# A Dynamic Contagion Process*

Angelos Dassios[†], Hongbiao Zhao[‡]

[†‡] *Department of Statistics, London School of Economics, Houghton Street, London WC2A 2AE, UK*
[†] A.Dassios@lse.ac.uk, [‡] H.Zhao1@lse.ac.uk

August 22, 2011

**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 by Davis (1984), and the extension of the martingale methodology used by 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 the 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.

**Mathematics Subject Classification (2010)** Primary: 60J75; Secondary: 60G55 · 60G44 · 91G40

**JEL Classification** C16

# 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 to 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 one Errais, Giesecke and Goldberg (2009). Jarrow and Yu (2001) pointed out that, a model with the default intensity only depending 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, Giesecke and Goldberg (2009) found that, by using the self-excited Hawkes process, originally introduced by Hawkes (1971) (see also Hawkes and Oakes (1974), Oakes (1975) ), the clustering of defaults observed from real financial data 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

*An edited version to be published by the journal *Advances in Applied Probability* 43(3), ISSN 0001-8678, September 2011

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 by 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 by Jarrow and Yu (2001), to have a richer set of parameters, capable to capture 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 by Davis (1984) (and see also Davis (1993)), and then extend the martingale methodology introduced by Dassios and Jang (2003) (and 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. Section 3 as the main section, analyses and derives 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, an application to credit risk and the algorithm for simulating the process in Section 5. Section 6 concludes this paper and suggests 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) (see also Hawkes and Oakes (1974)) gives 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* $\{D_m\}_{m=1,2,\ldots} \in (0,\infty)$ *and shot noise stochastic intensity process*

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

*where*

- $a \geq 0$ *is the constant reversion level,*

- $\lambda_0 > 0$ *is a constant as the initial value of the stochastic intensity process (defined later by (1)),*
- $\delta > 0$ *is the constant rate of exponential decay,*
- $\{Y_i\}_{i=1,2,...}$ *is a sequence of independent identical distributed positive (externally excited) jumps with distribution function* $H(y), y > 0$, *at the corresponding random times* $\left\{T_i^{(1)}\right\}_{i=1,2,...}$ *following a homogeneous Poisson process* $M_t$ *with constant intensity* $\rho > 0$,
- $\mathbb{I}$ *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* $0, 1, 2, ...$ *with the following branching structure:*

*the immigrant* $D_m$ *is said to be of generation* $0$. *Given generations* $0, 1, ..., j$ *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* $Ze^{-\delta\left(\cdot - T^{(2)}\right)}$ *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* $0, 1, ..., j$.

**(c)** *Given the immigrants, the centered clusters*

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

*are independent identical distributed, and independent of* $A$.

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

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

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

$$\lambda_t = a + (\lambda_0 - a)\, e^{-\delta t} + \sum_{i\geq 1} Y_i e^{-\delta(t-T_i^{(1)})}\mathbb{I}\left\{T_i^{(1)} \leq t\right\} + \sum_{k\geq 1} Z_k e^{-\delta(t-T_k^{(2)})}\mathbb{I}\left\{T_k^{(2)} \leq t\right\}, \tag{1}$$

*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,*
- $\{Z_k\}_{k=1,2,...}$ *is a sequence of independent identical distributed positive (self-excited) jumps with distribution function* $G(z), z > 0$, *at the corresponding random times* $\left\{T_k^{(2)}\right\}_{k=1,2,...}$,
- *the sequences* $\{Y_i\}_{i=1,2,...}$, $\left\{T_i^{(1)}\right\}_{i=1,2,...}$ *and* $\{Z_k\}_{k=1,2,...}$ *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

$$\begin{aligned}\mathcal{A}f(\lambda, n, t) \quad = \quad & \frac{\partial f}{\partial t} - \delta(\lambda - a)\frac{\partial f}{\partial \lambda} + \rho\left(\int_0^\infty f(\lambda + y, n, t)\mathrm{d}H(y) - f(\lambda, n, t)\right) \\ & +\lambda\left(\int_0^\infty f(\lambda + z, n+1, t)\mathrm{d}G(z) - f(\lambda, n, t)\right),\end{aligned} \tag{2}$$

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

$$\begin{aligned} \left|\int_0^\infty f(\lambda+y,n,t)\mathrm{d}H(y) - f(\lambda,n,t)\right| &< \infty, \\ \left|\int_0^\infty f(\lambda+z,n+1,t)\mathrm{d}G(z) - f(\lambda,n,t)\right| &< \infty. \end{aligned}$$

**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 by Duffie, Filipović and Schachermayer (2003), with the infinitesimal generator specified by (2).

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

$$\begin{aligned} P\left\{N_{t+\Delta t} - N_t = 1 \middle| N_t\right\} &= \lambda_t \Delta t + o(\Delta t), \\ P\left\{N_{t+\Delta t} - N_t > 1 \middle| N_t\right\} &= o(\Delta t), \end{aligned}$$

where $\Delta t$ is a sufficient small time interval and $\lambda_t$ is given by (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 is the following: 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 the generation 0 of the dynamic contagion process, i.e. the point process $A$ (as described by **(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 branching structure (as described by **(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* for illustrating how the externally excited jumps $\{Y_i\}_{i=1,2,\dots}$ (marked by single arrow $\downarrow$) and self-excited jumps $\{Z_k\}_{k=1,2,\dots}$ (marked by double arrow $\updownarrow$) in the intensity process $\lambda_t$ interact with its dynamic contagion point process $N_t$.

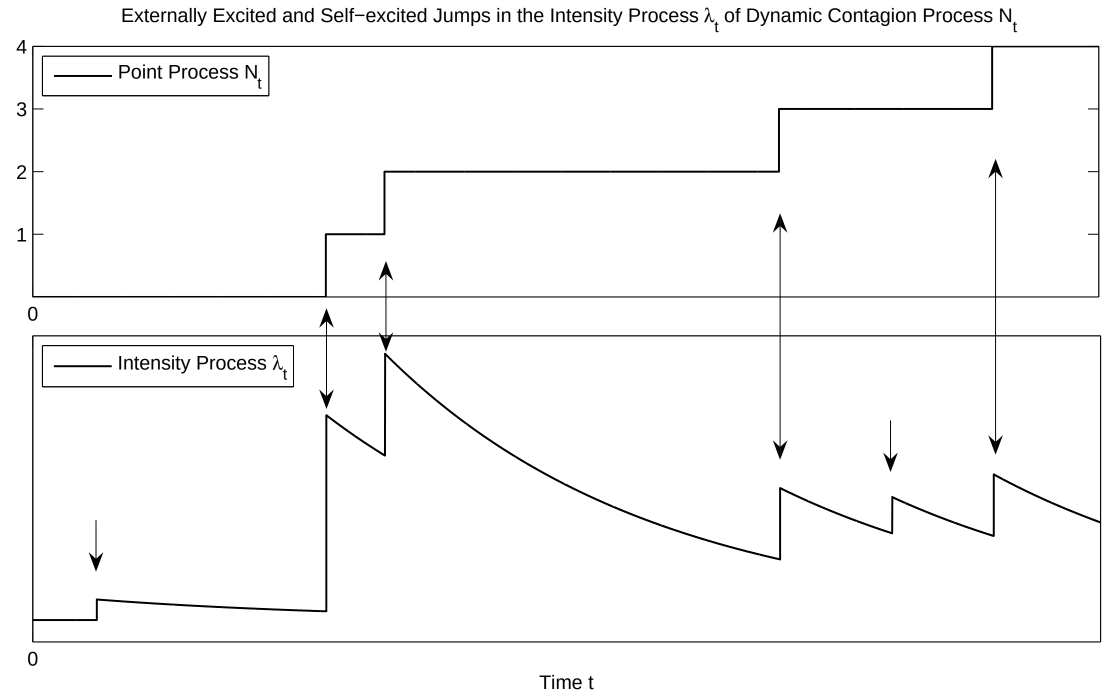


**Figure 1:** Externally Excited and Self-excited Jumps in Intensity Process $\lambda_t$ of Dynamic Contagion Process $N_t$

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

# 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, 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*

$$\mathbb{E}\left[\theta^{(N_T - N_t)} e^{-v\lambda_T} \middle| \mathcal{F}_t\right] = e^{-\big(c(T)-c(t)\big)} e^{-B(t)\lambda_t}, \tag{3}$$

*where $B(t)$ is determined by the non-linear ODE*

$$-B'(t) + \delta B(t) + \theta \hat{g}\big(B(t)\big) - 1 = 0, \tag{4}$$

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

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

$$c(t) = a\delta \int_0^t B(s)\mathrm{d}s + \rho \int_0^t \Big[1 - \hat{h}\big(B(s)\big)\Big]\,\mathrm{d}s, \tag{5}$$

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

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

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

substitute into $\mathcal{A}f = 0$ in (2); we then have

$$\frac{A'(t)}{A(t)}n + \Big( - B'(t) + \delta B(t) + A(t)\hat{g}(B(t)) - 1\Big)\lambda + \Big(c'(t) + \rho\hat{h}(B(t)) - \rho - a\delta B(t)\Big) = 0. \qquad (6)$$

