

Angelos Dassios and Hongbiao Zhao

A generalised contagion process with an application to credit risk

**Article (Accepted version)
(Refereed)**

Original citation:

Dassios, Angelos and Zhao, Hongbiao (2017) *A generalised contagion process with an application to credit risk*. International Journal of Theoretical and Applied Finance . ISSN 0219-0249

DOI: [10.1142/S0219024917500030](https://doi.org/10.1142/S0219024917500030)

© 2016 World Scientific Publishing Company

This version available at: <http://eprints.lse.ac.uk/68558/>

Available in LSE Research Online: December 2016

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's final accepted version of the journal article. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

A Generalised Contagion Process with an Application to Credit Risk*

Angelos Dassios[†] Hongbiao Zhao[‡]
London School of Economics *Xiamen University*

December 6, 2016

Abstract

We introduce a class of analytically tractable jump processes with contagion effects by generalising the classical Hawkes process. This model framework combines the characteristics of three popular point processes in the literature: (1) Cox process with CIR intensity; (2) Cox process with Poisson shot-noise intensity; (3) Hawkes process with exponentially decaying intensity. Hence, it can be considered as a self-exciting and externally-exciting point process with mean-reverting stochastic intensity. Essential probabilistic properties such as moments, Laplace transform of intensity process, and probability generating function of point process as well as some important asymptotics have been derived. Some special cases and a method for change of measure are discussed. This point process may be applicable to modelling contagious arrivals of events for various circumstances (such as jumps, transactions, losses, defaults, catastrophes) in finance, insurance and economics with both endogenous and exogenous risk factors within one framework. More specifically, these exogenous factors could contain relatively short-lived shocks and long-lasting risk drivers. We make a simple application to calculate the default probability for credit risk and price defaultable zero-coupon bonds.

Keywords: Credit risk; Contagion risk; Stochastic intensity model; Jump process; Point process; Self-exciting process; Hawkes process; Cox process; CIR process; Dynamic contagion process; Dynamic contagion process with diffusion.

*To appear in *International Journal of Theoretical and Applied Finance*

[†]Department of Statistics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom. Email: a.dassios@lse.ac.uk

[‡]Corresponding Author, Department of Finance, School of Economics & Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen, Fujian 361005, China; MOE Key Laboratory of Econometrics, Xiamen University, Xiamen, Fujian 361005, China. Email: h.zhao1@lse.ac.uk

1 Introduction

Contagion risk in finance and economics has become much more prevalent, particularly after the global financial crisis 2007-2008 and the recent sovereign debt crisis in the eurozone. It is important to analyse and quantify the contagion feature of event arrivals. However, there are not plenty of continuous-time models available for it in the literature that could go beyond the simple measure of correlation. Hawkes (1971a,b)¹ early introduced a *self-exciting* point process, and its stochastic intensity process is a function of the past of point process itself, and jumps occur simultaneously in the point process and its intensity process. It now has been widely adopted for modelling contagion effects in finance and insurance, such as trade arrivals in market microstructure, defaults in credit market, jumps in returns of investments, and loss claims in insurance portfolios, see Chavez-Demoulin et al. (2005), Bowsher (2007), Large (2007), Stabile and Torrisi (2010), Embrechts et al. (2011), Bacry et al. (2013), Zhu (2013b), and more recently, Aït-Sahalia et al. (2014) and Aït-Sahalia et al. (2015)². The theoretical framework was later extended by Brémaud and Massoulié (1996, 2002), Zhu (2013a) and Boumezoued et al. (2016). It also has various applications in many other fields, see Vere-Jones (1978), Chornoboy et al. (1988), Ogata (1988), Crane and Sornette (2008), Marsan and Lengline (2008), Veen and Schoenberg (2008), Mohler et al. (2011), Xu et al. (2014), Zadeh and Sharda (2014) and Hall and Willett (2016).

Although the framework has been set up, the exact mathematical properties have not been fully understood, as pointed by Errais et al. (2010). Dassios and Zhao (2011) analysed some key probabilistic properties in details for *dynamic contagion process* (DCP), a generalised univariate Hawkes process with extra *externally-excited* components³. In this paper, we further extend the DCP, and allow the stochastic intensity process being perturbed by an additional independent diffusion⁴. The resulting process named *dynamic contagion process with diffusion* (DCPD) here in fact is a *self-excited*⁵ and *externally-excited*

¹See also a series of pioneering work in Hawkes (1971a,b), Hawkes and Oakes (1974) and Oakes (1975).

²Bacry et al. (2015) provide a very good survey of the recent academic literature devoted to the applications of Hawkes processes in finance.

³The associated applications to ruin problem in insurance can also be found in Dassios and Zhao (2012).

⁴It is also an extension of *Hawkes process with general immigrants* (Brémaud and Massoulié, 2002).

⁵The term "self-excited" is treated equivalently as "self-exciting" throughout this paper.

point process with mean-reverting stochastic intensity. More precisely, it is a hybrid of three popular point processes:

- (1) a *Cox process with CIR intensity*;
- (2) a *Cox process with Poisson shot-noise intensity*;
- (3) a *Hawkes process with exponentially decaying intensity*.

These three separate models have been widely applied to finance, insurance and economics, particularly, for risk management and asset pricing. Now we consider combining all of them together within one framework⁶.

Our main contribution of this paper is that, with the aid of martingale approach (Dassios and Embrechts, 1989) and infinitesimal generator analysis (also known as *Dynkin's formula*), we fundamentally investigate the DCPD's distributional properties of intensity process and point process. This extension from the DCP is nontrivial, as the DCPD is a point process that acts very differently from a DCP: it could not be classified neither as a *piecewise-deterministic Markov process* (Davis, 1984) nor as a branching process; the trajectory between two successive jumps in intensity process is no longer deterministic, due to the oscillation character of the additional component of independent Brownian motion; moreover, the intensity process is possible to go down below the mean-reverting level and even reach zero. Hence, some methods of deriving the distributional properties for the DCPD are not the same as the ones for the DCP in Dassios and Zhao (2011), for instance, the Laplace transform of stationary distribution of intensity process as later given by *Theorem 3.2*. Additionally, we also investigate the asymptotics of stationary distribution of the intensity around zero. Our motivation of this extension for potential applications in finance is that, the DCPD equipped with all these three components could provide a more realistic model, for instance, the default intensity (or frequency) could be influenced by some internal and external risk shocks (e.g. financial reports, crises, earthquakes) in the economy, as well as some additional certain degree of external risks or noises (e.g. GDP, CPI, stock indexes) persistently driving in the market. These two types of relatively short-lived shocks are modelled by our jump components, and the long-lasting external factors could be captured by the component of mean-reverting diffusion. By further introducing the addition of this supplementary diffusion, risk factors with different characteristics of short-lived and long-lasting effects could be more specifically

⁶Our model is also the generalisation of so-called *generalised Hawkes process* used in Zhang et al. (2009) (see also Giesecke and Kim (2007), Giesecke et al. (2011) and Zhu (2014)) by adding another series of externally-excited jumps in the underlying intensity process.

distinguished and captured respectively.

The paper is organised as follows. Section 2 provides a mathematical definition of the DCPD. In Section 3, we derive its key distributional properties, such as the moments, Laplace transform of asymptotic and stationary distribution of intensity process, and the probability generating function of point process; some special cases of exponential distribution are discussed. A method for change of measure via Esscher transform is also presented in Section 4. We apply our model to study the probability of default for credit risk and price defaultable zero-coupon bonds with numerical examples in Section 5. Finally, Section 6 makes a brief conclusion for this paper.

2 Definition

We provide a mathematical definition for the DCPD via the stochastic intensity representation in *Definition 2.1*.

Definition 2.1 (Dynamic Contagion Process with Diffusion). **Dynamic contagion process with diffusion (DCPD)** is a point process $N \equiv \{T_k^{(2)}\}_{k=1,2,\dots}$ with the non-negative \mathcal{F}_t -stochastic (conditional) intensity

$$\begin{aligned} \lambda_t = & a + (\lambda_0 - a) e^{-\delta t} + \sigma \int_0^t e^{-\delta(t-s)} \sqrt{\lambda_s} dW_s \\ & + \sum_{0 \leq T_i^{(1)} < t} Y_i^{(1)} e^{-\delta(t-T_i^{(1)})} + \sum_{0 \leq T_k^{(2)} < t} Y_k^{(2)} e^{-\delta(t-T_k^{(2)})}, \quad t \geq 0, \end{aligned} \quad (2.1)$$

where

- $\{\mathcal{F}_t\}_{t \geq 0}$ is a history of the process N_t , with respect to which $\{\lambda_t\}_{t \geq 0}$ is adapted;
- $\lambda_0 > 0$ is the initial intensity at time $t = 0$;
- $a \geq 0$ is the constant *mean-reverting level*;
- $\delta > 0$ is the constant *mean-reverting rate*;
- $\sigma > 0$ is the constant *volatility of intensity diffusion* (i.e. the volatility of diffusion part of intensity process);
- $\{W_t\}_{t \geq 0}$ is a standard Brownian motion;
- $\{Y_i^{(1)}\}_{i=1,2,\dots}$ are the sizes of *externally-excited jumps*, a sequence of *i.i.d.* positive random variables with distribution function $H(y), y > 0$;

- $\{T_i^{(1)}\}_{i=1,2,\dots}$ are the arrival times of a Poisson process M_t with constant rate $\varrho > 0$;
- $\{Y_k^{(2)}\}_{k=1,2,\dots}$ are the sizes of *self-excited jumps*, a sequence of *i.i.d.* positive random variables with distribution function $G(y), y > 0$;
- the sequences $\{Y_i^{(1)}\}_{i=1,2,\dots}, \{Y_k^{(2)}\}_{k=1,2,\dots}, \{T_i^{(1)}\}_{i=1,2,\dots}$ and $\{W_t\}_{t \geq 0}$ are assumed to be independent of each other.

In fact, $\{N_t\}_{t \geq 0}$ is a simple point process so that there is no double jumps at any particular time. More precisely, it can be defined by $N_t := \sum_{k=1}^{\infty} \mathbb{1}\{T_k^{(2)} \leq t\}$ where $\mathbb{1}\{\cdot\}$ is the indicator function, and λ_t is a conventional intensity of point process that satisfies

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

where Δt is a sufficiently small time interval, and $o(\Delta t)/\Delta t \rightarrow 0$ when $\Delta t \rightarrow 0$. The joint process of $\{(\lambda_t, N_t)\}_{t \geq 0}$ is a Markov process in the state space $\mathbb{R} \times (\mathbb{N} \cup \{0\})$. By Markov property, the infinitesimal generator of process (λ_t, N_t, t) acting on a function $f(\lambda, n, t)$ within its domain $\Omega(\mathcal{A})$ is given by

$$\begin{aligned} \mathcal{A}f(\lambda, n, t) &= \frac{\partial f}{\partial t} - \delta(\lambda - a) \frac{\partial f}{\partial \lambda} + \frac{1}{2} \sigma^2 \lambda \frac{\partial^2 f}{\partial \lambda^2} + \varrho \left[\int_0^\infty f(\lambda + y, n, t) dH(y) - f(\lambda, n, t) \right] \\ &\quad + \lambda \left[\int_0^\infty f(\lambda + y, n + 1, t) dG(y) - f(\lambda, n, t) \right]. \end{aligned} \quad (2.2)$$

Note that, this point process is not a classical *doubly stochastic Poisson process* or Cox process (Cox, 1955), since N_t conditional on λ_t is not of Poisson type and does not satisfy the fundamental definition, more precisely, for any time $t \in [0, T]$,

$$\mathbb{E}[\theta^{(N_T - N_t)} \mid \mathcal{F}_t] \neq \mathbb{E}[e^{-(1-\theta)(\Lambda_T - \Lambda_t)} \mid \mathcal{F}_t], \quad \theta \in [0, 1], \quad (2.3)$$

where $\Lambda_t =: \int_0^t \lambda_s ds$ is the aggregated intensity process (or, the *compensator* of point process N_t).

If there is no externally-excited jumps and diffusion, and all the sizes of self-excited jumps are fixed to be the same, then, it recovers the classical Hawkes process. The dynamic contagion process with diffusion is a generalised Hawkes process which is still within the general framework of affine processes, see Duffie et al. (2000), Duffie et al. (2003) and Glasserman and Kim (2010). A sample path of simulated intensity process λ_t based on *discretisation scheme*⁷ is plotted in Figure 1.

