LIMIT THEOREMS FOR A COX–INGERSOLL–ROSS PROCESS WITH HAWKES JUMPS

LINGJIONG ZHU,* *New York University*

Abstract

In this paper we propose a stochastic process, which is a Cox–Ingersoll–Ross process with Hawkes jumps. It can be seen as a generalization of the classical Cox–Ingersoll–Ross process and the classical Hawkes process with exponential exciting function. Our model is a special case of the affine point processes. We obtain Laplace transforms and limit theorems, including the law of large numbers, central limit theorems, and large deviations.

Keywords: Cox–Ingersoll–Ross process; point process; Hawkes process; self-exciting process; central limit theorem; large deviations

2010 Mathematics Subject Classification: Primary 60G07; 60G55; 60F05; 60F10

1. Introduction and main results

1.1. Cox–Ingersoll–Ross process

A Cox–Ingersoll–Ross process is a stochastic process r_t satisfying the stochastic differential equation

$$\mathrm{d}r_t = b(c - r_t)\,\mathrm{d}t + \sigma\sqrt{r_t}\,\mathrm{d}W_t,$$

where W_t is a standard Brownian motion, $b, c, \sigma > 0$, and $2bc \geq \sigma^2$. The process was proposed by Cox *et al.* [5] to model the short-term interest rate. Under the assumption that $2bc \geq \sigma^2$, Feller [10] proved that the process is nonnegative. Given r_0, it is well known that $4br_t/\sigma^2(1 - \mathrm{e}^{-bt})$ follows a noncentral χ^2 distribution with degree of freedom $4bc/\sigma^2$ and noncentrality parameter $4br_0\mathrm{e}^{-bt}/\sigma^2(1 - \mathrm{e}^{-bt})$. As $t \to \infty$, $r_t \to r_\infty$, where r_∞ follows a gamma distribution with shape parameter $2bc/\sigma^2$ and scale parameter $\sigma^2/2b$. The conditional first and second moments are given by, for $s > t$,

$$\mathbb{E}[r_s \mid r_t] = r_t\mathrm{e}^{-b(s-t)} + c(1 - \mathrm{e}^{-b(s-t)}), \tag{1}$$

$$\mathbb{E}[r_s^2 \mid r_t] = r_t\left(2c + \frac{\sigma^2}{b}\right)\mathrm{e}^{-b(s-t)} + \left(r_t^2 - r_t\frac{\sigma^2}{b} - 2r_tc\right)\mathrm{e}^{-2b(s-t)}$$
$$+ \left(\frac{c\sigma^2}{2b} + c^2\right)\left(1 - \mathrm{e}^{-b(s-t)}\right)^2. \tag{2}$$

The Cox–Ingersoll–Ross process has been widely applied in finance, mostly for the short-term interest rate (see, e.g. [5]) and the Heston stochastic volatility model (see, e.g. [14]). Other applications include the modelling of mortality intensities (see, e.g. the extended Cox–Ingersoll–Ross process used in [6]) and of default intensities in credit risk models (see, e.g. the special affine process case in [8]).

A natural generalization of the classical Cox–Ingersoll–Ross process takes into account the jumps, i.e.

$$dr_t = b(c - r_t)\,dt + \sigma\sqrt{r_t}\,dW_t + a\,dN_t,$$

where N_t is a homogeneous Poisson process with constant intensity $\lambda > 0$. However, in the real world, the occurrence of events may not be time homogeneous, but should be time dependent. Errais *et al.* [9] pointed out

> The collapse of Lehman Brothers brought the financial system to the brink of a breakdown. The dramatic repercussions point to the existence of feedback phenomena that are channeled through the complex web of informational and contractual relationships in the economy …. This and related episodes motivate the design of models of correlated default timing that incorporate the feedback phenomena that plague credit markets.

According to Kou and Peng [15], 'We need better models to incorporate the default clustering effect, i.e., one default event tends to trigger more default ….'

In this respect, it is natural to replace the Poisson process with a simple point process which can describe the time dependence in a natural way. The Hawkes process, a simple point process that has a self-exciting property and clustering effect becomes a natural choice.

1.2. Hawkes process

A Hawkes process is a simple point process N admitting an intensity

$$\lambda_t := \lambda\left(\int_{-\infty}^{t} h(t - s)N(ds)\right), \tag{3}$$

where $\lambda(\cdot)\colon \mathbb{R}^+ \to \mathbb{R}^+$ is locally integrable, left continuous, $h(\cdot)\colon \mathbb{R}^+ \to \mathbb{R}^+$, and we always assume that $\|h\|_{L^1} = \int_0^\infty h(t)\,dt < \infty$. In (3), $\int_{-\infty}^t h(t-s)N(ds)$ stands for $\int_{(-\infty,t)} h(t-s)N(ds) = \sum_{\tau < t} h(t-\tau)$, where τ are the occurrences of the points before time t.

In the literature, $h(\cdot)$ and $\lambda(\cdot)$ are often referred to as the exciting function and the rate function, respectively. An important observation is that a Hawkes process is Markovian if and only if $h(\cdot)$ is an exponential function. One usually assumes that $\lambda(\cdot)$ is increasing and $h(\cdot)$ is decreasing.

A Hawkes process is linear if $\lambda(\cdot)$ is linear and it is nonlinear otherwise. The linear Hawkes process, i.e. the classical Hawkes process, is named after Hawkes, who was the first to present the model in [12]. The nonlinear Hawkes process was first introduced by Brémaud and Massoulié [4].

By the definition of the Hawkes process, it has the self-exciting property, i.e. the intensity λ_t increases when a jump is observed. It therefore creates a clustering effect, which is used to model the default clustering in finance. When no new jumps are observed, the intensity λ_t decreases as $h(\cdot)$ decays.

The law of large numbers and central limit theorem for the linear Hawkes process have been obtained in [13]. The law of large numbers and central limit theorem have also been studied in [2] as a special case of multivariate Hawkes processes. The large deviation principle for linear Hawkes process was obtained in [3]. The moderate deviation principle for linear Hawkes process was obtained in [19]. For nonlinear Hawkes process, the central limit theorem was obtained in [18], and the large deviations have been studied in [21] and [22].