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

$$\begin{cases} \frac{A'(t)}{A(t)} = 0 & (7.1) \\ -B'(t) + \delta B(t) + A(t)\hat{g}(B(t)) - 1 = 0 & (7.2) \\ c'(t) + \rho\hat{h}(B(t)) - \rho - a\delta B(t) = 0 & (7.3) \end{cases}. \qquad (7)$$

We have $A(t) = \theta$ immediately from (7.1); and substitute into (7.2) by adding the boundary condition $B(T) = v$, we have the ODE as (4); then, by (7.3) with boundary condition $c(0) = 0$, the integration as (5) follows. Since $e^{c(t)}\theta^{N_t}e^{-B(t)\lambda_t}$ is a $\mathcal{F}-$martingale by the property of the infinitesimal generator, we have

$$\mathbb{E}\left[e^{c(T)}\theta^{N_T}e^{-B(T)\lambda_T}\middle|\mathcal{F}_t\right] = e^{c(t)}\theta^{N_t}e^{-B(t)\lambda_t}. \qquad (8)$$

Then, by the boundary condition $B(T) = v$, (3) 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 $\delta > \mu_{1_G}$, is given by*

$$\mathbb{E}\left[e^{-v\lambda_T}\middle|\lambda_0\right] = \exp\left(-\int_{\mathcal{G}^{-1}_{v,1}(T)}^{v} \frac{a\delta u + \rho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1}\mathrm{d}u\right) \times \exp\left(-\mathcal{G}^{-1}_{v,1}(T)\lambda_0\right), \qquad (9)$$

*where*

$$\mu_{1_G} =: \int_0^\infty z\mathrm{d}G(z),$$

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

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

$$\mathbb{E}\left[e^{-v\lambda_T}\middle|\mathcal{F}_0\right] = e^{-c(T)}e^{-B(0)\lambda_0}, \qquad (10)$$

where $B(0)$ is uniquely determined by the non-linear ODE

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

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

1. Set $B(t) = L(T - t)$ and $\tau = T - t$, it is equivalent to the initial value problem

$$\frac{\mathrm{d}L(\tau)}{\mathrm{d}\tau} = 1 - \delta L(\tau) - \hat{g}(L(\tau)) =: f_1(L), \qquad (11)$$

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

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

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

$$\frac{\partial f_1(L)}{\partial L} = \int_0^{\infty} y e^{-Lz} \mathrm{d}G(z) - \delta \leq \int_0^{\infty} z \mathrm{d}G(z) - \delta = \mu_{1_G} - \delta < 0, \quad \text{for } L \geq 0,$$

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

3. Rewrite (11) as

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

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

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

where $L \geq 0$, we define the function on left hand side as

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

then,

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

obviously $L \to v$ when $\tau \to 0$; by convergence test,

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

and we know that $\int_0^v \frac{1}{u} \mathrm{d}u = \infty$, then,

$$\int_0^v \frac{\mathrm{d}u}{\delta u + \hat{g}(u) - 1} = \infty,$$

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

$$c(T) = a\delta \int_0^T \mathcal{G}_{v,1}^{-1}(\tau) \mathrm{d}\tau + \rho \int_0^T \left[1 - \hat{h}\left(\mathcal{G}_{v,1}^{-1}(\tau)\right)\right] \mathrm{d}\tau, \tag{12}$$

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

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

similarly,

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

Finally, substitute $B(0)$ and $c(T)$ into (10), and *Theorem 3.2* follows. $\square$

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

$$\lim_{T\to\infty} \mathbb{E}\left[e^{-v\lambda_T}\middle|\lambda_0\right] = \exp\left(-\int_0^v \frac{a\delta u + \rho[1-\hat{h}(u)]}{\delta u + \hat{g}(u) - 1}\mathrm{d}u\right), \tag{13}$$

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

*Proof.* Let $T \to \infty$ in *Theorem 3.2*, then $\mathcal{G}_{v,1}^{-1}(T) \to 0$ and the Laplace transform of the asymptotic distribution follows immediately as given by (13).

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

$$\int_E \mathcal{A}f(\lambda)\Pi(\lambda)\mathrm{d}\lambda = 0, \tag{14}$$

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.

$$\mathcal{A}f(\lambda) = -\delta\,(\lambda - a)\frac{\mathrm{d}f(\lambda)}{\mathrm{d}\lambda} + \rho\left(\int_0^\infty f(\lambda+y)\mathrm{d}H(y) - f(\lambda)\right) + \lambda\left(\int_0^\infty f(\lambda+z)\mathrm{d}G(z) - f(\lambda)\right), \tag{15}$$

and $\Pi(\lambda)$ is the density function of $\lambda$ with the Laplace transform given by (13).

We will now try to solve equation (14). For the first term of (14), we have

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

or,

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

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

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

For the second term of (14), by change variable $\lambda + y = s$ $(y \leq s)$ in the double integral,

$$\begin{aligned}
& \int_E\left[\rho\int_0^\infty f(\lambda+y)\mathrm{d}H(y)\right]\Pi(\lambda)\mathrm{d}\lambda \\
= & \ \rho\int_{\lambda=a}^\infty \Pi(\lambda)\int_{y=0}^\infty f(\lambda+y)\mathrm{d}H(y)\mathrm{d}\lambda = \rho\int_{s=a}^\infty f(s)\int_{y=0}^s \Pi(s-y)\mathrm{d}H(y)\mathrm{d}s,
\end{aligned}$$

or,

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

For the third term of (14), by change variable $\lambda + z = s$ $(z \leq s)$ in the double integral,

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

or,

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

Therefore,

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

Set

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

for any function $f(\lambda)\in\Omega(\mathcal{A})$, then,

$$\delta\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big((\lambda-a)\Pi(\lambda)\Big)+\rho\left(\int_0^\lambda \Pi(\lambda-y)\mathrm{d}H(y)-\Pi(\lambda)\right)+\left(\int_0^\lambda(\lambda-z)\Pi(\lambda-z)\mathrm{d}G(z)-\lambda\Pi(\lambda)\right)=0,$$

by the Laplace transform

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

we have

$$\begin{aligned}\mathcal{L}\left\{\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big((\lambda-a)\Pi(\lambda)\Big)\right\} &= v\mathcal{L}\{(\lambda-a)\Pi(\lambda)\} = v\left(-\frac{\mathrm{d}\hat{\Pi}(v)}{\mathrm{d}v}-a\hat{\Pi}(v)\right),\\ \mathcal{L}\left\{\int_0^\lambda \Pi(\lambda-y)\mathrm{d}H(y)\right\} &= \mathcal{L}\left\{\int_0^\lambda \Pi(\lambda-y)h(y)\mathrm{d}y\right\} = \hat{\Pi}(v)\hat{h}(v),\\ \mathcal{L}\left\{\int_0^\lambda (\lambda-z)\Pi(\lambda-z)\mathrm{d}G(z)\right\} &= \mathcal{L}\left\{\int_0^\lambda (\lambda-z)\Pi(\lambda-z)g(z)\mathrm{d}z\right\}\\ &= \mathcal{L}\{\lambda\Pi(\lambda)\}\hat{g}(v) = -\frac{\mathrm{d}\hat{\Pi}(v)}{\mathrm{d}v}\hat{g}(v),\end{aligned}$$

then,

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

or,

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

which is an ODE with the solution given by

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

Note that, given the initial condition

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

we have the unique solution

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

which is exactly given by (13).

Since $\Pi$ is the unique solution to (14), we have the stationarity for the intensity process $\{\lambda_t\}_{t\geq 0}$. $\square$

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 by Hawkes (1971), and applied to risk theory by Chavez-Demoulin, Davison and Mc Neil (2005), and then only very recently applied to credit risk for modelling the default contagion by Errais, Giesecke and Goldberg (2009). It can be considered as an important special case under this more general framework of 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*. The probability generating function of the Hawkes point process $N_t$ will be given by *Corollary 3.2* of Section 3.3.

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

$$\lim_{T\to\infty} \mathbb{E}\left[e^{-v\lambda_T}\middle|\lambda_0\right] = \exp\left(-a\delta\int_0^v \frac{u}{\delta u + \hat{g}(u) - 1}\mathrm{d}u\right), \tag{16}$$

*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. $\square$

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 $\delta > \mu_{1_G}$, is given by*

$$\mathbb{E}\big[\theta^{N_T}\big|\lambda_0\big] = \exp\left(-\int_0^{\mathcal{G}^{-1}_{0,\theta}(T)} \frac{a\delta u + \rho[1-\hat{h}(u)]}{1 - \delta u - \theta\hat{g}(u)}\mathrm{d}u\right) \times \exp\Big(-\mathcal{G}^{-1}_{0,\theta}(T)\lambda_0\Big),$$

*where*

$$\mathcal{G}_{0,\theta}(L) =: \int_0^L \frac{\mathrm{d}u}{1 - \delta u - \theta\hat{g}(u)}, \qquad 0 \leq \theta < 1. \tag{17}$$

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

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

where $B(0)$ is uniquely determined by the non-linear ODE

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

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

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

$$\frac{\mathrm{d}L(\tau)}{\mathrm{d}\tau} = 1 - \delta L(\tau) - \theta \hat{g}(L(\tau)) =: f_2(L), \qquad 0 \le \theta < 1, \tag{18}$$

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

2. 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 $0 < \theta < 1$, 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 $\theta = 0$, there is only one singular point $v^* = \frac{1}{\delta} > 0$; and for both cases,

$$v^* = \frac{1}{\delta}\Big(1 - \theta \hat{g}(v^*)\Big) \ge \frac{1-\theta}{\delta} > 0; \tag{19}$$

then, we have $f_2(L) > 0$ for $0 \le L < v^*$ and $f_2(L) < 0$ for $L > v^*$.