⁷The numerical algorithm of *exact* Monte Carlo simulation for generating this point process N_t is devel-

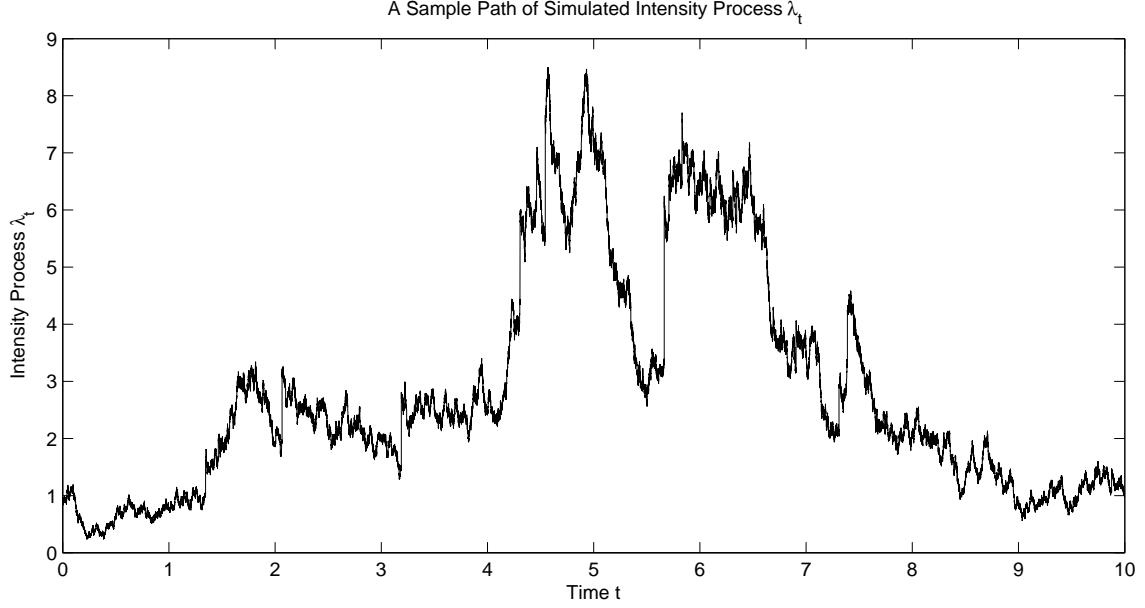


Figure 1: A sample path of simulated intensity process λ_t based on the standard *discretisation scheme* with the parameters $(a, q, \delta; \alpha, \beta; \sigma; \lambda_0) = (0.9, 0.1, 1.0; 10, 1.2; 1.0; 0.9)$ where the jump sizes of two types are assumed to follow exponential distributions, i.e. $H \sim \text{Exp}(\alpha)$ and $G \sim \text{Exp}(\beta)$

Remark 2.1. Externally-excited jumps $\{(Y_i^{(1)}, T_i^{(1)})\}_{i=1,2,\dots}$ and self-excited jumps $\{(Y_k^{(2)}, T_k^{(2)})\}_{k=1,2,\dots}$ are designated to capture the relatively short-lived endogenous and exogenous risk shocks, respectively. The diffusion driven by $\{W_t\}_{t \geq 0}$ is for modelling certain external risk always persisting in the market. δ controls the time decay of impacts. We assume the same decay rate of δ for the diffusion process, self-excited and externally-excited jumps, as this assumption makes our model analytically tractable.

3 Distributional Properties

To simplify notations, for the two types of jump sizes $Y^{(1)}$ and $Y^{(2)}$ in λ_t of (2.1), the first, second moments and Laplace transforms are denoted respectively by

$$\begin{aligned} \mu_{1_H} &:= \int_0^\infty y dH(y), & \mu_{2_H} &:= \int_0^\infty y^2 dH(y), & \hat{h}(u) &:= \int_0^\infty e^{-uy} dH(y), \\ \mu_{1_G} &:= \int_0^\infty y dG(y), & \mu_{2_G} &:= \int_0^\infty y^2 dG(y), & \hat{g}(u) &:= \int_0^\infty e^{-uy} dG(y), \end{aligned}$$

and the constant $\kappa := \delta - \mu_{1_G}$. The moments and Laplace transforms above are all assumed to be finite.

oped in Dassios and Zhao (2015).

3.1 Joint Laplace Transform – Probability Generating Function of (λ_T, N_T)

We first look at the joint distributional property of intensity process and point process via their joint transform function.

Lemma 3.1. *For constants $0 \leq \theta \leq 1$, $v \geq 0$ and time $0 \leq t \leq T$, the conditional joint Laplace transform – probability generating function of (λ_T, N_T) is given by*

$$\mathbb{E} \left[\theta^{(N_T - N_t)} e^{-v\lambda_T} \mid \mathcal{F}_t \right] = e^{-\left(c(T) - c(t)\right)} \times e^{-B(t)\lambda_t}, \quad t \in [0, T], \quad (3.1)$$

where $B(t)$ is determined by the ODE

$$-B'(t) + \delta B(t) + \theta \hat{g}(B(t)) - 1 + \frac{1}{2} \sigma^2 B^2(t) = 0, \quad (3.2)$$

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

$$c(T) - c(t) = a\delta \int_t^T B(s) ds + \varrho \int_t^T [1 - \hat{h}(B(s))] ds. \quad (3.3)$$

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 (2.2), we then have

$$\begin{aligned} \frac{A'(t)}{A(t)} n + \left[-B'(t) + \delta B(t) + A(t) \hat{g}(B(t)) - 1 + \frac{1}{2} \sigma^2 B^2(t) \right] \lambda \\ + \left[c'(t) + \varrho \hat{h}(B(t)) - \varrho - a\delta B(t) \right] = 0. \end{aligned} \quad (3.4)$$

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

$$\begin{cases} \frac{A'(t)}{A(t)} = 0, & (.1) \\ -B'(t) + \delta B(t) + A(t) \hat{g}(B(t)) - 1 + \frac{1}{2} \sigma^2 B^2(t) = 0, & (.2) \\ c'(t) + \varrho \hat{h}(B(t)) - \varrho - a\delta B(t) = 0. & (.3) \end{cases} \quad (3.5)$$

We have $A(t) = \theta$ immediately from (3.5.1); and substitute it into (3.5.2) by adding the boundary condition $B(T) = v$, we have the ODE as (3.2). Then, by (3.5.3), the integration of (3.3) follows. By the property of infinitesimal generator, $e^{c(t)} \theta^{N_t} e^{-B(t)\lambda_t}$ is a martingale, and we have

$$\mathbb{E} \left[e^{c(T)} \theta^{N_T} e^{-B(T)\lambda_T} \mid \mathcal{F}_t \right] = e^{c(t)} \theta^{N_t} e^{-B(t)\lambda_t}. \quad (3.6)$$

Then, with the boundary condition $B(T) = v$, (3.1) follows. \square

3.2 Laplace Transform of λ_T

Based on Lemma 3.1, we then investigate the distributional properties of intensity process $\{\lambda_t\}_{t \geq 0}$ in detail as follows.

Theorem 3.1. For $\kappa > 0$, the Laplace transform λ_T conditional on λ_0 is given by

$$\mathbb{E} [e^{-v\lambda_T} \mid \lambda_0] = \exp \left(- \int_{\mathcal{G}_{v,1}^{-1}(T)}^v \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du \right) \times e^{-\mathcal{G}_{v,1}^{-1}(T)\lambda_0}, \quad (3.7)$$

where

$$\mathcal{G}_{v,1}(L) := \int_L^v \frac{du}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2}, \quad L \in (0, v]. \quad (3.8)$$

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

$$\mathbb{E} [e^{-v\lambda_T} \mid \mathcal{F}_0] = e^{-\left(c(T) - c(0)\right)} e^{-B(0)\lambda_0}, \quad (3.9)$$

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

$$-B'(t) + \delta B(t) + \hat{g}(B(t)) - 1 + \frac{1}{2}\sigma^2 B^2(t) = 0,$$

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

$$\frac{dL(\tau)}{d\tau} = 1 - \delta L(\tau) - \hat{g}(L(\tau)) - \frac{1}{2}\sigma^2 L^2(\tau), \quad (3.10)$$

with the initial condition $L(0) = v > 0$; we define the right-hand side of (3.10) as

$$f_1(L) := 1 - \delta L - \hat{g}(L) - \frac{1}{2}\sigma^2 L^2.$$

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

$$\frac{\partial f_1(L)}{\partial L} = \int_0^\infty z e^{-Lz} dG(z) - \delta - \sigma^2 L \leq \int_0^\infty z dG(z) - \delta = \mu_{1_G} - \delta = \kappa < 0, \quad \text{for } L \geq 0,$$

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

3. Rewrite (3.10) as

$$\frac{dL}{\delta L + \hat{g}(L) - 1 + \frac{1}{2}\sigma^2 L^2} = -d\tau,$$

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

$$\int_L^v \frac{du}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} = \tau.$$

Define the function on the left-hand side as (3.8), then, $\mathcal{G}_{v,1}(L) = \tau$. Obviously $L \rightarrow v$ when $\tau \rightarrow 0$; by convergence test,

$$\lim_{u \rightarrow 0} \frac{\frac{1}{u}}{\frac{1}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2}} = \delta + \lim_{u \rightarrow 0} \frac{\hat{g}(u) - 1}{u} = \delta - \mu_{1_G} = \kappa > 0,$$

and we know that $\int_0^v \frac{1}{u} du = \infty$, then, $\lim_{L \downarrow 0} \mathcal{G}_{v,1}(L) = \infty$, hence, $L \rightarrow 0$ when $\tau \rightarrow \infty$; the integrand of (3.8) is positive in the domain $u \in (0, \infty)$ and also for $0 < L \leq v$, $\mathcal{G}_{v,1}(L)$ is a strictly decreasing function; therefore, $\mathcal{G}_{v,1}(L) : (0, v] \rightarrow [0, \infty)$ is a well defined (monotone) function, and its inverse function $\mathcal{G}_{v,1}^{-1}(\tau) : [0, \infty) \rightarrow (0, v]$ exists.

4. The unique solution is found by $L(\tau) = \mathcal{G}_{v,1}^{-1}(\tau)$, or $B(t) = \mathcal{G}_{v,1}^{-1}(T - t)$.

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

Then, $c(T) - c(0)$ is determined by

$$c(T) - c(0) = a\delta \int_0^T \mathcal{G}_{v,1}^{-1}(\tau) d\tau + \varrho \int_0^T \left[1 - \hat{h}(\mathcal{G}_{v,1}^{-1}(\tau)) \right] d\tau, \quad (3.11)$$

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

$$\int_0^T \left[1 - \hat{h}(\mathcal{G}_{v,1}^{-1}(\tau)) \right] d\tau = \int_{\mathcal{G}_{v,1}^{-1}(0)}^{\mathcal{G}_{v,1}^{-1}(T)} [1 - \hat{h}(u)] \frac{\partial \tau}{\partial u} du = \int_{\mathcal{G}_{v,1}^{-1}(T)}^v \frac{1 - \hat{h}(u)}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du,$$

similarly,

$$\int_0^T \mathcal{G}_{v,1}^{-1}(\tau) d\tau = \int_{\mathcal{G}_{v,1}^{-1}(T)}^v \frac{u}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du.$$

Finally, substitute $B(0)$ and $c(T) - c(0)$ into (3.9), and *Theorem 3.1* follows. \square

Corollary 3.1. For $\kappa > 0$, the Laplace transform of asymptotic distribution of λ_T conditional on λ_0 is given by

$$\lim_{T \rightarrow \infty} \mathbb{E} \left[e^{-v\lambda_T} \mid \lambda_0 \right] = \hat{\Pi}(v),$$

where

$$\hat{\Pi}(v) := \mathcal{L}\{\Pi(\lambda)\} = \exp\left(-\int_0^v \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du\right). \quad (3.12)$$

Proof. Let $T \rightarrow \infty$ in Theorem 3.1, then $\mathcal{G}_{v,1}^{-1}(T) \rightarrow 0$ which largely simplifies the expression (3.7), and the Laplace transform of asymptotic distribution follows immediately as given by (3.12). \square

Π is denoted as the distribution determined by its Laplace transform of (3.12), and $\Pi(\lambda)$ is denoted as the associated density function.

Corollary 3.2. For $\kappa > 0$ and any time $T \geq 0$, if $\lambda_0 \sim \Pi$, then $\lambda_T \sim \Pi$.

Proof. By Theorem 3.1 and given the Laplace transform of distribution Π by (3.12), we have

$$\begin{aligned} & \mathbb{E}[e^{-v\lambda_T}] \\ &= \mathbb{E}[\mathbb{E}[e^{-v\lambda_T} \mid \lambda_0]] \\ &= \exp\left(-\int_{\mathcal{G}_{v,1}^{-1}(T)}^v \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du\right) \mathbb{E}[e^{-\mathcal{G}_{v,1}^{-1}(T)\lambda_0}] \\ &= \exp\left(-\int_{\mathcal{G}_{v,1}^{-1}(T)}^v \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du\right) \exp\left(-\int_0^{\mathcal{G}_{v,1}^{-1}(T)} \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du\right) \\ &= \exp\left(-\int_0^v \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du\right) \\ &= \hat{\Pi}(v). \end{aligned}$$

\square

The stationarity property revealed in Corollary 3.2 can be formally proved in Theorem 3.2 as below. To rigorously prove the existence and uniqueness of this stationary process, one can easily follow the same approach as adopted in the proof for Theorem 3.3. in Dassios and Zhao (2011).

Theorem 3.2. For $\kappa > 0$, (3.12) is also the Laplace transform of stationary distribution of $\{\lambda_t\}_{t \geq 0}$.