The central limit theorem for the Hawkes process has been applied to the study of high-frequency trading and the microstructure in finance (see, e.g. [2] and [1]), and the large deviations result has been applied to the study of ruin probabilities in insurance (see, e.g. [16] and [20]).

1.3. A Cox–Ingersoll–Ross process with Hawkes jumps

In this paper we propose a stochastic process r_t that satisfies the stochastic differential equation

$$\mathrm{d}r_t = b(c - r_t)\,\mathrm{d}t + a\,\mathrm{d}N_t + \sigma\sqrt{r_t}\,\mathrm{d}W_t, \tag{4}$$

where W_t is a standard Brownian motion and N_t is a simple point process with intensity $\lambda_t := \alpha + \beta r_t$ at time t. We assume that $a, b, c, \alpha, \beta, \sigma > 0$ and that the following conditions hold.

- $b > a\beta$. This condition is needed to guarantee that there exists a unique stationary process r_∞ which satisfies (4).

- $2bc \geq \sigma^2$. This condition is needed to guarantee that $r_t \geq 0$ with probability 1. Indeed, we know that r_t stochastically dominates the classical Cox–Ingersoll–Ross process and, hence, $2bc \geq \sigma^2$ is enough to guarantee $r_t \geq 0$. On the other hand, on any given time interval, the probability that there is no jump is always positive, which implies that $2bc \geq \sigma^2$ is needed to guarantee positivity.

The Cox–Ingersoll–Ross process with Hawkes jumps preserves the mean-reverting and nonnegative properties of the classical Cox–Ingersoll–Ross process. In addition, it contains the Hawkes jumps, which have the self-exciting property that create a clustering effect.

Clearly, the Cox–Ingersoll–Ross process proposed in (4) includes the classical Cox–Ingersoll–Ross process and the classical linear Hawkes process with exponential exciting function. We summarize this in the following.

(i) When $a = 0$ or $\alpha = \beta = 0$, it reduces to the classical Cox–Ingersoll–Ross process, i.e.

$$\mathrm{d}r_t = b(c - r_t)\,\mathrm{d}t + \sigma\sqrt{r_t}\,\mathrm{d}W_t.$$

(ii) When $\beta = 0$ and $a, \alpha > 0$, it reduces to the Cox–Ingersoll–Ross process with Poisson jumps, i.e.

$$\mathrm{d}r_t = b(c - r_t)\,\mathrm{d}t + \sigma\sqrt{r_t}\,\mathrm{d}W_t + a\,\mathrm{d}N_t,$$

where N_t is a homogeneous Poisson process with intensity α.

(iii) When $c = 0$ and $\sigma = 0$, N_t reduces to a Hawkes process with intensity $\lambda_t = \alpha + \beta r_t$, where

$$\mathrm{d}r_t = -br_t\,\mathrm{d}t + a\,\mathrm{d}N_t,$$

and it is easy to see that the intensity λ_t indeed satisfies

$$\lambda_t = \alpha + \beta \int_0^t a\mathrm{e}^{-b(t-s)} N(\mathrm{d}s),$$

which implies that N_t is a classical linear Hawkes process with $\lambda(z) = \alpha + \beta z$ and $h(t) = a\mathrm{e}^{-bt}$.

It is easy to observe that r_t is Markovian with generator

$$\mathcal{A}f(r) = bc\frac{\partial f}{\partial r} - br\frac{\partial f}{\partial r} + \frac{1}{2}\sigma^2 r\frac{\partial^2 f}{\partial r^2} + (\alpha + \beta r)[f(r+a) - f(r)].$$

1.4. Main results

In this section we summarize the main results of this paper. We will start with conditional first and second moments of r_t, and then move onto the limit theorems, i.e. the law of large numbers, central limit theorems, and large deviations. Next, we show that there exists a unique stationary probability measure for r_t and we obtain the Laplace transform of r_t and r_∞. Finally, we consider a short-rate interest model. The proofs will be given in Section 2.

In the following proposition we give the formulae for the conditional first and second moments of the Cox–Ingersoll–Ross process with Hawkes jumps.

Proposition 1. (i) *For any $s > t$, we have the conditional expectation*

$$\mathbb{E}[r_s \mid r_t] = \frac{bc + a\alpha}{b - a\beta} - e^{-(b-a\beta)(s-t)}\left[\frac{bc + a\alpha}{b - a\beta} - r_t\right]. \tag{5}$$

(ii) *For any $s > t$, we have the conditional expectation*

$$\mathbb{E}[r_s^2 \mid r_t] = r_t^2 e^{-2(b-a\beta)(s-t)}$$
$$+ \left[(2bc + \sigma^2 + 2a\alpha + a^2\beta)\frac{bc + a\alpha}{2(b - a\beta)^2} + \frac{a^2\alpha}{2(b - a\beta)}\right][1 - e^{-2(b-a\beta)(s-t)}]$$
$$- (2bc + \sigma^2 + 2a\alpha + a^2\beta)\frac{bc + a\alpha}{(b - a\beta)^2}[e^{-(b-a\beta)(s-t)} - e^{-2(b-a\beta)(s-t)}]$$
$$+ (2bc + \sigma^2 + 2a\alpha + a^2\beta)\frac{r_t}{b - a\beta}[e^{-(b-a\beta)(s-t)} - e^{-2(b-a\beta)(s-t)}]. \tag{6}$$

Remark 1. Let $a = 0$ in (5). Then $\mathbb{E}[r_s \mid r_t] = c - e^{-b(s-t)}(c - r_t) = r_t e^{-b(s-t)} + c(1 - e^{-b(s-t)})$, which recovers (1). Similarly, by letting $a = 0$ in (6), we recover (2).