3. Rewrite (18) as

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

and integrate,

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

where $0 \le L < v^*$, we define the function on left-hand side as

$$\mathcal{G}_{0,\theta}(L) =: \int_0^L \frac{\mathrm{d}u}{1 - \delta u - \theta \hat{g}(u)} \tag{20}$$

then,

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

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

4. The unique solution is found by

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

5. $B(0)$ is obtained,

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

Then, $c(T)$ is determined by

$$c(T) = a\delta \int_0^T \mathcal{G}_{0,\theta}^{-1}(\tau)\mathrm{d}\tau + \rho \int_0^T \left[1 - \hat{h}\left(\mathcal{G}_{0,\theta}^{-1}(\tau)\right)\right] \mathrm{d}\tau, \tag{21}$$

where, by the change of variable,

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

Finally, substitute $B(0)$ and $c(T)$ into (10), and the result follows. □

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

$$\mathbb{E}\big[\theta^{N_T}\big|\lambda_0\big] = \exp\left(-a\delta\int_0^{\mathcal{G}_{0,\theta}^{-1}(T)}\frac{u}{1-\delta u-\theta\hat{g}(u)}\mathrm{d}u\right)\times e^{-\mathcal{G}_{0,\theta}^{-1}(T)\lambda_0}.$$

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

The probability $P\{N_T = 0\big|\lambda_0\}$ 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 $\delta > \mu_{1_G}$, is given by*

$$P\left\{N_T = 0\big|\lambda_0\right\} = \exp\left(-\int_0^{u_T}\frac{a\delta u+\rho[1-\hat{h}(u)]}{1-\delta u}\mathrm{d}u\right)\times e^{-u_T\lambda_0}, \tag{22}$$

*where*

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

*Proof.* Since

$$P\left\{N_T = 0\big|\lambda_0\right\} = \mathbb{E}\left[\theta^{N_T}\big|\lambda_0\right]\Bigg|_{\theta=0},$$

and

$$\mathcal{G}_{0,0}(L) =: \mathcal{G}_{0,\theta}(L)\Bigg|_{\theta=0} = \int_0^L\frac{1}{1-\delta u}\mathrm{d}u = -\frac{1}{\delta}\ln\left(1-\delta L\right),$$

then, the inverse function

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

by letting $\theta = 0$ in *Theorem 3.4*, (22) 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 $P\left\{N_T = 0\big|\lambda_0\right\}$ is not dependent on the distribution of the self-excited jumps, and the result is similar to the non-self-excited case by Dassios and Jang (2003).

Theoretically, the probability $P\left\{N_T = n\big|\lambda_0\right\}$ for any natural number $n \in \mathbb{N}$ can be derived by

$$P\left\{N_T = n\big|\lambda_0\right\} = \frac{\partial^n}{\partial\theta^n}\mathbb{E}\left[\theta^{N_T}\big|\lambda_0\right],$$

here, we derive the result of $P\left\{N_T = 1\big|\lambda_0\right\}$ in *Corollary 3.4*, for instance.

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

$$\begin{aligned}P\left\{N_T = 1\big|\lambda_0\right\} \;=\; & P\left\{N_T = 0\big|\lambda_0\right\}\times\Bigg\{\Big[a\left(1-e^{-\delta T}\right)+\rho[1-\hat{h}(u_T)]+\lambda_0 e^{-\delta T}\Big] \\ & \times\int_0^{u_T}\frac{\hat{g}(u)}{(1-\delta u)^2}\mathrm{d}u-\int_0^{u_T}\frac{\hat{g}(u)}{(1-\delta u)^2}\Big(a\delta u+\rho[1-\hat{h}(u)]\Big)\mathrm{d}u\Bigg\},\end{aligned}$$

*where*

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

*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,

$$\begin{aligned}
P\left\{N_T = 1\middle|\lambda_0\right\} &= \frac{\partial}{\partial\theta}\exp\left[-\int_0^{\mathcal{G}_{0,\theta}^{-1}(T)}\varphi(u,\theta)\mathrm{d}u - \mathcal{G}_{0,\theta}^{-1}(T)\lambda_0\right]\Bigg|_{\theta=0}\\
&= P\left\{N_T = 0\middle|\lambda_0\right\}\times(-1)\left[\int_0^{\mathcal{G}_{0,\theta}^{-1}(T)}\frac{\partial\varphi(u,\theta)}{\partial\theta}\mathrm{d}u + \left(\varphi\left(\mathcal{G}_{0,\theta}^{-1}(T),\theta\right)+\lambda_0\right)\frac{\partial}{\partial\theta}\mathcal{G}_{0,\theta}^{-1}(T)\right]\Bigg|_{\theta=0}\\
&= P\left\{N_T = 0\middle|\lambda_0\right\}\times(-1)\left[\int_0^{u_T}\frac{\partial\varphi(u,\theta)}{\partial\theta}\bigg|_{\theta=0}\mathrm{d}u + \left(\varphi\left(u_T,0\right)+\lambda_0\right)\frac{\partial}{\partial\theta}\mathcal{G}_{0,\theta}^{-1}(T)\bigg|_{\theta=0}\right],
\end{aligned}$$

where

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

and $\frac{\partial}{\partial\theta}\mathcal{G}_{0,\theta}^{-1}(T)\Big|_{\theta=0}$ can be derived as below. Since $L(T;\theta) = \mathcal{G}_{0,\theta}^{-1}(T)$, we have the non-linear ODE of $L(\tau;\theta)$,

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

with the initial condition $L(0;\theta)=0$, differentiate both sides with respect to $\theta$,

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

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 $L(\tau;0)=\frac{1}{\delta}\left(1-e^{-\delta\tau}\right)$, then, $L^{(1)}(\tau;0)$ can be uniquely solved,

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

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

$$\int_0^T\hat{g}\left(\frac{1-e^{-\delta s}}{\delta}\right)e^{\delta s}\mathrm{d}s = \int_0^{u_T}\frac{\hat{g}(u)}{(1-\delta u)^2}\mathrm{d}u.$$

□

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, $\widetilde{N}_t$, under the condition $\delta>\mu_{1_G}$, we have*

$$\begin{aligned}
\mathbb{E}\left[\theta^{\widetilde{N}_T}\middle|\widetilde{\lambda}_0\right] &= e^{-\mathcal{G}_{0,\theta}^{-1}(T)\widetilde{\lambda}_0}, \qquad (23)\\
\mathbb{E}\left[\theta^{\widetilde{N}_\infty}\middle|\widetilde{\lambda}_0\right] &= e^{-v^*\widetilde{\lambda}_0},
\end{aligned}$$

*where $\mathcal{G}_{0,\theta}(\cdot)$ and $v^*$ are given by (17) and (19), respectively, and $\widetilde{\lambda}_0$ 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*

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

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

$$\mathbb{E}\big[\theta^{\widetilde{N}_\infty}\big] = \frac{1-\delta v^*}{\theta}. \tag{24}$$

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

$$\mathcal{A}f(\widetilde{\lambda}, \widetilde{n}, t) = \frac{\partial f}{\partial t} - \delta\widetilde{\lambda}\frac{\partial f}{\partial \widetilde{\lambda}} + \widetilde{\lambda}\left(\int_0^\infty f(\widetilde{\lambda}+z, \widetilde{n}+1, t)\mathrm{d}G(z) - f(\widetilde{\lambda}, \widetilde{n}, t)\right),$$

as it is just a special case of *Theorem 3.1* and *Theorem 3.4*, we can derive (23) immediately. By the proof of *Theorem 3.4*, we know that

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

then,

$$\mathbb{E}\left[\theta^{\widetilde{N}_\infty}\middle|\widetilde{\lambda}_0\right] = \lim_{T\to\infty}\mathbb{E}\left[\theta^{\widetilde{N}_T}\middle|\widetilde{\lambda}_0\right] = \lim_{T\to\infty} e^{-\mathcal{G}_{0,\theta}^{-1}(T)\widetilde{\lambda}_0} = e^{-v^*\widetilde{\lambda}_0}.$$

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

$$\mathbb{E}[\theta^{\widetilde{N}_\infty}] = \mathbb{E}\left[\mathbb{E}\left[\theta^{\widetilde{N}_\infty}\middle|\widetilde{\lambda}_0\right]\right] = \mathbb{E}[e^{-v^*\widetilde{\lambda}_0}] = \mathbb{E}[e^{-v^*Y_1}] = \hat{h}(v^*);$$

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

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

□

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

**Remark 3.3.** Alternatively, (24) can be derived from the perspective of the cluster process definition given by *Definition 2.1*, and we observe that each subcluster has the same distribution $\mathcal{E}(\theta) = \mathbb{E}\big[\theta^{\widetilde{N}_\infty}\big]$ 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 (24).

