

## A DYNAMIC CONTAGION PROCESS

ANGELOS DASSIOS \* \*\* AND

HONGBIAO ZHAO,\* *London School of Economics*

### Abstract

We introduce a new point process, the dynamic contagion process, by generalising the Hawkes process and the Cox process with shot noise intensity. Our process includes both self-excited and externally excited jumps, which could be used to model the dynamic contagion impact from endogenous and exogenous factors of the underlying system. We have systematically analysed the theoretical distributional properties of this new process, based on the piecewise-deterministic Markov process theory developed in Davis (1984), and the extension of the martingale methodology used in Dassios and Jang (2003). The analytic expressions of the Laplace transform of the intensity process and the probability generating function of the point process have been derived. An explicit example of specified jumps with exponential distributions is also given. The object of this study is to produce a general mathematical framework for modelling the dependence structure of arriving events with dynamic contagion, which has the potential to be applicable to a variety of problems in economics, finance, and insurance. We provide an application of this process to credit risk, and a simulation algorithm for further industrial implementation and statistical analysis.

*Keywords:* Dynamic contagion process; Cox process with shot noise intensity; piecewise-deterministic Markov process; cluster point process; self-exciting point process; Hawkes process

2010 Mathematics Subject Classification: Primary 60J75

Secondary 60G55; 60G44; 91G40

### 1. Introduction

The behavior of default contagion through business links is more obvious during the recent financial crisis, especially after the collapse of Lehman Brothers in September 2008. More recently, the Greek debt crisis in 2010 has the contagion impact spreading to EU members, such as Portugal, Spain, and even the United Kingdom. A point process with its intensity dependent on the point process itself could provide a more effective model to capture this contagion phenomenon. However, only a few examples exist in the literature. These include the pioneering work of Jarrow and Yu (2001), and the more recent work of Errais *et al.* (2009). Jarrow and Yu (2001) pointed out that a model with the default intensity depending only linearly on a set of macroeconomic variables is not sufficient to explain the phenomena of clustering defaults around an economic recession; therefore, they introduced the concept of credit contagion, whereby, upon default of a given name, the contagion jump shocks will impact immediately to the counterpart's default intensity. Furthermore, Errais *et al.* (2009) found that, by using the self-excited Hawkes process, originally introduced in Hawkes (1971) (see also Hawkes and Oakes (1974) and Oakes (1975)), the clustering of defaults observed from real financial data

---

Received 5 August 2010; revision received 13 March 2011.

\* Postal address: Department of Statistics, London School of Economics, Houghton Street, London WC2A 2AE, UK.

\*\* Email address: a.dassios@lse.ac.uk

could be modelled more consistently. On the other hand, there are plenty of papers, including Duffie and Gârleanu (2001) and Longstaff and Rajan (2008), suggesting that the default intensity could be impacted exogenously by multiple common factors, such as idiosyncratic, sector specific, or market-wide events.

In this paper we combine both ideas and introduce a new point process, the dynamic contagion process, by generalising the Hawkes process (with exponential decay) and the Cox process with shot noise intensity (with exponential decay) used in Dassios and Jang (2003), to include both the self-excited and externally excited jumps. We use it to model the dynamic contagion impact from both endogenous (self-excited) and exogenous (externally excited) factors of the underlying system. This approach also extends the idea of default contagion in Jarrow and Yu (2001), to include a richer set of parameters, capable of capturing some key aspects of the behavior of arriving events, such as the frequency, magnitude of the impact, and the decay with time.

To define and characterise the dynamic contagion process mathematically, we give a cluster process representation, implement the piecewise-deterministic Markov process theory developed in Davis (1984) (see also Davis (1993)), and then extend the martingale methodology introduced in Dassios and Jang (2003) (see also Dassios and Jang (2005)) to obtain the distributional properties for this new process. This process is analysed by deriving the first and second moments, and then more importantly the Laplace transform of the intensity process and the probability generating function of the point process, respectively. Furthermore, an explicit example of jumps with exponential distributions and an application in credit risk are also given. The simulation algorithm is provided for further industrial implementation and statistical analysis.

The paper is organised as follows. Section 2 gives the mathematical definition of the process. In Section 3, the main section, we analyse and derive some key distributional properties. The joint Laplace transform, probability generating function of the intensity process, and the point process is derived in Section 3.1. The Laplace transform of the intensity process and the probability generating function of the point process are obtained in Section 3.2 and Section 3.3, respectively; the Hawkes process with exponential decay is included as an important special case and a brief summary of its distributional properties is also given. In Section 3.4 we obtain the first and second moments of the intensity process and the point process. We also provide an explicit example of jumps with exponential distributions in Section 4, and an application to credit risk and the algorithm for simulating the process in Section 5. In Section 6 we conclude this paper and suggest some further potential applications.

## 2. Definition

The dynamic contagion process includes both the self-excited jumps, which are distributed according to the branching structure of a Hawkes process with exponential fertility rate, and the externally excited jumps, which are distributed according to a particular shot noise Cox process.

Daley and Vere-Jones (2003, pp. 175–193) (see also Hawkes and Oakes (1974)) gave a cluster process representation for a general Hawkes process; now we extend it to represent the mathematical definition for our process in Definition 2.1 as a cluster point process, additionally characterised by the stochastic intensity representation and infinitesimal generator.

**Definition 2.1.** The *dynamic contagion process* is a cluster point process  $\mathbb{D}$  on  $\mathbb{R}_+$ : the number of points in the time interval  $(0, t]$  is defined by  $N_t = N_{\mathbb{D}(0,t]}$ . The cluster centers of  $\mathbb{D}$  are

the particular points called immigrants, the other points are called offspring. They have the following structure.

- (a) The immigrants are distributed according to a Cox process  $A$  with points  $\{D_m\}_{m=1,2,\dots} \in (0, \infty)$  and shot noise stochastic intensity process

$$a + (\lambda_0 - a)e^{-\delta t} + \sum_{i \geq 1} Y_i e^{-\delta(t - T_i^{(1)})} \mathbf{1}\{T_i^{(1)} \leq t\},$$

where

- $a \geq 0$  is the constant reversion level,
  - $\lambda_0 > 0$  is a constant and the initial value of the stochastic intensity process (defined later by (2.1)),
  - $\delta > 0$  is the constant rate of exponential decay,
  - $\{Y_i\}_{i=1,2,\dots}$  is a sequence of independent, identically distributed positive (externally excited) jumps with distribution function  $H(y)$ ,  $y > 0$ , at the corresponding random times  $\{T_i^{(1)}\}_{i=1,2,\dots}$  following a homogeneous Poisson process  $M_t$  with constant intensity  $\rho > 0$ ,
  - $\mathbf{1}\{\cdot\}$  is the indicator function.
- (b) Each immigrant  $D_m$  generates a cluster  $C_m = C_{D_m}$ , which is the random set formed by the points of generations  $0, 1, 2, \dots$  with the following branching structure. The immigrant  $D_m$  is said to be of generation 0. Given generations  $0, 1, \dots, j$  in  $C_m$ , each point  $T^{(2)} \in C_m$  of generation  $j$  generates a Cox process on  $(T^{(2)}, \infty)$  of offspring of generation  $j + 1$  with the stochastic intensity  $Ze^{-\delta(\cdot - T^{(2)})}$ , where  $Z$  is a positive (self-excited) jump at time  $T^{(2)}$  with distribution function  $G(z)$ ,  $z > 0$ , independent of the points of generation  $0, 1, \dots, j$ .
  - (c) Given the immigrants, the centered clusters

$$C_m - D_m = \{T^{(2)} - D_m : T^{(2)} \in C_m\}, \quad D_m \in A,$$

are independent, identically distributed, and independent of  $A$ .

- (d)  $\mathbb{D}$  consists of the union of all clusters, i.e.

$$\mathbb{D} = \bigcup_{m=1,2,\dots} C_{D_m}.$$

Therefore, the dynamic contagion process can also be defined as a point process  $N_t \equiv \{T_k^{(2)}\}_{k \geq 1}$  on  $\mathbb{R}_+$ , with the nonnegative  $\mathcal{F}_t$ -stochastic intensity process  $\lambda_t$  following the piecewise-deterministic dynamics with positive jumps, i.e.

$$\lambda_t = a + (\lambda_0 - a)e^{-\delta t} + \sum_{i \geq 1} Y_i e^{-\delta(t - T_i^{(1)})} \mathbf{1}\{T_i^{(1)} \leq t\} + \sum_{k \geq 1} Z_k e^{-\delta(t - T_k^{(2)})} \mathbf{1}\{T_k^{(2)} \leq t\}, \quad (2.1)$$

where

- $\{\mathcal{F}_t\}_{t \geq 0}$  is a history of the process  $N_t$ , with respect to which  $\{\lambda_t\}_{t \geq 0}$  is adapted,

- $\{Z_k\}_{k=1,2,\dots}$  is a sequence of independent, identically distributed positive (self-excited) jumps with distribution function  $G(y)$ ,  $y > 0$ , at the corresponding random times  $\{T_k^{(2)}\}_{k=1,2,\dots}$ ,
- the sequences  $\{Y_i\}_{i=1,2,\dots}$ ,  $\{T_i^{(1)}\}_{i=1,2,\dots}$ , and  $\{Z_k\}_{k=1,2,\dots}$  are assumed to be independent of each other.

From the definition above and because of the exponential decay, we can see that  $\lambda_t$  is a Markov process. In particular, it decreases with rate  $\delta(\lambda_t - a)$ , and incurs additive upward (externally excited) jumps that have distribution function  $H$  with rate  $\rho$ , and additive upward (self-excited) jumps that have distribution function  $G$  with rate  $\lambda_t$ . Moreover, when jumps of the latter type occur,  $N_t$  increases by 1. Hence,  $(N_t, \lambda_t)$  is also a Markov process.

With the aid of piecewise-deterministic Markov process theory and using the results in Davis (1984), the infinitesimal generator of the dynamic contagion process  $(\lambda_t, N_t, t)$  acting on a function  $f(\lambda, n, t)$  within its domain  $\Omega(\mathcal{A})$  is given by

$$\begin{aligned} \mathcal{A}f(\lambda, n, t) = & \frac{\partial f}{\partial t} + \delta(a - \lambda)\frac{\partial f}{\partial \lambda} + \rho \left( \int_0^\infty f(\lambda + y, n, t) dH(y) - f(\lambda, n, t) \right) \\ & + \lambda \left( \int_0^\infty f(\lambda + z, n + 1, t) dG(z) - f(\lambda, n, t) \right), \end{aligned} \tag{2.2}$$

where  $\Omega(\mathcal{A})$  is the domain for the generator  $\mathcal{A}$  such that  $f(\lambda, n, t)$  is differentiable with respect to  $\lambda$  and  $t$  for all  $\lambda, n$ , and  $t$ , and

$$\begin{aligned} \left| \int_0^\infty f(\lambda + y, n, t) dH(y) - f(\lambda, n, t) \right| < \infty, \\ \left| \int_0^\infty f(\lambda + z, n + 1, t) dG(z) - f(\lambda, n, t) \right| < \infty. \end{aligned}$$

**Remark 2.1.** We could alternatively define the dynamic contagion process as a special case (without the diffusion terms) of the general affine point processes in Duffie *et al.* (2003), with the infinitesimal generator specified by (2.2).

**Remark 2.2.** Note that the dynamic contagion process is a point process  $N_t$  such that

$$P\{N_{t+\Delta t} - N_t = 1 \mid N_t\} = \lambda_t \Delta t + o(\Delta t), \quad P\{N_{t+\Delta t} - N_t > 1 \mid N_t\} = o(\Delta t),$$

where  $\Delta t$  is a sufficiently small time interval and  $\lambda_t$  is given by (2.1).

**Remark 2.3.** Note that the intensity process  $\lambda_t$  is always above the level  $a$ , i.e.  $\lambda_t \in E = [a, \infty)$  for any time  $t$ .

**Remark 2.4.** An economic interpretation from the perspective of the cluster process representation for the dynamic contagion process as follows. For a certain company, there are two classes of economic shocks: the primary shocks directly to this company and the common market-wide shocks. The arrivals of these primary shocks to this company are modelled by generation 0 of the dynamic contagion process, i.e. the point process  $A$  (as described in Definition 2.1(a)) with the intensity process modelled based on the external economic evolution including a stream of market-wide shocks: a shock at time  $T_i^{(1)}$  has the magnitude of impact  $Y_i$  with distribution  $H$  and decays exponentially with rate  $\delta$ . In the aftermath of each primary shock to this company, it could further trigger a series of subsidiary internal turbulences in this company following the

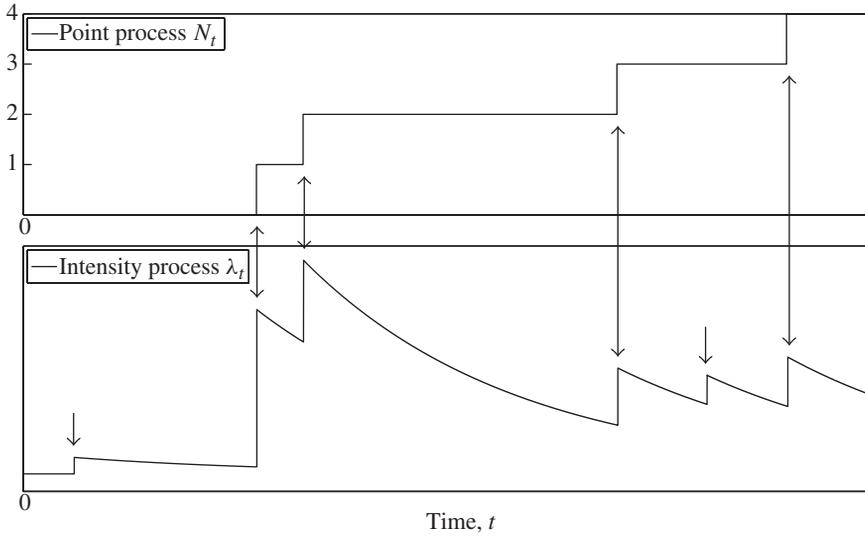


FIGURE 1: Externally excited and self-excited jumps in the intensity process  $\lambda_t$  of the dynamic contagion process  $N_t$ .

branching structure (as described in Definition 2.1(b)). Similarly, a turbulence at time  $T_k^{(2)}$  has the magnitude of impact  $Z_k$  with distribution  $G$  and decays exponentially with rate  $\delta$ .

To give an intuitive picture of this process from the perspective of the stochastic intensity representation, we present Figure 1 to illustrate how the externally excited jumps  $\{Y_i\}_{i=1,2,\dots}$  (marked by downwards arrows) and self-excited jumps  $\{Z_k\}_{k=1,2,\dots}$  (marked by up-down arrows) in the intensity process  $\lambda_t$  interact with its dynamic contagion point process  $N_t$ .

