \ldots &\quad \ldots. \end{aligned}$$

Summing these diffusions and using

$$\sum_i \sigma_i\,\mathrm{d}B_i = \sqrt{\sum_i \sigma_i^2}\,\mathrm{d}B \tag{3.23}$$

(see the remark following the proof) we have

$$\begin{aligned} \sum \mathrm{d}G_i &= p_1\left(\sum\sqrt{G_i}\,\mathrm{d}B_i\right) + \left[q_1\sum G_i + (\mu_1+\eta_1)\left(\nu_1 + \frac{\eta_2}{\mu_2+\eta_2}\sum H_i\right)\right]\mathrm{d}t \\ &= p_1\sqrt{\sum G_i}\,\mathrm{d}\tilde{B} + \left[q_1\sum G_i + (\mu_1+\eta_1)\left(\nu_1 + \frac{\eta_2}{\mu_2+\eta_2}\sum H_i\right)\right]\mathrm{d}t \\ \sum \mathrm{d}H_i &= p_2\left(\sum\sqrt{H_i}\,\mathrm{d}B_i'\right) + \left[q_1\sum H_i + (\mu_2+\eta_2)\left(\nu_2 + \frac{\eta_1}{\mu_1+\eta_1}\sum G_i\right)\right]\mathrm{d}t \\ &= p_2\sqrt{\sum H_i}\,\mathrm{d}\tilde{B}' + \left[q_2\sum H_i + (\mu_2+\eta_2)\left(\nu_2 + \frac{\eta_1}{\mu_1+\eta_1}\sum G_i\right)\right]\mathrm{d}t \end{aligned}$$

which means

$$\left(\sum G_i, \sum H_i\right) \overset{\text{dist}}{=} (Z_1, Z_2).$$

Hence

$$(C_1', C_2') \sim (C_1, C_2)$$

which completes the proof. □

Remark. Note that we have used the fact that the summation of independent generalized Feller diffusions is again a generalized Feller diffusion. This claim is not generally true for other diffusions. The reason we could perform such an operation ‘legally’ is the linearity of the infinitesimal parameters.

The generalization of the above results to more complex networks such as figure 11 is straightforward.

In figure 11 the lines connecting populations represent that populations interact by immigration in one direction or two directions. Similar explanations apply to the lines connecting Feller diffusions. We suppose that D_i is the death sequence of Y_i, C_i is the Cox process with intensity $\propto Z_i$. Equivalent relationships between two systems can be established accordingly.

4. Conclusions and discussions

We have explored the relationships between point processes generated by birth and death processes and Feller diffusion processes. Although the birth and death process and the Feller process seem quite different, surprisingly we have asserted that the generated point processes (the death sequence from the birth and death process and the Cox process from the Feller diffusion) are equivalent. The conclusion is not only true for the processes generated by a single birth and death process and single Feller process, but also the equivalent relationship remains valid for interacting populations and interacting Feller processes. The relationship provides us with a valuable tool to carry out further statistical inferences on various parameters of the model.

Here are some notes on technical aspects:

- In all our proofs, the rates of death and emigration have to be constants as the necessity of everlasting compatibility.
- The pgf is a powerful tool to investigate point processes, as demonstrated in this paper. With respect to point processes, they are analogues of moment generating functions or moments of random variables. In appendices A and B, some further applications of pgf are included.
- The heuristic reasoning concerning interacting diffusions is transparent because of the linearity of the infinitesimal coefficients of Feller diffusions, even for the inhomogeneous situations. Based upon this, we doubt whether there are other Cox processes driven by diffusions which can be analysed similarly.

Acknowledgment

This work is partially supported by EPSRC.

Appendix A. Increasing time sequence

We have found the pgf for death sequences of a birth and death process. What is the pgf for an increasing time sequence of a birth and death process? And furthermore, what is the relationship between the increasing time sequence and the death sequence? We answer these questions here.