Proof. By the martingale property of infinitesimal generator of (2.2), we have a martingale $f(\lambda_t, N_t, t) - f(\lambda_0, N_0, 0) - \int_0^t \mathcal{A}(\lambda_s, N_s, s) ds$. Set $f(\lambda, n, t) = e^{-v\lambda}$, we have

$$\mathcal{A}(e^{-v\lambda}) = e^{-v\lambda} \left[-a\delta v + \varrho[\hat{h}(v) - 1] + \left(\delta v + \hat{g}(v) - 1 + \frac{1}{2}\sigma^2 v^2 \right) \lambda \right],$$

then,

$$\begin{aligned}
& \mathbb{E} [e^{-v\lambda_t} \mid \mathcal{F}_0] \\
&= \int_0^t \mathbb{E} [\mathcal{A}(e^{-v\lambda_s}) \mid \mathcal{F}_0] ds + e^{-v\lambda_0} \\
&= \int_0^t \left[\left(-a\delta v + \varrho[\hat{h}(v) - 1] \right) \mathbb{E} [e^{-v\lambda_s} \mid \mathcal{F}_0] + \left(\delta v + \hat{g}(v) - 1 + \frac{1}{2}\sigma^2 v^2 \right) \mathbb{E} [\lambda_s e^{-v\lambda_s} \mid \mathcal{F}_0] \right] ds + e^{-v\lambda_0}.
\end{aligned}$$

Differentiate two sides with respect to t , as

$$\frac{\partial}{\partial t} \int_0^t \mathbb{E} [\lambda_s e^{-v\lambda_s} \mid \mathcal{F}_0] = -\frac{\partial}{\partial v} \mathbb{E} [e^{-v\lambda_t} \mid \mathcal{F}_0],$$

we have

$$\frac{\partial \mathbb{E} [e^{-v\lambda_t} \mid \mathcal{F}_0]}{\partial t} = \left(-a\delta v + \varrho[\hat{h}(v) - 1] \right) \mathbb{E} [e^{-v\lambda_t} \mid \mathcal{F}_0] - \left(\delta v + \hat{g}(v) - 1 + \frac{1}{2}\sigma^2 v^2 \right) \frac{\partial}{\partial v} \mathbb{E} [e^{-v\lambda_t} \mid \mathcal{F}_0].$$

Denote $\hat{\Pi}(v, t) := \mathbb{E} [e^{-v\lambda_t} \mid \mathcal{F}_0]$, then, we have the first-order PDE

$$\frac{\partial \hat{\Pi}(v, t)}{\partial t} = \left(-a\delta v + \varrho[\hat{h}(v) - 1] \right) u(v, t) - \left(\delta v + \hat{g}(v) - 1 + \frac{1}{2}\sigma^2 v^2 \right) \frac{\partial \hat{\Pi}(v, t)}{\partial v},$$

with the boundary conditions $\hat{\Pi}(0, t) = 1$ and $\hat{\Pi}(v, 0) = e^{-v\lambda_0}$. Because of the stationarity, $\hat{\Pi}(v, t)$ should be independent of time t , i.e. $\hat{\Pi}(v, t) = \hat{\Pi}(v)$ for any t , so $\frac{\partial}{\partial t} \hat{\Pi}(v) = 0$, then, we have the ODE

$$\left(-a\delta v + \varrho[\hat{h}(v) - 1] \right) \hat{\Pi}(v) - \left(\delta v + \hat{g}(v) - 1 + \frac{1}{2}\sigma^2 v^2 \right) \frac{d\hat{\Pi}(v)}{dv} = 0. \quad (3.13)$$

Given the initial condition $\hat{\Pi}(0) = \int_0^\infty \Pi(\lambda) d\lambda = 1$, we have the solution (3.12). Since Π is the unique solution to the ODE (3.13), we have the stationarity of intensity process $\{\lambda_t\}_{t \geq 0}$. \square

Now, we investigate the asymptotics of distribution Π via its Laplace transform (3.12).

Theorem 3.3. *We have the asymptotics of stationary distribution of intensity around zero,*

$$\Pi(\lambda) \sim \frac{\epsilon^{\frac{2\delta}{\sigma^2}a}}{\Gamma\left(\frac{2\delta}{\sigma^2}a\right)} \hat{\Pi}(\epsilon) \mathcal{E}(\epsilon) \lambda^{\left(\frac{2\delta}{\sigma^2}a-1\right)}, \quad \lambda \rightarrow 0, \quad (3.14)$$

where ϵ is any positive constant and

$$\mathcal{E}(\epsilon) := \exp \left(- \int_\epsilon^\infty \frac{\varrho[1 - \hat{h}(u)]u - \frac{2\delta}{\sigma^2}a[\delta u - \hat{g}(u) - 1]}{(\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2)u} du \right) < \infty, \quad \epsilon > 0.$$

Proof. By convergence test

$$\lim_{u \rightarrow \infty} \frac{\frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2}}{\frac{a\delta}{\delta + \frac{1}{2}\sigma^2 u}} = \lim_{u \rightarrow \infty} \frac{u + \frac{\varrho}{a\delta}[1 - \hat{h}(u)]}{u + \frac{\hat{g}(u) - 1}{\delta + \frac{1}{2}\sigma^2 u}} = \lim_{u \rightarrow \infty} \frac{\left(u + \frac{\varrho}{a\delta}[1 - \hat{h}(u)]\right)'}{\left(u + \frac{\hat{g}(u) - 1}{\delta + \frac{1}{2}\sigma^2 u}\right)'} = 1,$$

since

$$\lim_{v \rightarrow \infty} \int_0^v \frac{a\delta}{\delta + \frac{1}{2}\sigma^2 u} du = \infty,$$

we have

$$\lim_{v \rightarrow \infty} \int_0^v \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du = \infty.$$

For $\epsilon > 0$, we have

$$\begin{aligned} & \epsilon^{-\frac{2\delta}{\sigma^2}a} \lim_{v \rightarrow \infty} v^{\frac{2\delta}{\sigma^2}a} \hat{\Pi}(v) \\ &= \lim_{v \rightarrow \infty} \exp \left(- \int_0^v \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du + \frac{2\delta}{\sigma^2}a \int_\epsilon^v \frac{1}{u} du \right) \\ &= \exp \left(- \int_0^\epsilon \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du \right) \\ &\quad \times \lim_{v \rightarrow \infty} \exp \left(- \int_\epsilon^v \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du + \frac{2\delta}{\sigma^2}a \int_\epsilon^v \frac{1}{u} du \right) \\ &= \hat{\Pi}(\epsilon) \lim_{v \rightarrow \infty} \exp \left(- \int_\epsilon^v \left[\frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} - \frac{\frac{2\delta}{\sigma^2}a}{u} \right] du \right) \\ &= \hat{\Pi}(\epsilon) \lim_{v \rightarrow \infty} \exp \left(- \int_\epsilon^v \frac{\varrho[1 - \hat{h}(u)]u - \frac{2\delta}{\sigma^2}a[\delta u - \hat{g}(u) - 1]}{(\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2)u} du \right) \\ &= \hat{\Pi}(\epsilon) \mathcal{E}(\epsilon). \end{aligned}$$

Hence,

$$\lim_{v \rightarrow \infty} v^{\frac{2\delta}{\sigma^2}a} \hat{\Pi}(v) = \epsilon^{\frac{2\delta}{\sigma^2}a} \hat{\Pi}(\epsilon) \mathcal{E}(\epsilon),$$

i.e.

$$\hat{\Pi}(v) \sim \epsilon^{\frac{2\delta}{\sigma^2}a} \hat{\Pi}(\epsilon) \mathcal{E}(\epsilon) v^{-\frac{2\delta}{\sigma^2}a}, \quad v \rightarrow \infty,$$

and by *Tauberian Theorem* (Feller, 1971), we have (3.14). \square

Remark 3.1. If $\frac{2\delta}{\sigma^2}a > 1$, then, $\lim_{u \rightarrow 0} \Pi(u) = 0$ and there is no mass at zero for distribution Π . The Feller's condition $\frac{2\delta}{\sigma^2}a > 1$ is also a well known condition that guarantees CIR process positive with probability one (Feller, 1951).

If the sizes of two types of jumps follow exponential distributions, the explicit expressions for the Laplace transforms of asymptotic/stationary λ_t can be derived, and for

some special cases, the exact distributions can further be identified.

Corollary 3.3. *For the special case of pure diffusion, i.e. without externally excited and self-excited jumps, we have*

$$\{\lambda_t\}_{t \geq 0} \sim \text{Gamma}\left(\frac{2\delta}{\sigma^2}a, \frac{2\delta}{\sigma^2}\right).$$

Proof. By Theorem 3.2, we have

$$\hat{\Pi}(v) = \exp\left(-\int_0^v \frac{a\delta u}{\delta u + \frac{1}{2}\sigma^2 u^2} du\right) = \exp\left(-\frac{2\delta}{\sigma^2}a \int_0^v \frac{1}{\frac{2\delta}{\sigma^2} + u} du\right) = \left(\frac{\frac{2\delta}{\sigma^2}}{v + \frac{2\delta}{\sigma^2}}\right)^{\frac{2\delta}{\sigma^2}a}.$$

□

Corollary 3.4. *For the special case without self-excited jumps, assume $H \sim \text{Exp}(\alpha)$, we have*

$$\hat{\Pi}(v) = \left(\frac{\frac{2\delta}{\sigma^2}}{v + \frac{2\delta}{\sigma^2}}\right)^{\left(\frac{2\delta}{\sigma^2}a - \frac{2\varrho}{2\delta - \alpha\sigma^2}\right)} \left(\frac{\alpha}{\alpha + v}\right)^{\frac{2\varrho}{2\delta - \alpha\sigma^2}}.$$

Proof. By Theorem 3.2, we have

$$\begin{aligned} \hat{\Pi}(v) &= \exp\left(-\int_0^v \frac{a\delta u}{\delta u + \frac{1}{2}\sigma^2 u^2} du\right) \exp\left(-\int_0^v \frac{\varrho\left(1 - \frac{\alpha}{\alpha + u}\right)}{\delta u + \frac{1}{2}\sigma^2 u^2} du\right) \\ &= \left(\frac{\frac{2\delta}{\sigma^2}}{v + \frac{2\delta}{\sigma^2}}\right)^{\frac{2\delta}{\sigma^2}a} \left(\frac{\frac{2\delta}{\sigma^2}}{v + \frac{2\delta}{\sigma^2}}\right)^{-\frac{2\varrho}{2\delta - \alpha\sigma^2}} \left(\frac{\alpha}{\alpha + v}\right)^{\frac{2\varrho}{2\delta - \alpha\sigma^2}}. \end{aligned}$$

□

Corollary 3.5. *For the special case without externally excited jumps, assume $G \sim \text{Exp}(\beta)$ and $\delta\beta > 1$, we have*

$$\{\lambda_t\}_{t \geq 0} \sim \text{Gamma}\left(\frac{2a\delta}{\sigma^2}w_1, -u_-\right) + \text{Gamma}\left(\frac{2a\delta}{\sigma^2}w_2, -u_+\right),$$

where constants $w_1, w_2 > 0, u_-, u_+ < 0$,

$$w_1 := \frac{u_- + \beta}{u_- - u_+}, \quad w_2 := -\frac{u_+ + \beta}{u_- - u_+}, \quad u_{\pm} := \frac{-\left(\frac{2\delta}{\sigma^2} + \beta\right) \pm \sqrt{\left(\frac{2\delta}{\sigma^2} - \beta\right)^2 + \frac{8}{\sigma^2}}}{2}.$$

Proof. By Theorem 3.2, we have

$$\hat{\Pi}(v) = \exp\left(-\int_0^v \frac{a\delta u}{\delta u + \frac{\beta}{\beta + u} - 1 + \frac{1}{2}\sigma^2 u^2} du\right) = \left(\frac{-u_-}{v - u_-}\right)^{\frac{2a\delta}{\sigma^2}w_1} \left(\frac{-u_+}{v - u_+}\right)^{\frac{2a\delta}{\sigma^2}w_2}.$$

Denote $f(u) = u^2 + \left(\frac{2\delta}{\sigma^2} + \beta\right)u + \frac{2}{\sigma^2}(\delta\beta - 1)$. u_- and u_+ are the two solutions to $f(u) = 0$

under the condition $\delta\beta > 1$, and it is easy to check they are both negative. Also $f(-\beta) = -\frac{2}{\sigma^2} < 0$, we have $u_- < -\beta < u_+ < 0$, then $w_1, w_2 > 0$. Note that, $w_1 + w_2 = 1$. \square

Remark 3.2. For the special case without diffusion, i.e. $\sigma = 0$, assume $H \sim \text{Exp}(\alpha)$, $G \sim \text{Exp}(\beta)$ and $\delta\beta > 1$, we have

$$\{\lambda_t\}_{t \geq 0} \sim \begin{cases} a + \tilde{\Gamma}_1 + \tilde{\Gamma}_2, & \text{for } \alpha \geq \beta, \\ a + \tilde{\Gamma}_3 + \tilde{B}, & \text{for } \alpha < \beta \text{ and } \alpha \neq \beta - \frac{1}{\delta}, \\ a + \tilde{\Gamma}_4 + \tilde{P}, & \text{for } \alpha = \beta - \frac{1}{\delta}, \end{cases}$$

where $\tilde{\Gamma}_1, \tilde{\Gamma}_2, \tilde{\Gamma}_3, \tilde{\Gamma}_4$ are different gamma random variables; \tilde{B} follows a compound negative binomial distribution with underlying exponential jumps; \tilde{P} follows a compound Poisson distribution with underlying exponential jumps. They are all independent of each other. This interesting result of explicit distributional decomposition in detail together with the associated proof is provided as Theorem 4.1. in Dassios and Zhao (2011).