Now, in this more general framework of the dynamic contagion process, the classic Cox process with shot noise intensity (with exponential decay), used in Dassios and Jang (2003) for pricing catastrophe reinsurance and derivatives, can be recovered, by setting the reversion level  $a = 0$  and eliminating the self-excited jumps  $\{Z_k\}_{k=1,2,\dots}$ ; the Hawkes process (with exponential decay), used in Errais *et al.* (2009) for modelling the portfolio credit risk, can be recovered, by setting the intensity  $\rho = 0$  of the externally excited jumps  $\{Y_i\}_{i=1,2,\dots}$ .

### 3. Dynamic contagion process

#### 3.1. Joint Laplace transform, probability generating function of $(\lambda_T, N_T)$

We derive the joint Laplace transform, probability generating function of  $(\lambda_T, N_T)$  for a fixed time  $T$  in Theorem 3.1 below, which leads to the key results of this paper, and the Laplace transform of  $\lambda_T$  and probability generating function of  $N_T$  in Section 3.2 and Section 3.3, respectively.

**Theorem 3.1.** *For the constants  $0 \leq \theta \leq 1$ ,  $v \geq 0$ , and time  $0 \leq t \leq T$ , the conditional joint Laplace transform, probability generating function for the process  $\lambda_t$  (defined in Definition 2.1) and the point process  $N_t$  is given by*

$$E[\theta^{(N_T - N_t)} e^{-v\lambda_T} \mid \mathcal{F}_t] = e^{-(c(T) - c(t))} e^{-B(t)\lambda_t}, \tag{3.1}$$

where  $B(t)$  is determined by the nonlinear ordinary differential equation (ODE)

$$-B'(t) + \delta B(t) + \theta \hat{g}(B(t)) - 1 = 0, \tag{3.2}$$

$$\hat{g}(u) := \int_0^\infty e^{-uz} dG(z),$$

with boundary condition  $B(T) = v$ , and  $c(t)$  is determined by

$$c(t) = a\delta \int_0^t B(s) ds + \rho \int_0^t [1 - \hat{h}(B(s))] ds, \tag{3.3}$$

$$\hat{h}(u) := \int_0^\infty e^{-uy} dH(y).$$

*Proof.* Consider a function  $f(\lambda, n, t)$  with the exponential affine form

$$f(\lambda, n, t) = e^{c(t)} A^n(t) e^{-B(t)\lambda},$$

and substitute this into  $\mathcal{A}f = 0$  in (2.2); we then have

$$\frac{A'(t)}{A(t)} n + (-B'(t) + \delta B(t) + A(t)\hat{g}(B(t)) - 1)\lambda + (c'(t) + \rho\hat{h}(B(t)) - \rho - a\delta B(t)) = 0.$$

Since this equation holds for any  $n$  and  $\lambda$ , it is equivalent to solving three separate equations:

$$\frac{A'(t)}{A(t)} = 0, \tag{3.4a}$$

$$-B'(t) + \delta B(t) + A(t)\hat{g}(B(t)) - 1 = 0, \tag{3.4b}$$

$$c'(t) + \rho\hat{h}(B(t)) - \rho - a\delta B(t) = 0. \tag{3.4c}$$

We have  $A(t) = \theta$  immediately from (3.4a), which when substituted into (3.4b), adding the boundary condition  $B(T) = v$ , gives the ODE (3.2); then, by (3.4c), with the boundary condition  $c(0) = 0$ , (3.3) follows. Since  $e^{c(t)}\theta^{N_t}e^{-B(t)\lambda_t}$  is an  $\mathcal{F}$ -martingale by the property of the infinitesimal generator, we have

$$E[e^{c(T)}\theta^{N_T}e^{-B(T)\lambda_T} \mid \mathcal{F}_t] = e^{c(t)}\theta^{N_t}e^{-B(t)\lambda_t}.$$

Then, by the boundary condition  $B(T) = v$ , (3.1) follows.

### 3.2. Laplace transform of $\lambda_T$

**Theorem 3.2.** *The conditional Laplace transform  $\lambda_T$  given  $\lambda_0$  at time  $t = 0$ , under the condition  $\delta > \mu_{1G}$ , is given by*

$$E[e^{-v\lambda_T} \mid \lambda_0] = \exp\left(-\int_{\mathcal{G}_{v,1}^{-1}(T)}^v \frac{a\delta u + \rho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1} du\right) \exp(-\mathcal{G}_{v,1}^{-1}(T)\lambda_0),$$

where

$$\mu_{1G} := \int_0^\infty z dG(z), \quad \mathcal{G}_{v,1}(L) := \int_L^v \frac{du}{\delta u + \hat{g}(u) - 1}.$$

(It will be clear in the proof that  $\mathcal{G}_{v,1}(L)$  is a one by one function of  $L$  and, hence, its inverse function  $\mathcal{G}_{v,1}^{-1}(T)$  exists.)

*Proof.* By setting  $t = 0$  and  $\theta = 1$  in Theorem 3.1, we have

$$E[e^{-v\lambda_T} \mid \mathcal{F}_0] = e^{-c(T)}e^{-B(0)\lambda_0}, \tag{3.5}$$

where  $B(0)$  is uniquely determined by the nonlinear ODE

$$-B'(t) + \delta B(t) + \hat{g}(B(t)) - 1 = 0,$$

with boundary condition  $B(T) = v$ . It can be solved, under the condition  $\delta > \mu_{1G}$ , by the following steps.

1. Set  $B(t) = L(T - t)$  and  $\tau = T - t$ . Then we have the initial value problem

$$\frac{dL(\tau)}{d\tau} = 1 - \delta L(\tau) - \hat{g}(L(\tau)) =: f_1(L), \tag{3.6}$$

with initial condition  $L(0) = v$ ; we define the right-hand side to be the function  $f_1(L)$ .

2. Under the condition  $\delta > \mu_{1G}$ , we have

$$\frac{\partial f_1(L)}{\partial L} = \int_0^\infty ye^{-Lz} dG(z) - \delta \leq \int_0^\infty z dG(z) - \delta = \mu_{1G} - \delta < 0 \quad \text{for } L \geq 0;$$

then,  $f_1(L) < 0$  for  $L > 0$ .

3. Rewrite (3.6) as

$$\frac{dL}{\delta L + \hat{g}(L) - 1} = -d\tau.$$

By integrating both sides from time 0 to  $\tau$  with initial condition  $L(0) = v > 0$ , we obtain

$$\int_L^v \frac{du}{\delta u + \hat{g}(u) - 1} = \tau,$$

where  $L \geq 0$ . Let

$$\mathcal{G}_{v,1}(L) := \int_L^v \frac{du}{\delta u + \hat{g}(u) - 1}.$$

Then

$$\mathcal{G}_{v,1}(L) = \tau.$$

Obviously,  $L \rightarrow v$  when  $\tau \rightarrow 0$ . By the convergence test,

$$\lim_{u \rightarrow 0} \frac{1/u}{1/(\delta u + \hat{g}(u) - 1)} = \delta + \lim_{u \rightarrow 0} \frac{\hat{g}(u) - 1}{u} = \delta - \mu_{1G} > 0,$$

and we know that  $\int_0^v (1/u) du = \infty$ . Then,

$$\int_0^v \frac{du}{\delta u + \hat{g}(u) - 1} = \infty;$$

hence,  $L \rightarrow 0$  when  $\tau \rightarrow \infty$ . The integrand is positive in the domain  $u \in (0, v]$  and also for  $L \leq v$ . The function  $\mathcal{G}_{v,1}(L)$  is strictly decreasing; therefore,  $\mathcal{G}_{v,1}(L): (0, v] \rightarrow [0, \infty)$  is a well-defined (monotone) function, and its inverse function  $\mathcal{G}_{v,1}^{-1}(\tau): [0, \infty) \rightarrow (0, v]$  exists.

4. The unique solution is found by

$$L(\tau) = \mathcal{G}_{v,1}^{-1}(\tau) \quad \text{or} \quad B(t) = \mathcal{G}_{v,1}^{-1}(T - t).$$

5.  $B(0)$  is obtained:

$$B(0) = L(T) = \mathcal{G}_{v,1}^{-1}(T).$$

Then,  $c(T)$  is determined by

$$c(T) = a\delta \int_0^T \mathcal{G}_{v,1}^{-1}(\tau) \, d\tau + \rho \int_0^T [1 - \hat{h}(\mathcal{G}_{v,1}^{-1}(\tau))] \, d\tau.$$

By the change of variable  $\mathcal{G}_{v,1}^{-1}(\tau) = u$  we have  $\tau = \mathcal{G}_{v,1}(u)$  and

$$\int_0^T [1 - \hat{h}(\mathcal{G}_{v,1}^{-1}(\tau))] \, d\tau = \int_{\mathcal{G}_{v,1}^{-1}(0)}^{\mathcal{G}_{v,1}^{-1}(T)} [1 - \hat{h}(u)] \frac{\partial \tau}{\partial u} \, du = \int_{\mathcal{G}_{v,1}^{-1}(T)}^v \frac{1 - \hat{h}(u)}{\delta u + \hat{g}(u) - 1} \, du.$$

Similarly,

$$\int_0^T \mathcal{G}_{v,1}^{-1}(\tau) \, d\tau = \int_{\mathcal{G}_{v,1}^{-1}(T)}^v \frac{u}{\delta u + \hat{g}(u) - 1} \, du.$$

Finally, substitute  $B(0)$  and  $c(T)$  into (3.5), and Theorem 3.2 follows.

**Theorem 3.3.** *If  $\delta > \mu_{1G}$  then the Laplace transform of the asymptotic distribution of  $\lambda_T$  is given by*

$$\lim_{T \rightarrow \infty} E[e^{-v\lambda_T} \mid \lambda_0] = \exp\left(- \int_0^v \frac{a\delta u + \rho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1} \, du\right), \tag{3.7}$$

and this is also the Laplace transform of the stationary distribution of the process  $\{\lambda_t\}_{t \geq 0}$ .

*Proof.* Let  $T \rightarrow \infty$  in Theorem 3.2. Then  $\mathcal{G}_{v,1}^{-1}(T) \rightarrow 0$  and the Laplace transform of the asymptotic distribution follows immediately as given by (3.7).

To further prove the stationarity, by Proposition 9.2 of Ethier and Kurtz (1986) (see also Costa (1990)), we need to prove that, for any function  $f$  within its domain  $\Omega(\mathcal{A})$ , we have

$$\int_E \mathcal{A}f(\lambda) \Pi(\lambda) \, d\lambda = 0, \tag{3.8}$$

where  $E = [a, \infty)$  is the domain for  $\lambda$ ,  $\mathcal{A}f(\lambda)$  is the infinitesimal generator of the dynamic contagion process acting on  $f(\lambda)$ , i.e.

$$\begin{aligned} \mathcal{A}f(\lambda) = & -\delta(\lambda - a) \frac{df(\lambda)}{d\lambda} + \rho \left( \int_0^\infty f(\lambda + y) \, dH(y) - f(\lambda) \right) \\ & + \lambda \left( \int_0^\infty f(\lambda + z) \, dG(z) - f(\lambda) \right), \end{aligned}$$

and  $\Pi(\lambda)$  is the density function of  $\lambda$  with Laplace transform given by (3.7).

We will now try to solve (3.8). For the first term of (3.8), we have

$$\begin{aligned} \int_E \left( -\delta(\lambda - a) \frac{df(\lambda)}{d\lambda} \right) \Pi(\lambda) d\lambda &= -\delta \int_a^\infty (\lambda - a) f'(\lambda) \Pi(\lambda) d\lambda \\ &= -\delta \int_{\lambda=a}^\infty f'(\lambda) \int_{u=a}^\lambda [(u - a)\Pi(u)]' du d\lambda \\ &= -\delta \int_{u=a}^\infty \int_{\lambda=u}^\infty f'(\lambda) [(u - a)\Pi(u)]' d\lambda du \\ &= \delta \int_a^\infty f(u) [(u - a)\Pi(u)]' du \end{aligned}$$

or

$$\int_E \left( -\delta(\lambda - a) \frac{df(\lambda)}{d\lambda} \right) \Pi(\lambda) d\lambda = \delta \int_a^\infty f(\lambda) [(\lambda - a)\Pi(\lambda)]' d\lambda,$$

since, for a density function  $\Pi$ , we obviously have

$$\lim_{y \rightarrow a} \Pi(y)(y - a) = 0.$$

For the second term of (3.8), by the change of variable  $\lambda + y = s$  ( $y \leq s$ ) in the double integral,

$$\begin{aligned} \int_E \left( \rho \int_0^\infty f(\lambda + y) dH(y) \right) \Pi(\lambda) d\lambda &= \rho \int_{\lambda=a}^\infty \Pi(\lambda) \int_{y=0}^\infty f(\lambda + y) dH(y) d\lambda \\ &= \rho \int_{s=a}^\infty f(z) \int_{y=0}^s \Pi(s - y) dH(y) ds \end{aligned}$$

or

$$\int_E \left( \rho \int_0^\infty f(\lambda + y) dH(y) \right) \Pi(\lambda) d\lambda = \rho \int_{\lambda=a}^\infty f(\lambda) \int_{y=0}^\lambda \Pi(\lambda - y) dH(y) d\lambda.$$

For the third term of (3.8), by the change of variable  $\lambda + z = s$  ( $z \leq s$ ) in the double integral,

$$\begin{aligned} \int_E \left[ \lambda \left( \int_0^\infty f(\lambda + z) dG(z) \right) \right] \Pi(\lambda) d\lambda &= \int_{\lambda=a}^\infty \lambda \Pi(\lambda) \int_{z=0}^\infty f(\lambda + z) dG(z) d\lambda \\ &= \int_{s=a}^\infty f(s) \int_{z=0}^s (s - z) \Pi(s - z) dG(z) ds \end{aligned}$$

or

$$\int_E \left[ \lambda \left( \int_0^\infty f(\lambda + z) dG(z) \right) \right] \Pi(\lambda) d\lambda = \int_{\lambda=a}^\infty f(\lambda) \int_{z=0}^\lambda (\lambda - z) \Pi(\lambda - z) dG(z) d\lambda.$$

Therefore,

$$\begin{aligned} \int_E \mathcal{A} f(\lambda) \Pi(\lambda) d\lambda &= \int_a^\infty f(\lambda) \left[ \delta \frac{d}{d\lambda} ((\lambda - a)\Pi(\lambda)) + \rho \left( \int_0^\lambda \Pi(\lambda - y) dH(y) - \Pi(\lambda) \right) \right. \\ &\quad \left. + \left( \int_0^\lambda (\lambda - z) \Pi(\lambda - z) dG(z) - \lambda \Pi(\lambda) \right) \right] d\lambda. \end{aligned}$$