We also provide an explicit example for *Theorem 3.5* in *Theorem 4.3* by assuming the jumps with the 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 zero, respectively. Alternatively, we can obtain the first and second moments of $\lambda_t$ and $N_t$ directly by solving ODEs, and also this method is slightly easier to generalise to derive higher moments beyond the condition $\delta > \mu_{1_G}$, 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*

$$\mathbb{E}\left[\lambda_t \middle| \lambda_0\right] = \frac{\mu_{1_H}\rho + a\delta}{\delta - \mu_{1_G}} + \left(\lambda_0 - \frac{\mu_{1_H}\rho + a\delta}{\delta - \mu_{1_G}}\right) e^{-(\delta - \mu_{1_G})t}, \quad \textit{for } \delta \neq \mu_{1_G}, \tag{25}$$

$$\mathbb{E}\big[\lambda_t \big| \lambda_0\big] = \lambda_0 + \left(\mu_{1_H}\rho + a\delta\right) t, \quad \textit{for } \delta = \mu_{1_G}, \tag{26}$$

*where*

$$\mu_{1_H} =: \int_0^\infty y \mathrm{d}H(y).$$

*Proof.* By the martingale property of the infinitesimal generator as given in (2), we have a $\mathcal{F}-$martingale

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

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

$$\mathcal{A}\lambda = -(\delta - \mu_{1_G})\lambda + \mu_{1_H}\rho + a\delta,$$

then, $\lambda_t - \lambda_0 - \int_0^t \mathcal{A}\lambda_s \mathrm{d}s$ is a $\mathcal{F}-$martingale, and we have

$$\mathbb{E}\left[\lambda_t - \int_0^t \mathcal{A}\lambda_s \mathrm{d}s \middle| \lambda_0\right] = \lambda_0.$$

Hence,

$$\mathbb{E}\left[\lambda_t \middle| \lambda_0\right] = \lambda_0 + \mathbb{E}\left[\int_0^t \mathcal{A}\lambda_s \mathrm{d}s \middle| \lambda_0\right] = \lambda_0 - (\delta - \mu_{1_G}) \int_0^t \mathbb{E}\left[\lambda_s \middle| \lambda_0\right] \mathrm{d}s + \left(\mu_{1_H}\rho + a\delta\right) t,$$

by differentiating with respect to $t$, we obtain the non-linear inhomogeneous ODE,

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = -\left(\delta - \mu_{1_G}\right) u(t) + \mu_{1_H}\rho + a\delta,$$

where $u(t) = \mathbb{E}\left[\lambda_t \middle| \lambda_0\right]$, with the initial condition $u(0) = \lambda_0$. This ODE has the solution given by (25) and (26). ☐

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

$$\begin{aligned}
&\mathbb{E}\left[\lambda_t^2 \middle| \lambda_0\right] = \lambda_0^2 e^{-2(\delta - \mu_{1_G})t} \\
&+ \frac{2(\mu_{1_H}\rho + a\delta) + \mu_{2_G}}{\delta - \mu_{1_G}} \left(\lambda_0 - \frac{\mu_{1_H}\rho + a\delta}{\delta - \mu_{1_G}}\right) \left(e^{-(\delta - \mu_{1_G})t} - e^{-2(\delta - \mu_{1_G})t}\right) \\
&+ \left(\frac{(2(\mu_{1_H}\rho + a\delta) + \mu_{2_G})(\mu_{1_H}\rho + a\delta)}{2(\delta - \mu_{1_G})^2} + \frac{\mu_{2_H}\rho}{2(\delta - \mu_{1_G})}\right) \left(1 - e^{-2(\delta - \mu_{1_G})t}\right), \quad \textit{for } \delta \neq \mu_{1_G},
\end{aligned} \tag{27}$$

$$\mathbb{E}\big[\lambda_t^2 \big| \lambda_0\big] = \lambda_0^2 + \Big(2(\mu_{1_H}\rho + a\delta) + \mu_{2_G}\Big) \left(\lambda_0 t + \frac{1}{2}\left(\mu_{1_H}\rho + a\delta\right) t^2\right) + \mu_{2_H}\rho t, \quad \textit{for } \delta = \mu_{1_G}, \tag{28}$$

*where*

$$\mu_{2_H} =: \int_0^\infty y^2 \mathrm{d}H(y); \quad \mu_{2_G} =: \int_0^\infty z^2 \mathrm{d}G(z).$$

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

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

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

$$\mathbb{E}\left[\lambda_t^2 - \int_0^t \mathcal{A}\lambda_s^2 \mathrm{d}s \middle| \lambda_0\right] = \lambda_0^2.$$

Hence,

$$\mathbb{E}\left[\lambda_t^2\big|\lambda_0\right] = \lambda_0^2 - 2(\delta-\mu_{1_G})\int_0^t \mathbb{E}\left[\lambda_s^2\big|\lambda_0\right]\mathrm{d}s + \big(2(\mu_{1_H}\rho + a\delta) + \mu_{2_G}\big)\int_0^t \mathbb{E}\left[\lambda_s\big|\lambda_0\right]\mathrm{d}s + \mu_{2_H}\rho t,$$

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

$$\begin{aligned}\frac{\mathrm{d}u(t)}{\mathrm{d}t} + 2(\delta-\mu_{1_G})u(t) &= \Big(2(\mu_{1_H}\rho + a\delta) + \mu_{2_G}\Big)\left(\lambda_0 - \frac{\mu_{1_H}\rho + a\delta}{\delta-\mu_{1_G}}\right)e^{-(\delta-\mu_{1_G})t} \\ &\quad + \frac{\left(2(\mu_{1_H}\rho + a\delta) + \mu_{2_G}\right)(\mu_{1_H}\rho + a\delta)}{\delta-\mu_{1_G}} + \mu_{2_H}\rho,\end{aligned}$$

where $u(t) = \mathbb{E}\left[\lambda_t^2\big|\lambda_0\right]$, with the initial condition $u(0) = \lambda_0^2$. This ODE has the solution given by (27) and (28). □

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

$$\begin{aligned}\mathrm{Var}\left[\lambda_t\big|\lambda_0\right] &= \frac{1}{2(\delta-\mu_{1_G})}\left(\frac{\mu_{2_G}(\mu_{1_H}\rho + a\delta)}{\delta-\mu_{1_G}} - \mu_{2_H}\rho - 2\mu_{2_G}\lambda_0\right)e^{-2(\delta-\mu_{1_G})t} \\ &\quad + \frac{\mu_{2_G}}{\delta-\mu_{1_G}}\left(\lambda_0 - \frac{\mu_{1_H}\rho + a\delta}{\delta-\mu_{1_G}}\right)e^{-(\delta-\mu_{1_G})t} \\ &\quad + \frac{1}{2(\delta-\mu_{1_G})}\left(\mu_{2_H}\rho + \frac{\mu_{2_G}(\mu_{1_H}\rho + a\delta)}{\delta-\mu_{1_G}}\right), \quad \textit{for } \delta \neq \mu_{1_G}, \qquad (29)\\ \mathrm{Var}\left[\lambda_t\big|\lambda_0\right] &= \frac{1}{2}\mu_{2_G}\left(\mu_{1_H}\rho + a\delta\right)t^2 + \left(\mu_{2_G}\lambda_0 + \mu_{2_H}\rho\right)t, \quad \textit{for } \delta = \mu_{1_G}. \qquad (30)\end{aligned}$$

*Proof.* By $\mathrm{Var}\left[\lambda_t\big|\lambda_0\right] = \mathbb{E}\left[\lambda_t^2\big|\lambda_0\right] - \left(\mathbb{E}\left[\lambda_t\big|\lambda_0\right]\right)^2$ based on *Theorem 3.6* and *Lemma 3.1*, the result follows. □

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

$$\begin{aligned}\mathbb{E}\left[\lambda_t\right] &= \frac{\mu_{1_H}\rho + a\delta}{\delta-\mu_{1_G}}, \qquad (31)\\ \mathbb{E}\left[\lambda_t^2\right] &= \frac{(2(\mu_{1_H}\rho + a\delta) + \mu_{2_G})(\mu_{1_H}\rho + a\delta)}{2(\delta-\mu_{1_G})^2} + \frac{\mu_{2_H}\rho}{2(\delta-\mu_{1_G})}, \qquad (32)\\ \mathrm{Var}\left[\lambda_t\right] &= \frac{1}{2(\delta-\mu_{1_G})}\left(\mu_{2_H}\rho + \frac{\mu_{2_G}(\mu_{1_H}\rho + a\delta)}{\delta-\mu_{1_G}}\right).\end{aligned}$$

*Proof.* By setting time $t \to \infty$ in (25), (26), (27), (28), and (29), (30), respectively, then the results follow. □

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

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

$$\mathbb{E}\left[N_t\right] = \frac{\mu_{1_H}\rho + a\delta}{\delta-\mu_{1_G}}t. \qquad (33)$$