Corollary 3.6. *For the general case, assume $H \sim \text{Exp}(\alpha)$, $G \sim \text{Exp}(\beta)$ and $\delta\beta > 1$, we have*

$$\hat{\Pi}(v) = \begin{cases} \left(\frac{\alpha}{v+\alpha} \right)^{\frac{2a\delta}{\sigma^2}(\omega_1 a_1 + \omega_2 b_1)} \times \left(\frac{-u_-}{v-u_-} \right)^{\frac{2a\delta}{\sigma^2}\omega_1 a_2} \times \left(\frac{-u_+}{v-u_+} \right)^{\frac{2a\delta}{\sigma^2}\omega_2 b_2}, & \alpha \neq -u_-, -u_+, \\ \left(\frac{\alpha}{v+\alpha} \right)^{\frac{2\delta a}{\sigma^2}(\omega_1 + \omega_2 b_1)} \times \left(\frac{-u_+}{v-u_+} \right)^{\frac{2a\delta}{\sigma^2}\omega_2 b_2} \times \exp\left(-\frac{2\delta a}{\sigma^2 \alpha} \omega_1 (\beta - \alpha) \frac{v}{\alpha+v}\right), & \alpha = -u_-, \\ \left(\frac{\alpha}{v+\alpha} \right)^{\frac{2a\delta}{\sigma^2}(\omega_1 a_1 + \omega_2)} \times \left(\frac{-u_-}{v-u_-} \right)^{\frac{2a\delta}{\sigma^2}\omega_1 a_2} \times \exp\left(-\frac{2\delta a}{\sigma^2 \alpha} \omega_2 (\beta - \alpha) \frac{v}{\alpha+v}\right), & \alpha = -u_+, \end{cases}$$

where

$$\omega_1 := -\frac{u_- + \alpha + \frac{\rho}{a\delta}}{u_+ - u_-}, \quad \omega_2 := \frac{u_+ + \alpha + \frac{\rho}{a\delta}}{u_+ - u_-}, \\ a_1 := \frac{\alpha - \beta}{\alpha + u_-}, \quad a_2 = \frac{u_- + \beta}{\alpha + u_-}, \quad b_1 := \frac{\alpha - \beta}{\alpha + u_+}, \quad b_2 = \frac{u_+ + \beta}{\alpha + u_+}.$$

Proof. By Theorem 3.2, we have

$$\hat{\Pi}(v) = \exp\left(-\frac{2\delta a}{\sigma^2} \omega_1 \int_0^v \frac{u + \beta}{(u + \alpha)(u - u_-)} du\right) \exp\left(-\frac{2\delta a}{\sigma^2} \omega_2 \int_0^v \frac{u + \beta}{(u + \alpha)(u - u_+)} du\right).$$

If $\alpha \neq -u_-$ and $\alpha \neq -u_+$, then,

$$\hat{\Pi}(v) = \left(\frac{\alpha}{v+\alpha} \right)^{\frac{2a\delta}{\sigma^2}\omega_1 a_1} \left(\frac{-u_-}{v-u_-} \right)^{\frac{2a\delta}{\sigma^2}\omega_1 a_2} \left(\frac{\alpha}{v+\alpha} \right)^{\frac{2a\delta}{\sigma^2}\omega_2 b_1} \left(\frac{-u_+}{v-u_+} \right)^{\frac{2a\delta}{\sigma^2}\omega_2 b_2}.$$

If $\alpha = -u_-$ or $\alpha = -u_+$, we have

$$\int_0^v \frac{u + \beta}{(u + \alpha)^2} du = \ln\left(\frac{v + \alpha}{\alpha}\right) + (\beta - \alpha) \left(\frac{1}{\alpha} - \frac{1}{v + \alpha}\right),$$

so, if $\alpha = -u_-$, then,

$$\begin{aligned}\hat{\Pi}(v) &= \exp\left(-\frac{2\delta a}{\sigma^2}\omega_1 \int_0^v \frac{u+\beta}{(u+\alpha)^2} du\right) \exp\left(-\frac{2\delta a}{\sigma^2}\omega_2 \int_0^v \frac{u+\beta}{(u+\alpha)(u-u_+)} du\right) \\ &= \left(\frac{\alpha}{v+\alpha}\right)^{\frac{2\delta a}{\sigma^2}\omega_1} \exp\left(-\frac{2\delta a}{\sigma^2\alpha}\omega_1(\beta-\alpha)\left(1-\frac{\alpha}{\alpha+v}\right)\right) \left(\frac{\alpha}{v+\alpha}\right)^{\frac{2\delta a}{\sigma^2}\omega_2 b_1} \left(\frac{-u_+}{v-u_+}\right)^{\frac{2\delta a}{\sigma^2}\omega_2 b_2};\end{aligned}$$

if $\alpha = -u_+$, then,

$$\begin{aligned}\hat{\Pi}(v) &= \exp\left(-\frac{2\delta a}{\sigma^2}\omega_1 \int_0^v \frac{u+\beta}{(u+\alpha)(u-u_-)} du\right) \exp\left(-\frac{2\delta a}{\sigma^2}\omega_2 \int_0^v \frac{u+\beta}{(u+\alpha)^2} du\right) \\ &= \left(\frac{\alpha}{v+\alpha}\right)^{\frac{2\delta a}{\sigma^2}\omega_1 a_1} \left(\frac{-u_-}{v-u_-}\right)^{\frac{2\delta a}{\sigma^2}\omega_1 a_2} \left(\frac{\alpha}{v+\alpha}\right)^{\frac{2\delta a}{\sigma^2}\omega_2} \exp\left(-\frac{2\delta a}{\sigma^2\alpha}\omega_2(\beta-\alpha)\left(1-\frac{\alpha}{\alpha+v}\right)\right).\end{aligned}$$

□

Remark 3.3. For *Corollary 3.6*, in particular, if $\alpha = \beta$, then we have

$$\hat{\Pi}(v) = \left(\frac{-u_-}{v-u_-}\right)^{\frac{2\delta a}{\sigma^2}\omega_1} \left(\frac{-u_+}{v-u_+}\right)^{\frac{2\delta a}{\sigma^2}\omega_2}.$$

3.3 Probability Generating Function of N_T

Based on *Lemma 3.1*, we can derive the distributional properties of point process $\{N_t\}_{t \geq 0}$.

Theorem 3.4. For $\kappa > 0$, the probability generating function of N_T conditional on λ_0 and $N_0 = 0$ is given by

$$\phi_T(\theta) := \mathbb{E}[\theta^{N_T} \mid \lambda_0] = \exp\left(-\int_0^{\mathcal{G}_{0,\theta}^{-1}(T)} \frac{a\delta u + \varrho[1-\hat{h}(u)]}{1-\delta u - \theta\hat{g}(u) - \frac{1}{2}\sigma^2 u^2} du\right) \times e^{-\mathcal{G}_{0,\theta}^{-1}(T)\lambda_0}, \quad (3.15)$$

where

$$\mathcal{G}_{0,\theta}(L) := \int_0^L \frac{du}{1-\delta u - \theta\hat{g}(u) - \frac{1}{2}\sigma^2 u^2}, \quad \theta \in [0, 1). \quad (3.16)$$

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

$$\mathbb{E}[\theta^{N_T} \mid \mathcal{F}_0] = e^{-\left(c(T)-c(0)\right)} e^{-B(0)\lambda_0}, \quad (3.17)$$

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

$$-B'(t) + \delta B(t) + \theta\hat{g}(B(t)) - 1 + \frac{1}{2}\sigma^2 B^2(t) = 0,$$

with the boundary condition $B(T) = 0$. Under the condition $\delta > \mu_{1_G}$, it can be solved for

$\sigma > 0$ by the following steps:

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

$$\frac{dL(\tau)}{d\tau} = 1 - \delta L(\tau) - \theta \hat{g}(L(\tau)) - \frac{1}{2}\sigma^2 L^2(\tau), \quad \theta \in [0, 1), \quad (3.18)$$

with the initial condition $L(0) = 0$; we define the right-hand side of (3.18) as

$$f_2(L) := 1 - \delta L - \theta \hat{g}(L) - \frac{1}{2}\sigma^2 L^2. \quad (3.19)$$

2. There is only one positive singular point to the equation $f_2(L) = 0$, which is denoted by $v^* = v^*(\theta) > 0$. This is because,

- for the case $0 < \theta < 1$, the equation $f_2(L) = 0$ is equivalent to

$$\hat{g}(u) = \frac{1}{\theta} \left(1 - \delta u - \frac{1}{2}\sigma^2 u^2 \right), \quad \theta \in (0, 1),$$

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 $\theta = 0$, the equation $f_2(L) = 0$ is equivalent to

$$1 - \delta u - \frac{1}{2}\sigma^2 u^2 = 0,$$

which has only one positive solution

$$v^* := \frac{-\delta + \sqrt{\delta^2 + 2\sigma^2}}{\sigma^2} > 0;$$

and for both cases,

$$0 < \frac{-\delta + \sqrt{\delta^2 + 2\sigma^2(1 - \theta)}}{\sigma^2} < v^* \leq \frac{-\delta + \sqrt{\delta^2 + 2\sigma^2}}{\sigma^2}; \quad (3.20)$$

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

3. Rewrite (3.18) as

$$\frac{dL}{1 - \delta L - \theta \hat{g}(L) - \frac{1}{2}\sigma^2 L^2} = d\tau,$$

and integrate, we have

$$\int_0^L \frac{du}{1 - \delta u - \theta \hat{g}(u) - \frac{1}{2}\sigma^2 u^2} = \tau, \quad L \in [0, v^*).$$

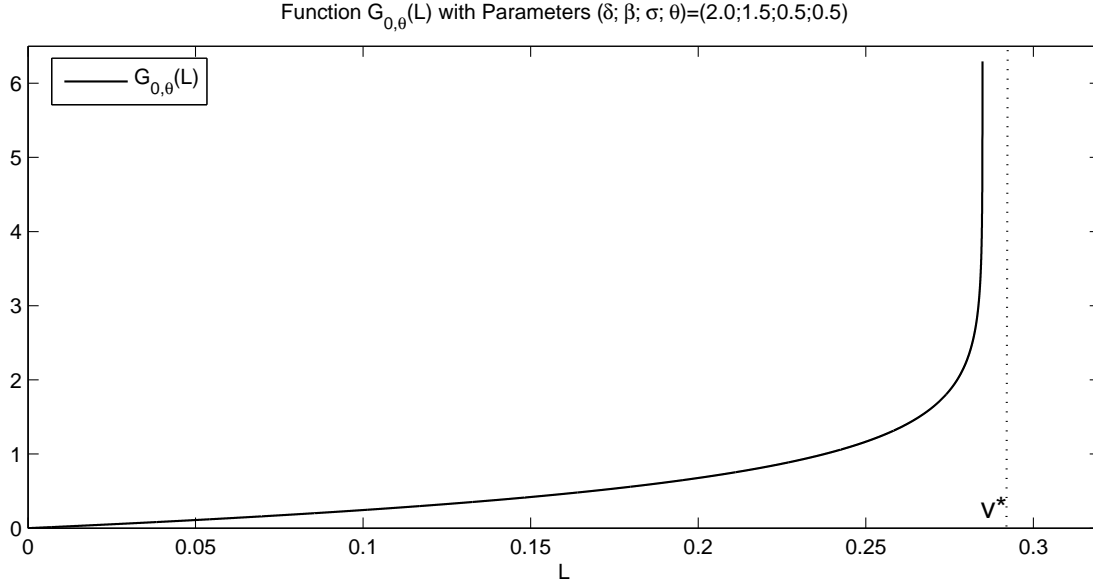


Figure 2: Numerical example: the function $\mathcal{G}_{0,\theta}(L)$ of (3.16) when $G \sim \text{Exp}(\beta)$ with parameters $(\delta; \beta; \sigma; \theta) = (2.0; 1.5; 0.5; 0.5)$ and $v^* = 0.2848$

Define the function on the left-hand side as (3.16), then, $\mathcal{G}_{0,\theta}(L) = \tau$. Note that, $\int_0^{v^*} \frac{1}{u - v^*} du = \infty$, $\lim_{L \uparrow v^*} \mathcal{G}_{0,\theta}(L) = \infty$; hence, as $L \rightarrow 0$ when $\tau \rightarrow 0$, and $L \rightarrow v^*$ when $\tau \rightarrow \infty$; the integrand is positive in the domain $u \in (0, v^*)$, and $\mathcal{G}_{0,\theta}(L)$ is a strictly increasing function of L , $L \in (0, v^*)$, see a numerical example of function $\mathcal{G}_{0,\theta}(L)$ in Figure 2; therefore, $\mathcal{G}_{0,\theta}(L) : [0, v^*) \rightarrow [0, \infty)$ is a well defined function, and its inverse function $\mathcal{G}_{0,\theta}^{-1}(\tau) : [0, \infty) \rightarrow [0, v^*)$ exists.

4. The unique solution is found by $L(\tau) = \mathcal{G}_{0,\theta}^{-1}(\tau)$, or, $B(t) = \mathcal{G}_{0,\theta}^{-1}(T - t)$.
5. $B(0)$ is obtained by $B(0) = L(T) = \mathcal{G}_{0,\theta}^{-1}(T)$.