Set

$$\int_E \mathcal{A}f(\lambda)\Pi(\lambda) \, d\lambda = 0$$

for any function  $f(\lambda) \in \Omega(\mathcal{A})$ . Then

$$\begin{aligned} & \delta \frac{d}{d\lambda}((\lambda - a)\Pi(\lambda)) + \rho \left( \int_0^\lambda \Pi(\lambda - y) \, dH(y) - \Pi(\lambda) \right) \\ & + \left( \int_0^\lambda (\lambda - z)\Pi(\lambda - z) \, dG(z) - \lambda\Pi(\lambda) \right) = 0. \end{aligned}$$

By the Laplace transform,

$$\hat{\Pi}(v) =: \mathcal{L}\{\Pi(\lambda)\} = \int_E \Pi(\lambda)e^{-v\lambda} \, d\lambda.$$

Thus, we have

$$\begin{aligned} \mathcal{L} \left\{ \frac{d}{d\lambda}((\lambda - a)\Pi(\lambda)) \right\} &= v\mathcal{L}\{(\lambda - a)\Pi(\lambda)\} = v \left( -\frac{d\hat{\Pi}(v)}{dv} - a\hat{\Pi}(v) \right), \\ \mathcal{L} \left\{ \int_0^\lambda \Pi(\lambda - y) \, dH(y) \right\} &= \mathcal{L} \left\{ \int_0^\lambda \Pi(\lambda - y)h(y) \, dy \right\} = \hat{\Pi}(v)\hat{h}(v), \\ \mathcal{L} \left\{ \int_0^\lambda (\lambda - z)\Pi(\lambda - z) \, dG(z) \right\} &= \mathcal{L} \left\{ \int_0^\lambda (\lambda - z)\Pi(\lambda - z)g(z) \, dz \right\} \\ &= \mathcal{L} \left\{ \lambda\Pi(\lambda) \right\} \hat{g}(v) \\ &= -\frac{d\hat{\Pi}(v)}{dv} \hat{g}(v). \end{aligned}$$

Then

$$\delta v \left( -\frac{d\hat{\Pi}(v)}{dv} - a\hat{\Pi}(v) \right) + \rho(\hat{h}(v) - 1)\hat{\Pi}(v) + (1 - \hat{g}(v))\frac{d\hat{\Pi}(v)}{dv} = 0,$$

or

$$(1 - \delta v - \hat{g}(v))\frac{d\hat{\Pi}(v)}{dv} + (-a\delta v + \rho(\hat{h}(v) - 1))\hat{\Pi}(v) = 0,$$

which is an ODE with the solution

$$\hat{\Pi}(v) = \hat{\Pi}(0) \exp \left( -\int_0^v \frac{a\delta u + \rho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1} \, du \right).$$

Note that, given the initial condition

$$\hat{\Pi}(0) = \int_E \Pi(\lambda) \, d\lambda = 1,$$

we have the unique solution

$$\hat{\Pi}(v) = \exp \left( -\int_0^v \frac{a\delta u + \rho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1} \, du \right),$$

which is exactly given by (3.7).

Since  $\Pi$  is the unique solution to (3.8), we have the stationarity for the intensity process  $\{\lambda_t\}_{t \geq 0}$ .

Alternative approaches for proving the stationarity for the special case of the Hawkes process and other related processes can be found in Hawkes and Oakes (1974), Brémaud and Massoulié (1996), and Massoulié (1998).

The self-excited Hawkes process was introduced theoretically in Hawkes (1971) and applied to risk theory in Chavez-Demoulin *et al.* (2005), and only recently applied to credit risk for modelling the default contagion in Errais *et al.* (2009). It can be considered as an important special case under this more general framework of the dynamic contagion process; all of the counterpart results can be obtained by eliminating the impact from the externally excited jumps, i.e. setting its intensity  $\rho = 0$  in the corresponding results. Here we give the Laplace transform of the stationary distribution of the intensity process  $\lambda_t$  for the Hawkes process with exponential decay in Corollary 3.1 below. The probability generating function of the Hawkes point process  $N_t$  will be given in Corollary 3.2 of Section 3.3.

**Corollary 3.1.** *If  $\delta > \mu_{1G}$  then the Laplace transform of the asymptotic distribution of  $\lambda_T$  for the Hawkes process with exponential decay is given by*

$$\lim_{T \rightarrow \infty} E[e^{-v\lambda_T} \mid \lambda_0] = \exp\left(-a\delta \int_0^v \frac{u}{\delta u + \hat{g}(u) - 1} du\right),$$

and this is also the Laplace transform of the stationary distribution of the process  $\{\lambda_t\}_{t \geq 0}$ .

*Proof.* By setting the intensity of the externally excited jumps  $\rho = 0$  in Theorem 3.3, the result follows immediately.

The limit of the log-Laplace transform for Hawkes processes with a general fertility rate can be found in Bordenave and Torrisi (2007) and Stabile and Torrisi (2010).

**3.3. Probability generating function of  $N_T$**

**Theorem 3.4.** *The conditional probability generating function of  $N_T$  given  $\lambda_0$  and  $N_0 = 0$  at time  $t = 0$ , under the condition  $\delta > \mu_{1G}$ , is given by*

$$E[\theta^{N_T} \mid \lambda_0] = \exp\left(-\int_0^{\mathcal{G}_{0,\theta}^{-1}(T)} \frac{a\delta u + \rho[1 - \hat{h}(u)]}{1 - \delta u - \theta \hat{g}(u)} du\right) \exp(-\mathcal{G}_{0,\theta}^{-1}(T)\lambda_0),$$

where

$$\mathcal{G}_{0,\theta}(L) := \int_0^L \frac{du}{1 - \delta u - \theta \hat{g}(u)}, \quad 0 \leq \theta < 1. \tag{3.9}$$

*Proof.* By setting  $t = 0$  and  $v = 0$ , and assuming that  $N_0 = 0$  in Theorem 3.1, we have

$$E[\theta^{N_T} \mid \mathcal{F}_0] = e^{-c(T)} e^{-B(0)\lambda_0},$$

where  $B(0)$  is uniquely determined by the nonlinear ODE

$$-B'(t) + \delta B(t) + \theta \hat{g}(B(t)) - 1 = 0,$$

with boundary condition  $B(T) = 0$ . It can be solved, under the condition  $\delta > \mu_{1G}$ , by the following steps.

1. Set  $B(t) = L(T - t)$  and  $\tau = T - t$ . Then

$$\frac{dL(\tau)}{d\tau} = 1 - \delta L(\tau) - \theta \hat{g}(L(\tau)) =: f_2(L), \quad 0 \leq \theta < 1, \tag{3.10}$$

with initial condition  $L(0) = 0$ ; we define the right-hand side to be the function  $f_2(L)$ .

- There is only one positive singular point, denoted by  $v^* > 0$ , obtained by solving the equation  $f_2(L) = 0$ . This is because, for the case in which  $0 < \theta < 1$ , the equation  $f_2(L) = 0$  is equivalent to

$$\hat{g}(u) = \frac{1}{\theta}(1 - \delta u), \quad 0 < \theta < 1.$$

Note that  $\hat{g}(\cdot)$  is a convex function. Then it is clear that there is only one positive solution to this equation. For the case in which  $\theta = 0$ , there is only one singular point,  $v^* = 1/\delta > 0$ . For both cases,

$$v^* = \frac{1}{\delta}(1 - \theta \hat{g}(v^*)) \geq \frac{1 - \theta}{\delta} > 0. \tag{3.11}$$

Then we have  $f_2(L) > 0$  for  $0 \leq L < v^*$  and  $f_2(L) < 0$  for  $L > v^*$ .

- Rewrite (3.10) as

$$\frac{dL}{1 - \delta L - \theta \hat{g}(L)} = d\tau.$$

Integrating yields

$$\int_0^L \frac{du}{1 - \delta u - \theta \hat{g}(u)} = \tau,$$

where  $0 \leq L < v^*$ . Let

$$\mathcal{G}_{0,\theta}(L) := \int_0^L \frac{du}{1 - \delta u - \theta \hat{g}(u)}.$$

Then

$$\mathcal{G}_{0,\theta}(L) = \tau$$

as  $L \rightarrow 0$  when  $\tau \rightarrow 0$ , and  $L \rightarrow v^*$  when  $\tau \rightarrow \infty$ . The integrand is positive in the domain  $u \in [0, v^*)$  and  $L \geq 0$ , and  $\mathcal{G}_{0,\theta}(L)$  is a strictly increasing function; therefore,  $\mathcal{G}_{0,\theta}(L) : [0, v^*) \rightarrow [0, \infty)$  is a well-defined function, and its inverse function  $\mathcal{G}_{0,\theta}^{-1}(\tau) : [0, \infty) \rightarrow [0, v^*)$  exists.

- The unique solution is found by

$$L(\tau) = \mathcal{G}_{0,\theta}^{-1}(\tau) \quad \text{or} \quad B(t) = \mathcal{G}_{0,\theta}^{-1}(T - t).$$

- $B(0)$  is obtained:

$$B(0) = L(T) = \mathcal{G}_{0,\theta}^{-1}(T).$$

Then,  $c(T)$  is determined by

$$c(T) = a\delta \int_0^T \mathcal{G}_{0,\theta}^{-1}(\tau) d\tau + \rho \int_0^T [1 - \hat{h}(\mathcal{G}_{0,\theta}^{-1}(\tau))] d\tau,$$

where, by a change of variable,

$$\begin{aligned} \int_0^T \mathcal{G}_{0,\theta}^{-1}(\tau) d\tau &= \int_0^{\mathcal{G}_{0,\theta}^{-1}(T)} \frac{u}{1 - \delta u - \theta \hat{g}(u)} du, \\ \int_0^T [1 - \hat{h}(\mathcal{G}_{0,\theta}^{-1}(\tau))] d\tau &= \int_0^{\mathcal{G}_{0,\theta}^{-1}(T)} \frac{1 - \hat{h}(u)}{1 - \delta u - \theta \hat{g}(u)} du. \end{aligned}$$

Finally, substitute  $B(0)$  and  $c(T)$  into (3.5), and the result follows.

**Corollary 3.2.** *The conditional probability generating function of  $N_T$  for the Hawkes process with exponential decay, under the condition  $\delta > \mu_{1G}$  given  $\lambda_0$  and  $N_0 = 0$ , is given by*

$$E[\theta^{N_T} \mid \lambda_0] = \exp\left(-a\delta \int_0^{\mathcal{G}_{0,\theta}^{-1}(T)} \frac{u}{1 - \delta u - \theta \hat{g}(u)} du\right) \exp[-\mathcal{G}_{0,\theta}^{-1}(T)\lambda_0],$$

where

$$\mathcal{G}_{0,\theta}(L) := \int_0^L \frac{du}{1 - \delta u - \theta \hat{g}(u)}, \quad 0 \leq \theta < 1.$$

*Proof.* By setting the intensity of the externally excited jumps  $\rho = 0$  in Theorem 3.4, the result follows immediately.

The probability  $P\{N_T = 0 \mid \lambda_0\}$  can be derived by simply letting  $\theta = 0$  in the probability generating function of  $N_T$  in Theorem 3.4.

**Corollary 3.3.** *The conditional probability of no jump given  $\lambda_0$  and  $N_0 = 0$ , under the condition  $\delta > \mu_{1G}$ , is given by*

$$P\{N_T = 0 \mid \lambda_0\} = \exp\left[-\int_0^{u_T} \frac{a\delta u + \rho[1 - \hat{h}(u)]}{1 - \delta u} du\right] e^{-u_T \lambda_0}, \tag{3.12}$$

where

$$u_T := \frac{1}{\delta}(1 - e^{-\delta T}).$$

*Proof.* Since

$$P\{N_T = 0 \mid \lambda_0\} = E[\theta^{N_T} \mid \lambda_0] \Big|_{\theta=0}$$

and

$$\mathcal{G}_{0,0}(L) := \mathcal{G}_{0,\theta}(L) \Big|_{\theta=0} = \int_0^L \frac{1}{1 - \delta u} du = -\frac{1}{\delta} \ln(1 - \delta L),$$

then the inverse function

$$u_T = \mathcal{G}_{0,0}^{-1}(T) = \frac{1}{\delta}(1 - e^{-\delta T}).$$

By letting  $\theta = 0$  in Theorem 3.4, (3.12) follows.

**Remark 3.1.** Note that, since there is no jump in the point process  $N_t$  from time  $t = 0$  to  $t = T$ , the conditional probability  $P\{N_T = 0 \mid \lambda_0\}$  is not dependent on the distribution of the self-excited jumps, and the result is similar to the nonself-excited case in Dassios and Jang (2003).

Theoretically, the probability  $P\{N_T = n \mid \lambda_0\}$  for any natural number  $n \in \mathbb{N}$  can be evaluated as

$$P\{N_T = n \mid \lambda_0\} = \frac{\partial^n}{\partial \theta^n} E[\theta^{N_T} \mid \lambda_0] \Big|_{\theta=0}.$$

In Corollary 3.4 below we derive the result for  $P\{N_T = 1 \mid \lambda_0\}$ .