Theorem 1. (Law of large numbers.) *For any $r_0 := r \in \mathbb{R}^+$,*

(i) $\dfrac{1}{t}\displaystyle\int_0^t r_s \, ds \to \dfrac{bc + a\alpha}{b - a\beta}$ *in $L^2(\mathbb{P})$ as $t \to \infty$,*

(ii) $\dfrac{N_t}{t} \to \dfrac{b(\alpha + \beta c)}{b - a\beta}$ *in $L^2(\mathbb{P})$ as $t \to \infty$.*

Theorem 2. (Central Limit Theorem.) *For any $r_0 := r \in \mathbb{R}^+$,*

(i) $\dfrac{\int_0^t r_s \, ds - (bc + a\alpha)t/(b - a\beta)}{\sqrt{t}} \to N\left(0, \dfrac{a^2\alpha(b - a\beta) + (a^2\beta + \sigma^2)(bc + a\alpha)}{(b - a\beta)^3}\right)$ *in distribution as $t \to \infty$,*

(ii) $\dfrac{N_t - b(\alpha + \beta c)t/(b - a\beta)}{\sqrt{t}} \to N\left(0, \dfrac{b^3 a^2(\alpha + \beta c) + 4\sigma^2 b^2(bc + a\alpha)}{a^2(b - a\beta)^3}\right)$ *in distribution as $t \to \infty$.*

Before we proceed, recall that a sequence $(P_n)_{n \in \mathbb{N}}$ of probability measures on a topological space X satisfies the large deviation principle with rate function $I : X \to \mathbb{R}$ if I is nonnegative and lower semicontinuous, and, for any measurable set A, we have

$$-\inf_{x \in A^o} I(x) \le \liminf_{n \to \infty} \frac{1}{n} \log P_n(A) \le \limsup_{n \to \infty} \frac{1}{n} \log P_n(A) \le -\inf_{x \in \bar{A}} I(x).$$

Here A^o denotes the interior of A and $\bar{A}$ denotes its closure. We refer the reader to [7] and [17] for general background on the theory and applications of large deviations.

Theorem 3. (Large deviation principle.) *For any* $r_0 := r \in \mathbb{R}^+$,

(i) $((1/t) \int_0^t r_s \, ds \in \cdot)$ *satisfies a large deviation principle with rate function*

$$I(x) = \sup_{\theta \le \theta_c} \{\theta x - bcy(\theta) - \alpha(e^{ay(\theta)} - 1)\},$$

where, for $\theta \le \theta_c$, $y = y(\theta)$ *is the smaller solution of*

$$-by + \tfrac{1}{2}\sigma^2 y^2 + \beta(e^{ay} - 1) + \theta = 0,$$

and

$$\theta_c = by_c - \tfrac{1}{2}\sigma^2 y_c^2 - \beta(e^{ay_c} - 1)$$

with y_c *the unique positive solution to the equation* $b = \sigma^2 y_c + \beta a e^{ay_c}$,

(ii) $(N_t/t \in \cdot)$ *satisfies a large deviation principle with rate function*

$$I(x) = \sup_{\theta \le \theta_c} \{\theta x - bcy(\theta) - \alpha(e^{ay(\theta)+\theta} - 1)\},$$

where, for $\theta \le \theta_c$, $y(\theta)$ *is the smaller solution of*

$$-by(\theta) + \tfrac{1}{2}\sigma^2 y^2(\theta) + \beta(e^{ay(\theta)+\theta} - 1) = 0,$$

and

$$\theta_c = \log\left(\frac{\sqrt{\sigma^4 + a^2 b^2 + 2a^2\sigma^2\beta} - \sigma^2}{a^2\beta}\right) - \frac{\sigma^2 + ab - \sqrt{\sigma^4 + a^2 b^2 + 2a^2\sigma^2\beta}}{\sigma^2}.$$

Remark 2. It is easy to see that, when $c = 0$ and $\sigma = 0$, Theorem 1(ii), Theorem 2(ii), and Theorem 3(ii) are consistent with the law of large numbers and central limit theorem results for linear Hawkes process with exponential exciting function given in [2], and the large deviation principle given in [3].

Proposition 2. *Assume that* $b > a\beta$ *and* $2bc \ge \sigma^2$. *Then there exists a unique invariant probability measure for* r_t.

Proposition 3. *For any* $\theta > 0$, *the Laplace transform of* r_t *satisfies* $\mathbb{E}[e^{-r_t} \mid r_0 = r] = e^{A(t)r + B(t)}$, *where* $A(t)$ *and* $B(t)$ *satisfy the ordinary differential equations*

$$A'(t) = -bA(t) + \tfrac{1}{2}\sigma^2 A(t)^2 + \beta(e^{aA(t)} - 1),$$
$$B'(t) = bcA(t) + \alpha(e^{aA(t)} - 1),$$
$$A(0) = -\theta, \qquad B(0) = 0.$$

In particular, $\mathbb{E}[e^{-\theta r_\infty}] = \exp[\int_0^\infty bcA(t) + \alpha(e^{aA(t)} - 1) \, dt]$.

We can use r_t as a stochastic model for the short-rate term structure. We are interested in the value of a default-free discount bond paying one unit at time T, i.e.

$$P(t, T, r) := \mathbb{E}\left[\exp\left[-\int_t^T r_s \, ds\right] \, \bigg| \, r_t = r\right].$$

Proposition 4. (i) $P(t, T, r) = e^{A(t)r + b(t)}$, *where $A(t)$ and $B(t)$ satisfy the ordinary differential equations*

$$A'(t) - bA(t) + \tfrac{1}{2}\sigma^2 A(t)^2 + \beta(e^{aA(t)} - 1) - 1 = 0,$$

$$B'(t) + bcA(t) + \alpha(e^{aA(t)} - 1) = 0,$$

$$A(T) = B(T) = 0.$$

(ii) *We have the asymptotic result*

$$\lim_{T \to \infty} \frac{1}{T} \log P(t, T, r) = bcx_* + \alpha(e^{ax_*} - 1),$$

where x_* is the unique negative solution to the equation

$$-bx + \tfrac{1}{2}\sigma^2 x^2 + \beta(e^{ax} - 1) - 1 = 0.$$

Remark 3. A natural way to generalize the Cox–Ingersoll–Ross process with Hawkes jumps is to allow the jump size to be random, i.e.

$$dr_t = b(c - r_t)\, dt + \sigma \sqrt{r_t}\, dW_t + dJ_t,$$

where $J_t = \sum_{i=1}^{N_{t-}} a_i$, and the a_i are independent and identically distributed positive random variables, independent of the past history, with probability distribution $Q(da)$. Here N_t is a simple point process with intensity $\lambda_t = \alpha + \beta r_t$ at time $t > 0$. We assume that $a, b, c, \alpha, \beta, \sigma > 0$, $b > \int_{\mathbb{R}^+} a\, Q(da)\beta$, and $2bc \geq \sigma^2$.