Theorem 6. *For the* IBD (ν, λ, μ) *process, the pgf of the times of population increase is given by*

$$G_k(\xi) = \left[\frac{\mu}{\mu + \lambda(1 - f(0))}\right]^{k+\delta} \exp\left\{\nu \int_0^\infty (f(t) - 1)\,\mathrm{d}t\right\} \tag{A.1}$$

where k is the initial population size and the function f was defined in (2.4).

Proof. We first consider the pgf of the increases in the simple birth–death process BD(λ, μ) starting from a single individual. Let

$$g(t) \overset{\text{def}}{=} \mathbf{E}\left(\prod_i \xi(t+s_i)|Y(0)=1\right)$$

where s_i is the ith birth instance.

The time of the first event (birth or death) after time 0 has exponential density with parameter $(\lambda+\mu)$. The first event is birth or death with probability ratio $\lambda:\mu$. If it is a death the process dies out, if it is a birth the process splits into two i.i.d. point processes starting at that time instance. Hence we have the following,

$$\begin{aligned} g(t) &= \int_0^\infty \mathrm{e}^{-(\lambda+\mu)s}(\lambda g^2(t+s)\xi(t+s)+\mu)\,\mathrm{d}s \\ &= \int_t^\infty \mathrm{e}^{(\lambda+\mu)t}\,\mathrm{e}^{-(\lambda+\mu)s}(\lambda g^2(s)\xi(s)+\mu)\,\mathrm{d}s \end{aligned}$$

so that

$$g' - (\lambda+\mu)g + (\lambda g^2\xi + \mu) = 0 \tag{A.2}$$

with the boundary condition $g(t) \to 1$ as $t \to \infty$. It is therefore straightforward to verify that

$$g(t) = \frac{1}{1+\lambda(1-f(t))/\mu} \tag{A.3}$$

with f as defined in (2.4).

Now consider the conditional pgf of the increase times of the IBD(ν, λ, μ) process starting from zero population size, that is

$$\mathbf{E}\left(\prod_i \xi(s_i)|Y(0)=0\right)$$

where s_i is the ith increase. Since the population process is strongly Markovian and all the immigration times T_j, $j=1,2,\ldots$ are stopping times, we have

$$\begin{aligned} \mathbf{E}\left[\prod_i \xi(s_i)|Y(0)=0\right] &= \mathbf{E}\left[\prod_j g(T_j)\xi(T_j)\right] \\ &= \exp\left\{-\int_0^\infty (1-g\xi)\nu\,\mathrm{d}t\right\}. \end{aligned}$$

Thus conditioning on $Y(0)=k$

$$\mathbf{E}\left(\prod_i \xi(s_i)|Y(0)=k\right) = g_0^k \exp\left\{-\int_0^\infty (1-g\xi)\nu\,\mathrm{d}t\right\}. \tag{A.4}$$

From (A.2) we have

$$1 - g\xi = \frac{1}{\lambda}\left[\frac{g'}{g} - \mu\left(1-\frac{1}{g}\right)\right].$$

so that

$$-\int_0^\infty (1-g\xi)\nu\,\mathrm{d}t = \frac{\nu}{\lambda}\log g_0 + \mu\delta\int_0^\infty \left(1-\frac{1}{g}\right)\mathrm{d}t.$$

Finally, from (A.4) and (A.3) we have

$$G_k(\xi) = \left[\frac{\mu}{\mu+\lambda(1-f(0))}\right]^{k+\delta} \exp\left\{\nu\int_0^\infty (f(t)-1)\,\mathrm{d}t\right\} \tag{A.5}$$

which completes the proof. □

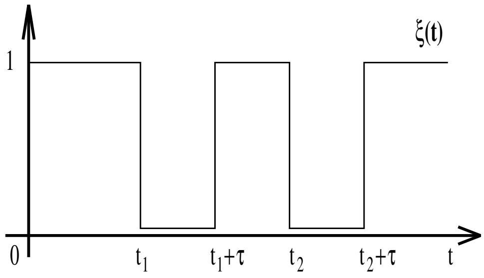

Figure B1. Well-like 0–1 step function.