**Corollary 3.4.** *The conditional probability of exactly one jump given  $\lambda_0$  and  $N_0 = 0$ , under the condition  $\delta > \mu_{1G}$ , is given by*

$$\begin{aligned} \mathbb{P}\{N_T = 1 \mid \lambda_0\} &= \mathbb{P}\{N_T = 0 \mid \lambda_0\} \\ &\times \left\{ [a(1 - e^{-\delta T}) + \rho[1 - \hat{h}(u_T)] + \lambda_0 e^{-\delta T}] \int_0^{u_T} \frac{\hat{g}(u)}{(1 - \delta u)^2} du \right. \\ &\quad \left. - \int_0^{u_T} \frac{\hat{g}(u)}{(1 - \delta u)^2} (a\delta u + \rho[1 - \hat{h}(u)]) du \right\}, \end{aligned}$$

where

$$u_T = \frac{1}{\delta}(1 - e^{-\delta T}).$$

*Proof.* To simplify the notation, we define

$$\varphi(u, \theta) := \frac{a\delta u + \rho[1 - \hat{h}(u)]}{1 - \delta u - \theta \hat{g}(u)}.$$

Then

$$\begin{aligned} \mathbb{P}\{N_T = 1 \mid \lambda_0\} &= \frac{\partial}{\partial \theta} \exp \left[ - \int_0^{\mathfrak{G}_{0,\theta}^{-1}(T)} \varphi(u, \theta) du - \mathfrak{G}_{0,\theta}^{-1}(T) \lambda_0 \right] \Big|_{\theta=0} \\ &= \mathbb{P}\{N_T = 0 \mid \lambda_0\} (-1) \\ &\times \left[ \int_0^{\mathfrak{G}_{0,\theta}^{-1}(T)} \frac{\partial \varphi(u, \theta)}{\partial \theta} du + (\varphi(\mathfrak{G}_{0,\theta}^{-1}(T), \theta) + \lambda_0) \frac{\partial}{\partial \theta} \mathfrak{G}_{0,\theta}^{-1}(T) \right] \Big|_{\theta=0} \\ &= \mathbb{P}\{N_T = 0 \mid \lambda_0\} (-1) \\ &\times \left[ \int_0^{u_T} \frac{\partial \varphi(u, \theta)}{\partial \theta} \Big|_{\theta=0} du + (\varphi(u_T, 0) + \lambda_0) \frac{\partial}{\partial \theta} \mathfrak{G}_{0,\theta}^{-1}(T) \Big|_{\theta=0} \right], \end{aligned}$$

where

$$\begin{aligned} \frac{\partial \varphi(u, \theta)}{\partial \theta} \Big|_{\theta=0} &= \frac{\hat{g}(u)(a\delta u + \rho[1 - \hat{h}(u)])}{(1 - \delta u - \theta \hat{g}(u))^2} \Big|_{\theta=0} = \frac{\hat{g}(u)(a\delta u + \rho[1 - \hat{h}(u)])}{(1 - \delta u)^2}, \\ \varphi(u_T, 0) &= e^{\delta T} (a(1 - e^{-\delta T}) + \rho(1 - \hat{h}(u_T))), \end{aligned}$$

and  $\partial \mathfrak{G}_{0,\theta}^{-1}(T) / \partial \theta |_{\theta=0}$  can be derived as below. Since  $L(T; \theta) = \mathfrak{G}_{0,\theta}^{-1}(T)$ , we have the nonlinear ODE of  $L(\tau; \theta)$ :

$$L(\tau; \theta)' = 1 - \delta L(\tau; \theta) - \theta \hat{g}(L(\tau; \theta)), \quad 0 \leq \theta < 1,$$

with the initial condition  $L(0; \theta) = 0$ . Differentiate both sides with respect to  $\theta$  to obtain

$$L^{(1)}(\tau; \theta)' = -\delta L^{(1)}(\tau; \theta) - [\hat{g}(L(\tau; \theta)) + \theta \hat{g}^{(1)}(L(\tau; \theta))], \quad 0 \leq \theta < 1,$$

where

$$L^{(1)}(\tau; \theta) = \frac{\partial}{\partial \theta} L(\tau; \theta), \quad \hat{g}^{(1)}(L(\tau; \theta)) = \frac{\partial}{\partial \theta} \hat{g}(L(\tau; \theta)).$$

By setting  $\theta = 0$ , we have the ODE for  $L^{(1)}(\tau; 0)$ :

$$L^{(1)}(\tau; 0)' = -\delta L^{(1)}(\tau; 0) - \hat{g}(L(\tau; 0)),$$

with the initial condition  $L^{(1)}(0; 0) = 0$ . Given  $L(\tau; 0) = (1 - e^{-\delta\tau})/\delta$ ,  $L^{(1)}(\tau; 0)$  can be uniquely solved, i.e.

$$\frac{\partial}{\partial\theta} \mathcal{G}_{0,\theta}^{-1}(T) \Big|_{\theta=0} = L^{(1)}(\tau; 0) = -e^{-\delta T} \int_0^T \hat{g} \left( \frac{1 - e^{-\delta s}}{\delta} \right) e^{\delta s} ds < 0;$$

equivalently, by the change of variable  $u = (1 - e^{-\delta s})/\delta$ ,

$$\int_0^T \hat{g} \left( \frac{1 - e^{-\delta s}}{\delta} \right) e^{\delta s} ds = \int_0^{uT} \frac{\hat{g}(u)}{(1 - \delta u)^2} du.$$

Similarly to the point process  $N_t$ , the probability generating function of the size of a cluster generated by a point of any generation can also be derived as follows.

**Theorem 3.5.** *For the size of a cluster generated by a point of any generation,  $\tilde{N}_t$ , under the condition  $\delta > \mu_{1G}$ , we have*

$$\begin{aligned} E[\theta^{\tilde{N}_t} \mid \tilde{\lambda}_0] &= \exp[-\mathcal{G}_{0,\theta}^{-1}(T)\tilde{\lambda}_0], \\ E[\theta^{\tilde{N}_\infty} \mid \tilde{\lambda}_0] &= e^{-v^*\tilde{\lambda}_0}, \end{aligned} \tag{3.13}$$

where  $\mathcal{G}_{0,\theta}(\cdot)$  and  $v^*$  are given by (3.9) and (3.11), respectively, and  $\tilde{\lambda}_0$  is the value of one of the associated externally excited or self-excited jumps. In particular, for a cluster generated by a point of generation 0, we have

$$E[\theta^{\tilde{N}_\infty}] = \hat{h}(v^*);$$

for a cluster generated by a point of subsequent generations, we have

$$E[\theta^{\tilde{N}_\infty}] = \frac{1 - \delta v^*}{\theta}. \tag{3.14}$$

*Proof.* For the size of a cluster generated by a point of any generation, the infinitesimal generator of the process  $(\tilde{\lambda}_t, \tilde{N}_t, t)$  acting on a function  $f(\tilde{\lambda}, \tilde{n}, t)$  within its domain  $\Omega(\mathcal{A})$  is given by

$$\mathcal{A}f(\tilde{\lambda}, \tilde{n}, t) = \frac{\partial f}{\partial t} - \delta \tilde{\lambda} \frac{\partial f}{\partial \tilde{\lambda}} + \tilde{\lambda} \left( \int_0^\infty f(\tilde{\lambda} + z, \tilde{n} + 1, t) dG(z) - f(\tilde{\lambda}, \tilde{n}, t) \right).$$

As this is just a special case of Theorem 3.1 and Theorem 3.4, we can derive (3.13) immediately. By the proof of Theorem 3.4, we know that

$$\lim_{T \rightarrow \infty} \mathcal{G}_{0,\theta}^{-1}(T) = v^*.$$

Then

$$E[\theta^{\tilde{N}_\infty} \mid \tilde{\lambda}_0] = \lim_{T \rightarrow \infty} E[\theta^{\tilde{N}_T} \mid \tilde{\lambda}_0] = \lim_{T \rightarrow \infty} \exp[-\mathcal{G}_{0,\theta}^{-1}(T)\tilde{\lambda}_0] = e^{-v^*\tilde{\lambda}_0}.$$

In particular, for a cluster generated by a point of generation 0, we have

$$E[\theta^{\tilde{N}_\infty}] = E[E[\theta^{\tilde{N}_\infty} \mid \tilde{\lambda}_0]] = E[e^{-v^*\tilde{\lambda}_0}] = E[e^{-v^*Y_1}] = \hat{h}(v^*);$$

for a cluster generated by a point of subsequent generations, we have

$$E[\theta^{\tilde{N}_\infty}] = E[e^{-v^*Z_1}] = \hat{g}(v^*) = \frac{1 - \delta v^*}{\theta}.$$

**Remark 3.2.** The size of a cluster generated by a point of any generation is actually a pure Hawkes process with reversion level  $a = 0$ , a special case of the dynamic contagion process. As time  $t \rightarrow \infty$ , the distribution of  $\lambda_t$  converges to the distribution of a degenerate random variable at 0.

**Remark 3.3.** Alternatively, (3.14) can be derived from the perspective of the cluster process definition given in Definition 2.1, and we observe that each subcluster has the same distribution,  $\mathcal{E}(\theta) = E[\theta^{\tilde{N}_\infty}]$ , as its ancestor (for a cluster generated by a point of subsequent generation 1, 2, ...), and, hence,  $\mathcal{E}(\theta)$  satisfies the functional equation

$$\mathcal{E}(\theta) = \hat{g}\left(\frac{1 - \theta\mathcal{E}(\theta)}{\delta}\right),$$

which also leads to (3.14).

We also provide an explicit example of Theorem 3.5 in Theorem 4.3 below by assuming that the jumps have exponential distributions.

**3.4. Moments of  $\lambda_t$  and  $N_t$**

Any moment of  $\lambda_t$  and  $N_t$  can be obtained by differentiating the Laplace transform of  $\lambda_t$  and the probability generating function of  $N_t$  with respect to  $v$  and  $\theta$ , and then setting  $v$  and  $\theta$  equal to 0, respectively. Alternatively, we can obtain the first and second moments of  $\lambda_t$  and  $N_t$  directly by solving ODEs, a method which is slightly easier to generalise to derive higher moments beyond the condition  $\delta > \mu_{1G}$ ; therefore, we will proceed with this method here.

**Theorem 3.6.** *The conditional expectation of the process  $\lambda_t$  given  $\lambda_0$  at time  $t = 0$  is given by*

$$E[\lambda_t \mid \lambda_0] = \frac{\mu_{1H}\rho + a\delta}{\delta - \mu_{1G}} + \left(\lambda_0 - \frac{\mu_{1H}\rho + a\delta}{\delta - \mu_{1G}}\right)e^{-(\delta - \mu_{1G})t} \text{ for } \delta \neq \mu_{1G}, \tag{3.15}$$

$$E[\lambda_t \mid \lambda_0] = \lambda_0 + (\mu_{1H}\rho + a\delta)t \text{ for } \delta = \mu_{1G}, \tag{3.16}$$

where

$$\mu_{1H} := \int_0^\infty y \, dH(y).$$

*Proof.* By the martingale property of the infinitesimal generator given in (2.2), we have an  $\mathcal{F}$ -martingale

$$f(\lambda_t, N_t, t) - f(\lambda_0, N_0, 0) - \int_0^t \mathcal{A}(\lambda_s, N_s, s) \, ds$$

for  $f \in \Omega(\mathcal{A})$ . Now, by setting  $f(\lambda, n, t) = \lambda$  we have

$$\mathcal{A}\lambda = -(\delta - \mu_{1G})\lambda + \mu_{1H}\rho + a\delta.$$

Then  $\lambda_t - \lambda_0 - \int_0^t \mathcal{A}\lambda_s \, ds$  is an  $\mathcal{F}$ -martingale, and we have

$$E\left[\lambda_t - \int_0^t \mathcal{A}\lambda_s \, ds \mid \lambda_0\right] = \lambda_0.$$

Hence,

$$E[\lambda_t \mid \lambda_0] = \lambda_0 + E\left[\int_0^t \mathcal{A}\lambda_s \, ds \mid \lambda_0\right] = \lambda_0 - (\delta - \mu_{1G}) \int_0^t E[\lambda_s \mid \lambda_0] \, ds + (\mu_{1H}\rho + a\delta)t.$$

By differentiating with respect to  $t$ , we obtain the nonlinear inhomogeneous ODE

$$\frac{du(t)}{dt} = -(\delta - \mu_{1G})u(t) + \mu_{1H}\rho + a\delta,$$

where  $u(t) = E[\lambda_t \mid \lambda_0]$ , with the initial condition  $u(0) = \lambda_0$ . This ODE has a solution given by (3.15) and (3.16).

**Lemma 3.1.** *The second moment of the process  $\lambda_t$  given  $\lambda_0$  at time  $t = 0$  is given by*

$$\begin{aligned} E[\lambda_t^2 \mid \lambda_0] &= \lambda_0^2 e^{-2(\delta - \mu_{1G})t} \\ &+ \frac{2(\mu_{1H}\rho + a\delta) + \mu_{2G}}{\delta - \mu_{1G}} \left( \lambda_0 - \frac{\mu_{1H}\rho + a\delta}{\delta - \mu_{1G}} \right) (e^{-(\delta - \mu_{1G})t} - e^{-2(\delta - \mu_{1G})t}) \\ &+ \left( \frac{(2(\mu_{1H}\rho + a\delta) + \mu_{2G})(\mu_{1H}\rho + a\delta)}{2(\delta - \mu_{1G})^2} + \frac{\mu_{2H}\rho}{2(\delta - \mu_{1G})} \right) \\ &\times (1 - e^{-2(\delta - \mu_{1G})t}) \quad \text{for } \delta \neq \mu_{1G}, \end{aligned} \tag{3.17}$$

$$\begin{aligned} E[\lambda_t^2 \mid \lambda_0] &= \lambda_0^2 + (2(\mu_{1H}\rho + a\delta) + \mu_{2G})(\lambda_0 t + \frac{1}{2}(\mu_{1H}\rho + a\delta)t^2) \\ &+ \mu_{2H}\rho t \quad \text{for } \delta = \mu_{1G}, \end{aligned} \tag{3.18}$$

where

$$\mu_{2H} := \int_0^\infty y^2 dH(y), \quad \mu_{2G} := \int_0^\infty z^2 dG(z).$$

*Proof.* By setting  $f(\lambda, n, t) = \lambda^2$  in (2.2), we have

$$\mathcal{A}\lambda^2 = -2(\delta - \mu_{1G})\lambda^2 + (2(\mu_{1H}\rho + a\delta) + \mu_{2G})\lambda + \mu_{2H}\rho.$$

Since  $\lambda_t^2 - \lambda_0^2 - \int_0^t \mathcal{A}\lambda_s^2 ds$  is an  $\mathcal{F}$ -martingale by the martingale property of the generator, we have

$$E \left[ \lambda_t^2 - \int_0^t \mathcal{A}\lambda_s^2 ds \mid \lambda_0 \right] = \lambda_0^2.$$

Hence,

$$\begin{aligned} E[\lambda_t^2 \mid \lambda_0] &= \lambda_0^2 - 2(\delta - \mu_{1G}) \int_0^t E[\lambda_s^2 \mid \lambda_0] ds + (2(\mu_{1H}\rho + a\delta) + \mu_{2G}) \int_0^t E[\lambda_s \mid \lambda_0] ds \\ &+ \mu_{2H}\rho t. \end{aligned}$$

By differentiating with respect to  $t$ , we have the ODE

$$\begin{aligned} \frac{du(t)}{dt} + 2(\delta - \mu_{1G})u(t) &= (2(\mu_{1H}\rho + a\delta) + \mu_{2G}) \left( \lambda_0 - \frac{\mu_{1H}\rho + a\delta}{\delta - \mu_{1G}} \right) e^{-(\delta - \mu_{1G})t} \\ &+ \frac{(2(\mu_{1H}\rho + a\delta) + \mu_{2G})(\mu_{1H}\rho + a\delta)}{\delta - \mu_{1G}} + \mu_{2H}\rho, \end{aligned}$$

where  $u(t) = E[\lambda_t^2 \mid \lambda_0]$ , with the initial condition  $u(0) = \lambda_0^2$ . This ODE has a solution given by (3.17) and (3.18).