We can write the generator as

$$\mathcal{A}f(r) = bc\frac{\partial f}{\partial r} - br\frac{\partial f}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 f}{\partial r^2} + (\alpha + \beta r)\int_{\mathbb{R}^+} [f(r + a) - f(r)]Q(da).$$

All the results in this paper can be generalized to this model after a careful analysis.

Remark 4. Another possibility to generalize the Cox–Ingersoll–Ross process with Hawkes jumps is to allow the jumps to follow a nonlinear Hawkes process, i.e. r_t satisfies the dynamics in (4) and N_t is a simple point process with intensity $\lambda(r_t)$, where $\lambda(\cdot) \colon \mathbb{R}^+ \to \mathbb{R}^+$ is in general a nonlinear function. This can be considered as a generalization to the classical nonlinear Hawkes process with exponential exciting function. Because of the nonlinearity, we will not be able to get a closed expression in the limit for the limit theorems or a set of ordinary differential equations which are related to the Laplace transform of the process.

2. Proofs

Proof of Proposition 1. (i) Taking the expectation of both sides of (4), we have

$$d\mathbb{E}[r_t] = b(c - \mathbb{E}[r_t])\, dt + a(\alpha + \beta\mathbb{E}[r_t])\, dt,$$

which implies that for any $s > t$, we have the following conditional expectation,

$$\mathbb{E}[r_s \mid r_t] = \frac{bc + a\alpha}{b - a\beta} - e^{-(b - a\beta)(s - t)}\left[\frac{bc + a\alpha}{b - a\beta} - r_t\right].$$

(ii) By Itô's formula, we have

$$d(r_t^2) = 2r_t[b(c - r_t)\, dt + \sigma \sqrt{r_t}\, dW_t] + \sigma^2 r_t\, dt + 2r_t a\, dN_t + a^2\, dN_t.$$

Taking the expectation of both sides, we obtain

$$\frac{d\mathbb{E}[r_t^2]}{dt} = 2bc\mathbb{E}[r_t] - 2b\mathbb{E}[r_t^2] + \sigma^2\mathbb{E}[r_t] + 2a(\alpha\mathbb{E}[r_t] + \beta\mathbb{E}[r_t^2]) + a^2\alpha + a^2\beta\mathbb{E}[r_t].$$

This implies that

$$\mathbb{E}[r_s^2 \mid r_t]e^{2(b-a\beta)s} - r_t^2 e^{2(b-a\beta)t}$$
$$= (2bc + \sigma^2 + 2a\alpha + a^2\beta)\int_t^s e^{2(b-a\beta)u}\mathbb{E}[r_u \mid r_t]\,du + a^2\alpha\int_t^s e^{2(b-a\beta)u}\,du$$
$$= \left[(2bc + \sigma^2 + 2a\alpha + a^2\beta)\frac{bc + a\alpha}{2(b - a\beta)^2} + \frac{a^2\alpha}{2(b - a\beta)}\right][e^{2(b-a\beta)s} - e^{2(b-a\beta)t}]$$
$$- (2bc + \sigma^2 + 2a\alpha + a^2\beta)\frac{bc + a\alpha}{b - a\beta}\frac{e^{(b-a\beta)t}}{b - a\beta}[e^{(b-a\beta)s} - e^{(b-a\beta)t}]$$
$$+ (2bc + \sigma^2 + 2a\alpha + a^2\beta)r_t\frac{e^{(b-a\beta)t}}{b - a\beta}[e^{(b-a\beta)s} - e^{(b-a\beta)t}],$$

which yields (6).

Proof of Theorem 1. (i) To prove the convergence in the $L^2(\mathbb{P})$ norm, we need to show that

$$\mathbb{E}\left(\frac{1}{t}\int_0^t r_s\,ds - \frac{bc + a\alpha}{b - a\beta}\right)^2$$
$$= \frac{1}{t^2}\mathbb{E}\left(\int_0^t r_s\,ds\right)^2 - \frac{2}{t}\int_0^t \mathbb{E}[r_s]\,ds\frac{bc + a\alpha}{b - a\beta} + \left(\frac{bc + a\alpha}{b - a\beta}\right)^2$$
$$\to 0,$$

as $t \to \infty$. From (5), it is clear that

$$\frac{1}{t}\int_0^t \mathbb{E}[r_s]\,ds \to \frac{bc + a\alpha}{b - a\beta} \quad \text{as } t \to \infty.$$

Therefore, it suffices to show that

$$\frac{1}{t^2}\mathbb{E}\left(\int_0^t r_s\,ds\right)^2 \to \frac{bc + a\alpha}{b - a\beta} \quad \text{as } t \to \infty.$$

Applying (5), we obtain

$$\frac{1}{t^2}\mathbb{E}\left(\int_0^t r_s\,ds\right)^2 = \frac{2}{t^2}\iint_{0<s_1<s_2<t}\mathbb{E}[r_{s_1}\mathbb{E}[r_{s_2} \mid r_{s_1}]]\,ds_1\,ds_2$$
$$= \frac{2}{t^2}\iint_{0<s_1<s_2<t}\frac{bc + a\alpha}{b - a\beta}\mathbb{E}[r_{s_1}]\,ds_1\,ds_2$$
$$- \frac{2}{t^2}\iint_{0<s_1<s_2<t}e^{-(b-a\beta)(s_2-s_1)}\left[\frac{bc + a\alpha}{b - a\beta}\mathbb{E}[r_{s_1}] - \mathbb{E}[r_{s_1}^2]\right]ds_1\,ds_2.$$