*Proof.* By setting $f(\lambda, n, t) = n$ in (2), we have $\mathcal{A}n = \lambda$. Since $N_t - N_0 - \int_0^t \lambda_s \mathrm{d}s$ is a martingale by the martingale property of the intensity process $\lambda_t$ of the point process $N_t$ given by the definition (1), we have

$$\mathbb{E}\left[N_t - N_0\big|\mathcal{F}_0\right] = \mathbb{E}\left[\int_0^t \lambda_s \mathrm{d}s\bigg|\mathcal{F}_0\right],$$

and also we know $\mathbb{E}\left[\lambda_t\right]$ from *Corollary 3.5*, then, by assuming $N_0 = 0$, we have

$$\mathbb{E}\left[N_t\right] = \mathbb{E}\left[N_t - N_0\right] = \int_0^t \mathbb{E}\left[\lambda_s\right]\mathrm{d}s = \frac{\mu_{1_H}\rho + a\delta}{\delta-\mu_{1_G}}t.$$

□

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

$$\mathbb{E}\left[\lambda_t N_t\right] = \bar{k}\left(1 - e^{-\left(\delta-\mu_{1_G}\right)t}\right) + \left(\frac{\mu_{1_H}\rho + a\delta}{\delta - \mu_{1_G}}\right)^2 t, \tag{34}$$

*where*

$$\bar{k} =: \frac{2\mu_{1_G}\left(\mu_{1_H}\rho + a\delta\right) + \mu_{2_H}\rho}{2\left(\delta - \mu_{1_G}\right)^2} + \frac{\mu_{2_G}\left(\mu_{1_H}\rho + a\delta\right)}{2\left(\delta - \mu_{1_G}\right)^3}. \tag{35}$$

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

$$\mathcal{A}\left(\lambda n\right) = -(\delta - \mu_{1_G})\lambda n + \left(\mu_{1_H}\rho + a\delta\right)n + \lambda^2 + \mu_{1_G}\lambda.$$

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

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = -\left(\delta - \mu_{1_G}\right)u(t) + \left(\mu_{1_H}\rho + a\delta\right)\mathbb{E}\left[N_t\right] + \mathbb{E}\left[\lambda_t^2\right] + \mu_{1_G}\mathbb{E}\left[\lambda_t\right],$$

where $u(t) = \mathbb{E}\left[\lambda_t N_t\right]$, with the initial condition $u(0) = 0$. Note that, $\mathbb{E}\left[N_t\right]$, $\mathbb{E}\left[\lambda_t^2\right]$ and $\mathbb{E}\left[\lambda_t\right]$ are already given by (33), (32) and (31), respectively, therefore, this ODE has the solution given by (34). ☐

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

$$\begin{aligned}
\mathbb{E}\left[N_t^2\right] &= \frac{2}{\delta - \mu_{1_G}}\left(e^{-\left(\delta-\mu_{1_G}\right)t} - 1\right) + 2\bar{k}t + \left(\frac{\mu_{1_H}\rho + a\delta}{\delta - \mu_{1_G}}\right)^2 t^2, \\
\mathrm{Var}\left[N_t\right] &= \frac{2}{\delta - \mu_{1_G}}\left(e^{-\left(\delta-\mu_{1_G}\right)t} - 1\right) + 2\bar{k}t,
\end{aligned}$$

*where constant $\bar{k}$ is given by (35).*

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

$$\mathbb{E}\left[N_t^2\right] = 2\int_0^t \mathbb{E}\left[\lambda_s N_s\right]\mathrm{d}s + \int_0^t \mathbb{E}\left[\lambda_s\right]\mathrm{d}s,$$

where $\mathbb{E}\left[\lambda_t N_t\right]$ and $\mathbb{E}\left[\lambda_t\right]$ are given by(34) and (31), respectively, then $\mathbb{E}\left[N_t^2\right]$ follows. Since $\mathrm{Var}\left[N_t\right] = \mathbb{E}\left[N_t^2\right] - \mathbb{E}\left[N_t\right]^2$ given $\mathbb{E}\left[N_t\right]$ in (33), $\mathrm{Var}\left[N_t\right]$ 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, Massoulié and Ridolfi (2002).

# 4 Example: Jumps with Exponential Distributions

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

$$h(y) = \alpha e^{-\alpha y}; \quad g(z) = \beta e^{-\beta z}, \quad \text{where } y, z; \alpha, \beta > 0, \tag{36}$$

the Laplace transforms have the explicit forms

$$\hat{h}(u) = \frac{\alpha}{\alpha + u}; \quad \hat{g}(u) = \frac{\beta}{\beta + u}. \tag{37}$$

Then the corresponding Laplace transform of $\lambda_T$, conditional probability generating function of $N_T$, conditional probability $P\left\{N_T = 0 \middle| \lambda_0\right\}$ and $P\left\{N_T = 1 \middle| \lambda_0\right\}$ are obtained respectively as below. 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 (36), then the conditional Laplace transform of $\lambda_T$ given $\lambda_0$ at time $t = 0$, under the condition $\delta\beta > 1$, is given by*

$$\mathbb{E}\left[e^{-v\lambda_T} \middle| \lambda_0\right] = e^{-\left(\mathcal{C}_1(v)-\mathcal{C}_1\left(\mathcal{G}_{v,1}^{-1}(T)\right)\right)} e^{-\mathcal{G}_{v,1}^{-1}(T)\lambda_0},$$

*where*

$$\mathcal{C}_1(u) =: \begin{cases} au + \frac{\rho(\alpha-\beta)}{\delta(\alpha-\beta)+1}\ln(\alpha+u) + \frac{1}{\delta}\left(a + \frac{\rho}{\delta(\alpha-\beta)+1}\right)\ln\left(u + \frac{\delta\beta-1}{\delta}\right) & \text{for } \alpha \neq \beta - \frac{1}{\delta} \\ au + \frac{\rho\beta}{\delta\beta-1}\ln(\alpha+u) - \frac{\alpha\rho}{\delta(\delta\beta-1)}\frac{1}{\alpha+u} + \frac{1}{\delta}\left(a - \frac{\rho}{\delta\beta-1}\right)\ln\left(u + \frac{\delta\beta-1}{\delta}\right) & \text{for } \alpha = \beta - \frac{1}{\delta} \end{cases}, \tag{38}$$

*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 $\mu_{1_G} = \frac{1}{\beta}$, the condition is $\delta > \frac{1}{\beta}$; and substitute (37), into *Theorem 3.2*, we have

$$\mathcal{G}_{v,1}(L) = \int_L^v \frac{u+\beta}{\delta u\left(u + \frac{\delta\beta-1}{\delta}\right)}\mathrm{d}u,$$

and

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

Note that, when calculating the integral, we need consider the special case when $\alpha = \beta - \frac{1}{\delta}$. 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 (36), then, under the condition $\delta\beta > 1$, the stationary distribution of the process $\{\lambda_t\}_{t\geq 0}$ is given by*

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

*where independent random variables*

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

*$\tilde{B}$ follows a compound negative binomial distribution with underlying exponential jumps; $\tilde{P}$ follows a compound Poisson distribution with underlying exponential jumps.*

*Proof.* By *Lemma 4.1*, *Theorem 3.3*, and as $\mathcal{G}_{v,1}^{-1}(T) \to 0$ when $T \to \infty$, we use the explicit function $C_1(u)$ in (38) to derive the Laplace transform of the stationary distribution of the process $\{\lambda_t\}_{t\geq 0}$ by $\hat{\Pi}(v) = e^{-(C_1(v)-C_1(0))}$, then,

$$\hat{\Pi}(v) = \begin{cases} e^{-va}\left(\frac{\alpha}{\alpha+v}\right)^{\frac{\rho(\alpha-\beta)}{\delta(\alpha-\beta)+1}}\left(\frac{\frac{\delta\beta-1}{\delta}}{v+\frac{\delta\beta-1}{\delta}}\right)^{\frac{1}{\delta}\left(a+\frac{\rho}{\delta(\alpha-\beta)+1}\right)} & \text{for } \alpha \geq \beta \\ e^{-va}\left(\frac{\frac{\delta\beta-1}{\delta}}{v+\frac{\delta\beta-1}{\delta}}\right)^{\frac{a+\rho}{\delta}}\left(\frac{\frac{\gamma_2}{\gamma_1}}{1-\left(1-\frac{\gamma_2}{\gamma_1}\right)\frac{\gamma_1}{\gamma_1+v}}\right)^{\frac{\rho}{\delta}\frac{\beta-\alpha}{\gamma_1-\gamma_2}} & \text{for } \alpha < \beta \text{ and } \alpha \neq \beta - \frac{1}{\delta} \\ e^{-va}\left(\frac{\alpha}{\alpha+v}\right)^{\frac{\rho+a}{\delta}}\exp\left[\frac{\rho}{\delta^2\alpha}\left(\frac{\alpha}{\alpha+v}-1\right)\right] & \text{for } \alpha = \beta - \frac{1}{\delta} \end{cases} . \tag{39}$$

If $\alpha \geq \beta$, it is obvious that, (39) 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 $\alpha \neq \beta - \frac{1}{\delta}$, then always $\gamma_1 > \gamma_2$, and the second term is the Laplace transform of Gamma distribution with two parameters $\frac{a+\rho}{\delta}$ and $\frac{\delta\beta-1}{\delta}$; the third term is the Laplace transform of a compound negative binomial distribution with two parameters $\frac{\rho}{\delta}\frac{\beta-\alpha}{\gamma_1-\gamma_2}$ and $\frac{\gamma_2}{\gamma_1}$, and the underlying jumps follows an exponential distribution with parameter $\gamma_1$, since we know that the Laplace transform of negative binomial distribution $N_1$ with two parameters $(r, p)$ is

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

Then

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

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 also easy to identify the corresponding Laplace transforms for the case when $\alpha = \beta - \frac{1}{\delta}$. □

We discuss some important special cases below.