**Theorem 3.7.** *The conditional variance of the process  $\lambda_t$  given  $\lambda_0$  at time  $t = 0$  is given by*

$$\begin{aligned} \text{var}[\lambda_t \mid \lambda_0] &= \frac{1}{2(\delta - \mu_{1G})} \left( \frac{\mu_{2G}(\mu_{1H}\rho + a\delta)}{\delta - \mu_{1G}} - \mu_{2H}\rho - 2\mu_{2G}\lambda_0 \right) e^{-2(\delta - \mu_{1G})t} \\ &\quad + \frac{\mu_{2G}}{\delta - \mu_{1G}} \left( \lambda_0 - \frac{\mu_{1H}\rho + a\delta}{\delta - \mu_{1G}} \right) e^{-(\delta - \mu_{1G})t} \\ &\quad + \frac{1}{2(\delta - \mu_{1G})} \left( \mu_{2H}\rho + \frac{\mu_{2G}(\mu_{1H}\rho + a\delta)}{\delta - \mu_{1G}} \right) \quad \text{for } \delta \neq \mu_{1G}, \end{aligned} \tag{3.19}$$

$$\text{var}[\lambda_t \mid \lambda_0] = \frac{1}{2}\mu_{2G}(\mu_{1H}\rho + a\delta)t^2 + (\mu_{2G}\lambda_0 + \mu_{2H}\rho)t \quad \text{for } \delta = \mu_{1G}. \tag{3.20}$$

*Proof.* By  $\text{var}[\lambda_t \mid \lambda_0] = E[\lambda_t^2 \mid \lambda_0] - (E[\lambda_t \mid \lambda_0])^2$ , based on Theorem 3.6 and Lemma 3.1, the result follows.

**Corollary 3.5.** *Assume that  $\delta > \mu_{1G}$ . Then the first and second moments and the variance of the stationary distribution of the process  $\lambda_t$  are given by*

$$E[\lambda_t] = \frac{\mu_{1H}\rho + a\delta}{\delta - \mu_{1G}}, \tag{3.21}$$

$$E[\lambda_t^2] = \frac{(2(\mu_{1H}\rho + a\delta) + \mu_{2G})(\mu_{1H}\rho + a\delta)}{2(\delta - \mu_{1G})^2} + \frac{\mu_{2H}\rho}{2(\delta - \mu_{1G})}, \tag{3.22}$$

$$\text{var}[\lambda_t] = \frac{1}{2(\delta - \mu_{1G})} \left( \mu_{2H}\rho + \frac{\mu_{2G}(\mu_{1H}\rho + a\delta)}{\delta - \mu_{1G}} \right).$$

*Proof.* By setting time  $t \rightarrow \infty$  in (3.15), (3.16), (3.17), (3.18), (3.19), and (3.20), respectively, the results follow.

We will now derive the moments for the point process  $N_t$ , assuming that  $\delta > \mu_{1G}$ .

**Theorem 3.8.** *For the stationary distribution of the process  $\lambda_t$ , given the condition  $\delta > \mu_{1G}$  and  $N_0 = 0$ , the expectation of the point process  $N_t$  is given by*

$$E[N_t] = \frac{\mu_{1H}\rho + a\delta}{\delta - \mu_{1G}}t. \tag{3.23}$$

*Proof.* By setting  $f(\lambda, n, t) = n$  in (2.2), we have  $\mathcal{A}n = \lambda$ . Since  $N_t - N_0 - \int_0^t \lambda_s \, ds$  is a martingale by the martingale property of the intensity process  $\lambda_t$  of the point process  $N_t$  given by definition (2.1), we have

$$E[N_t - N_0 \mid \mathcal{F}_0] = E \left[ \int_0^t \lambda_s \, ds \mid \mathcal{F}_0 \right],$$

and we also know  $E[\lambda_t]$  from Corollary 3.5. Then, by assuming that  $N_0 = 0$ , we have

$$E[N_t] = E[N_t - N_0] = \int_0^t E[\lambda_s] \, ds = \frac{\mu_{1H}\rho + a\delta}{\delta - \mu_{1G}}t.$$

**Lemma 3.2.** *For the stationary distribution of the process  $\lambda_t$ , given the condition  $\delta > \mu_{1G}$  and  $N_0 = 0$ , we have*

$$E[\lambda_t N_t] = \bar{k}(1 - e^{-(\delta - \mu_{1G})t}) + \left( \frac{\mu_{1H}\rho + a\delta}{\delta - \mu_{1G}} \right)^2 t, \tag{3.24}$$

where

$$\bar{k} := \frac{2\mu_{1G}(\mu_{1H}\rho + a\delta) + \mu_{2H}\rho}{2(\delta - \mu_{1G})^2} + \frac{\mu_{2G}(\mu_{1H}\rho + a\delta)}{2(\delta - \mu_{1G})^3}. \tag{3.25}$$

*Proof.* By setting  $f(\lambda, n, t) = \lambda n$  in (2.2), we have

$$\mathcal{A}(\lambda n) = -(\delta - \mu_{1G})\lambda n + (\mu_{1H}\rho + a\delta)n + \lambda^2 + \mu_{1G}\lambda.$$

Since  $\lambda_t N_t - \lambda_0 N_0 - \int_0^t \mathcal{A}(\lambda_s N_s) ds$  is an  $\mathcal{F}$ -martingale by the martingale property of the generator, given  $N_0 = 0$ , we have the ODE

$$\frac{du(t)}{dt} = -(\delta - \mu_{1G})u(t) + (\mu_{1H}\rho + a\delta) E[N_t] + E[\lambda_t^2] + \mu_{1G} E[\lambda_t],$$

where  $u(t) = E[\lambda_t N_t]$ , with the initial condition  $u(0) = 0$ . Note that  $E[N_t]$ ,  $E[\lambda_t^2]$ , and  $E[\lambda_t]$  are already given by (3.23), (3.22), and (3.21), respectively; therefore, this ODE has a solution given by (3.24).

**Theorem 3.9.** *For the stationary distribution of the process  $\lambda_t$ , given the condition  $\delta > \mu_{1G}$  and  $N_0 = 0$ , the second moment and the variance of the point process  $N_t$  are given by*

$$E[N_t^2] = \frac{2}{\delta - \mu_{1G}}(e^{-(\delta - \mu_{1G})t} - 1) + 2\bar{k}t + \left(\frac{\mu_{1H}\rho + a\delta}{\delta - \mu_{1G}}\right)^2 t^2,$$

$$\text{var}[N_t] = \frac{2}{\delta - \mu_{1G}}(e^{-(\delta - \mu_{1G})t} - 1) + 2\bar{k}t,$$

where the constant  $\bar{k}$  is given by (3.25).

*Proof.* By setting  $f(\lambda, n, t) = n^2$  in (2.2), we have  $\mathcal{A}(n^2) = (2n + 1)\lambda$ . Since  $N_t^2 - N_0^2 - \int_0^t (2N_s + 1)\lambda_s ds$  is an  $\mathcal{F}$ -martingale by the martingale property of the generator, given  $N_0 = 0$ , we have

$$E[N_t^2] = 2 \int_0^t E[\lambda_s N_s] ds + \int_0^t E[\lambda_s] ds,$$

where  $E[\lambda_t N_t]$  and  $E[\lambda_t]$  are given by (3.24) and (3.21), respectively. Then  $E[N_t^2]$  follows. Since  $\text{var}[N_t] = E[N_t^2] - E[N_t]^2$  given  $E[N_t]$  in (3.23),  $\text{var}[N_t]$  follows.

The moments for the special case Hawkes process and other similar processes can also be found in Oakes (1975) and Azizpour and Giesecke (2008), and, more generally, in Brémaud *et al.* (2002).

#### 4. Example: jumps with exponential distributions

To give an explicit example for the key distributional properties derived above, in this section we assume that both externally excited and self-excited jumps follow exponential distributions, i.e. the density functions are given by

$$h(y) = \alpha e^{-\alpha y}, \quad g(z) = \beta e^{-\beta z}, \quad \text{where } y, z; \alpha, \beta > 0, \tag{4.1}$$

and the Laplace transforms have the explicit forms

$$\hat{h}(u) = \frac{\alpha}{\alpha + u}, \quad \hat{g}(u) = \frac{\beta}{\beta + u}. \tag{4.2}$$

In the next two subsections we obtain the corresponding Laplace transform of  $\lambda_T$ , the conditional probability generating function of  $N_T$ , and the conditional probabilities  $P\{N_T = 0 \mid \lambda_0\}$  and  $P\{N_T = 1 \mid \lambda_0\}$ . We will use these results to model the credit default risk in Section 5. Note that there are parameters  $(a, \rho, \delta; \alpha, \beta; \lambda_0)$  for the general dynamic contagion process and  $(a, \delta; \beta; \lambda_0)$  for the Hawkes process.

**4.1. Laplace transform of  $\lambda_T$**

**Lemma 4.1.** *If both the self-excited and externally excited jumps follow exponential distributions, i.e. the density functions are specified by (4.1), then the conditional Laplace transform of  $\lambda_T$  given  $\lambda_0$  at time  $t = 0$ , under the condition  $\delta\beta > 1$ , is given by*

$$E[e^{-v\lambda_T} \mid \lambda_0] = \exp[-(\mathcal{C}_1(v) - \mathcal{C}_1(\mathcal{G}_{v,1}^{-1}(T)))] \exp[-\mathcal{G}_{v,1}^{-1}(T)\lambda_0],$$

where

$$\mathcal{C}_1(u) := \begin{cases} au + \frac{\rho(\alpha - \beta)}{\delta(\alpha - \beta) + 1} \ln(\alpha + u) \\ \quad + \frac{1}{\delta} \left( a + \frac{\rho}{\delta(\alpha - \beta) + 1} \right) \ln\left(u + \frac{\delta\beta - 1}{\delta}\right) & \text{for } \alpha \neq \beta - \frac{1}{\delta}, \\ au + \frac{\rho\beta}{\delta\beta - 1} \ln(\alpha + u) - \frac{\alpha\rho}{\delta(\delta\beta - 1)} \frac{1}{\alpha + u} \\ \quad + \frac{1}{\delta} \left( a - \frac{\rho}{\delta\beta - 1} \right) \ln\left(u + \frac{\delta\beta - 1}{\delta}\right) & \text{for } \alpha = \beta - \frac{1}{\delta}, \end{cases} \tag{4.3}$$

and

$$\mathcal{G}_{v,1}(L) = \frac{1}{\delta(\delta\beta - 1)} \left[ \delta\beta \ln\left(\frac{v}{L}\right) - \ln\left(\frac{\delta v + (\delta\beta - 1)}{\delta L + (\delta\beta - 1)}\right) \right].$$

*Proof.* By Theorem 3.2 and  $\mu_{1G} = 1/\beta$ , the condition is  $\delta > 1/\beta$ . Substituting (4.2) into Theorem 3.2 yields

$$\mathcal{G}_{v,1}(L) = \int_L^v \frac{u + \beta}{\delta u(u + (\delta\beta - 1)/\delta)} du$$

and

$$\mathcal{C}_1(v) - \mathcal{C}_1(\mathcal{G}_{v,1}^{-1}(T)) = \int_{\mathcal{G}_{v,1}^{-1}(T)}^v \frac{(a + \rho/\delta(u + \alpha))(\beta + u)}{u + (\delta\beta - 1)/\delta} du.$$

Note that, when calculating the integral, we need to consider the special case  $\alpha = \beta - 1/\delta$ . Then the result follows.

**Theorem 4.1.** *If both the externally excited and self-excited jumps follow exponential distributions, i.e. the density functions are specified by (4.1), then, under the condition  $\delta\beta > 1$ , the stationary distribution of the process  $\{\lambda_t\}_{t \geq 0}$  is given by*

$$\begin{aligned} a + \tilde{\Gamma}_1 + \tilde{\Gamma}_2 & \text{ for } \alpha \geq \beta, \\ a + \tilde{\Gamma}_3 + \tilde{B} & \text{ for } \alpha < \beta \text{ and } \alpha \neq \beta - 1/\delta, \\ a + \tilde{\Gamma}_4 + \tilde{P} & \text{ for } \alpha = \beta - 1/\delta, \end{aligned}$$

where

$$\begin{aligned}
 \tilde{\Gamma}_1 &\sim \text{Gamma}\left(\frac{1}{\delta}\left(a + \frac{\rho}{\delta(\alpha - \beta) + 1}\right), \frac{\delta\beta - 1}{\delta}\right), & \tilde{\Gamma}_2 &\sim \text{Gamma}\left(\frac{\rho(\alpha - \beta)}{\delta(\alpha - \beta) + 1}, \alpha\right), \\
 \tilde{\Gamma}_3 &\sim \text{Gamma}\left(\frac{a + \rho}{\delta}, \frac{\delta\beta - 1}{\delta}\right), & \tilde{\Gamma}_4 &\sim \text{Gamma}\left(\frac{a + \rho}{\delta}, \alpha\right), \\
 N_1 &\sim \text{NegBin}\left(\frac{\rho}{\delta} \frac{\beta - \alpha}{\gamma_1 - \gamma_2}, \frac{\gamma_2}{\gamma_1}\right), & N_2 &\sim \text{Poisson}\left(\frac{\rho}{\delta^2\alpha}\right), \\
 X_i^{(1)} &\sim \text{Exp}(\gamma_1), & X_i^{(2)} &\sim \text{Exp}(\alpha), \\
 \gamma_1 &= \max\left\{\alpha, \frac{\delta\beta - 1}{\delta}\right\}, & \gamma_2 &= \min\left\{\alpha, \frac{\delta\beta - 1}{\delta}\right\}, \\
 \tilde{B} &\stackrel{D}{=} \sum_{i=1}^{N_1} X_i^{(1)}, & \tilde{P} &\stackrel{D}{=} \sum_{i=1}^{N_2} X_i^{(2)},
 \end{aligned}$$

are independent random variables;  $\tilde{B}$  follows a compound negative binomial distribution with underlying exponential jumps and  $\tilde{P}$  follows a compound Poisson distribution with underlying exponential jumps.

