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

LINGJIONG ZHU

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. Laplace transforms and limit theorems have been obtained, including law of large numbers, central limit theorems and large deviations.

1. Introduction and Main Results

1.1. **Cox-Ingersoll-Ross Process.** A Cox-Ingersoll-Ross process is a stochastic process r_t satisfying the following stochastic differential equation,

$$dr_t = b(c - r_t)dt + \sigma\sqrt{r_t}dW_t, \tag{1.1}$$

where W_t is a standard Brownian motion, $b, c, \sigma > 0$ and $2bc \geq \sigma^2$. The process is proposed by Cox, Ingersoll and Ross in Cox et al. [5] to model the short term interest rate. Under the assumption $2bc \geq \sigma^2$, Feller [10] proved that the process is non-negative. Given r_0, it is well known that $\frac{4b}{\sigma^2(1-e^{-bt})}r_t$ follows a non-central χ^2 distribution with degree of freedom $\frac{4bc}{\sigma^2}$ and non-centrality parameter $\frac{4b}{\sigma^2(1-e^{-bt})}r_0e^{-bt}$. As $t \to \infty$, $r_t \to r_\infty$, where r_∞ follows a Gamma distribution with shape parameter $\frac{2bc}{\sigma^2}$ and scale parameter $\frac{\sigma^2}{2b}$. The conditional first and second moments are given by, $s > t$,

$$\mathbb{E}[r_s|r_t] = r_te^{-b(s-t)} + c(1 - e^{-b(s-t)}) \tag{1.2}$$

$$\begin{aligned}\mathbb{E}[r_s^2|r_t] &= r_t\left(2c + \frac{\sigma^2}{b}\right)e^{-b(s-t)} + \left(r_t^2 - r_t\frac{\sigma^2}{b} - 2r_tc\right)e^{-2b(s-t)} \\ &\quad + \left(\frac{c\sigma^2}{2b} + c^2\right)\left(1 - e^{-b(s-t)}\right)^2.\end{aligned} \tag{1.3}$$

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

Date: 19 April 2013. *Revised:* 28 August 2013.

2000 *Mathematics Subject Classification.* 60G07, 60G55, 60F05,60F10.

Key words and phrases. Cox-Ingersoll-Ross process, point processes, Hawkes processes, self-exciting processes, central limit theorem, large deviations.

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

$$dr_t = b(c - r_t)dt + \sigma\sqrt{r_t}dW_t + adN_t, \tag{1.4}$$

where N_t is a homogeneous Poisson process with constant intensity $\lambda > 0$. But in the real world, the occurence of events may not be time-homogeneous and it should have dependence over time. 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 exisence 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 Poisson process by a simple point process which can describe the time dependence in a natural way. The Hawkes process, a simple point process that has 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

$$\lambda_t := \lambda\left(\int_{-\infty}^{t} h(t-s)N(ds)\right), \tag{1.5}$$

where $\lambda(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ is locally integrable, left continuous, $h(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ and we always assume that $\|h\|_{L^1} = \int_0^\infty h(t)dt < \infty$. In (1.5), $\int_{-\infty}^t h(t-s)N(ds)$ stands for $\int_{(-\infty,t)} h(t-s)N(ds) = \sum_{\tau<t} h(t-\tau)$, where τ are the occurences of the points before time t.

In the literature, $h(\cdot)$ and $\lambda(\cdot)$ are often referred to as exciting function and 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. Linear Hawkes process, i.e. the classical Hawkes process, is named after Hawkes, who first invented the model in Hawkes [12]. Nonlinear Hawkes process was first introduced by Brémaud and Massoulié [4].

By the definition of Hawkes process, it has the self-exciting property, i.e. the intensity λ_t increases when you witness a jump. It therefore creates a clustering effect, which is to model the default clustering in finance. When you do not witness new jumps, the intensity λ_t decreases as $h(\cdot)$ decays.

The law of large numbers and central limit theorems for linear Hawkes process have been obtained in Hawkes and Oakes [13]. The law of large numbers and central limit theorem have also been studied in Bacry et al. [1] as a special case of multivariate Hawkes processes. The large deviation principle for linear Hawkes process was obtained in Bordenave and Torrisi [3]. The moderate deviation principle for linear Hawkes process was obtained in Zhu [21]. For nonlinear Hawkes process, the central limit thereom was obtained in Zhu [20] and the large deviations have been studied in Zhu [18] and Zhu [19].

The central limit theorem of Hawkes process has been applied to study the high frequency trading and the microstructure in finance, see e.g. Bacry et al. [1] and Bacry et al. [2] and the large deviations result has been applied to study the ruin probabilities in insurance, see e.g. Stabile and Torrisi [16] and Zhu [22].

1.3. **A Cox-Ingersoll-Ross Process with Hawkes Jumps.** In this paper, we propose a stochastic process r_t that satisfies the following stochastic differential equation,

$$dr_t = b(c - r_t)dt + adN_t + \sigma\sqrt{r_t}dW_t, \tag{1.6}$$

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

- $b > a\beta$. This condition is needed to guarantee that there exists a unique stationary process r_∞ which satisfies the dynamics (1.6).
- $2bc \geq \sigma^2$. 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 non-negative properties of the classical Cox-Intersoll-Ross process. In addition, it contains the Hawkes jumps, which have the self-exciting property create a clustering effect.

Clearly, the Cox-Ingersoll-Ross process we proposed in (1.6) includes the classical Cox-Ingersoll-Ross process and the classical linear Hawkes process with exponential exciting function. We summarize this in the following.