**Remark 4.1.** If both jumps follows 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 non-self-excited case, i.e. when $\beta = \infty$, we have the Laplace transform of the stationary distribution of the process $\{\lambda_t\}_{t\geq 0}$ given by

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

then, $\{\lambda_t\}_{t\geq 0}$ follows a shifted Gamma distribution,

$$\{\lambda_t\}_{t\geq 0} \stackrel{\mathcal{D}}{=} a + \tilde{\Gamma}_5,$$

where

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

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

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

$$\hat{\Pi}(v) = e^{-va}\left(\frac{\frac{\delta\beta-1}{\delta}}{v+\frac{\delta\beta-1}{\delta}}\right)^{\frac{a}{\delta}}, \tag{40}$$

then, $\{\lambda_t\}_{t\geq 0}$ follows a shifted Gamma distribution,

$$\{\lambda_t\}_{t\geq 0} \stackrel{\mathcal{D}}{=} a + \tilde{\Gamma}_6,$$

where

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

The result for the particular case $\alpha = \beta - \frac{1}{\delta}$ is actually the limit version of the result for the case when $\alpha < \beta$ and $\alpha \neq \beta - \frac{1}{\delta}$. In the following sections, we only focus on the main case when $\alpha \neq \beta - \frac{1}{\delta}$, with the Laplace transform of the stationary distribution of the process $\{\lambda_t\}_{t\geq 0}$ specified by (39).

## 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 as (36), 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*

$$\mathbb{E}\big[\theta^{N_T}\big|\lambda_0\big] = e^{-\left(\mathcal{C}_2\left(\mathcal{G}_{0,\theta}^{-1}(T)\right)-\mathcal{C}_2(0)\right)}e^{-\mathcal{G}_{0,\theta}^{-1}(T)\lambda_0}, \qquad \alpha \neq -v_-^*,$$

*where*

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

*and*

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

*where*

$$K(u) =: -\frac{1}{\delta\left(v^*-v_-^*\right)}\Big[(v^*+\beta)\ln\left(v^*-u\right) - \left(v_-^*+\beta\right)\ln\left(u-v_-^*\right)\Big], \qquad 0 \leq u < v^*,$$

$$\begin{aligned}
v^* \quad &= \quad \frac{\sqrt{\Delta}-(\delta\beta-1)}{2\delta} > 0; \\
-\beta \leq v_-^* \quad &= \quad -\frac{\sqrt{\Delta}+(\delta\beta-1)}{2\delta} < 0, \\
\Delta \quad &= \quad (\delta\beta+1)^2 - 4\theta\delta\beta > 0, \qquad 0 \leq \theta < 1.
\end{aligned} \tag{41}$$

*Proof.* Since $0 < u < v^*$, by substituting the explicit results of (37) into *Theorem 3.4*, we have

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

and

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

and also,

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

where $v^*_- = -\beta$ only when $\theta = 0$. $\square$

**Remark 4.4.** We need to assume $\alpha \neq -v^*_-$ 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 the unique $v^*_-$, where

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

Therefore, if $\alpha \in \left(\beta - \frac{1}{\delta}, \beta\right]$, there exists the unique $\theta \in [0, 1)$, such that $\alpha + v^*_- = 0$.

$\alpha = -v^*_-$ is a very particular case, and we will not consider it here and assume $\alpha \neq -v^*_-$ in the sequel.

Now we derive the probability $P\left\{N_T = 0 \middle| \lambda_0\right\}$ in *Corollary 4.1*, and $P\left\{N_T = 1 \middle| \lambda_0\right\}$ for case $\alpha \neq \beta$ in *Corollary 4.2*, a discussion for the special case $\alpha = \beta$ is given in *Remark 4.5*.

**Corollary 4.1.** *If both the externally excited and self-excited jumps follow exponential distributions, i.e. the density functions are specified by (36), then the conditional probability of no jump given $\lambda_0$ and $N_0 = 0$, under the condition $\delta\beta > 1$, is given by*

$$P\left\{N_T = 0 \middle| \lambda_0\right\} = e^{-\left(a + \frac{\rho}{1+\delta\alpha}\right)T} e^{\frac{a-\lambda_0}{\delta}\left(1 - e^{-\delta T}\right)} \left(\frac{1 - e^{-\delta T} + \delta\alpha}{\delta\alpha}\right)^{\frac{\alpha\rho}{1+\delta\alpha}}.$$

*Proof.* By *Theorem 4.2* and setting $\theta = 0$, then, $\Delta = (\delta\beta + 1)^2$, $v^* = \frac{1}{\delta}$, $v^*_- = -\beta$,

$$\begin{aligned} \mathcal{G}^{-1}_{0,0}(T) &= \frac{1}{\delta}\left(1 - e^{-\delta T}\right), \\ K(u) &= -\frac{1}{\delta}\ln(1 - \delta u), \qquad 0 \leq u < \frac{1}{\delta}, \end{aligned}$$

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

and the result follows. $\square$

**Corollary 4.2.** *If both the externally excited and self-excited jumps follow exponential distributions, i.e. the density functions are specified by (36) ($\alpha \neq \beta$), then the conditional probability of exactly one jump given $\lambda_0$ and $N_0 = 0$, under the condition $\delta\beta > 1$, is given by*

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

*where*

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

*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 have $Q_T$ by

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

and

$$\int_0^{u_T}\frac{\hat{g}(u)u}{(1-\delta u)^2}\mathrm{d}u=\frac{\beta}{\delta}\left(e^{\delta T}-1\right)-\beta Q_T,$$

also, when $\alpha\neq\beta$,

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

then, the result follows. $\square$

**Remark 4.5.** In particular, if $\alpha=\beta$, then,

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

where

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

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

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

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

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

We will state and prove the results for the size of clusters based on *Theorem 3.5* for this exponential distribution case as below.

**Theorem 4.3.** *If both the externally excited and self-excited jumps follow exponential distributions, i.e. the density functions are specified as (36), then for the size of a cluster generated by a point of any generation, $\widetilde{N}_t$, under the condition $\delta\beta > 1$, we have*

$$\mathbb{E}\left[\theta^{\widetilde{N}_\infty}\middle|\widetilde{\lambda}_0\right] = \exp\left(-\frac{\sqrt{(\delta\beta-1)^2+4\delta\beta(1-\theta)}-(\delta\beta-1)}{2\delta}\widetilde{\lambda}_0\right); \tag{42}$$

*and $\widetilde{N}_\infty$ conditional on $\widetilde{\lambda}_0$ actually follows a mixed Poisson distribution,*

$$P\left\{\widetilde{N}_\infty = k\middle|\widetilde{\lambda}_0\right\} = \int_0^\infty \frac{v^k e^{-v}}{k!}m(v)\mathrm{d}v, \qquad (k = 0,1,2,...) \tag{43}$$

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

$$m(v) = e^{\frac{\delta\beta-1}{2\delta}\widetilde{\lambda}_0}e^{-\left(\frac{\delta\beta-1}{2\delta}\right)^2\frac{\delta}{\beta}v}\frac{\sqrt{\frac{\beta}{2\delta}}\widetilde{\lambda}_0}{\sqrt{2\pi}v^{\frac{3}{2}}}e^{-\frac{\frac{\beta}{2\delta}\widetilde{\lambda}_0^2}{2v}}, \tag{44}$$

*which is an inverse Gaussian distribution with parameters $\left(\frac{\beta}{\delta\beta-1}\widetilde{\lambda}_0, \frac{\beta}{2\delta}\widetilde{\lambda}_0^2\right)$.*

*Proof.* By substituting the explicit exponential distribution functions of (37) and the constant $v^*$ of (41) into Theorem *3.5*, we obtain (42) immediately.