Taking the expectation of the initial population from the equilibrium distribution NegBin(c, δ), the theorem leads to the same pgf as that of the death sequences.

However, the pgf (A.5) is of particular interest in itself because of the following corollary.

Corollary 3. *The increase time sequence of the* IBD (ν, λ, μ) *process with initial population size k is equivalent to the death time sequence of the* IBD (ν, λ, μ) *process with initial population sampling from* NegBin$(\lambda/(\lambda + \mu), k + \nu/\lambda)$.

We have demonstrated a number of equivalences between various point processes. But it is important to note that these equivalence claims are concerned only with the overall behaviour. The characteristics of the conditional death process would not necessarily have analogues for the increase process. For example, we have shown that when the initial population $Y(0)$ is sampled from Poisson$(Z(0)/\mu)$ or when $Y(0) = Z(0) = 0$, the death process is equivalent to the Cox process. But this is not true for the increase process. When conditioned on $Y(0) = 0$, the first increase is characterized by the immigration process Poisson(ν), while the first death is only partially dependent on it.

Appendix B. Calculating coincidence density from pgf

In [3] we gave explicit formulae for the coincidence densities of the Poisson process with Feller intensity (2.14). Now we will explore the feasibility of calculating coincidence densities from the pgf. We will need the following lemma.

Lemma 6. *When considering the Ricatti equation*

$$f' + af + bf^2 = c \tag{B.1}$$

where a, b, and c are continuous functions, the transformation

$$f = f_1 + 1/v \tag{B.2}$$

where f_1 is a particular solution, leads to a linear differential equation for v.

We have shown that the conditional pgf $f(t)$ satisfies the simple Ricatti equation

$$f' - (\lambda + \mu)f + \lambda f^2 = -\mu\xi(t) \tag{B.3}$$

where $\xi(t) \in \mathcal{V}$. In order to obtain coincidence densities, we will have to consider the *well-like* functions shown in figure B1 for which there are two types of interval to consider, those in which $\xi(t) = 0$ and those in which $\xi(t) = 1$.

Following the classical procedures of solving the Ricatti equations and together with lemma 6 we have the following solutions.

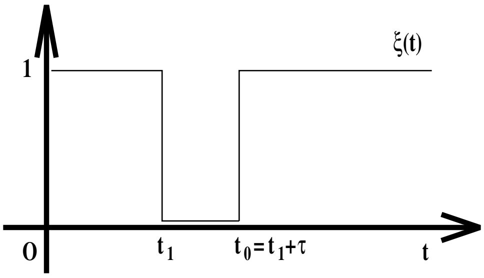

Figure B2. The well-like step function for the mean intensity.

Case (i). Suppose that $\xi(t) = 0$ for $t_1 < t < t_0$ then equation (B.3) has the solution

$$f(t) = \frac{f(t_0)(\lambda+\mu)}{f(t_0)\lambda + (\lambda+\mu-\lambda f(t_0))\,\mathrm{e}^{-(\lambda+\mu)(t-t_0))}}. \tag{B.4}$$

Case (ii). Similarly when $\xi(t) = 1$ for $t_1 < t < t_0$, we have

$$f(t) = \frac{\mu(1-f(t_0)) + (\lambda f(t_0)-\mu)\,\mathrm{e}^{-(\mu-\lambda)(t-t_0)}}{\lambda(1-f(t_0)) + (\lambda f(t_0)-\mu)\,\mathrm{e}^{-(\mu-\lambda)(t-t_0)}}. \tag{B.5}$$

To demonstrate how to obtain the coincidence densities from the pgf, we will derive the first-order density h_1 for the stationary case, which we have already shown to be $\mu\nu/(\mu-\lambda)$ in [3].