Then, $c(T) - c(0)$ is determined by

$$c(T) - c(0) = a\delta \int_0^T \mathcal{G}_{0,\theta}^{-1}(\tau) d\tau + \varrho \int_0^T [1 - \hat{h}(\mathcal{G}_{0,\theta}^{-1}(\tau))] d\tau, \quad (3.21)$$

where, by change of variable,

$$\begin{aligned} \int_0^T \mathcal{G}_{0,\theta}^{-1}(\tau) d\tau &= \int_0^{\mathcal{G}_{0,\theta}^{-1}(T)} \frac{u}{1 - \delta u - \theta \hat{g}(u) - \frac{1}{2}\sigma^2 u^2} du, \\ \int_0^T [1 - \hat{h}(\mathcal{G}_{0,\theta}^{-1}(\tau))] d\tau &= \int_0^{\mathcal{G}_{0,\theta}^{-1}(T)} \frac{1 - \hat{h}(u)}{1 - \delta u - \theta \hat{g}(u) - \frac{1}{2}\sigma^2 u^2} du. \end{aligned}$$

Finally, substitute $B(0)$ and $c(T) - c(0)$ into (3.17), and the result follows. \square

Corollary 3.7. *For the special case of zero reversion level and no external excitement, i.e. $a = \varrho = 0$, we have*

$$\mathbb{E}[\theta^{N_T} \mid \lambda_0] = e^{-\mathcal{G}_{0,\theta}^{-1}(T)\lambda_0}, \quad \mathbb{E}[\theta^{N_\infty} \mid \lambda_0] = e^{-v^*\lambda_0}.$$

where the function $\mathcal{G}_{0,\theta}(\cdot)$ is given by (3.16) and v^* is the unique positive solution to (3.19).

Proof. Set $a = \varrho = 0$ in Theorem 3.4 and $T \rightarrow \infty$, and the results follow immediately. \square

3.4 Moments of λ_T and N_T

The moments of λ_t and N_t can be derived by differentiating the Laplace transform of λ_t in Section 3.2 and the probability generating function of N_t in Section 3.3. Alternatively, they can be obtained by solving the ODEs as below.

3.4.1 Moments of Intensity Process λ_t

Theorem 3.5. *The expectation of λ_t conditional on λ_0 is given by*

$$\mu_1(t; \lambda_0) := \mathbb{E}[\lambda_t \mid \lambda_0] = \begin{cases} \frac{\mu_{1_H}\varrho + a\delta}{\kappa} + \left(\lambda_0 - \frac{\mu_{1_H}\varrho + a\delta}{\kappa}\right) e^{-\kappa t}, & \kappa \neq 0, \\ \lambda_0 + (\mu_{1_H}\varrho + a\delta)t, & \kappa = 0. \end{cases} \quad (3.22)$$

Proof. By the martingale property of infinitesimal generator (2.2), we have a \mathcal{F} -martingale $f(\lambda_t, N_t, t) - f(\lambda_0, N_0, 0) - \int_0^t \mathcal{A}f(\lambda_s, N_s, s)ds$ for $f \in \Omega(\mathcal{A})$. Set $f(\lambda, n, t) = \lambda$, we have

$$\mathcal{A}\lambda = -\kappa\lambda + \mu_{1_H}\varrho + a\delta,$$

and $\lambda_t - \lambda_0 - \int_0^t \mathcal{A}\lambda_s ds$ is a \mathcal{F} -martingale, then we have $\mathbb{E}\left[\lambda_t - \lambda_0 - \int_0^t \mathcal{A}\lambda_s ds \mid \lambda_0\right] = 0$, and

$$\mathbb{E}[\lambda_t \mid \lambda_0] = \lambda_0 + \mathbb{E}\left[\int_0^t \mathcal{A}\lambda_s ds \mid \lambda_0\right] = \lambda_0 - \kappa \int_0^t \mathbb{E}[\lambda_s \mid \lambda_0] ds + (\mu_{1_H}\varrho + a\delta)t.$$

Differentiate it with respect to t , we obtain the non-linear inhomogeneous ODE

$$\mu_1'(t; \lambda_0) = -\kappa\mu_1(t; \lambda_0) + \mu_{1_H}\varrho + a\delta,$$

with the initial condition $\mu_1(0; \lambda_0) = \lambda_0$. This ODE has the solution of (3.22). \square

Theorem 3.6. *The second moment of λ_t conditional λ_0 is given by*

$$\mu_2(t; \lambda_0) := \mathbb{E} [\lambda_t^2 | \lambda_0] = \begin{cases} \lambda_0^2 e^{-2\kappa t} + \frac{2(\mu_{1_H} \varrho + a\delta) + \mu_{2_G} + \sigma^2}{\kappa} \left(\lambda_0 - \frac{\mu_{1_H} \varrho + a\delta}{\kappa} \right) (e^{-\kappa t} - e^{-2\kappa t}) \\ \quad + \left[\frac{2(\mu_{1_H} \varrho + a\delta) + \mu_{2_G} + \sigma^2}{2\kappa^2} (\mu_{1_H} \varrho + a\delta) + \frac{\mu_{2_H} \varrho}{2\kappa} \right] (1 - e^{-2\kappa t}), & \kappa \neq 0, \\ \lambda_0^2 + \left[2(\mu_{1_H} \varrho + a\delta) + \mu_{2_G} + \sigma^2 \right] \left(\lambda_0 t + \frac{1}{2} (\mu_{1_H} \varrho + a\delta) t^2 \right) + \mu_{2_H} \varrho t, & \kappa = 0. \end{cases} \quad (3.23)$$

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

$$\mathcal{A}\lambda^2 = -2\kappa\lambda^2 + [2(\mu_{1_H} \varrho + a\delta) + \mu_{2_G} + \sigma^2] \lambda + \mu_{2_H} \varrho.$$

Since $\lambda_t^2 - \lambda_0^2 - \int_0^t \mathcal{A}(\lambda_s^2) ds$ is a \mathcal{F} -martingale, we have $\mathbb{E} \left[\lambda_t^2 - \int_0^t \mathcal{A}(\lambda_s^2) ds | \lambda_0 \right] = \lambda_0^2$, and

$$\mathbb{E} [\lambda_t^2 | \lambda_0] = \lambda_0^2 - 2\kappa \int_0^t \mathbb{E} [\lambda_s^2 | \lambda_0] ds + \left[2(\mu_{1_H} \varrho + a\delta) + \mu_{2_G} + \sigma^2 \right] \int_0^t \mathbb{E} [\lambda_s | \lambda_0] ds + \mu_{2_H} \varrho t.$$

Differentiate it with respect to t , we have the ODE

$$\begin{aligned} \mu_2'(t; \lambda_0) + 2\kappa\mu_2(t; \lambda_0) &= \left[2(\mu_{1_H} \varrho + a\delta) + \mu_{2_G} + \sigma^2 \right] \left(\lambda_0 - \frac{\mu_{1_H} \varrho + a\delta}{\delta - \mu_{1_G}} \right) e^{-\kappa t} \\ &\quad + \frac{\left[2(\mu_{1_H} \varrho + a\delta) + \mu_{2_G} + \sigma^2 \right] (\mu_{1_H} \varrho + a\delta)}{\kappa} + \mu_{2_H} \varrho, \end{aligned}$$

with the initial condition $\mu_2(0; \lambda_0) = \lambda_0^2$. This ODE has the solution of (3.23). \square

Corollary 3.8. *The variance of λ_t conditional on λ_0 is given by*

$$\text{Var} [\lambda_t | \lambda_0] = \begin{cases} \frac{1}{2\kappa} \left[\frac{(\mu_{2_G} + \sigma^2)(\mu_{1_H} \varrho + a\delta)}{\kappa} - \mu_{2_H} \varrho - 2(\mu_{2_G} + \sigma^2) \lambda_t \right] e^{-2\kappa t} \\ \quad + \frac{\mu_{2_G} + \sigma^2}{\kappa} \left(\lambda_0 - \frac{\mu_{1_H} \varrho + a\delta}{\kappa} \right) e^{-\kappa t} + \frac{1}{2\kappa} \left[\mu_{2_H} \varrho + \frac{(\mu_{2_G} + \sigma^2)(\mu_{1_H} \varrho + a\delta)}{\kappa} \right], & \kappa \neq 0, \\ \frac{1}{2} (\mu_{2_G} + \sigma^2) (\mu_{1_H} \varrho + a\delta) t^2 + [(\mu_{2_G} + \sigma^2) \lambda_0 + \mu_{2_H} \varrho] t, & \kappa = 0. \end{cases}$$

Proof. $\text{Var} [\lambda_t | \lambda_0] = \mathbb{E} [\lambda_t^2 | \lambda_0] - (\mathbb{E} [\lambda_t | \lambda_0])^2$ where $\mathbb{E} [\lambda_t^2 | \lambda_0]$ and $\mathbb{E} [\lambda_t | \lambda_0]$ are given by Theorem 3.6 and Theorem 3.5. \square

Corollary 3.9. *Under the condition $\kappa > 0$, the asymptotic first and second moments of intensity*

level λ_t are given by

$$\begin{aligned}\mu_1 &:= \mathbb{E}[\lambda_t] = \lim_{t \rightarrow \infty} \mathbb{E}[\lambda_t \mid \lambda_0] = \frac{\mu_{1_H} \varrho + a\delta}{\kappa}, \\ \mu_2 &:= \mathbb{E}[\lambda_t^2] = \lim_{t \rightarrow \infty} \mathbb{E}[\lambda_t^2 \mid \lambda_0] = \frac{[2(\mu_{1_H} \varrho + a\delta) + \mu_{2_G} + \sigma^2](\mu_{1_H} \varrho + a\delta)}{2\kappa^2} + \frac{\mu_{2_H} \varrho}{2\kappa}.\end{aligned}$$

3.4.2 Moments of Point Process N_t

Theorem 3.7. *The expectation of N_t conditional on $N_0 = 0$ and λ_0 is given by*

$$\nu_1(t; \lambda_0) := \mathbb{E}[N_t \mid \lambda_0] = \begin{cases} \mu_1 t + (\lambda_0 - \mu_1) \frac{1}{\kappa} (1 - e^{-\kappa t}), & \kappa \neq 0, \\ \lambda_0 t + \frac{1}{2} (\mu_{1_H} \varrho + a\delta) t^2, & \kappa = 0. \end{cases} \quad (3.24)$$

Proof. By setting $f(\lambda, n, t) = n$ in (2.2), we have $\mathcal{A}n = \lambda$. Since $N_t - N_0 - \int_0^t \lambda_s ds$ is a martingale, we have

$$\mathbb{E}[N_t \mid \lambda_0] = \mathbb{E}\left[\int_0^t \lambda_s ds \mid \lambda_0\right] = \int_0^t \mathbb{E}[\lambda_s \mid \lambda_0] ds,$$

where $\mathbb{E}[\lambda_s \mid \lambda_0]$ is given by Theorem 3.5. □

Lemma 3.2. *The expectation of $N_t \lambda_t$ conditional on $N_0 = 0$ and λ_0 is given by*

$$\vartheta(t; \lambda_0) := \mathbb{E}[N_t \lambda_t \mid \lambda_0] = e^{-\kappa t} \int_0^t e^{\kappa s} p(s; \lambda_0) ds, \quad (3.25)$$

where

$$p(t; \lambda_0) := (\mu_{1_H} \varrho + a\delta) \nu_1(t; \lambda_0) + \mu_2(t; \lambda_0) + \mu_{1_G} \mu_1(t; \lambda_0).$$

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

$$\mathcal{A}(n\lambda) = -\kappa \lambda n + (\mu_{1_H} \varrho + a\delta) n + \lambda^2 + \mu_{1_G} \lambda,$$

then,

$$\mathbb{E}[N_t \lambda_t \mid \lambda_0] = \mathbb{E}\left[\int_0^t \left(-\kappa \lambda_u N_u + (\mu_{1_H} \varrho + a\delta) N_u + \lambda_u^2 + \mu_{1_G} \lambda_u\right) du \mid \lambda_0\right].$$

Differentiate it w.r.t. t , we have the ODE

$$\frac{d}{dt} \vartheta(t; \lambda_0) = -\kappa \vartheta(t; \lambda_0) + (\mu_{1_H} \varrho + a\delta) \mathbb{E}[N_t \mid \lambda_0] + \mathbb{E}[\lambda_t^2 \mid \lambda_0] + \mu_{1_G} \mathbb{E}[\lambda_t \mid \lambda_0],$$

with the initial condition $\vartheta(0; \lambda_0) = 0$. More concisely, it can be expressed by

$$\frac{d}{dt}\vartheta(t; \lambda_0) = -\kappa\vartheta(t; \lambda_0) + p(t; \lambda_0).$$

Solve this ODE, we have the solution (3.25). □

Theorem 3.8. *The second moment of N_t conditional on $N_0 = 0$ and λ_0 is given by*

$$v_2(t; \lambda_0) := \mathbb{E}[N_t^2 | \lambda_0] = 2 \int_0^t \vartheta(s; \lambda_0) ds + \int_0^t \mu_1(s; \lambda_0) ds. \quad (3.26)$$

Proof. Set $f(\lambda, n, t) = n^2$ in (2.2), we have $\mathcal{A}(n^2) = (2n+1)\lambda$. Since $N_t^2 - N_0^2 - \int_0^t (2N_s + 1)\lambda_s ds$, we have

$$\mathbb{E}[N_t^2 | \lambda_0] = 2 \int_0^t \mathbb{E}[N_s \lambda_s | \lambda_0] ds + \int_0^t \mathbb{E}[\lambda_s | \lambda_0] ds,$$

which can be expressed by (3.26). □

Based on Theorem 3.5 and Theorem 3.8, it is straightforward to derive the variance of N_t by

$$\text{Var}[N_t | \lambda_0] = v_2(t; \lambda_0) - v_1^2(t; \lambda_0).$$

All of the moments of N_t given the formulas above can be easily calculated explicitly, but their expressions would be very long with various simple exponential functions. To save the space, we just leave their concise expressions there.

4 Change of Measure

In this section, we develop a simple method of change of measure for the joint process (λ_t, N_t) via *Esscher transform* (Gerber and Shiu, 1994) (or *exponential tilting*) and scaling the jump-size distributions. By appropriately choosing a set of parameters, it might be useful for pricing under alternative probability measures or improving simulation efficiency via importance sampling.