*Proof.* By Lemma 4.1 and Theorem 3.3, and as  $\mathcal{G}_{v,1}^{-1}(T) \rightarrow 0$  when  $T \rightarrow \infty$ , the Laplace transform of the stationary distribution of the process  $\{\lambda_t\}_{t \geq 0}$  is given by  $\hat{\Pi}(v) = e^{-(C_1(v) - C_1(0))}$ , where  $C_1(u)$  is the explicit function given in (4.3). Then

$$\hat{\Pi}(v) = \begin{cases} e^{-va} \left(\frac{\alpha}{\alpha + v}\right)^{\rho(\alpha - \beta)/(\delta(\alpha - \beta) + 1)} \\ \quad \times \left(\frac{(\delta\beta - 1)/\delta}{v + (\delta\beta - 1)/\delta}\right)^{(1/\delta)(a + \rho/(\delta(\alpha - \beta) + 1))} & \text{for } \alpha \geq \beta, \\ e^{-va} \left(\frac{(\delta\beta - 1)/\delta}{v + (\delta\beta - 1)/\delta}\right)^{(a + \rho)/\delta} \\ \quad \times \left(\frac{\gamma_2/\gamma_1}{1 - (1 - \gamma_2/\gamma_1)(\gamma_1/(\gamma_1 + v))}\right)^{(\rho/\delta)((\beta - \alpha)/(\gamma_1 - \gamma_2))} & \text{for } \alpha < \beta \text{ and} \\ & \alpha \neq \beta - 1/\delta, \\ e^{-va} \left(\frac{\alpha}{\alpha + v}\right)^{(\rho + a)/\delta} \exp\left[\frac{\rho}{\delta^2\alpha} \left(\frac{\alpha}{\alpha + v} - 1\right)\right] & \text{for } \alpha = \beta - 1/\delta. \end{cases} \tag{4.4}$$

If  $\alpha \geq \beta$ , it is obvious that (4.4) is the Laplace transform of two independent gamma distributions  $\tilde{\Gamma}_1$  and  $\tilde{\Gamma}_2$  shifted by a constant  $a$ . If  $\alpha < \beta$  and  $\alpha \neq \beta - 1/\delta$ , then  $\gamma_1 > \gamma_2$  always, and the second term is the Laplace transform of a gamma distribution with parameters  $(a + \rho)/\delta$  and  $(\delta\beta - 1)/\delta$ . The third term is the Laplace transform of a compound negative binomial distribution with parameters  $(\rho/\delta)((\beta - \alpha)/(\gamma_1 - \gamma_2))$  and  $\gamma_2/\gamma_1$ , and the underlying jumps follow an exponential distribution with parameter  $\gamma_1$ , since we know that the Laplace transform of a negative binomial distribution  $N_1$  with parameters  $(r, p)$  is

$$E[e^{-vN_1}] = \left(\frac{p}{1 - (1 - p)e^{-v}}\right)^r.$$

Then

$$\begin{aligned} E[e^{-v\tilde{B}}] &= E\left[E\left[\exp\left[-v\sum_{i=1}^{N_1} X_i^{(1)}\right]\right] \mid N_1\right] \\ &= E\left[\left(\frac{\gamma_1}{\gamma_1 + v}\right)^{N_1}\right] \\ &= E\left[\exp\left[-\ln\left(\frac{\gamma_1 + v}{\gamma_1}\right)N_1\right]\right] \\ &= \left(\frac{p}{1 - (1 - p)\exp[-\ln((\gamma_1 + v)/\gamma_1)]}\right)^r \\ &= \left(\frac{p}{1 - (1 - p)(\gamma_1/(\gamma_1 + v))}\right)^r, \end{aligned}$$

where

$$p = \frac{\gamma_1}{\gamma_2} \in (0, 1), \quad r = \frac{\rho}{\delta} \frac{\beta - \alpha}{\gamma_1 - \gamma_2} \in \mathbb{R}^+.$$

Also, it is easy to identify the corresponding Laplace transforms for the case when  $\alpha = \beta - 1/\delta$ .

We discuss some important special cases below.

**Remark 4.1.** If both jumps follow the same exponential distribution, i.e.  $\alpha = \beta$ , then  $\tilde{\Gamma}_1$  and  $\tilde{\Gamma}_2$  combine as one single gamma random variable  $\tilde{\Gamma}_3$ .

**Remark 4.2.** For the nonself-excited case, i.e. when  $\beta = \infty$ , the Laplace transform of the stationary distribution of the process  $\{\lambda_t\}_{t \geq 0}$  is given by

$$\hat{\Pi}(v) = e^{-va} \left(\frac{\alpha}{\alpha + v}\right)^{\rho/\delta}.$$

Then  $\lambda_t$  follows a shifted gamma distribution, i.e.

$$\lambda_t \stackrel{D}{=} a + \tilde{\Gamma}_5,$$

where

$$\tilde{\Gamma}_5 \sim \text{Gamma}\left(\frac{\rho}{\delta}, \alpha\right),$$

which recovers the result in Dassios and Jang (2003) by setting  $a = 0$ .

**Remark 4.3.** For the Hawkes process, i.e. the nonexternally excited case when  $\alpha = \infty$ , or  $\rho = 0$ , the Laplace transform of the stationary distribution of the process  $\{\lambda_t\}_{t \geq 0}$  is given by

$$\hat{\Pi}(v) = e^{-va} \left(\frac{(\delta\beta - 1)/\delta}{v + (\delta\beta - 1)/\delta}\right)^{a/\delta}.$$

Then  $\lambda_t$  follows a shifted Gamma distribution, i.e.

$$\lambda_t \stackrel{D}{=} a + \tilde{\Gamma}_6,$$

where

$$\tilde{\Gamma}_6 \sim \text{Gamma}\left(\frac{a}{\delta}, \frac{\delta\beta - 1}{\delta}\right).$$

The result for the particular case  $\alpha = \beta - 1/\delta$  is actually the limit version of the result for the case when  $\alpha < \beta$  and  $\alpha \neq \beta - 1/\delta$ . In the following sections, we focus only on the main case when  $\alpha \neq \beta - 1/\delta$ , with the Laplace transform of the stationary distribution of the process  $\{\lambda_t\}_{t \geq 0}$  specified by (4.4).

**4.2. Probability generating function of  $N_T$**

**Theorem 4.2.** *If both the externally excited and self-excited jumps follow exponential distributions, i.e. the density functions are specified by (4.1), then the conditional probability generating function of  $N_T$  given  $\lambda_0$  and  $N_0 = 0$  at time  $t = 0$ , under the condition  $\delta\beta > 1$ , is given by*

$$E[\theta^{N_T} \mid \lambda_0] = \exp[-(\mathcal{C}_2(\mathcal{G}_{0,\theta}^{-1}(T)) - \mathcal{C}_2(0))] \exp[-\mathcal{G}_{0,\theta}^{-1}(T)\lambda_0], \quad \alpha \neq -v_-^*,$$

where

$$\begin{aligned} \mathcal{C}_2(u) := & -au + \frac{\alpha(\beta - \alpha)\rho}{\delta(\alpha + v_-^*)(\alpha + v^*)} \ln(u + \alpha) \\ & + \frac{1}{\delta(v^* - v_-^*)} \left\{ \left[ a(v_-^* + (1 - \theta)\beta) + \rho v_-^* \frac{\beta + v_-^*}{\alpha + v_-^*} \right] \ln(u - v_-^*) \right. \\ & \left. - \left[ a(v^* + (1 - \theta)\beta) + \rho v^* \frac{\beta + v^*}{\alpha + v^*} \right] \ln(v^* - u) \right\} \end{aligned}$$

and

$$\mathcal{G}_{0,\theta}(L) = K(L) - K(0), \quad 0 \leq L < v^*,$$

where

$$\begin{aligned} K(u) := & -\frac{1}{\delta(v^* - v_-^*)} [(v^* + \beta) \ln(v^* - u) - (v_-^* + \beta) \ln(u - v_-^*)], \quad 0 \leq u < v^*, \\ v^* = & \frac{\sqrt{\Delta} - (\delta\beta - 1)}{2\delta} > 0, \\ -\beta \leq v_-^* = & -\frac{\sqrt{\Delta} + (\delta\beta - 1)}{2\delta} < 0, \\ \Delta = & (\delta\beta + 1)^2 - 4\theta\delta\beta > 0, \quad 0 \leq \theta < 1. \end{aligned} \tag{4.5}$$

*Proof.* Since  $0 < u < v^*$ , by substituting the explicit results of (4.2) into Theorem 3.4, we have

$$\mathcal{G}_{0,\theta}(L) = \int_0^L \frac{\beta + u}{-\delta u^2 - (\delta\beta - 1)u + (1 - \theta)\beta} du = K(L) - K(0)$$

and

$$\begin{aligned} \mathcal{C}_2(u) = & -a \left\{ u - K(u) - \frac{\theta\beta}{\delta} \frac{1}{v^* - v_-^*} \ln \frac{v^* - u}{u - v_-^*} \right\} \\ & + \rho \left\{ K(u) + \frac{\alpha}{\delta} \frac{1}{v^* - v_-^*} \left[ \ln \frac{v^* - u}{u - v_-^*} \right. \right. \\ & \left. \left. + (\beta - \alpha) \left( \frac{1}{\alpha + v^*} \ln \frac{v^* - u}{u + \alpha} - \frac{1}{\alpha + v_-^*} \ln \frac{u - v_-^*}{u + \alpha} \right) \right] \right\}; \end{aligned}$$

furthermore,

$$\begin{aligned} v^* &= \sqrt{\Delta} - \frac{\delta\beta - 1}{2\delta} \\ &= \frac{\sqrt{(\delta\beta - 1)^2 + 4(1 - \theta)\delta\beta} - (\delta\beta - 1)}{2\delta} \\ &> \frac{(\delta\beta - 1) - (\delta\beta - 1)}{2\delta} \\ &= 0, \end{aligned}$$

and

$$-v_-^* = \sqrt{\Delta} + \frac{\delta\beta - 1}{2\delta} = \frac{\sqrt{(\delta\beta + 1)^2 - 4\theta\delta\beta} + (\delta\beta - 1)}{2\delta} \leq \frac{(\delta\beta + 1) + (\delta\beta - 1)}{2\delta} = \beta,$$

where  $v_-^* = -\beta$  only when  $\theta = 0$ .

**Remark 4.4.** We need to assume that  $\alpha \neq -v_-^*$  in Theorem 4.2, since

$$-v_-^* = \frac{\sqrt{(\delta\beta + 1)^2 - 4\theta\delta\beta} + (\delta\beta - 1)}{2\delta},$$

and, for each  $\theta \in [0, 1)$ , we have a unique  $v_-^*$ , where

$$-v_-^* \in \left( \beta - \frac{1}{\delta}, \beta \right].$$

Therefore, if  $\alpha \in (\beta - 1/\delta, \beta]$ , there exists a unique  $\theta \in [0, 1)$  such that  $\alpha + v_-^* = 0$ .

We will not consider the particular case  $\alpha = -v_-^*$  here, choosing in the sequel to assume that  $\alpha \neq -v_-^*$ .

Now we derive the probabilities  $P\{N_T = 0 \mid \lambda_0\}$  in Corollary 4.1 and  $P\{N_T = 1 \mid \lambda_0\}$  for  $\alpha \neq \beta$  in Corollary 4.2; a discussion for the special case  $\alpha = \beta$  is given in Remark 4.5 below.

**Corollary 4.1.** *If both the externally excited and self-excited jumps follow exponential distributions, i.e. the density functions are specified by (4.1), then the conditional probability of no jump given  $\lambda_0$  and  $N_0 = 0$ , under the condition  $\delta\beta > 1$ , is given by*

$$\begin{aligned} P\{N_T = 0 \mid \lambda_0\} &= \exp\left[-\left(a + \frac{\rho}{1 + \delta\alpha}\right)T\right] \exp\left[\frac{a - \lambda_0}{\delta}(1 - e^{-\delta T})\right] \\ &\quad \times \left(\frac{1 - e^{-\delta T} + \delta\alpha}{\delta\alpha}\right)^{\alpha\rho/(1 + \delta\alpha)}. \end{aligned}$$

*Proof.* By Theorem 4.2 and setting  $\theta = 0$ , we have  $\Delta = (\delta\beta + 1)^2$ ,  $v^* = 1/\delta$ ,  $v_-^* = -\beta$ ,

$$g_{0,0}^{-1}(T) = \frac{1}{\delta}(1 - e^{-\delta T}), \quad K(u) = -\frac{1}{\delta}\ln(1 - \delta u), \quad 0 \leq u < 1/\delta,$$

$$\begin{aligned} \mathcal{C}_2(u) &= -au + \frac{\alpha(\beta - \alpha)\rho}{\delta(\alpha + v_-^*)(\alpha + v^*)} \ln(u + \alpha) \\ &\quad - \frac{1}{\delta(v^* - v_-^*)} \left(a + \frac{\rho v^*}{v^* + \alpha}\right) (v^* + \beta) \ln(v^* - u) \\ &= -au - \frac{\alpha\rho}{\delta\alpha + 1} \ln(u + \alpha) - \frac{1}{\delta} \left(a + \frac{\rho}{\delta\alpha + 1}\right) \ln\left(\frac{1}{\delta} - u\right), \end{aligned}$$

and the result follows.

**Corollary 4.2.** *If both the externally excited and self-excited jumps follow exponential distributions, i.e. the density functions are specified by (4.1) ( $\alpha \neq \beta$ ), then the conditional probability of exactly one jump given  $\lambda_0$  and  $N_0 = 0$ , under the condition  $\delta\beta > 1$ , is given by*

$$\begin{aligned} P\{N_T = 1 \mid \lambda_0\} &= P\{N_T = 0 \mid \lambda_0\} \\ &\times \left[ (H_T + a\delta\beta - \rho)Q_T - a\beta(e^{\delta T} - 1) \right. \\ &\quad \left. + \rho \frac{\alpha\beta}{1 + \delta\beta} \left( \bar{a} \ln\left(\frac{\alpha + u_T}{\alpha}\right) - \bar{b} \ln\left(\frac{\beta + u_T}{\beta}\right) + \bar{c}T + \bar{d}(e^{\delta T} - 1) \right) \right], \end{aligned}$$

where

$$\begin{aligned} H_T &= \left( a + \frac{\rho}{\delta\alpha + 1 - e^{-\delta T}} \right) (1 - e^{-\delta T}) + \lambda_0 e^{-\delta T}, \\ Q_T &= \frac{\beta}{1 + \delta\beta} \left[ \frac{1}{1 + \delta\beta} \ln\left(\frac{\beta + u_T}{\beta}\right) + \delta T + (e^{\delta T} - 1) \right], \\ u_T &= \frac{1}{\delta} (1 - e^{-\delta T}), \\ \bar{a} &= \frac{1}{1 + \delta\beta} \frac{1}{\beta - \alpha} + \frac{\delta}{1 + \delta\alpha} \left( \frac{1}{1 + \delta\beta} + \frac{1}{1 + \delta\alpha} \right), \\ \bar{b} &= \frac{1}{1 + \delta\beta} \frac{1}{\beta - \alpha}, \\ \bar{c} &= \frac{\delta^2}{1 + \delta\alpha} \left( \frac{1}{1 + \delta\beta} + \frac{1}{1 + \delta\alpha} \right), \\ \bar{d} &= \frac{\delta}{1 + \delta\alpha}. \end{aligned}$$

*Proof.* By Corollary 3.4 and

$$\frac{1}{(\beta + u)(1 - \delta u)^2} = \frac{1}{1 + \delta\beta} \left[ \frac{1}{1 + \delta\beta} \left( \frac{1}{\beta + u} + \frac{\delta}{1 - \delta u} \right) + \frac{\delta}{(1 - \delta u)^2} \right],$$

we obtain  $Q_T$  by

$$\begin{aligned} \int_0^{u_T} \frac{\hat{g}(u)}{(1 - \delta u)^2} du &= \beta \int_0^{u_T} \frac{1}{(\beta + u)(1 - \delta u)^2} du \\ &= \frac{\beta}{1 + \delta\beta} \left\{ \frac{1}{1 + \delta\beta} \left[ \ln\left(\frac{\beta + u_T}{\beta}\right) + \delta T \right] + e^{\delta T} - 1 \right\}, \end{aligned}$$

and

$$\int_0^{u_T} \frac{\hat{g}(u)u}{(1 - \delta u)^2} du = \frac{\beta}{\delta} (e^{\delta T} - 1) - \beta Q_T.$$

Furthermore, when  $\alpha \neq \beta$ ,

$$\begin{aligned} \int_0^{u_T} \frac{\hat{g}(u)\hat{h}(u)}{(1 - \delta u)^2} du &= \alpha\beta \int_0^{u_T} \frac{1}{(\alpha + u)(\beta + u)(1 - \delta u)^2} du \\ &= \frac{\alpha\beta}{1 + \delta\beta} \left( \bar{a} \ln\left(\frac{\alpha + u_T}{\alpha}\right) - \bar{b} \ln\left(\frac{\beta + u_T}{\beta}\right) + \bar{c}T + \bar{d}(e^{\delta T} - 1) \right), \end{aligned}$$

completing the proof.