Let $\xi(t) \in \mathcal{G}$ be the well-like function as depicted in figure B2, i.e.

$$\xi(t) = 1 - \mathbf{I}(t \in (t_1, t_0)).$$

Let N_τ be the random number of points in the interval (t_1, t_0), where $\tau = t_0 - t_1$, then

$$G(\xi) = 1 - P(N_\tau > 0) = 1 - P(N_\tau = 1) + o(\tau)$$

since $P(N_\tau > 1) = o(\tau)$. The first-order density is therefore given by

$$h_1(t_1) = \lim_{\tau\to 0} \frac{1-G(\xi)}{\tau}. \tag{B.6}$$

Since

$$G(\xi) = \left[\frac{\mu-\lambda}{\mu-\lambda f(0)}\right]^\delta \exp\left\{-\nu\int_0^\infty (1-f(t))\,\mathrm{d}t\right\} \tag{B.7}$$

we now consider the approximation of $f(t)$ when τ is small. Writing $\epsilon(t)$ for $1-f(t)$, we immediately have $\epsilon(t) = 0$ for $t \geqslant t_0$. From (B.4) we have

$$\epsilon(t) = \mu(t_0 - t) + o(\tau) \qquad t_1 < \tau < t_0$$

so that $\epsilon(t_1) = \mu\tau + o(\tau)$.

From (B.5), substituting t_1 for t_0 and using the fact that $\epsilon(t_1)$ is small, we have

$$\epsilon(t) = \mathrm{e}^{-(\mu-\lambda)(t_1-t)}\epsilon(t_1) + o(\tau) \qquad \text{for} \quad 0 < t < t_1.$$

It follows that

$$\begin{aligned}\int_0^\infty (1-f(t))\,\mathrm{d}t &= \int_0^{t_0} \epsilon(x)\,\mathrm{d}x = \int_0^{t_1}\epsilon(x)\,\mathrm{d}x + \int_{t_1}^{t_0}\epsilon(x)\,\mathrm{d}x \\ &= \frac{\mu\tau}{\mu-\lambda}(1-\mathrm{e}^{-(\mu-\lambda)t_1}) + o(\tau).\end{aligned} \tag{B.8}$$

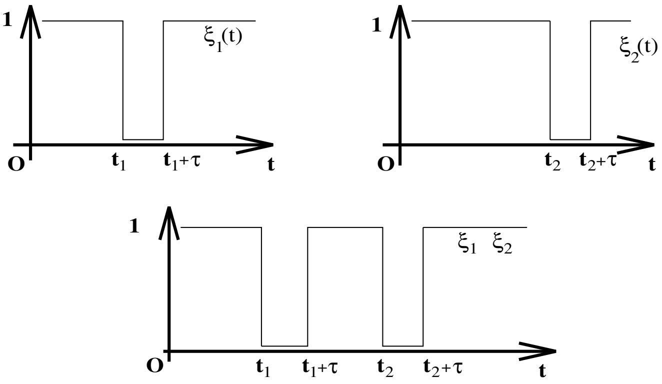

Figure B3. Well-like functions for second-order coincidence densities.

Simple algebra leads us to

$$\begin{aligned}
\left[\frac{\mu-\lambda}{\mu-\lambda f(0)}\right]^{\delta} &= \left[\frac{\mu-\lambda}{\mu-\lambda(1-\mathrm{e}^{-(\mu-\lambda)t_1}\mu\tau+o(\tau))}\right]^{\delta} \\
&= 1-\frac{\mu\nu}{\mu-\lambda}\mathrm{e}^{-(\mu-\lambda)t_1}+o(\tau) \\
\exp\left\{-\nu\int_0^\infty(1-f(t))\,\mathrm{d}t\right\} &= \exp\left\{-\frac{\mu\nu}{\mu-\lambda}(1-\mathrm{e}^{-(\mu-\lambda)t_1})\tau\right\}+o(\tau) \\
&= 1-\frac{\mu\nu}{\mu-\lambda}(1-\mathrm{e}^{-(\mu-\lambda)t_1})\tau+o(\tau)
\end{aligned} \tag{B.9}$$

Thus

$$\begin{aligned}
G(\xi) &= 1-\frac{\mu\nu}{\mu-\lambda}+o(\tau) \\
1-G(\xi) &= \frac{\mu\nu}{\mu-\lambda}+o(\tau)
\end{aligned} \tag{B.10}$$

$$h_1(t_1)=\frac{\mu\nu}{\mu-\lambda}. \tag{B.11}$$

Next, we consider the well-like functions ξ_1 and ξ_2 as in figure B3.