By Lemma 3.1, Theorem 3.1 and Theorem 3.4, we have a $\mathcal{F}_t^{\mathbb{P}}$ -martingale

$$e^{c(t)} \theta^{N_t} e^{-B(t)\lambda_t}, \quad (4.1)$$

where parameters $c(t)$ and $B(t)$ satisfy the equations

$$\begin{cases} -B'(t) + \delta B(t) + \theta \hat{g}(B(t)) - 1 + \frac{1}{2}\sigma^2 B^2(t) = 0, & (.1) \\ c'(t) + \varrho \hat{h}(B(t)) - \varrho - a\delta B(t) = 0. & (.2) \end{cases} \quad (4.2)$$

It can be uniquely determined for the following two cases (I, II) under the condition $\delta > \mu_{1_G}$ for $0 \leq t \leq T$:

I. $\theta = 1, B(T) = v > 0$:

$$B(t) \in (0, v], \quad c(t) \in \left[0, \int_0^v \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{\delta u + \hat{g}(u) - 1 + \frac{1}{2}\sigma^2 u^2} du\right), \quad t \in [0, T = \infty), \quad (4.3)$$

II. $0 \leq \theta < 1, B(T) = v = 0$:

$$B(t) \in [0, v^*), \quad c(t) \in \left[0, \int_0^{v^*} \frac{a\delta u + \varrho[1 - \hat{h}(u)]}{1 - \delta u - \theta \hat{g}(u) - \frac{1}{2}\sigma^2 u^2} du\right), \quad t \in [0, T = \infty), \quad (4.4)$$

where v^* is the unique positive solution to the equation

$$1 - \delta u - \theta \hat{g}(u) - \frac{1}{2}\sigma^2 u^2 = 0.$$

Theorem 4.1. Define an equivalent probability measure $\widetilde{\mathbb{P}}$, via the Radon-Nikodym derivative

$$\left. \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} \right|_{\mathcal{F}_t} := e^{c(t) - c(0)} \theta^{N_t - N_0} e^{-(B(t)\lambda_t - B(0)\lambda_0)}, \quad \theta \in [0, 1],$$

then, under the condition $\kappa > 0$, we have the parameter transformation for (N_t, λ_t) from $\mathbb{P} \rightarrow \widetilde{\mathbb{P}}$ by

- $a \rightarrow a\theta \hat{g}(B(t)) \left[1 + \frac{\sigma^2}{\delta} B(t)\right],$
- $\delta \rightarrow \frac{\delta}{1 + \frac{\sigma^2}{\delta} B(t)},$
- $\varrho \rightarrow \hat{h}(B(t))\varrho,$
- $h(u) \rightarrow \frac{\tilde{h}\left(\frac{1}{\theta \hat{g}(B(t))} u\right)}{\theta \hat{g}(B(t))},$
- $g(u) \rightarrow \frac{\tilde{g}\left(\frac{1}{\theta \hat{g}(B(t))} u\right)}{\theta \hat{g}(B(t))},$
- $\sigma \rightarrow \sigma,$

where

$$\tilde{h}(y) := \frac{e^{-B(t)y}}{\hat{h}(B(t))} \frac{1}{dy}, \quad \tilde{g}(y) := \frac{e^{-B(t)y}}{\hat{g}(B(t))} \frac{1}{dy}.$$

Proof. We use the martingale of (4.1) to define an equivalent martingale probability measure $\tilde{\mathbb{P}}$ via the Radon-Nikodym derivative

$$L_t := \left. \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} \right|_{\mathcal{F}_t} := \frac{e^{c(t)}\theta^{N_t}e^{-B(t)\lambda_t}}{\mathbb{E}[e^{c(t)}\theta^{N_t}e^{-B(t)\lambda_t}]} = e^{c(t)-c(0)}\theta^{N_t-N_0}e^{-(B(t)\lambda_t-B(0)\lambda_0)},$$

which is a $\mathcal{F}_t^{\mathbb{P}}$ -martingale with mean value 1. Let $\tilde{\mathcal{A}}$ be the generator and $\tilde{\mathbb{E}}$ be the expectation under the new measure $\tilde{\mathbb{P}}$. Based on the definition of *infinitesimal generator* (Øksendal, 2003), we have

$$\tilde{\mathcal{A}}\tilde{f} = \lim_{\Delta t \rightarrow 0} \frac{\tilde{\mathbb{E}}[\tilde{f}(t + \Delta t) | \mathcal{F}_t] - \tilde{f}(t)}{\Delta t}.$$

By the change of measure, we have

$$\tilde{\mathcal{A}}\tilde{f} = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{E}[e^{g(t+\Delta t)-g(t)}\tilde{f}(t + \Delta t) | \mathcal{F}_t] - \tilde{f}(t)}{\Delta t} = e^{-g}\mathcal{A}\{e^g\tilde{f}\},$$

where

$$e^{g(t)} = e^{c(t)}\theta^{N_t}e^{-B(t)\lambda_t}.$$

Set $f(\lambda, n, t) = e^{c(t)}\theta^n e^{-B(t)\lambda}\tilde{f}(\lambda, n, t)$ in the original generator (2.2), we have

$$\begin{aligned} \tilde{\mathcal{A}}\tilde{f} &= (c'(t) - B'(t)\lambda)\tilde{f} + \frac{\partial \tilde{f}}{\partial t} + \delta(a - \lambda) \left(-B(t)\tilde{f} + \frac{\partial \tilde{f}}{\partial \lambda} \right) + \frac{1}{2}\sigma^2\lambda \left(\frac{\partial^2 \tilde{f}}{\partial \lambda^2} - 2B(t)\frac{\partial \tilde{f}}{\partial \lambda} + B^2(t)\tilde{f} \right) \\ &+ \varrho \left[\int_0^\infty \tilde{f}(\lambda + y, n, t)e^{-B(t)y}dH(y) - \tilde{f}(\lambda, n, t) \right] \\ &+ \lambda \left[\theta \int_0^\infty \tilde{f}(\lambda + y, n + 1, t)e^{-B(t)y}dG(y) - \tilde{f}(\lambda, n, t) \right]. \end{aligned}$$

Given the parameter relationship by (4.2) without explicitly solving the equations, we can implement the *Esscher Transform*

$$d\tilde{H}(y) := \frac{e^{-B(t)y}}{\hat{h}(B(t))}dH(y), \quad d\tilde{G}(y) := \frac{e^{-B(t)y}}{\hat{g}(B(t))}dG(y),$$

then,

$$\begin{aligned} \tilde{\mathcal{A}}\tilde{f} &= \frac{\partial \tilde{f}}{\partial t} + \delta \left[a - \left(1 + \frac{\sigma^2}{\delta} B(t) \right) \lambda \right] \frac{\partial \tilde{f}}{\partial \lambda} + \frac{1}{2}\sigma^2\lambda \frac{\partial^2 \tilde{f}}{\partial \lambda^2} + \hat{h}(B(t))\varrho \left[\int_0^\infty \tilde{f}(\lambda + y, n, t)d\tilde{H}(y) - \tilde{f}(\lambda, n, t) \right] \\ &+ \theta\hat{g}(B(t))\lambda \left[\int_0^\infty \tilde{f}(\lambda + y, n + 1, t)d\tilde{G}(y) - \tilde{f}(\lambda, n, t) \right]. \end{aligned}$$

Let $\tilde{\lambda} = \theta \hat{g}(B(t))\lambda$, we have

$$\begin{aligned}\tilde{\mathcal{A}}\tilde{f} &= \frac{\partial \tilde{f}}{\partial t} + \frac{\delta}{1 + \frac{\sigma^2}{\delta}B(t)} \left[a\theta \hat{g}(B(t)) \left(1 + \frac{\sigma^2}{\delta}B(t) \right) - \tilde{\lambda} \right] \frac{\partial \tilde{f}}{\partial \tilde{\lambda}} + \frac{1}{2}\sigma^2 \tilde{\lambda} \frac{\partial^2 \tilde{f}}{\partial \tilde{\lambda}^2} \\ &\quad + \hat{h}(B(t))\varrho \left[\int_0^\infty \tilde{f}(\tilde{\lambda} + \theta \hat{g}(B(t))y, n, t) d\tilde{H}(y) - \tilde{f}(\tilde{\lambda}, n, t) \right] \\ &\quad + \tilde{\lambda} \left[\int_0^\infty \tilde{f}(\tilde{\lambda} + \theta \hat{g}(B(t))y, n+1, t) d\tilde{G}(y) - \tilde{f}(\tilde{\lambda}, n, t) \right].\end{aligned}$$

Change the variable by $u = \theta \hat{g}(B(t))y$, we have

$$\begin{aligned}\tilde{\mathcal{A}}\tilde{f} &= \frac{\partial \tilde{f}}{\partial t} + \frac{\delta}{1 + \frac{\sigma^2}{\delta}B(t)} \left[a\theta \hat{g}(B(t)) \left(1 + \frac{\sigma^2}{\delta}B(t) \right) - \tilde{\lambda} \right] \frac{\partial \tilde{f}}{\partial \tilde{\lambda}} + \frac{1}{2}\sigma^2 \tilde{\lambda} \frac{\partial^2 \tilde{f}}{\partial \tilde{\lambda}^2} \\ &\quad + \hat{h}(B(t))\varrho \left[\int_0^\infty \tilde{f}(\tilde{\lambda} + u, n, t) d\tilde{H}(y) - \tilde{f}(\tilde{\lambda}, n, t) \right] \\ &\quad + \tilde{\lambda} \left[\int_0^\infty \tilde{f}(\tilde{\lambda} + u, n+1, t) d\tilde{G}(y) - \tilde{f}(\tilde{\lambda}, n, t) \right].\end{aligned}$$

Since $d\tilde{H}(y) = \tilde{h}(y)dy$ and $d\tilde{G}(y) = \tilde{g}(y)dy$, finally, we have

$$\begin{aligned}\tilde{\mathcal{A}}\tilde{f} &= \frac{\partial \tilde{f}}{\partial t} + \frac{\delta}{1 + \frac{\sigma^2}{\delta}B(t)} \left[a\theta \hat{g}(B(t)) \left(1 + \frac{\sigma^2}{\delta}B(t) \right) - \tilde{\lambda} \right] \frac{\partial \tilde{f}}{\partial \tilde{\lambda}} + \frac{1}{2}\sigma^2 \tilde{\lambda} \frac{\partial^2 \tilde{f}}{\partial \tilde{\lambda}^2} \\ &\quad + \hat{h}(B(t))\varrho \left[\int_0^\infty \tilde{f}(\tilde{\lambda} + u, n, t) \frac{\tilde{h}\left(\frac{1}{\theta \hat{g}(B(t))}u\right)}{\theta \hat{g}(B(t))} du - \tilde{f}(\tilde{\lambda}, n, t) \right] \\ &\quad + \tilde{\lambda} \left[\int_0^\infty \tilde{f}(\tilde{\lambda} + u, n+1, t) \frac{\tilde{g}\left(\frac{1}{\theta \hat{g}(B(t))}u\right)}{\theta \hat{g}(B(t))} du - \tilde{f}(\tilde{\lambda}, n, t) \right].\end{aligned}\tag{4.5}$$

Therefore, we can uniquely specify the dynamics of the process under $\tilde{\mathbb{P}}$ measure based on (4.5). By comparing the original generator (2.2) with this new generator (4.5), it is straightforward to identify the parameter transformation from the original measure \mathbb{P} to the new measure $\tilde{\mathbb{P}}$ as given by *Theorem 4.1*. \square

Corollary 4.1. *If the condition holds under the original measure \mathbb{P} , i.e. $\delta > \mu_{1_G}$, then, it still holds under the new measure $\tilde{\mathbb{P}}$, i.e. $\tilde{\delta} > \mu_{1_{\tilde{G}}}$.*

Proof. Under the new measure $\tilde{\mathbb{P}}$, by the parameter transformation given by *Theorem 4.1* and change variable $y = \frac{1}{\theta \hat{g}(B(t))}u$, we have

$$\begin{aligned}\mu_{1_{\tilde{G}}} &= \int_0^\infty u \frac{\tilde{g}\left(\frac{1}{\theta \hat{g}(B(t))}u\right)}{\theta \hat{g}(B(t))} du \\ &= \int_0^\infty u \frac{1}{\theta \hat{g}(B(t))} \frac{e^{-B(t)\frac{1}{\theta \hat{g}(B(t))}u}}{\hat{g}(B(t))} \tilde{g}\left(\frac{1}{\theta \hat{g}(B(t))}u\right) du \\ &= \theta \int_0^\infty ye^{-B(t)y} dG(y).\end{aligned}$$

Since $0 \leq \theta \leq 1$ and $B(T) = v \geq 0$ as given by (4.3) and (4.4) and the condition holds under the measure \mathbb{P} , we have

$$\tilde{\delta} = \delta > \mu_{1_G} = \int_0^\infty y dG(y) > \theta \int_0^\infty y e^{-B(t)y} dG(y) = \mu_{1_{\tilde{G}}}.$$

□

Remark 4.1. In particular, we assume the jump-sizes follows exponential distributions, say, $H \sim \text{Exp}(\alpha)$ and $G \sim \text{Exp}(\beta)$, and the condition $\delta\beta > 1$. By *Theorem 4.1*, we have a nice explicit transformation:

- $a \rightarrow \frac{1 + \frac{\sigma^2}{\delta} B(t)}{\beta + B(t)} \theta \beta a,$
- $\delta \rightarrow \frac{\delta}{1 + \frac{\sigma^2}{\delta} B(t)},$
- $\varrho \rightarrow \frac{\alpha}{\alpha + B(t)} \varrho,$
- $H \sim \text{Exp}(\alpha) \rightarrow \text{Exp}\left(\frac{(\alpha + B(t))(\beta + B(t))}{\theta\beta}\right),$
- $G \sim \text{Exp}(\beta) \rightarrow \text{Exp}\left(\frac{(\beta + B(t))^2}{\theta\beta}\right),$
- $\sigma \rightarrow \sigma.$

5 Application to Finance: Probability of Default

We propose a generalised intensity-based model for modelling default probabilities, and extend the credit model of Dassios and Zhao (2011)⁸. We assume a final default or bankruptcy is caused by a number of adverse events relevant to the underlying company. These bad events could be, for instance, credit rating downgrades announced by rating agencies, or worse-than-expected financial reports, and the arrivals of these events often present contagion effects, i.e. one of events tends to trigger more of them. The frequency or intensity of the arrivals of these events depends on three key factors:

1. *internal risk factor*: a series of past credit events from the underlying company itself;
2. *external risk factor*: a series of other exogenous adverse events in the past that are independent of the company but common to the entire market;
3. *independent market noises*: a certain degree of noises that are persistently existing in the market and are time-varying with small fluctuations.