**Remark 4.5.** If  $\alpha = \beta$  then

$$\begin{aligned} P\{N_T = 1 \mid \lambda_0\} &= P\{N_T = 0 \mid \lambda_0\} \\ &\times \left\{ (H_T + a\delta\beta - \rho)Z_T - a\beta(e^{\delta T} - 1) \right. \\ &\quad \left. + \rho \left( \frac{\beta}{1 + \delta\beta} \right)^2 \left[ \frac{u_T}{\beta(\beta + u_T)} + \delta(e^{\delta T} - 1) \right. \right. \\ &\quad \left. \left. + \frac{2\delta}{\delta\beta + 1} \left( \ln \left( \frac{\beta + u_T}{\beta} \right) + \delta T \right) \right] \right\}, \end{aligned}$$

where

$$\begin{aligned} H_T &= \left( a + \frac{\rho}{\delta\beta + 1 - e^{-\delta T}} \right) (1 - e^{-\delta T}) + \lambda_0 e^{-\delta T}, \\ Q_T &= \frac{\beta}{1 + \delta\beta} \left[ \frac{1}{1 + \delta\beta} \ln \left( \frac{\beta + u_T}{\beta} \right) + \delta T + (e^{\delta T} - 1) \right]. \end{aligned}$$

Note that, when  $\alpha = \beta$ ,

$$\begin{aligned} &\int_0^{u_T} \frac{\hat{g}(u)\hat{h}(u)}{(1 - \delta u)^2} du \\ &= \beta^2 \int_0^{u_T} \left( \frac{1}{(\beta + u)(1 - \delta u)} \right)^2 du \\ &= \left( \frac{\beta}{1 + \delta\beta} \right)^2 \left[ \frac{u_T}{\beta(\beta + u_T)} + \delta(e^{\delta T} - 1) + \frac{2\delta}{\delta\beta + 1} \left( \ln \left( \frac{\beta + u_T}{\beta} \right) + \delta T \right) \right]. \end{aligned}$$

**Remark 4.6.** For the Hawkes process, we obtain the conditional probabilities of no jump and exactly one jump by setting  $\rho = 0$  in Corollary 4.1 and Corollary 4.2, respectively:

$$\begin{aligned} P\{N_T = 0 \mid \lambda_0\} &= e^{-aT} \exp \left[ \frac{a - \lambda_0}{\delta} (1 - e^{-\delta T}) \right], \\ P\{N_T = 1 \mid \lambda_0\} &= P\{N_T = 0 \mid \lambda_0\} \\ &\times \beta \left[ \frac{a(1 - e^{-\delta T} + \delta\beta) + \lambda_0 e^{-\delta T}}{1 + \delta\beta} \right. \\ &\quad \left. \times \left( \frac{1}{1 + \delta\beta} \ln \left( \frac{\beta + u_T}{\beta} \right) + \delta T + (e^{\delta T} - 1) \right) - a(e^{\delta T} - 1) \right]. \end{aligned}$$

**Remark 4.7.** The corresponding moments of  $\lambda_t$  and  $N_t$  based on exponential jump distributions are omitted as they can be easily obtained using the results in Section 3.4.

We now state and prove the results for the size of clusters based on Theorem 3.5 for this exponential distribution case.

**Theorem 4.3.** *If both the externally excited and self-excited jumps follow exponential distributions, i.e. the density functions are specified by (4.1), then, for the size of a cluster generated by a point of any generation,  $\tilde{N}_t$ , under the condition  $\delta\beta > 1$ , we have*

$$E[\theta^{\tilde{N}_\infty} \mid \tilde{\lambda}_0] = \exp \left[ - \frac{\sqrt{(\delta\beta - 1)^2 + 4\delta\beta(1 - \theta)} - (\delta\beta - 1)}{2\delta} \tilde{\lambda}_0 \right], \tag{4.6}$$

and  $\tilde{N}_\infty$  conditional on  $\tilde{\lambda}_0$  actually follows a mixed Poisson distribution, i.e.

$$P\{\tilde{N}_\infty = k \mid \tilde{\lambda}_0\} = \int_0^\infty \frac{v^k e^{-v}}{k!} m(v) dv, \quad k = 0, 1, 2, \dots, \tag{4.7}$$

where  $m(v)$  is the density function of the mixing distribution,

$$m(v) = \exp\left[\frac{\delta\beta - 1}{2\delta} \tilde{\lambda}_0\right] \exp\left[-\left(\frac{\delta\beta - 1}{2\delta}\right)^2 \frac{\delta}{\beta} v\right] \frac{(\sqrt{\beta/2\delta}) \tilde{\lambda}_0}{\sqrt{2\pi} v^{3/2}} \exp\left[-\frac{\beta \tilde{\lambda}_0^2 / 2\delta}{2v}\right], \tag{4.8}$$

which is an inverse Gaussian distribution with parameters  $(\beta \tilde{\lambda}_0 / (\delta\beta - 1), \beta \tilde{\lambda}_0^2 / 2\delta)$ .

*Proof.* By substituting the explicit exponential distribution functions of (4.2) and the constant  $v^*$  of (4.5) into Theorem 3.5, we obtain (4.6) immediately.

To prove that  $\tilde{N}_\infty$  follows a mixed Poisson distribution, we rewrite (4.6) as

$$E[\theta^{\tilde{N}_\infty} \mid \tilde{\lambda}_0] = \exp\left[\frac{\delta\beta - 1}{2\delta} \tilde{\lambda}_0\right] e^{-\sqrt{2\xi} \tilde{\lambda}_0},$$

where  $\xi = \frac{1}{2}((\delta\beta - 1)/2\delta)^2 + \beta(1 - \theta)/2\delta$ , and identify that

$$e^{-\sqrt{2\xi} \tilde{\lambda}_0} = E[e^{-\xi \tilde{I}G}] = \int_0^\infty e^{-\xi u} \frac{(\tilde{\lambda}_0^2)^{1/2}}{\sqrt{2\pi} u^{3/2}} e^{-\tilde{\lambda}_0^2 / 2u} du,$$

where  $\tilde{I}G$  follows the (infinite mean) inverse Gaussian distribution with parameters  $(\infty, \tilde{\lambda}_0^2)$ . Then we have

$$\begin{aligned} E[\theta^{\tilde{N}_\infty} \mid \tilde{\lambda}_0] &= \exp\left[\frac{\delta\beta - 1}{2\delta} \tilde{\lambda}_0\right] \int_0^\infty e^{-\xi u} \frac{(\tilde{\lambda}_0^2)^{1/2}}{\sqrt{2\pi} u^{3/2}} e^{-\tilde{\lambda}_0^2 / 2u} du \\ &= \int_0^\infty \exp\left[-\left[\frac{1}{2}\left(\frac{\delta\beta - 1}{2\delta}\right)^2 + \frac{\beta}{2\delta}(1 - \theta)\right] u\right] \exp\left[\frac{\delta\beta - 1}{2\delta} \tilde{\lambda}_0\right] \\ &\quad \times \frac{(\tilde{\lambda}_0^2)^{1/2}}{\sqrt{2\pi} u^{3/2}} e^{-\tilde{\lambda}_0^2 / 2u} du. \end{aligned}$$

Let  $v = \beta u / 2\delta$ . Then

$$\begin{aligned} E[\theta^{\tilde{N}_\infty} \mid \tilde{\lambda}_0] &= \int_0^\infty e^{-(1-\theta)v} \exp\left[\frac{\delta\beta - 1}{2\delta} \tilde{\lambda}_0\right] \exp\left[-\left(\frac{\delta\beta - 1}{2\delta}\right)^2 \frac{\delta}{\beta} v\right] \\ &\quad \times \frac{(\beta \tilde{\lambda}_0^2 / 2\delta)^{1/2}}{\sqrt{2\pi} v^{3/2}} \exp\left[-\frac{\beta \tilde{\lambda}_0^2 / 2\delta}{2v}\right] dv \\ &= \int_0^\infty e^{-(1-\theta)v} m(v) dv \\ &= \hat{m}(\theta - 1), \end{aligned}$$

where  $\hat{m}(u) = \int_0^\infty e^{-uv} m(v) dv$ . Hence, by the definition of the mixed Poisson distribution, we have (4.7) and (4.8); setting  $u = 1 - \theta$ , we have

$$\begin{aligned} \hat{m}(u) &= \exp\left[-\frac{\sqrt{(\delta\beta - 1)^2 + 4\delta\beta u} - (\delta\beta - 1)}{2\delta} \tilde{\lambda}_0\right] \\ &= \exp\left[\frac{\beta \tilde{\lambda}_0^2 / 2\delta}{\beta \tilde{\lambda}_0 / (\delta\beta - 1)} \left(1 - \sqrt{1 + 2 \frac{(\beta \tilde{\lambda}_0 / (2\delta - 1))^2}{\beta \tilde{\lambda}_0^2 / 2\delta} u}\right)\right], \end{aligned}$$

TABLE 1: The probability  $P\{\tilde{N}_\infty = k\}$  for  $k = 0, 1, 2, \dots$  and  $(\delta; \alpha, \beta) = (2.0; 2.0, 1.5)$ .

$k$	$P\{\tilde{N}_\infty = k\}$ (%)	$k$	$P\{\tilde{N}_\infty = k\}$ (%)
0	80.000	13	0.012
1	12.000	14	0.008
2	4.050	15	0.006
3	1.789	16	0.004
4	0.904	17	0.003
5	0.496	18	0.002
6	0.287	19	0.001
7	0.172	20	0.001
8	0.106	21	0.001
9	0.067	22	0.000
10	0.043	23	0.000
11	0.028	24	0.000
12	0.019	25	0.000

which is exactly the Laplace transform of an inverse Gaussian distribution with parameters  $(\beta\tilde{\lambda}_0/(\delta\beta - 1), \beta\tilde{\lambda}_0^2/2\delta)$ .

**Corollary 4.3.** *For a cluster generated by a point of generation 0, we have*

$$E[\theta^{\tilde{N}_\infty}] = \frac{2\delta\alpha}{\delta(2\alpha - \beta) + 1 + \sqrt{(\delta\beta - 1)^2 + 4\delta\beta(1 - \theta)}}; \tag{4.9}$$

for a cluster generated by a point of subsequent generations, we have

$$E[\theta^{\tilde{N}_\infty}] = \frac{2\delta\beta}{1 + \sqrt{1 - 4\delta\beta\theta/(\delta\beta + 1)^2}} \tag{4.10}$$

and

$$P\{\tilde{N}_\infty = k\} = \frac{(\delta\beta)^{k+1}}{(\delta\beta + 1)^{2k}} \frac{(2k)!}{k!(k+1)!}, \quad k = 0, 1, \dots \tag{4.11}$$

*Proof.* By substituting the explicit exponential distribution functions of (4.2) and the constant  $v^*$  of (4.5) into Theorem 3.5, we obtain (4.9). Setting  $\alpha = \beta$  in (4.6) and expanding explicitly, we obtain (4.10) and (4.11).

**Remark 4.8.** We can also expand (4.6) explicitly for some other special cases. For instance, if  $2\delta\alpha + (1 - \delta\beta) = 0$ , we have

$$P\{\tilde{N}_\infty = k\} = \frac{\delta\beta - 1}{2\sqrt{\delta\beta}} \frac{(2k)!}{(k! 2^k)^2} \left[ \frac{(\delta\beta + 1)^2}{4\delta\beta} \right]^{-(k+1/2)}, \quad k = 0, 1, \dots$$

For the general case, we can expand (4.9) with respect to  $\theta$  using the Taylor expansion function in MATLAB<sup>®</sup>. An example with  $(\delta; \alpha, \beta) = (2.0; 2.0, 1.5)$  for  $P\{\tilde{N}_\infty = k\}$  is given in Table 1.

### 5. An application in credit risk

Our motivation for applying the dynamic contagion process to model the credit risk is the works of Duffie and Singleton (1999) and Lando (1998). Duffie and Singleton (1999) introduced the affine processes to model the default intensity. Lando (1998), the extension of Jarrow *et al.* (1997), used the state of credit ratings as an indicator of the likelihood of default, and

modelled the underlying credit rating migration driven by a probability transition matrix with Cox processes in a finite-state Markov process framework. However, we go beyond this and model the bad events that can possibly lead to credit default, and the number and the intensity of these events are modelled by the dynamic contagion process.

Based on this idea, we proceed with the following modification of the intensity models. We assume that the final default or bankruptcy is caused by a number of bad events relating to the underlying company. The bad events are not only restricted to the credit rating downgrades announced by rating agencies, but could be other bad news relevant to this company, such as bad corporate financial reports. The frequency of these bad events is dependent both on the common bad news in the market exogenously and the company’s bad events endogenously. Each company has a certain level of capability or resistance to overcome some of its bad events to avoid bankruptcy, for example, if we use the credit rating system as the indicator to quantify this level, usually the higher-rated companies have a higher capability level. We provide an application in credit risk for this idea by using the dynamic contagion process, based on the explicit results obtained in Section 4 for the case of exponential jumps.

The point process  $N_t$  models the number of bad events released from the underlying company. It is driven by a series of bad events  $\{Z_j\}_{j=1,2,\dots}$  from itself and the common bad events  $\{Y_i\}_{i=1,2,\dots}$  widely in the whole market via its intensity process  $\lambda_t$ . The impact of each event decays exponentially with constant rate  $\delta$ . We assume that each jump, or bad event, can result to default with a constant probability  $d$ ,  $0 < d \leq 1$ , which measures and quantifies the resistance level. Therefore, the survival probability, conditional on the (initial) current intensity  $\lambda_0$  at time  $T$ , is  $P_s(T) = E[(1 - d)^{N_T} | \lambda_0]$ , which can be calculated simply by letting  $\theta = 1 - d$  in the conditional probability generating function derived in Theorem 4.2. By setting the parameters  $(a, \rho, \delta; \alpha, \beta; \lambda_0) = (0.7, 0.5, 2.0; 2.0, 1.5; 0.7)$ , the term structure of the survival probabilities  $p_s(T)$  based on  $d = 2\%$ ,  $10\%$ ,  $20\%$ , and  $100\%$  are shown in Figure 2, with the corresponding numerical results given in Table 2.

As in Lando (1998), we could consider different values of  $d$ , corresponding to different credit ratings, by assuming that these bad events are all related to the company’s credit ratings.