From Proposition 1, given $r_0 = r$, $\mathbb{E}[r_{s_1}]$ and $\mathbb{E}[r_{s_1}^2]$ are uniformly bounded by some universal constant depending only on r, say $M(r)$. Therefore,

$$\left| \frac{2}{t^2} \iint_{0<s_1<s_2<t} e^{-(b-a\beta)(s_2-s_1)} \left[\frac{bc+a\alpha}{b-a\beta} \mathbb{E}[r_{s_1}] - \mathbb{E}[r_{s_1}^2] \right] ds_1 \, ds_2 \right|$$

$$\le \frac{2}{t^2} M(r) \left[\frac{bc+a\alpha}{b-a\beta} + 1 \right] \iint_{0<s_1<s_2<t} e^{-(b-a\beta)(s_2-s_1)} \, ds_1 \, ds_2$$

$$\to 0 \quad \text{as } t \to \infty.$$

Again, by (5), it is easy to check that

$$\frac{2}{t^2} \iint_{0<s_1<s_2<t} \frac{bc+a\alpha}{b-a\beta} \mathbb{E}[r_{s_1}] \, ds_1 \, ds_2 \to \left(\frac{bc+a\alpha}{b-a\beta} \right)^2 \quad \text{as } t \to \infty.$$

Hence, we have proved the law of large numbers.

(ii) Observe that $N_t - \int_0^t \lambda_s \, ds = N_t - \alpha t - \beta \int_0^t r_s \, ds$ is a martingale and

$$\mathbb{E}\left[\left(\frac{N_t - \int_0^t \lambda_s \, ds}{t} \right)^2 \right] = \frac{1}{t^2} \mathbb{E}\left[\int_0^t \lambda_s \, ds \right] = \frac{\alpha}{t} + \frac{\beta}{t^2} \int_0^t \mathbb{E}[r_s] \, ds \to 0 \quad \text{as } t \to \infty$$

by Proposition 1. Therefore, we have

$$\frac{N_t}{t} - \alpha - \frac{\beta}{t} \int_0^t r_s \, ds \to 0$$

in $L^2(\mathbb{P})$ as $t \to \infty$ and the conclusion follows from (i).

Remark 5. The L^2 convergence in Theorem 1 implies the convergence in probability. Indeed, the convergence in Theorem 1 also holds almost surely by using Proposition 2 and the ergodic theorem. For example, by the ergodic theorem, $(1/t) \int_0^t r_s \, ds \to \mathbb{E}[r_\infty]$ almost surely as $t \to \infty$. Let π be the unique invariant probability measure of r_t. Then we have $\int \mathcal{A} f(r) \pi(dr) = 0$ for any smooth function f. Consider $f(r) = r$. We have $\int (bc - br + (\alpha + \beta r)a) \pi(dr) = 0$, which implies that $\mathbb{E}[r_\infty] = (bc + a\alpha)/(b - a\beta)$. Similarly, we can show that $N_t/t \to b(\alpha + \beta c)/(b - a\beta)$ as $t \to \infty$ almost surely. Indeed, the almost-sure convergence also follows by applying the large deviation principle and the Borel–Cantelli lemma. The limit can be identified as the unique zero of the corresponding rate function for the large deviations.

Proof of Theorem 2. (i) Observe that $f(r_t) - f(r_0) - \int_0^t \mathcal{A} f(r_s) \, ds$ is a martingale for $f(r) = Kr$, where K is a constant to be determined. Let $f(r) = Kr$. Then

$$\mathcal{A} f(r) = K[(a\beta - b)r + (\alpha a + bc)].$$

Let us choose $K = 1/(b - a\beta) > 0$. Then we have

$$\int_0^t r_s \, ds - \frac{bc+a\alpha}{b-a\beta} t = \left[f(r_t) - f(r_0) - \int_0^t \mathcal{A} f(r_s) \, ds \right] - f(r_t) + f(r_0).$$

Since $f(r_0) = Kr_0$ is fixed, $f(r_0)/\sqrt{t} \to 0$ as $t \to \infty$. Also, we have

$$\frac{\mathbb{E}[f(r_t)]}{\sqrt{t}} = \frac{K\mathbb{E}[r_t]}{\sqrt{t}} = \frac{K}{\sqrt{t}} \left\{ \frac{bc+a\alpha}{b-a\beta} - e^{-(b-a\beta)t} \left[\frac{bc+a\alpha}{b-a\beta} - r_0 \right] \right\} \to 0 \quad \text{as } t \to \infty$$

by Proposition 1. Therefore, $f(r_t)/\sqrt{t} \to 0$ as $t \to \infty$ in probability. The quadratic variation of the martingale $f(r_t) - f(r_0) - \int_0^t \mathscr{A}f(r_s)\,ds$ is the same as the quadratic variation of $f(r_t) = r_t/(b - a\beta)$, which is the same as the quadratic variation of $(aN_t + \int_0^t \sigma\sqrt{r_s}\,dW_s)/(b - a\beta)$, which is equivalent to $[a^2 N_t + \sigma^2 \int_0^t r_s\,ds]/(b - a\beta)^2$. By the law of large numbers (Theorem 1), we have

$$\frac{1}{t}\frac{1}{(b-a\beta)^2}\left[a^2 N_t + \sigma^2 \int_0^t r_s\,ds\right] \to \frac{1}{(b-a\beta)^2}\left[a^2\alpha + \frac{(a^2\beta + \sigma^2)(bc + a\alpha)}{b - a\beta}\right]$$

as $t \to \infty$. Hence, by the usual central limit theorem for martingales, we conclude that

$$\frac{\int_0^t r_s\,ds - (bc + a\alpha)t/(b - a\beta)}{\sqrt{t}} \to N\left(0, \frac{a^2\alpha(b - a\beta) + (a^2\beta + \sigma^2)(bc + a\alpha)}{(b - a\beta)^3}\right)$$

in distribution as $t \to \infty$.