⁸This model, i.e. the general formula of (5.1) for survival probability, was also adopted by Aït-Sahalia et al. (2014) for pricing credit default swaps.

The arrivals of these events are modelled by our point process N_t with intensity λ_t of Definition 2.1. These three factors can be captured, as their impacts acting on its intensity are self-excited jumps $\{Y_k^{(2)}\}_{k=1,2,\dots}$, externally-excited jumps $\{Y_k^{(1)}\}_{k=1,2,\dots}$ and diffusion $\{W_t\}_{t \geq 0}$, respectively. We assume each jump, or bad event could cause default of a constant probability $d, 0 < d \leq 1$, as the company has a certain degree of resistance to survive through these bad events. Hence, $P_s(T)$, the survival probability of underlying company within the time period $[0, T]$ as seen from time 0, is simply

$$P_s(T) = \mathbb{E}[(1 - d)^{N_T} \mid \lambda_0], \quad d \in (0, 1]. \quad (5.1)$$

It can be calculated via (3.15) by setting $\theta = 1 - d$, i.e.

$$P_s(T) = \phi_T(1 - d). \quad (5.2)$$

This simple model goes beyond the standard credit models such as Duffie and Singleton (1999): the point process here is to model the arrivals of adverse credit events instead of defaults, and these credit events include defaults as the special cases. In particular, if we set $d = 100\%$ (which means that each bad event could cause default of 100% probability, i.e. each of the credit events is default), it recovers the standard model for credit defaults.

To illustrate its applications, we further assume the two types of jump-sizes both follow exponential distributions, i.e. $H \sim \text{Exp}(\alpha)$, $G \sim \text{Exp}(\beta)$ and $\delta\beta > 1$ with the parameter setting of

$$\Theta := (a, \varrho, \delta; \alpha, \beta, \sigma; \lambda_0) = (0.7, 0.5, 2.0; 2.0, 1.5, 0.5; 0.7).$$

Note that, the numerators and denominators of the integrand functions in (3.15) and (3.16) are simple polynomial functions and both of the integrands can be factorised by partial fraction expansion. Hence, we can explicitly calculate the survival probabilities $P_s(T)$ via (3.15). In fact, different levels of d correspond to different credit ratings, and they can be considered as the measures for the capability to avoid default. Hence, higher credit rating corresponds to lower value of d . For instance, the term structures of $P_s(T)$ for time from $T = 1$ to $T = 5$ and different levels of d are given by Table 1.

The sensitivity analysis is provided in Figure 3 for the survival probability at time $T = 1$ with fixed $d = 0.5$ against the varying parameters of

1. *mean-reverting level* $a \in (0, 2]$ for $\sigma \in \{1, 2, 3, 4\}$ with $(\varrho, \delta; \alpha, \beta; \lambda_0) = (0.5, 2.0; 2.0, 1.5; 0.7)$,
2. *mean-reverting rate* $\delta \in (0, 2]$ for $\sigma \in \{1, 2, 3, 4\}$ with $(a, \varrho; \alpha, \beta; \lambda_0) = (0.7, 0.5; 2.0, 5.0; 0.7)$,
3. *externally-excited jump rate* $\varrho \in (0, 2]$ for $\sigma \in \{1, 2, 3, 4\}$ with $(a, \delta; \alpha, \beta; \lambda_0) = (0.7, 2.0; 2.0, 1.5; 0.7)$,
4. *externally-excited jump mean* $\mu_{1_H} \in (0, 2]$ for $\sigma \in \{1, 2, 3, 4\}$ with $(a, \varrho, \delta; \beta; \lambda_0) = (0.7, 0.5, 2.0; 1.5; 0.7)$,
5. *self-excited jump mean* $\mu_{1_G} \in (0, 2]$ for $\sigma \in \{1, 2, 3, 4\}$ with $(a, \varrho, \delta; \alpha; \lambda_0) = (0.7, 0.5, 2.5; 2.0; 0.7)$,
6. *volatility of intensity diffusion* $\sigma \in (0, 5]$ for the *initial intensity* $\lambda_0 \in \{0.1, 0.5, 0.7, 1.0\}$ with $(a, \varrho, \delta; \alpha, \beta) = (0.7, 0.5, 2.0; 2.0, 1.5)$,

respectively. In particular, the *externally-excited jump mean* $\mu_{1_H} = 1/\alpha$ and the *self-excited jump mean* $\mu_{1_G} = 1/\beta$ can be used to measure the negative impacts from external and internal shocks, respectively. We should be aware that, the intensity process has two parts: diffusion and jumps. The parameter σ as defined in Definition 2.1 is not the volatility of the whole intensity process λ_t but the volatility of diffusion part only. From all of the subplots in Figure 3, we can observe the volatility σ poses positive effects, i.e. the survival probabilities are increasing functions of σ . This means that higher volatility in the intensity diffusion may lead to higher survival probability. This can be proved by the Taylor's expansion that, for a small d , we have the approximation

$$\mathbb{E}[(1-d)^{N_T} \mid \lambda_0] \approx 1 - d\mathbb{E}[N_T \mid \lambda_0] + \frac{1}{2}d^2\mathbb{E}[N_T^2 - N_T \mid \lambda_0].$$

Obviously, based on the moments derived in Section 3.4, the second term is independent of volatility σ , and the third term is a strictly increasing function of volatility σ . Hence, $\mathbb{E}[(1-d)^{N_T} \mid \lambda_0]$ is an increasing function of σ .

Table 1: Term Structure of survival probability $P_s(T); (a, \varrho, \delta; \alpha, \beta, \sigma; \lambda_0) = (0.7, 0.5, 2.0; 2.0, 1.5, 0.5; 0.7)$

Time T d	1	2	3	4	5
2%	98.15%	95.93%	93.65%	91.41%	89.22%
10%	91.27%	81.83%	73.07%	65.20%	58.15%
20%	83.70%	68.05%	55.01%	44.42%	35.85%
50%	66.09%	41.96%	26.49%	16.71%	10.54%
100%	47.13%	21.71%	9.99%	4.59%	2.11%

We could further relax the constant d to be variable, for instant, depending on N_t .

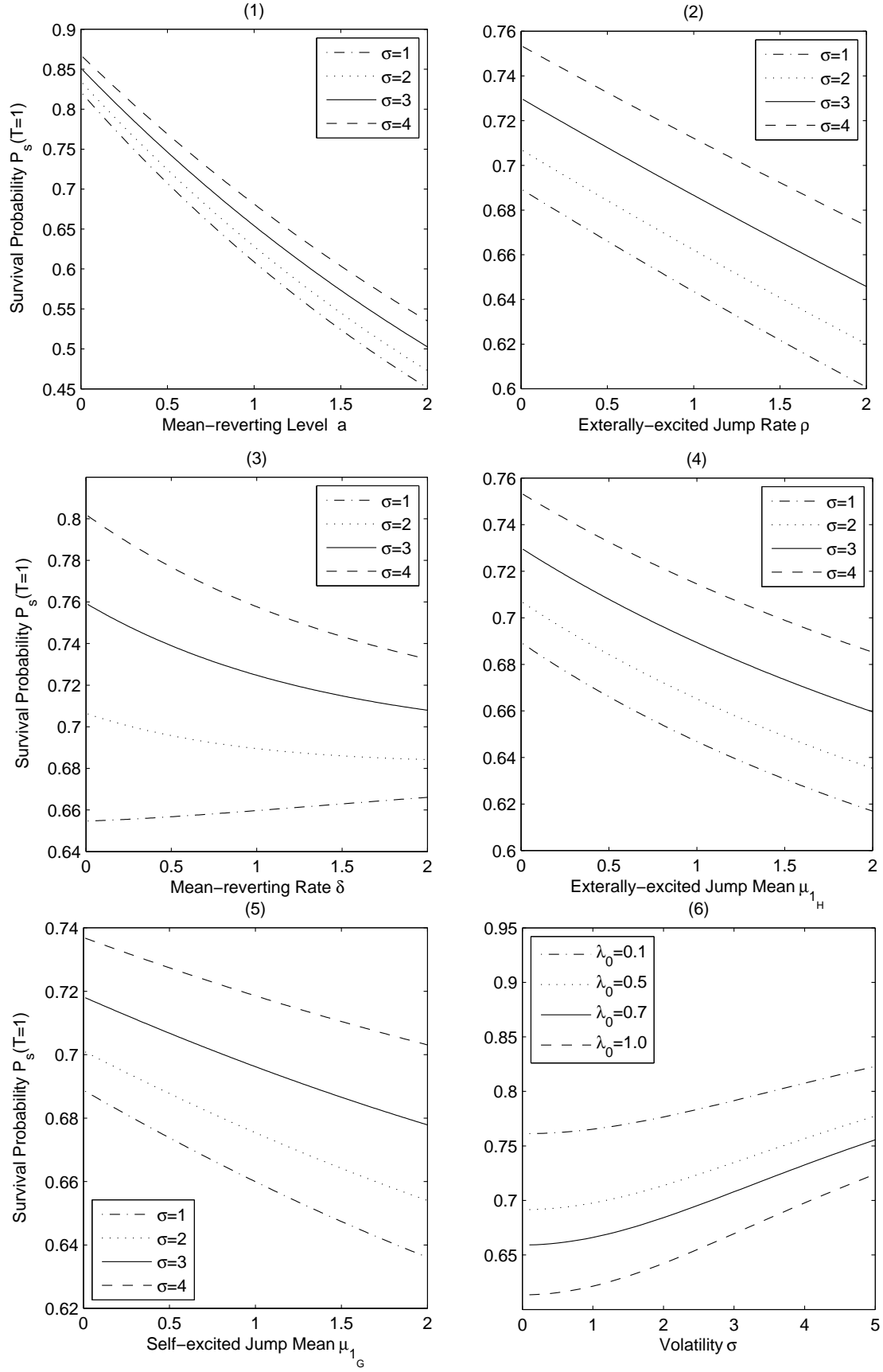


Figure 3: Sensitivity analysis for the survival probability at time $T = 1$ with fixed $d = 0.5$ against the varying parameters of $(a, \rho, \delta; \mu_{1H}, \mu_{1G}; \sigma, \lambda_0)$

Theorem 5.1. Suppose the k^{th} bad event could cause default of a constant probability $d_k \in (0, 1]$.

If

$$d_k = 1 - \frac{\hat{f}(k)}{\hat{f}(k-1)}, \quad k = 1, 2, \dots,$$

where

$$\hat{f}(k) := \int_0^\infty e^{-ku} f(u) du, \quad k = 0, 1, \dots,$$

and $f(u)$ is a (sub-density) function such that $\hat{f}(0) = \int_0^\infty f(u) du \leq 1$, then, the survival probability is given by

$$P_s(T) = \frac{1}{\hat{f}(0)} \int_0^\infty \phi(e^{-u}) f(u) du - \Pr\{N_T = 0 \mid \lambda_0\}, \quad (5.3)$$

where $\Pr\{N_T = 0 \mid \lambda_0\}$ is given by setting $\theta = 0$ in (3.15).

Proof. We have the survival probability

$$\begin{aligned} P_s(T) &= \mathbb{E} \left[\prod_{k=1}^{N_T} (1 - d_k) \mid \lambda_0 \right] \\ &= \sum_{n=1}^\infty \prod_{k=1}^n (1 - d_k) \Pr\{N_T = n \mid \lambda_0\} \\ &= \sum_{n=1}^\infty \prod_{k=1}^n \frac{\hat{f}(k)}{\hat{f}(k-1)} \Pr\{N_T = n \mid \lambda_0\} \\ &= \frac{1}{\hat{f}(0)} \int_0^\infty \sum_{n=1}^\infty e^{-nu} \Pr\{N_T = n \mid \lambda_0\} f(u) du \\ &= \frac{1}{\hat{f}(0)} \int_0^\infty \left(\mathbb{E}[e^{-uN_T} \mid \lambda_0] - \Pr\{N_T = 0 \mid \lambda_0\} \right) f(u) du \\ &= \frac{1}{\hat{f}(0)} \int_0^\infty \mathbb{E}[e^{-uN_T} \mid \lambda_0] f(u) du - \Pr\{N_T = 0 \mid \lambda_0\}. \end{aligned}$$

Set $\theta = e^{-u}$ in Theorem 3.4, then, we have (5.3). □

Remark 5.1. For example, if $f(u) = a_2 e^{-a_1 u}$, $a_2 \leq a_1$, then,

$$\hat{f}(k) = \frac{a_2}{a_1 + k}, \quad d_k = \frac{1}{a_1 + k}, \quad k = 0, 1, \dots,$$

and we have the survival probability

$$P_s(T) = \int_0^\infty \phi_T(e^{-u}) a_1 e^{-a_1 u} du - \Pr\{N_T = 0 \mid \lambda_0\} = \mathbb{E}[\phi_T(e^{-X})] - \Pr\{N_T = 0 \mid \lambda_0\},$$

where $X \sim \text{Exp}(a_1)$.