Using similar arguments as above, we can calculate the second-order coincidence densities by considering the corresponding pgf of ξ_1, ξ_2 and $\xi_1\xi_2$. And from the definition of coincidence densities

$$\begin{aligned}
h_2(t_1,t_2) &= \lim_{\tau\to 0}\frac{1}{\tau^2}\mathrm{P}[N(t_1,t_1+\tau)N(t_2,t_2+\tau)>0] \\
&= \lim_{\tau\to 0}\frac{1}{\tau^2}[1-G(\xi_1)-G(\xi_2)+G(\xi_1\xi_2)]
\end{aligned} \tag{B.12}$$

we can show that $G(\xi_1)$, $G(\xi_2)$ and $G(\xi_1\xi_2)$ can be calculated explicitly. Hence, in principle, we could obtain all coincidence densities by solving the Ricatti equations and calculating the pgfs of well-like functions.

References

[1] Bezrukov S M and Vodyanoy I 1997 Stochastic resonance in non-dynamical systems without response thresholds *Nature* **385** 319–21

[2] Chung K L 1974 *Elementary Probability Theory with Stochastic Processes* (New York: Springer)

[3] Clifford P and Wei G 1993 The equivalence of the Cox process with squared radial Ornstein–Uhlenbeck intensity and the death process in a simple population model *Ann. Appl. Probab.* **3** 863–73

[4] Clifford P, Wei G and Feng J F 2002 Reconstructing a population process from mortality data (in preparation)

[5] Cox D R and Lewis P A W 1966 *The Statistical Analysis of Series of Events* (London: Methuen)

[6] Cox D R and Miller H D 1965 *The Theory of Stochastic Process* (London: Spottiswoode, Ballantyne)

[7] Daley D J and Vere-Jones D 1988 *An Introduction to the Theory of Point Processes* (New York: Springer)

[8] Feng J F 2001 Is the integrate-and-fire model good enough? –a review *Neural Networks* **14** 955–75

[9] Feng J F 2002 *Computational Neuroscience: A Comprehensive Approached* ed J Feng (Boca Raton, FL: CRC Press)

[10] Karlin S and Taylor H M 1981 *A Second Course in Stochastic Processes* (New York: Academic)

[11] Karr A F 1991 *Point Processes and Their Statistical Inference* (New York: Dekker)

[12] Kendall D G 1948 On some modes of population growth leading to R A Fisher's logarithmic series distribution *Biometrika* **35** 6–15

[13] Kendall D G 1949 Stochastic processes and population growth *J. R. Stat. Soc.* B **11** 230–64

[14] Kendall M G and Stuart A 1977 *The Advanced Theory of Statistics* vol 1, 4th edn (London: Griffin)

[15] Lansky P, Fenton A A and Vaillant J 2001 The overdispersion in activity of place cells *Neurocomputing* **38** 1393–9

[16] Moran P A P 1968 *An Introduction to Probability Theory* (Oxford: Clarendon)

[17] Moyal J E 1962 The general theory of stochastic population processes *Acta Math.* **108** 1–31

[18] Snyder D L 1975 *Random Point Processes* (New York: Wiley)

[19] Srinivasan S K 1969 *Stochastic Theory and Cascade Processes* (New York: Elsevier)

[20] Srinivasan S K 1974 *Stochastic Point Processes and Their Applications* (London: Griffin)

[21] Srinivasan S K 1988 *Point Process Models of Cavity Radiation and Detection* (London: Griffin)

[22] Srinivasan S K, Sukavanam S and Sudarsham E 1973 Many photoncount distributions *J. Phys. A: Math. Gen.* **6** 1910–8

[23] Tuckwell H C 1988 *Stochastic Processes in the Neurosciences* (Philadelphia, PA: SIAM)