(ii) From (4), we have

$$N_t = \frac{r_t}{a} - \frac{r_0}{a} + \frac{bc}{a}t - \frac{b}{a}\int_0^t r_s\,ds - \frac{\sigma}{a}\int_0^t \sqrt{r_s}\,dW_s,$$

which implies that

$$
\begin{aligned}
N_t - \frac{b(\alpha + \beta c)}{b - a\beta} &= \frac{r_t}{a} - \frac{r_0}{a} - \frac{b}{a}\int_0^t\left(r_s - \frac{bc + a\alpha}{b - a\beta}\right)ds - \frac{\sigma}{a}\int_0^t \sqrt{r_s}\,dW_s \\
&= \frac{r_t}{a} - \frac{r_0}{a} - f(r_t) + f(r_0) \\
&\quad + \left[f(r_t) - f(r_0) - \int_0^t \mathscr{A}f(r_s)\,ds\right] - \frac{\sigma}{a}\int_0^t \sqrt{r_s}\,dW_s,
\end{aligned}
$$

where $f(r) = -b/a(b - a\beta)$ and we know that

$$\frac{1}{\sqrt{t}}\left[\frac{r_t}{a} - \frac{r_0}{a} - f(r_t) + f(r_0)\right] \to 0$$

as $t \to \infty$ in probability by the arguments in (i). Now,

$$\left[f(r_t) - f(r_0) - \int_0^t \mathscr{A}f(r_s)\,ds\right] - \frac{\sigma}{a}\int_0^t \sqrt{r_s}\,dW_s$$

is a martingale and it has the same quadratic variation as

$$-\frac{b}{b - a\beta}N_t - \frac{b\sigma}{a(b - a\beta)}\int_0^t \sqrt{r_s}\,dW_s - \frac{\sigma}{a}\int_0^t \sqrt{r_s}\,dW_s,$$

which has the quadratic variation $b^2/(b - a\beta)^2 N_t + 4\sigma^2 b^2/a^2(b - a\beta)^2 \int_0^t r_s\,ds$. By law of large numbers, i.e. Theorem 1, we have

$$
\begin{aligned}
\frac{1}{t}&\left[\frac{b^2}{(b-a\beta)^2}N_t + \frac{4\sigma^2 b^2}{a^2(b-a\beta)^2}\int_0^t r_s\,ds\right] \\
&\to \frac{b^2}{(b-a\beta)^2}\frac{b(a\alpha + \beta c)}{b - a\beta} + \frac{4\sigma^2 b^2}{a^2(b-a\beta)^2}\frac{bc + a\alpha}{b - a\beta} \quad \text{as } t \to \infty.
\end{aligned}
$$

Hence, by the usual central limit theorem for martingales, we conclude that

$$\frac{N_t - b(\alpha + \beta c)t/(b - a\beta)}{\sqrt{t}} \to N\left(0, \frac{b^3 a^2(\alpha + \beta c) + 4\sigma^2 b^2(bc + a\alpha)}{a^2(b - a\beta)^3}\right)$$

in distribution as $t \to \infty$.

Proof of Theorem 3. (i) Let $u(\theta, t, r) := \mathbb{E}[\exp[\theta \int_0^t r_s \, ds]]$. Then, by the Feynman–Kac formula, we have

$$\frac{\partial u}{\partial t} = bc\frac{\partial u}{\partial r} - br\frac{\partial u}{\partial r} + \frac{1}{2}\sigma^2 r\frac{\partial^2 u}{\partial r^2} + (\alpha + \beta r)[u(\theta, t, r + a) - u(\theta, t, r)] + \theta r u = 0,$$

$$u(\theta, 0, r) = 1.$$

Let us try $u(\theta, t, r) = e^{A(t)r + B(t)}$. Then $A(t)$ and $B(t)$ satisfy the ordinary differential equations

$$A'(t) = -bA(t) + \frac{1}{2}\sigma^2 A(t)^2 + \beta(e^{aA(t)} - 1) + \theta,$$

$$B'(t) = bcA(t) + \alpha(e^{aA(t)} - 1),$$

$$A(0) = B(0) = 0.$$

It is easy to see that $\lim_{t\to\infty} A(t) = y$, where y satisfies

$$-by + \frac{1}{2}\sigma^2 y^2 + \beta(e^{ay} - 1) + \theta = 0 \tag{7}$$

if the equation has a solution, and $\lim_{t\to\infty} A(t) = +\infty$ otherwise.

We claim that $y(\theta)$ is the smaller solution of (7) for $\theta \leq \theta_c$, where

$$\theta_c = \max_{y\in\mathbb{R}^+}\left\{by - \frac{1}{2}\sigma^2 y^2 - \beta(e^{ay} - 1)\right\} = by_c - \frac{1}{2}\sigma^2 y_c^2 - \beta(e^{ay_c} - 1), \tag{8}$$

where y_c is the unique positive solution to the equation $b = \sigma^2 y_c + \beta a e^{ay_c}$. This equation has a unique positive solution since $b > a\beta$.

Let us give more explanations here. The function $F(y) := -by + \frac{1}{2}\sigma^2 y^2 + \beta(e^{ay} - 1) + \theta$ is convex, and $F(y) = 0$ has two distinct solutions when $\theta < \theta_c$ and a unique positive solution when $\theta = \theta_c$. When $\theta < 0$, $y(\theta)$ is the unique negative solution of $F(y) = 0$ and, when $0 \leq \theta \leq \theta_c$, $y(\theta)$ is the smaller nonnegative solution of $F(y) = 0$. Hence, we have

$$\Gamma(\theta) := \lim_{t\to\infty}\frac{1}{t}\log u(\theta, t, r) = \begin{cases} bcy(\theta) + \alpha(e^{ay(\theta)} - 1) & \text{if } \theta \leq \theta_c, \\ +\infty & \text{otherwise.} \end{cases}$$

Since $b > a\beta$, for positive and sufficiently small y in (8), we have $by - \frac{1}{2}\sigma^2 y^2 - \beta(e^{ay} - 1) \sim by - \beta ay > 0$ and, thus, $\theta_c > 0$. Also, $\Gamma(\theta)$ is differentiable for $\theta < \theta_c$, and differentiating (7) with respect to θ we obtain

$$\frac{\partial y}{\partial \theta} = \frac{1}{b - \sigma^2 y - \beta a e^{ay}} \to +\infty$$

as $\theta \uparrow \theta_c$, since $y \uparrow y_c$ as $\theta \uparrow \theta_c$. Therefore, we have the essential smoothness and, by the Gärtner-Ellis theorem (for the definition of essential smoothness and a statement of the