The term structure of default (or survival) probability indeed is the most critical input to the risk management and asset pricing for credit risk. For example, based on our generalised contagion model in this paper and fundamental pricing formula in Jarrow and Turnbull (1995), we can easily price defaultable bonds: suppose the risk-free interest rate and default timing are assumed to be independent of each other, and the present value (at time 0) of a defaultable zero-coupon bond which pays \$1 at maturity T is then simply given by

$$B_C(T) = B_G(T) \left[R + (1 - R)P_s(T) \right], \quad (5.4)$$

where $B_G(T)$ is the present value (at time 0) of a default-free zero-coupon bond which pays \$1 at maturity T , and $R \in [0, 1]$ is the assumed constant recovery rate.

For numerical implementation, we assume that, the current price of this default-free bond is \$0.9 and the recovery rate is 40% with the parameter setting of $(a, \varrho, \delta; \sigma; \lambda_0) = (0.7, 0.5, 2.5; 0.5; 0.7)$. Then, the associated defaultable zero-coupon bond price $B_C(T)$ based on formulas of (5.2) and (5.4) can be exactly calculated. The plots of bond prices $B_C(T)$ against the externally-excited jump mean $\mu_{1_H} = 1/\alpha \in [0, 2]$ and the self-excited jump mean $\mu_{1_G} = 1/\beta \in [0, 2]$ are presented in Figure 4 for the maturities $T = 1, 2, 3, 4$, respectively. The negative impacts from external and internal risks to the underlying bond price become more evident, as the bond price declines when the mean of externally-excited or self-excited jumps increases.

6 Conclusion

In this paper, we introduce an analytically tractable point process with self-exciting, externally-exciting and mean-reverting stochastic intensity in a single framework. Key stochastic properties have been systematically analysed. This model may provide a very useful quantitative tool in finance for modelling contagious arrivals of events in a variety of circumstances in practice. In particular, relatively short-lived shocks and long-lasting fluctuations of exogenous risk factors can be separately modelled, and this is the major difference from most other models in the literature. A simple application to credit risk is demonstrated. The associated empirical work (such as calibration to real data) as well as other applications (such as credit derivative pricing) based on this model is proposed as future research.

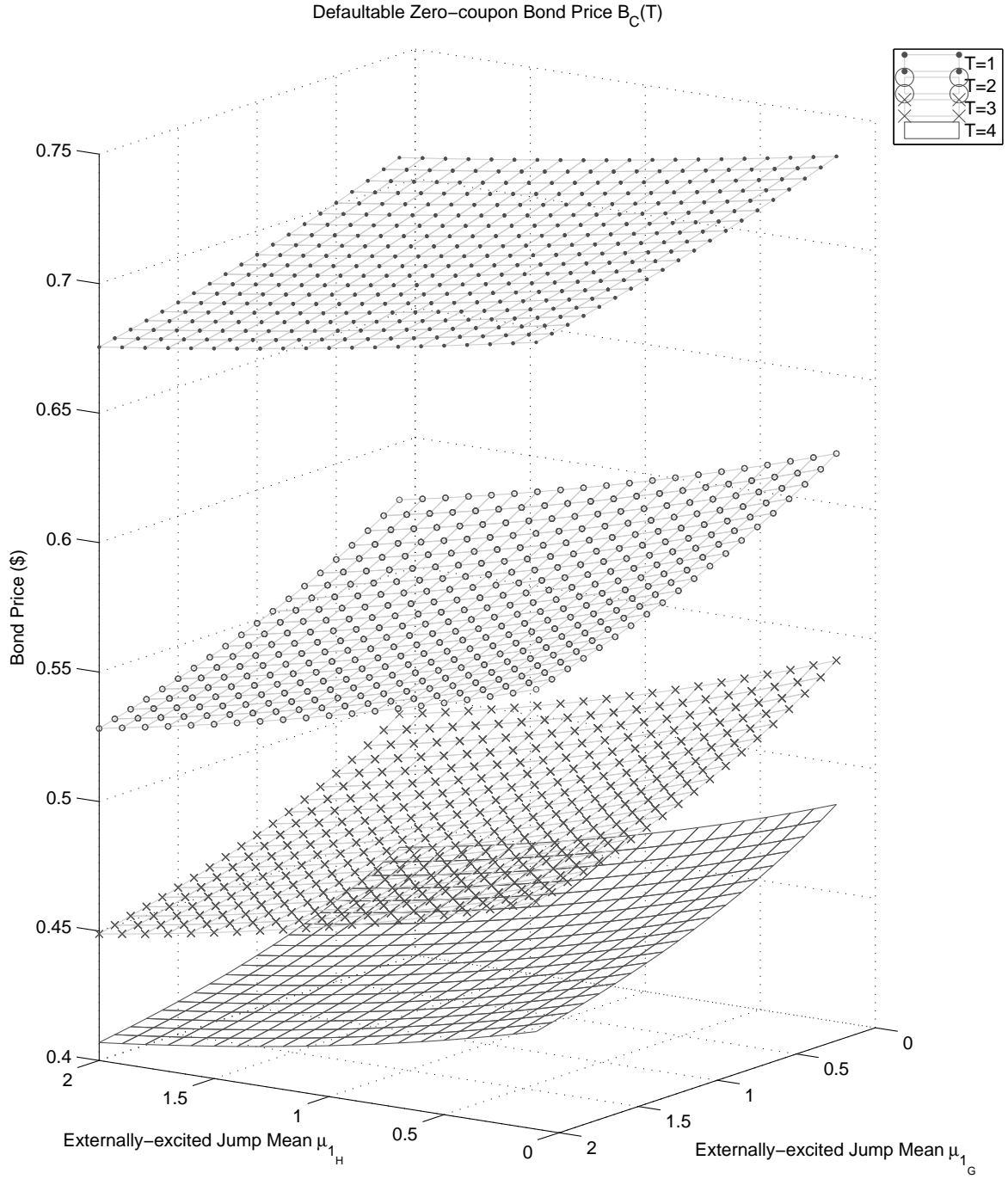


Figure 4: Plots of the defaultable zero-coupon bond prices $B_C(T)$ against the externally-excited jump mean $\mu_{1_H} = 1/\alpha \in [0, 2]$ and the self-excited jump mean $\mu_{1_G} = 1/\beta \in [0, 2]$ for the maturity $T = 1, 2, 3, 4$, respectively, based on the current default-free bond price \$0.9, recovery rate 40% and parameter setting of $(a, q, \delta; \sigma; \lambda_0) = (0.7, 0.5, 2.5; 0.5; 0.7)$

Acknowledgments

The authors would like to thank all reviewers for very helpful and constructive comments and suggestions. The corresponding author Hongbiao Zhao acknowledges the financial support from the National Natural Science Foundation of China (#71401147) and the research fund provided by Xiamen University.

References

- Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117(3):585–606.
- Aït-Sahalia, Y., Laeven, R. J., and Pelizzon, L. (2014). Mutual excitation in Eurozone sovereign CDS. *Journal of Econometrics*, 183(2):151–167.
- Bacry, E., Delattre, S., Hoffmann, M., and Muzy, J.-F. (2013). Modelling microstructure noise with mutually exciting point processes. *Quantitative Finance*, 13(1):65–77.
- Bacry, E., Mastromatteo, I., and Muzy, J.-F. (2015). Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(1):1550005.
- Boumezoued, A. et al. (2016). Population viewpoint on Hawkes processes. *Advances in Applied Probability*, 48(2):463–480.
- Bowsher, C. G. (2007). Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141(2):876–912.
- Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *The Annals of Probability*, 24(3):1563–1588.
- Brémaud, P. and Massoulié, L. (2002). Power spectra of general shot noises and Hawkes point processes with a random excitation. *Advances in Applied Probability*, 34(1):205–222.
- Chavez-Demoulin, V., Davison, A. C., and McNeil, A. J. (2005). Estimating value-at-risk: a point process approach. *Quantitative Finance*, 5(2):227–234.
- Chornoboy, E. S., Schramm, L. P., and Karr, A. F. (1988). Maximum likelihood identification of neural point process systems. *Biological Cybernetics*, 59(4-5):265–275.
- Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society. Series B (Methodological)*, 17(2):129–164.
- Crane, R. and Sornette, D. (2008). Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences*, 105(41):15649–15653.
- Dassios, A. and Embrechts, P. (1989). Martingales and insurance risk. *Stochastic Models*, 5(2):181–217.

- Dassios, A. and Zhao, H. (2011). A dynamic contagion process. *Advances in Applied Probability*, 43(3):814–846.
- Dassios, A. and Zhao, H. (2012). Ruin by dynamic contagion claims. *Insurance: Mathematics and Economics*, 51(1):93–106.
- Dassios, A. and Zhao, H. (2015). Efficient simulation of clustering jumps with CIR intensity. Working Paper. London School of Economics.
- Davis, M. H. (1984). Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 46(3):353–388.
- Duffie, D., Filipovic, D., and Schachermayer, W. (2003). Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1053.
- Duffie, D., Pan, J., and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376.
- Duffie, D. and Singleton, K. J. (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4):687–720.
- Embrechts, P., Liniger, T., and Lin, L. (2011). Multivariate Hawkes processes: an application to financial data. *Journal of Applied Probability*, 48A:367–378.
- Errais, E., Giesecke, K., and Goldberg, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, 1(1):642–665.
- Feller, W. (1951). Two singular diffusion problems. *The Annals of Mathematics*, 54(1):173–182.
- Feller, W. (1971). *An Introduction to Probability Theory and its Applications*, volume 2. John Wiley & Sons.
- Gerber, H. U. and Shiu, E. S. (1994). Option pricing by Esscher transforms. *Transactions of the Society of Actuaries*, 46:99–140.
- Giesecke, K. and Kim, B. (2007). Estimating tranche spreads by loss process simulation. In *Proceedings of the 2007 Winter Simulation Conference*, pages 967–975. IEEE Press.
- Giesecke, K., Kim, B., and Zhu, S. (2011). Monte Carlo algorithms for default timing problems. *Management Science*, 57(12):2115–2129.
- Glasserman, P. and Kim, K.-K. (2010). Moment explosions and stationary distributions in affine diffusion models. *Mathematical Finance*, 20(1):1–33.
- Hall, E. C. and Willett, R. M. (2016). Tracking dynamic point processes on networks. *IEEE Transactions on Information Theory*, 62(7):4327–4346.

- Hawkes, A. G. (1971a). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 33(3):438–443.
- Hawkes, A. G. (1971b). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90.
- Hawkes, A. G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11:493–503.
- Jarrow, R. A. and Turnbull, S. M. (1995). Pricing derivatives on financial securities subject to credit risk. *The Journal of Finance*, 50(1):53–85.
- Large, J. (2007). Measuring the resiliency of an electronic limit order book. *Journal of Financial Markets*, 10(1):1–25.
- Marsan, D. and Lengline, O. (2008). Extending earthquakes’ reach through cascading. *Science*, 319(5866):1076–1079.
- Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P., and Tita, G. E. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493):100–108.
- Oakes, D. (1975). The Markovian self-exciting process. *Journal of Applied Probability*, 12:69–77.
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83(401):9–27.
- Øksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications*. Springer.
- Stabile, G. and Torrisi, G. L. (2010). Risk processes with non-stationary Hawkes claims arrivals. *Methodology and Computing in Applied Probability*, 12(3):415–429.
- Veen, A. and Schoenberg, F. P. (2008). Estimation of space–time branching process models in seismology using an EM–type algorithm. *Journal of the American Statistical Association*, 103(482):614–624.
- Vere-Jones, D. (1978). Earthquake prediction – a statistician’s view. *Journal of Physics of the Earth*, 26(2):129–146.
- Xu, L., Duan, J. A., and Whinston, A. (2014). Path to purchase: A mutually exciting point process model for online advertising and conversion. *Management Science*, 60(6):1392–1412.
- Zadeh, A. H. and Sharda, R. (2014). Modeling brand post popularity dynamics in online social networks. *Decision Support Systems*, 65:59–68.
- Zhang, X., Glynn, P. W., Giesecke, K., and Blanchet, J. (2009). Rare event simulation for a generalized Hawkes process. In *Proceedings of the 2009 Winter Simulation Conference*, pages 1291–1298. IEEE.

- Zhu, L. (2013a). Central limit theorem for nonlinear Hawkes processes. *Journal of Applied Probability*, 50(3):760–771.
- Zhu, L. (2013b). Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. *Insurance: Mathematics and Economics*, 53(3):544–550.
- Zhu, L. (2014). Limit theorems for a Cox-Ingersoll-Ross process with Hawkes jumps. *Journal of Applied Probability*, 51(3):699–712.