To prove that $\widetilde{N}_\infty$ follows a mixed Poisson distribution, we rewrite (42) by

$$\mathbb{E}\left[\theta^{\widetilde{N}_\infty}\middle|\widetilde{\lambda}_0\right] = e^{\frac{\delta\beta-1}{2\delta}\widetilde{\lambda}_0}e^{-\sqrt{2\xi}\widetilde{\lambda}_0},$$

where

$$\xi = \frac{1}{2}\left(\frac{\delta\beta-1}{2\delta}\right)^2 + \frac{\beta}{2\delta}(1-\theta),$$

and identify that

$$e^{-\sqrt{2\xi}\widetilde{\lambda}_0} = \mathbb{E}\left[e^{-\xi\widetilde{\mathrm{IG}}}\right] = \int_0^\infty e^{-\xi u}\frac{\left(\widetilde{\lambda}_0^2\right)^{\frac{1}{2}}}{\sqrt{2\pi}u^{\frac{3}{2}}}e^{-\frac{\widetilde{\lambda}_0^2}{2u}}\mathrm{d}u,$$

where $\widetilde{\mathrm{IG}}$ follows the (infinite mean) inverse Gaussian distribution with parameters $\left(\infty, \widetilde{\lambda}_0^2\right)$, then, we have

$$\begin{aligned}\mathbb{E}\left[\theta^{\widetilde{N}_\infty}\middle|\widetilde{\lambda}_0\right] &= e^{\frac{\delta\beta-1}{2\delta}\widetilde{\lambda}_0}\int_0^\infty e^{-\xi u}\frac{\left(\widetilde{\lambda}_0^2\right)^{\frac{1}{2}}}{\sqrt{2\pi}u^{\frac{3}{2}}}e^{-\frac{\widetilde{\lambda}_0^2}{2u}}\mathrm{d}u\\ &= \int_0^\infty e^{-\left[\frac{1}{2}\left(\frac{\delta\beta-1}{2\delta}\right)^2+\frac{\beta}{2\delta}(1-\theta)\right]u}e^{\frac{\delta\beta-1}{2\delta}\widetilde{\lambda}_0}\frac{\left(\widetilde{\lambda}_0^2\right)^{\frac{1}{2}}}{\sqrt{2\pi}u^{\frac{3}{2}}}e^{-\frac{\widetilde{\lambda}_0^2}{2u}}\mathrm{d}u,\end{aligned}$$

and let $v = \frac{\beta}{2\delta}u$,

$$\begin{aligned}\mathbb{E}\big[\theta^{\widetilde{N}_\infty}\big|\widetilde{\lambda}_0\big] &= \int_0^\infty e^{-(1-\theta)v} e^{\frac{\delta\beta-1}{2\delta}\widetilde{\lambda}_0} e^{-\left(\frac{\delta\beta-1}{2\delta}\right)^2\frac{\delta}{\beta}v} \frac{\left(\frac{\beta}{2\delta}\widetilde{\lambda}_0^2\right)^{\frac{1}{2}}}{\sqrt{2\pi}v^{\frac{3}{2}}} e^{-\frac{\frac{\beta}{2\delta}\widetilde{\lambda}_0^2}{2v}}\mathrm{d}v \\ &= \int_0^\infty e^{-(1-\theta)v} m(v)\mathrm{d}v = \hat{m}(\theta-1),\end{aligned}$$

where

$$\hat{m}(u) = \int_0^\infty e^{-uv} m(v)\mathrm{d}v.$$

Hence, by the definition of the mixed Poisson distribution, we have (43) and (44); set $u = 1-\theta$, we have

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

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

**Corollary 4.3.** *In particular, for a cluster generated by a point of generation* 0*, we have*

$$\mathbb{E}\big[\theta^{\widetilde{N}_\infty}\big] = \frac{2\delta\alpha}{\delta(2\alpha-\beta)+1+\sqrt{(\delta\beta-1)^2+4\delta\beta(1-\theta)}}; \tag{45}$$

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

$$\mathbb{E}\big[\theta^{\widetilde{N}_\infty}\big] = \frac{2\delta\beta}{1+\sqrt{1-\frac{4\delta\beta}{(\delta\beta+1)^2}\theta}}, \tag{46}$$

*and*

$$P\left\{\widetilde{N}_\infty = k\right\} = \frac{(\delta\beta)^{k+1}}{(\delta\beta+1)^{2k}}\frac{(2k)!}{k!(k+1)!}, \qquad k = 0,1,.... \tag{47}$$

*Proof.* By substituting the explicit exponential distribution functions of (37) and the constant $v^*$ of (41) into Theorem *3.5*, we obtain (45). In particular, by setting $\alpha = \beta$ in (42) and expanding explicitly, we have (46) and (47). □

**Remark 4.7.** We can also expand (42) explicitly for some other special cases. For instance, if $2\delta\alpha + (1-\delta\beta) = 0$, we have

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

For the general case, we can expand (45) with respect to $\theta$ by Taylor expansion function in Matlab. An example with the parameter setting $(\delta;\alpha,\beta) = (2.0;2.0,1.5)$ for $P\{\widetilde{N}_\infty = k\}$ is given by *Table 1*.

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.

**Table 1:** Probability $P\{\widetilde{N}_\infty = k\}$ for $k = 0, 1, 2, ...$; $(\delta; \alpha, \beta) = (2.0; 2.0, 1.5)$

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $P\{\widetilde{N}_\infty = k\}$ | 80.0000% | 12.0000% | 4.0500% | 1.7888% | 0.9043% | 0.4956% | 0.2866% | 0.1722% | 0.1064% | 0.0672% | 0.0432% | 0.0282% | 0.0186% |
| $k$ | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| $P\{\widetilde{N}_\infty = k\}$ | 0.0124% | 0.0083% | 0.0056% | 0.0039% | 0.0026% | 0.0018% | 0.0013% | 0.0009% | 0.0006% | 0.0004% | 0.0003% | 0.0002% | 0.0001% |

# 5 An Application in Credit Risk

Our motivation of applying the dynamic contagion process to model the credit risk is a combination 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, Lando and Turnbull (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 also 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 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 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 $N_t$ is to model the number of bad events released from the underlying company. It is driven by a series of bad events $\{Z_j\}_{j=1,2,...}$ from itself and the common bad events $\{Y_i\}_{i=1,2,...}$ widely in the whole market via its intensity process $\lambda_t$. The impact of each event decays exponentially with constant rate $\delta$. We assume 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 $P_s(T) = \mathbb{E}\left[(1-d)^{N_T} \middle| \lambda_0\right]$, 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 in *Table 2*.

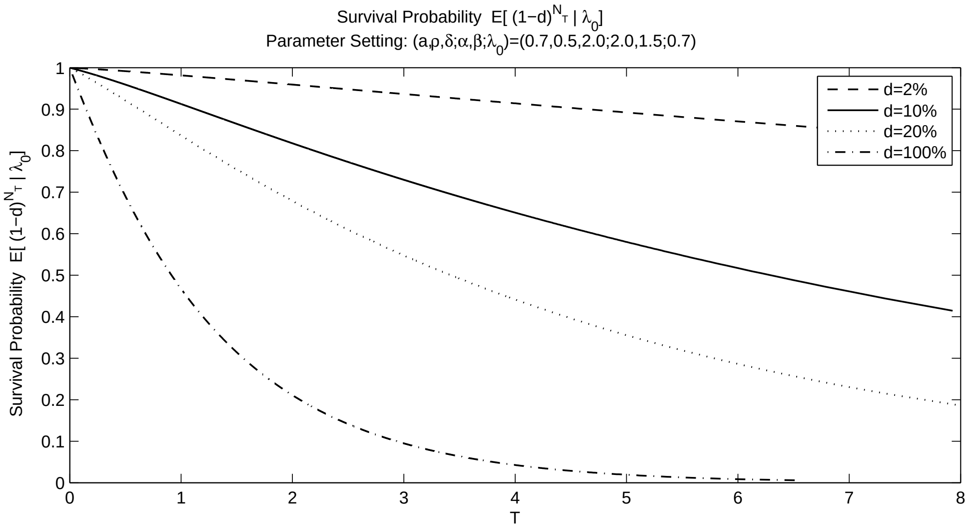


**Figure 2:** Survival Probability $P_s(T)$; $(a, \rho, \delta; \alpha, \beta; \lambda_0) = (0.7, 0.5, 2.0; 2.0, 1.5; 0.7)$

**Table 2:** Survival Probability $P_s(T)$; $(a, \rho, \delta; \alpha, \beta; \lambda_0) = (0.7, 0.5, 2.0; 2.0, 1.5; 0.7)$

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

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

We also provide a comparison for the survival probabilities based on three main processes discussed in this paper: dynamic contagion process, Hawkes process (by setting $\rho = 0$) and non-self-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*.

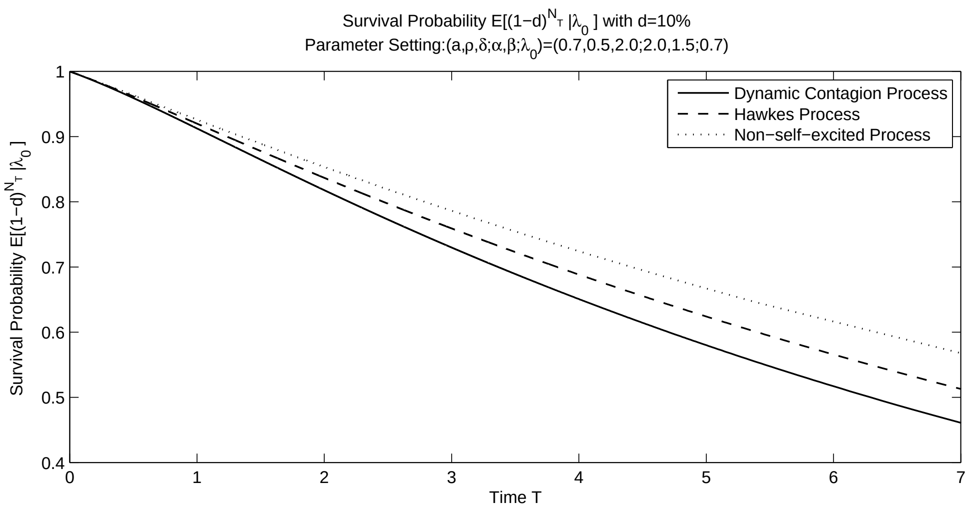


**Figure 3:** Survival Probability Comparison for the Dynamic Contagion, Hawkes and Non-self-excited Processes

**Table 3:** Survival Probability Comparison for the Dynamic Contagion, Hawkes and Non-self-excited Process

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

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 from the endogenous and exogenous factors respectively. This process is capable to capture more aspects of the risk, which is particularly useful for modelling the risks during the economic downturn involving more clusters of bad economic events.