Gärtner-Ellis theorem, we refer the reader to [7]), $((1/t) \int_0^t r_s \, ds \in \cdot)$ satisfies a large deviation principle with rate function

$$I(x) = \sup_{\theta \in \mathbb{R}} \{\theta x - bcy(\theta) - \alpha(e^{ay(\theta)} - 1)\}.$$

(ii) For a pair (r_t, N_t), the generator is given by

$$\mathscr{A} f(r, n) = bc \frac{\partial f}{\partial r} - br \frac{\partial f}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 f}{\partial r^2} + (\alpha + \beta r)[f(r + a, n + 1) - f(r, n)].$$

Let $u(t, r) := u(\theta, t, r) := \mathbb{E}[e^{\theta N_t} \mid r_0 = r]$. Consider $f(t, r_t, N_t) = \mathbb{E}[e^{\theta N_T} \mid r_t, N_t]$, with $f(t, r_t, N_t)_{t \leq T}$ a martingale only if $\partial f / \partial t + \mathscr{A} f = 0$ and $f(T, r_T, N_T) = e^{\theta N_T}$. Let $f(t, r, n) = u(t, r)e^{\theta n}$, and using $t \mapsto T - t$ to change the backward equation to the forward equation, we have

$$\frac{\partial u}{\partial t} = bc \frac{\partial u}{\partial r} - br \frac{\partial u}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 u}{\partial r^2} + (\alpha + \beta r)[u(t, r + a)e^{\theta} - u(t, r)], \qquad u(0, r) \equiv 1.$$

Now, trying $u(\theta, t, r) = e^{A(t)r + B(t)}$, we obtain

$$A'(t) = -bA(t) + \frac{1}{2}\sigma^2 A^2(t) + \beta(e^{aA(t)+\theta} - 1),$$
$$B'(t) = bcA(t) + \alpha(e^{aA(t)+\theta} - 1),$$
$$A(0) = B(0) = 0.$$

Hence, we have $\lim_{t \to \infty} A(t) = y(\theta)$, where $y(\theta)$ satisfies

$$-by(\theta) + \frac{1}{2}\sigma^2 y^2(\theta) + \beta(e^{ay(\theta)+\theta} - 1) = 0 \qquad (9)$$

if the equation has a solution, and $\lim_{t \to \infty} A(t) = +\infty$ otherwise. Similar to the arguments in (i), $y(\theta)$ is the smaller solution of (9) when $\theta \leq \theta_c$ and $+\infty$ otherwise. We determine θ_c as follows. We can rewrite (9) as

$$e^{\theta} = \left(by - \frac{1}{2}\sigma^2 y^2 + \beta\right) \frac{1}{\beta} e^{-ay}.$$

Let

$$\begin{aligned}
\theta_c &= \log \max_{y \in \mathbb{R}^+} \left\{ \left(by - \frac{1}{2}\sigma^2 y^2 + \beta\right) \frac{1}{\beta} e^{-ay} \right\} \\
&= \log \left\{ \left(by_c - \frac{1}{2}\sigma^2 y_c^2 + \beta\right) \frac{1}{\beta} e^{-ay_c} \right\} \\
&= \log \left(\frac{b - \sigma^2 y_c}{a\beta} \right) - ay_c \\
&= \log \left(\frac{\sqrt{\sigma^4 + a^2 b^2 + 2a^2 \sigma^2 \beta} - \sigma^2}{a^2 \beta} \right) - \frac{\sigma^2 + ab - \sqrt{\sigma^4 + a^2 b^2 + 2a^2 \sigma^2 \beta}}{\sigma^2}, \qquad (10)
\end{aligned}$$

where

$$y_c = \frac{(\sigma^2 + ab - \sqrt{(\sigma^2 + ab)^2 - 2a\sigma^2(b - a\beta)})}{a\sigma^2}.$$

Hence, we have

$$\Gamma(\theta) := \lim_{t \to \infty} \frac{1}{t} \log u(\theta, t, r) = \begin{cases} bcy(\theta) + \alpha(e^{ay(\theta)+\theta} - 1) & \text{if } \theta \leq \theta_c, \\ +\infty & \text{otherwise.} \end{cases}$$

Since $b > \beta a$, for positive and sufficiently small y in (10), we have

$$\left(by - \frac{1}{2}\sigma^2 y^2 + \beta\right)\frac{1}{\beta}e^{-ay} \sim \left(\frac{b}{\beta}y + 1\right)(1 - ay) \sim 1 + \left(\frac{b}{\beta} - 1\right)y > 1$$

and, thus, $\theta_c > 0$. Also, $\Gamma(\theta)$ is differentiable for $\theta < \theta_c$, and differentiating (9) with respect to θ we obtain

$$\frac{\partial y}{\partial \theta} = \frac{\beta e^{ay+\theta}}{b - \sigma^2 y - \beta a e^{ay+\theta}} \to +\infty$$

as $\theta \uparrow \theta_c$ since $y \uparrow y_c$ as $\theta \uparrow \theta_c$. By (10), we have $e^{\theta_c} = (b - \sigma^2 y_c)e^{-ay_c}/a\beta$. Therefore, we have the essential smoothness and by the Gärtner-Ellis theorem, $(N_t/t \in \cdot)$ satisfies a large deviation principle with rate function

$$I(x) = \sup_{\theta \in \mathbb{R}}\{\theta x - bcy(\theta) - \alpha(e^{ay(\theta)+\theta} - 1)\}.$$

Proof of Proposition 2. The lecture notes [11] by Hairer give the criterion for the existence and uniqueness of the invariant probability measure for Markov processes. Suppose that we have a jump diffusion process with generator $\mathcal{A}$. If we can find a u such that $u \geq 0$ and $\mathcal{A}u \leq C_1 - C_2 u$ for some constants $C_1, C_2 > 0$, then there exists an invariant probability measure. In our problem, recall that

$$\mathcal{A}u(r) = bc\frac{\partial u}{\partial r} - br\frac{\partial u}{\partial r} + \frac{1}{2}\sigma^2 r\frac{\partial^2 u}{\partial r^2} + (\alpha + \beta r)[u(r + a) - u(r)].$$