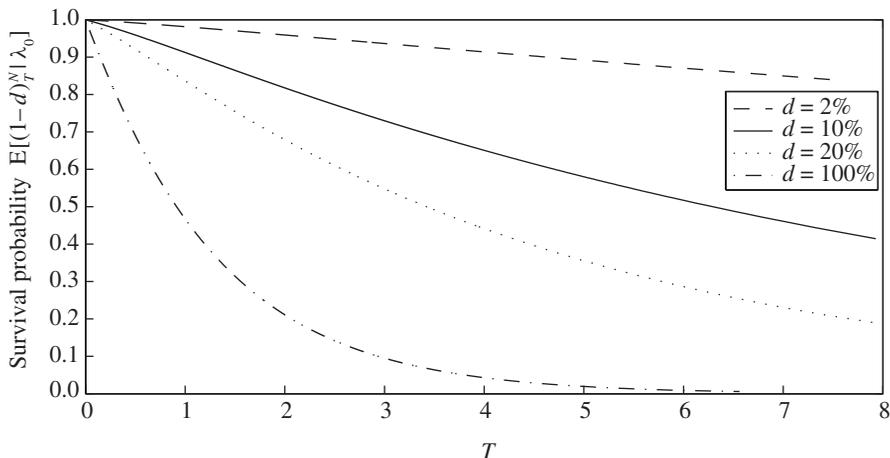


FIGURE 2: The survival probability  $P_s(T)$  for  $(a, \rho, \delta; \alpha, \beta; \lambda_0) = (0.7, 0.5, 2.0; 2.0, 1.5; 0.7)$ .

TABLE 2: The survival probability  $P_s(T)$  for  $(a, \rho, \delta; \alpha, \beta; \lambda_0) = (0.7, 0.5, 2.0; 2.0, 1.5; 0.7)$ .

$d$ (%)	Time, $T$ (%)					
	1	2	3	4	5	6
2	98.15	95.92	93.65	91.40	89.21	87.06
10	91.26	81.78	72.99	65.07	58.01	51.70
20	83.66	67.91	54.78	44.13	35.54	28.63
100	46.73	21.10	9.48	4.26	1.92	0.86

We also provide a comparison for the survival probabilities based on the three main processes discussed in this paper: the dynamic contagion process, the Hawkes process (by setting  $\rho = 0$ ), and the nonself-excited process (by setting  $\beta = \infty$ ), with the same parameter setting and fixed  $d = 10\%$ . The results are shown in Figure 3, with numerical output in Table 3.

We can see that the dynamic contagion process, as the most general case of the three processes, generates the lowest survival probability, and the differences between the other two processes explain the impact of the endogenous and exogenous factors, respectively. This process is capable of capturing more aspects of the risk, which is particularly useful for modelling the risks during the economic downturn involving clusters of bad economic events.

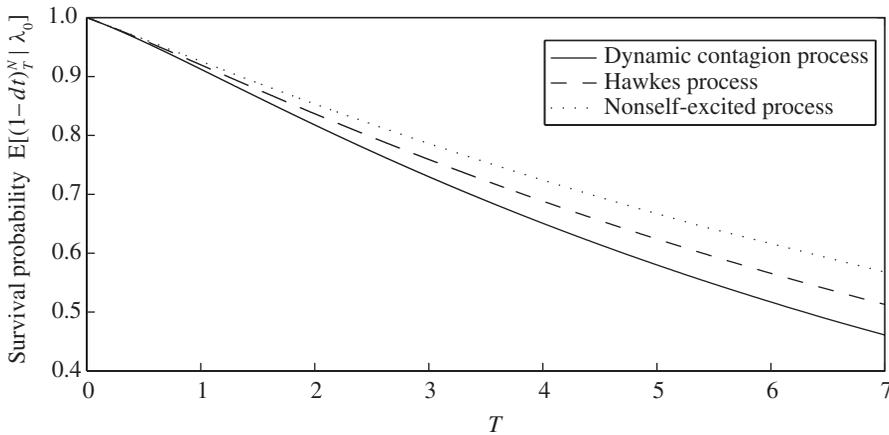


FIGURE 3: A survival probability comparison for the dynamic contagion, Hawkes, and nonself-excited processes.

TABLE 3: A survival probability comparison for the dynamic contagion, Hawkes, and nonself-excited processes.

Process	Time, $T$ (%)					
	1	2	3	4	5	6
Dynamic contagion	91.26	81.78	72.99	65.07	58.01	51.70
Hawkes	91.99	83.68	75.92	68.84	62.40	56.57
Non-self-excited	92.59	85.34	78.62	72.41	66.70	61.72

To make it easier for further industrial applications and statistical analysis, we also provide the following simulation algorithm for one sample path of the general dynamic contagion process  $(N_t, \lambda_t)$ , with  $m$  jump times  $\{T_1^*, T_2^*, \dots, T_m^*\}$  in the process  $\lambda_t$ .

**Algorithm 5.1.** Set the initial conditions  $T_0^* = 0$ ,  $\lambda_{T_0^*} = \lambda_0 > a$ , and  $i \in \{0, 1, 2, \dots, m - 1\}$ .

1. Simulate the  $(i + 1)$ th externally excited jump waiting time  $E_{i+1}^*$  by

$$E_{i+1}^* = -\frac{1}{\rho} \ln U, \quad U \sim U[0, 1].$$

2. Simulate the  $(i + 1)$ th self-excited jump waiting time  $S_{i+1}^*$  by

$$S_{i+1}^* = \begin{cases} S_{i+1}^{*(1)} \wedge S_{i+1}^{*(2)}, & d_{i+1} > 0, \\ S_{i+1}^{*(2)}, & d_{i+1} < 0, \end{cases}$$

where

$$d_{i+1} = 1 + \frac{\delta \ln U_1}{\lambda_{T_i^*} - a}, \quad U_1 \sim U[0, 1],$$

and

$$S_{i+1}^{*(1)} = -\frac{1}{\delta} \ln d_{i+1}, \quad S_{i+1}^{*(2)} = -\frac{1}{a} \ln U_2, \quad U_2 \sim U[0, 1].$$

3. Simulate the  $(i + 1)$ th jump time  $T_{i+1}^*$  in the process  $\lambda_t$  by

$$T_{i+1}^* = T_i^* + S_{i+1}^* \wedge E_{i+1}^*.$$

4. The change at jump time  $T_{i+1}^*$  in the process  $\lambda_t$  is given by

$$\lambda_{T_{i+1}^*} = \begin{cases} \lambda_{T_{i+1}^*} + Z_{i+1}, & Z_{i+1} \sim G(z), S_{i+1}^* \wedge E_{i+1}^* = S_{i+1}^*, \\ \lambda_{T_{i+1}^*} + Y_{i+1}, & Y_{i+1} \sim H(y), S_{i+1}^* \wedge E_{i+1}^* = E_{i+1}^*, \end{cases}$$

where

$$\lambda_{T_{i+1}^*} = (\lambda_{T_i^*} - a) \exp[-\delta(T_{i+1}^* - T_i^*)] + a.$$

5. The change at jump time  $T_{i+1}^*$  in the point process  $N_t$  is given by

$$N_{T_{i+1}^*} = \begin{cases} N_{T_{i+1}^*} + 1, & S_{i+1}^* \wedge E_{i+1}^* = S_{i+1}^*, \\ N_{T_{i+1}^*}, & S_{i+1}^* \wedge E_{i+1}^* = E_{i+1}^*. \end{cases}$$

Note that this simulation algorithm applies to the general distribution assumption for jump sizes  $H(y)$  and  $G(z)$  for externally and self-excited jumps, respectively.

By using the same parameter setting under the exponential distribution assumption for the jump sizes, we can regenerate the survival probabilities  $P_s(T)$  in Table 4 based on 10 000 simulated sample paths (truncated at time  $T$ ), which are very close to the analytical results in Table 2. For instance, one simulated sample path  $(N_t, \lambda_t)$  with  $T = 50$  is provided in Figure 4. For comparison, the theoretical expectations  $E[\lambda_t]$ ,  $E[\lambda_t | \lambda_0]$ , and  $E[N_t]$  (derived from Corollary 3.5, Theorem 3.6, and Theorem 3.8, respectively) are also plotted.

TABLE 4: The survival probability  $P_s(T)$  by 10 000 simulated sample paths.

$d$ (%)	Time, $T$ (%)					
	1	2	3	4	5	6
2	98.13	95.89	93.60	91.46	89.18	87.04
10	91.18	81.71	72.97	65.24	58.00	51.67
20	83.65	67.85	54.83	43.85	35.26	28.81
100	46.66	21.68	9.98	4.39	1.77	0.84

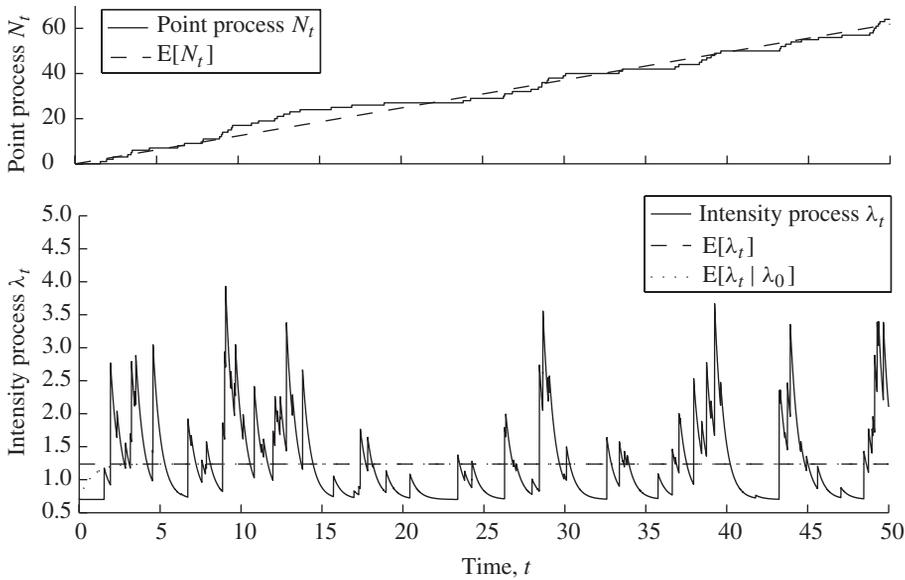


FIGURE 4: One simulated sample path of the dynamic contagion process  $(N_t, \lambda_t)$  against theoretical expectations. The parameter values are  $(a, \rho, \delta; \alpha, \beta; \lambda_0) = (0.7, 0.5, 2.0; 2.0, 1.5; 0.7)$ .

### 6. Conclusion

In this paper we have provided a general mathematical framework for modelling the dependence structure of arriving events with contagion dynamics, mainly based on generalising the Hawkes process (with exponential decay) and the Cox process with shot noise intensity (with exponential decay). The dynamic contagion process newly introduced here has been systemically studied by analysing its various distributional properties, and has the significant potential of being applicable to a variety of problems in economics, finance, and insurance. Here, we only look at one possible implementation in credit risk. However, other applications such as managing portfolio credit risk and pricing credit derivatives could be the object of further research work.

### Acknowledgement

We would like to thank the anonymous referee for pointing out the cluster process formulation, and for several other useful comments and suggestions.

## References

- AZIZPOUR, S. AND GIESECKE, K. (2008). Self-exciting corporate defaults: contagion vs. frailty. Working paper, Stanford University.
- BORDENAVE, C. AND TORRISI, G. L. (2007). Large deviations of Poisson cluster processes. *Stoch. Models* **23**, 593–625.
- BRÉMAUD, P. AND MASSOULIÉ, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Prob.* **24**, 1563–1588.
- BRÉMAUD, P. AND MASSOULIÉ, L. (2002). Power spectra of general shot noises and Hawkes processes with a random excitation. *Adv. Appl. Prob.* **34**, 205–222.
- CHAVEZ-DEMOULIN, V., DAVISON, A. C. AND McNEIL, A. J. (2005). Estimating value-at-risk: a point process approach. *Quant. Finance* **5**, 227–234.
- COSTA, O. L. V. (1990). Stationary distributions for piecewise-deterministic Markov processes. *J. Appl. Prob.* **27**, 60–73.
- DALEY, D. J. AND VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes*, Vol. 1, 2nd edn. Springer, New York.
- DASSIOS, A. AND JANG, J.-W. (2003). Pricing of catastrophe reinsurance and derivatives using the Cox process with shot noise intensity. *Finance Stoch.* **7**, 73–95.
- DASSIOS, A. AND JANG, H.-W. (2005). Kalman-Bucy filtering for linear systems driven by the Cox process with shot noise intensity and its application to the pricing of reinsurance contracts. *J. Appl. Prob.* **42**, 93–107.
- DAVIS, M. H. A. (1984). Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *J. R. Statist. Soc. B* **46**, 353–388.
- DAVIS, M. H. A. (1993). *Markov Models and Optimization*. Chapman and Hall, London.
- DUFFIE, D. AND GÂRLEANU, N. (2001). Risk and valuation of collateralized debt obligations. *Financial Analysts J.* **57**, 41–59.
- DUFFIE, D. AND SINGLETON, K. J. (1999). Modeling term structures of defaultable bonds. *Rev. Financial Studies* **12**, 687–720.
- DUFFIE, D., FILIPOVIĆ, D. AND SCHACHERMAYER, W. (2003). Affine processes and applications in finance. *Ann. Appl. Prob.* **13**, 984–1053.
- ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes*. John Wiley, New York.
- ERRAIS, E., GIESECKE, K. AND GOLDBERG, L. (2009). Affine point processes and portfolio credit risk. Working paper, Stanford University.
- HAWKES, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58**, 83–90.
- HAWKES, A. G. AND OAKES, D. (1974). A cluster representation of a self-exciting process. *J. Appl. Prob.* **11**, 493–503.
- JARROW, R. A. AND YU, F. (2001). Counterparty risk and the pricing of defaultable securities. *J. Finance* **56**, 1765–1799.
- JARROW, R. A., LANDO, D. AND TURNBULL, S. M. (1997). A Markov model for the term structure of credit risk spreads. *Rev. Financial Stud.* **10**, 481–523.
- LANDO, D. (1998). On Cox processes and credit risky securities. *Derivatives Res.* **2**, 99–120.
- LONGSTAFF, F. A. AND RAJAN, A. (2008). An empirical analysis of the pricing of collateralized debt obligations. *J. Finance* **63**, 529–563.
- MASSOULIÉ, L. (1998). Stability results for a general class of interacting point processes dynamics, and applications. *Stoch. Process. Appl.* **75**, 1–30.
- OAKES, D. (1975). The Markovian self-exciting process. *J. Appl. Prob.* **12**, 69–77.
- STABILE, G. AND TORRISI, G. L. (2010). Risk processes with non-stationary Hawkes claims arrivals. *Methodology Comput. Appl. Prob.* **12**, 415–429.