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

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

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

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

2. Simulate the $(i+1)^{\text{th}}$ 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

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

and

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

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

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

where

$$\lambda_{T_{i+1}^{*-}} = \left(\lambda_{T_i^{*+}} - a\right) e^{-\delta\left(T_{i+1}^* - T_i^*\right)} + a.$$

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

$$N_{T_{i+1}^{*+}} = \begin{cases} N_{T_{i+1}^{*-}} + 1 & \left(S_{i+1}^* \wedge E_{i+1}^* = S_{i+1}^*\right) \\ N_{T_{i+1}^{*-}} & \left(S_{i+1}^* \wedge E_{i+1}^* = E_{i+1}^*\right) \end{cases}.$$

Note that, this simulation procedure 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 10000 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 $\mathbb{E}[\lambda_t]$, $\mathbb{E}[\lambda_t|\lambda_0]$ and $\mathbb{E}[N_t]$ (derived by *Corollary 3.5*, *Theorem 3.6* and *Theorem 3.8*, respectively) are also plotted.

**Table 4:** Survival Probability $P_s(T)$ by 10000 Simulated Sample Paths

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

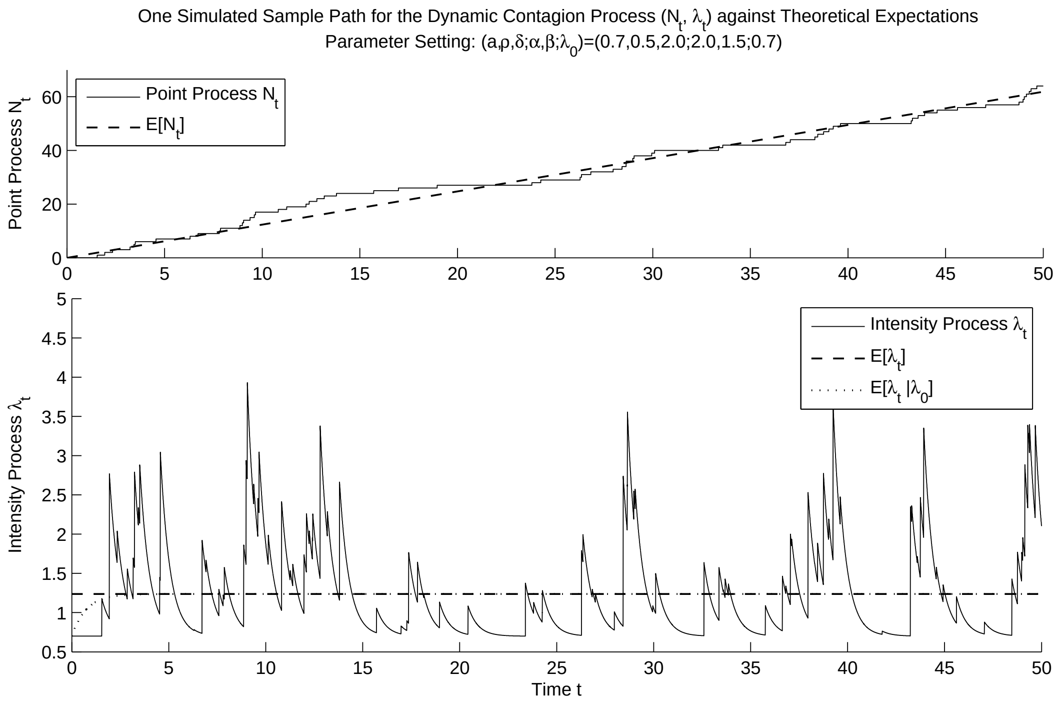


**Figure 4:** One Simulated Sample Path of the Dynamic Contagion Process $(N_t, \lambda_t)$

# 6 Conclusion

This paper produces a general mathematical framework for modelling the dependence structure of arriving events with contagion dynamics, mainly based on generalising the Hawkes process and the Cox process with shot noise intensity. The dynamic contagion process newly introduced here has been systemically studied by analysing its various distributional properties, and has the significant potential to be 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.

## References

[1] Azizpour, S., Giesecke, K. (2008). Self-exciting Corporate Defaults: Contagion vs. Frailty. Working Paper. Stanford University.

[2] Bordenave, C., Torrisi, G.L. (2007). Large Deviations of Poisson Cluster Processes. *Stoch Models.* **23**, 593-625.

[3] Brémaud, P. (1981). *Point Processes and Queues.* Springer, New York.

[4] Brémaud, P., Massoulié, L. (1996). Stability of Nonlinear Hawkes Processes. *Annals of Probability.* **24**, 1563-1588.

[5] Brémaud, P., Massoulié, L., Ridolfi, A. (2002). Power Spectra of General Shot Noises and Hawkes Processes with a Random Excitation. *Advances in Applied Probability.* **34**, 205-222.

[6] Chavez-Demoulin, V., Davison, A.C., Mc Neil, A.J. (2005). Estimating Value-at-risk: A Point Process Approach. *Quantitative Finance.* **5**, 227-234.

[7] Costa, O. L. V. (1990). Stationary Distributions for Piecewise-deterministic Markov Processes. *Journal of Applied Probability.* **27**, 60-73.

[8] Cox, D.R., Isham, V. (1980). *Point Processes.* Chapman and Hall, London.

[9] Daley, D.J., Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes (2nd edition).* Springer, New York.

[10] Dassios, A., Jang, J. (2003). Pricing of Catastrophe Reinsurance and Derivatives Using the Cox Process with Shot Noise Intensity. *Finance & Stochastics.* **7(1)**, 73-95.

[11] Dassios, A., Jang, J. (2005). Kalman-Bucy Filtering for Linear System Driven by the Cox Process with Shot Noise Intensity and its Application to the Pricing of Reinsurance Contracts. *Journal of Applied Probability.* **42/1**, 93-107.

[12] Davis, M.H.A. (1984). Piecewise Deterministic Markov Processes: A General Class of Nondiffusion Stochastic Models. *Journal of the Royal Statistical Society B.* **46**, 353-388.

[13] Davis, M.H.A. (1993). *Markov Model and Optimization.* Chapman and Hall.

[14] Duffie, D., Singleton, K.J. (1999). Modeling Term Structures of Defaultable Bonds. *Review of Financial Studies.* Vol. **12(4)**, 687-720.

[15] Duffie, D., Gârleanu, N. (2001). Risk and Valuation of Collateralized Debt Obligations. *Financial Analysts Journal.* **57(1)**, 42-59.

[16] Duffie, D., Filipović, D. Schachermayer, W. (2003). Affine Processes and Applications in Finance. *Annals of Applied Probability.* **13**, 984-1053.

[17] Ethier, S.N., Kurtz, T.G. (1986). *Markov Processes: Characterization and Convergence.* John Wiley and Sons, New York.

[18] Errais, E., Giesecke, K., Goldberg, L. (2009). Affine Point Processes and Portfolio Credit Risk. Working Paper. Stanford University.

[19] Hawkes, A.G. (1971). Spectra of Some Self-exciting and Mutually Exciting Point Processes. *Biometrika.* **58(1)**, 83-90.

[20] Hawkes, A.G., Oakes, D. (1974). A Cluster Representation of a Self-exciting Process. *Journal of Applied Probability.* **11**, 493-503.

[21] Jarrow, R., Lando, D., Turnbull, S. (1997). A Markov Model for the Term Structure of Credit Risk Spreads. *Review of Financial Studies.* **10**, 481-523.

[22] Jarrow, R., Yu, F. (2001). Counterparty Risk and the Pricing of Defaultable Securities. *Journal of Finance.* **56(5)**, 1765-1799.

[23] Lando, D. (1998). On Cox Processes and Credit Risky Securities. *Derivatives Research.* **2**, 99-120.

[24] Longstaff, F.A., Rajan, A. (2008). An Empirical Analysis of the Pricing of Collateralized Debt Obligations. *Journal of Finance.* **63(2)**, 529-563.

[25] Massoulié, L. (1998). Stability Results for a General Class of Interacting Point Processes Dynamics, and Applications. *Stochastic Processes and their Applications.* **75**, 1-30.

[26] Oakes, D. (1975). The Markovian Self-exciting Process. *Journal of Applied Probability.* **12**, 69-77.

[27] Stabile, G., Torrisi, G.L. (2010). Risk Processes with Non-stationary Hawkes Claims Arrivals. *Methodology and Computing in Applied Probability.* **12**, 415-429.