Let us try $u(r) = r$, and choose $0 < C_2 < b - a\beta$ and $C_1 > \alpha a + bc$. Then we have

$$\mathcal{A}u + C_2 u = bc - br + \alpha a + \beta ar + C_2 r = (bc + \alpha a) + (\beta a - b + C_2)r \leq bc + \alpha a \leq C_1.$$

Next, we will prove the uniqueness of the invariant probability measure. To obtain the uniqueness of the invariant probability measure, it is sufficient to prove that, for any $x, y > 0$, there exists some $T > 0$ such that $\mathcal{P}^x(T, \cdot)$ and $\mathcal{P}^y(T, \cdot)$ are not mutually singular. Here $\mathcal{P}^x(T, \cdot) = \mathbb{P}(r_T^x \in \cdot)$, where r_T^x is r_T starting at $r_0 = x$. For any $x, y > 0$, conditional on the event that r_t^x and r_t^y have no jumps during the time interval $(0, T)$, which has a positive probability, the laws of $\mathcal{P}^x(T, \cdot)$ and $\mathcal{P}^y(T, \cdot)$ are absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^+$, which implies that $\mathcal{P}^x(T, \cdot)$ and $\mathcal{P}^y(T, \cdot)$ are not mutually singular.

Proof of Proposition 3. By the Kolmogorov equation, $u(t, r) = \mathbb{E}[e^{-\theta r_t} \mid r_0 = r]$ satisfies

$$\frac{\partial u}{\partial t} = bc\frac{\partial u}{\partial r} - br\frac{\partial u}{\partial r} + \frac{1}{2}\sigma^2 r\frac{\partial^2 u}{\partial r^2} + (\alpha + \beta r)[u(t, r + a) - u(t, r)], \qquad u(0, r) = e^{-\theta r}.$$

Now, trying $u(t, r) = e^{A(t)r + B(t)}$, we obtain the desired results.

Proof of Proposition 4. (i) By the Feynman–Kac formula, $P(t, T, r)$ satisfies the integro-partial differential equation

$$\frac{\partial P}{\partial t} + bc\frac{\partial P}{\partial r} - br\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2 r\frac{\partial^2 P}{\partial r^2} + (\alpha + \beta r)[P(t, T, r + a) - P(t, T, r)] - rP(t, T, r)$$
$$= 0,$$
$$P(T, T, r) = 1.$$

Let us try $P(t, T, r) = e^{A(t)r + B(t)}$. We obtain

$$A'(t) - bA(t) + \tfrac{1}{2}\sigma^2 A(t)^2 + \beta(e^{aA(t)} - 1) - 1 = 0,$$
$$B'(t) + bcA(t) + \alpha(e^{aA(t)} - 1) = 0,$$
$$A(T) = B(T) = 0.$$

(ii) By using the same arguments as in the proof of Theorem 3, we obtain the asymptotic result

$$\lim_{T \to \infty} \frac{1}{T} \log P(t, T, r) = bcx_* + \alpha(e^{ax_*} - 1),$$

where x_* is the unique negative solution of

$$-bx + \tfrac{1}{2}\sigma^2 x^2 + \beta(e^{ax} - 1) - 1 = 0.$$

Acknowledgements

References

[1] BACRY, E., DELATTRE, S., HOFFMANN, M. AND MUZY, J. F. (2013). Modelling microstructure noise with mutually exciting point processes. *Quant. Finance* **13**, 65–77.
[2] BACRY, E., DELATTRE, S., HOFFMANN, M. AND MUZY, J. F. (2013). Scaling limits for Hawkes processes and application to financial statistics. *Stoch. Process. Appl.* **123**, 2475–2499.
[3] BORDENAVE, C. AND TORRISI, G. L. (2007). Large deviations of Poisson cluster processes. *Stoch. Models* **23**, 593–625.
[4] BRÉMAUD, P. AND MASSOULIÉ, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Prob.* **24**, 1563–1588.
[5] COX, J. C., INGERSOLL, J. E., JR. AND ROSS, S. A. (1985). A theory of the term structure of interest rates. *Econometrica* **53**, 385–407.
[6] DAHL, M. (2004). Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. *Insurance Math. Econom.* **35**, 113–136.
[7] DEMBO, A. AND ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd edn. Springer, New York.
[8] DUFFIE, D. (2005). Credit risk modeling with affine processes. *J. Banking Finance* **29**, 2751–2802.
[9] ERRAIS, E., GIESECKE, K. AND GOLDBERG, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM J. Financial Math.* **1**, 642–665.
[10] FELLER, W. (1951). Two singular diffusion problems. *Ann. Math. (2)* **54**, 173–182.
[11] HAIRER, M. (2010). Convergence of Markov processes. Lecture Notes, University of Warwick. Available at http://www.hairer.org/notes/Convergence.pdf.
[12] HAWKES, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58**, 83–90.
[13] HAWKES, A. G. AND OAKES, D. (1974). A cluster process representation of a self-exciting process. *J. Appl. Prob.* **11**, 493–503.

[14] HESTON, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Studies* **6,** 327–343.

[15] PENG, X. AND KOU, S. (2009). Default clustering and valuation of collateralized debt obligations. Working Paper, Columbia University.

[16] STABILE, G. AND TORRISI, G. L. (2010). Risk processes with non-stationary Hawkes claims arrivals. *Methodol. Comput. Appl. Prob.* **12,** 415–429.

[17] VARADHAN, S. R. S. (1984). *Large Deviations and Applications*. SIAM, Philadelphia, PA.

[18] ZHU, L. (2013). Central limit theorem for nonlinear Hawkes processes. *J. Appl. Prob.* **50,** 760–771.

[19] ZHU, L. (2013). Moderate deviations for Hawkes processes. *Statist. Prob. Lett.* **83,** 885–890.

[20] ZHU, L. (2013). Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. *Insurance Math. Econom.* **53,** 544–550.

[21] ZHU, L. (2014). Large deviations for Markovian nonlinear Hawkes processes. To appear in *Ann. Appl. Prob.*

[22] ZHU, L. (2014). Process-level large deviations for nonlinear Hawkes point processes. *Ann. Inst. H. Poincaré Prob. Statist.* **50,** 845–871.