(1) When $a = 0$ or $\alpha = \beta = 0$, it reduces to the classical Cox-Ingersoll-Ross process, i.e.
$$dr_t = b(c - r_t)dt + \sigma\sqrt{r_t}dW_t.$$
(2) When $\beta = 0$ and $a, \alpha > 0$, it reduces to the Cox-Ingersoll-Ross process with Poisson jumps, i.e.
$$dr_t = b(c - r_t)dt + \sigma\sqrt{r_t}dW_t + adN_t,$$
where N_t is a homogeneous Poisson process with intensity α.
(3) When $c = 0$ and $\sigma = 0$, N_t reduces to a Hawkes process with intensity $\lambda_t = \alpha + \beta r_t$, where
$$dr_t = -br_t dt + adN_t,$$
and it is easy to see that the intensity λ_t indeed satisfies
$$\lambda_t = \alpha + \beta \int_0^t ae^{-b(t-s)}N(ds),$$
which implies that N_t is a classical linear Hawkes process with $\lambda(z) = \alpha + \beta z$ and $h(t) = ae^{-bt}$.

It is easy to observe that r_t is Markovian with generator

$$\mathcal{A}f(r) = bc\frac{\partial f}{\partial r} - br\frac{\partial f}{\partial r} + \frac{1}{2}\sigma^2 r\frac{\partial^2 f}{\partial r^2} + (\alpha + \beta r)[f(r + a) - f(r)]. \tag{1.7}$$

1.4. **Main Results.** In this section, we will 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.

The following proposition gives the formulas for the conditional first moment and second moment of the Cox-Ingersoll-Ross process with Hawkes jumps.

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

$$\mathbb{E}[r_s|r_t] = \frac{bc + a\alpha}{b - a\beta} - e^{-(b-a\beta)(s-t)} \left[\frac{bc + a\alpha}{b - a\beta} - r_t\right]. \tag{1.8}$$

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

$$\begin{aligned}
&\mathbb{E}[r_s^2|r_t] \\
&= r_t^2 e^{-2(b-a\beta)(s-t)} \\
&\quad + \left[(2bc + \sigma^2 + 2a\alpha + a^2\beta)\frac{bc + a\alpha}{2(b - a\beta)^2} + \frac{a^2\alpha}{2(b - a\beta)}\right][1 - e^{-2(b-a\beta)(s-t)}] \\
&\qquad - (2bc + \sigma^2 + 2a\alpha + a^2\beta)\frac{bc + a\alpha}{(b - a\beta)^2}[e^{-(b-a\beta)(s-t)} - e^{-2(b-a\beta)(s-t)}] \\
&\qquad\quad + (2bc + \sigma^2 + 2a\alpha + a^2\beta)\frac{r_t}{b - a\beta}[e^{-(b-a\beta)(s-t)} - e^{-2(b-a\beta)(s-t)}].
\end{aligned} \tag{1.9}$$

Remark 2. *Let $a = 0$ in (1.8), we get $\mathbb{E}[r_s|r_t] = c - e^{-b(s-t)}(c - r_t) = r_t e^{-b(s-t)} + c(1 - e^{-b(s-t)})$, which recovers (1.2). Similarly, by letting $a = 0$ in (1.9), we recover (1.3).*

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

(i)

$$\frac{1}{t}\int_0^t r_s ds \to \frac{bc + a\alpha}{b - a\beta}, \quad \text{in } L^2(\mathbb{P}) \text{ as } t \to \infty. \tag{1.10}$$

(ii)

$$\frac{N_t}{t} \to \frac{b(\alpha + \beta c)}{b - a\beta}, \quad \text{in } L^2(\mathbb{P}) \text{ as } t \to \infty. \tag{1.11}$$

Theorem 4 (Central Limit Theorem)**.** *For any $r_0 := r \in \mathbb{R}^+$,*

(i)

$$\frac{\int_0^t r_s ds - \frac{bc + a\alpha}{b - a\beta}t}{\sqrt{t}} \to N\left(0, \frac{a^2\alpha(b - a\beta) + (a^2\beta + \sigma^2)(bc + a\alpha)}{(b - a\beta)^3}\right), \tag{1.12}$$

in distribution as $t \to \infty$.

(ii)

$$\frac{N_t - \frac{b(\alpha+\beta c)}{b - a\beta}t}{\sqrt{t}} \to N\left(0, \frac{b^3a^2(\alpha + \beta c) + 4\sigma^2b^2(bc + a\alpha)}{a^2(b - a\beta)^3}\right), \tag{1.13}$$

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 non-negative, lower semicontinuous and for any measurable set A, we have

$$-\inf_{x\in A^o} I(x) \leq \liminf_{n\to\infty} \frac{1}{n}\log P_n(A) \leq \limsup_{n\to\infty} \frac{1}{n}\log P_n(A) \leq -\inf_{x\in\overline{A}} I(x). \tag{1.14}$$

Here, A^o is the interior of A and $\overline{A}$ is its closure. We refer to Dembo and Zeitouni [7] and Varadhan [17] for general background of the theory and the applications of large deviations.

Theorem 5 (Large Deviation Principle)**.** *For any $r_0 := r \in \mathbb{R}^+$,*

(i) $(\frac{1}{t}\int_0^t r_s ds \in \cdot)$ satisfies a large deviation principle with rate function

$$I(x) = \sup_{\theta\leq\theta_c}\left\{\theta x - bcy(\theta) - \alpha(e^{ay(\theta)}-1)\right\}, \tag{1.15}$$

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

$$-by + \frac{1}{2}\sigma^2 y^2 + \beta(e^{ay}-1) + \theta = 0, \tag{1.16}$$

and

$$\theta_c = by_c - \frac{1}{2}\sigma^2 y_c^2 - \beta(e^{ay_c}-1), \tag{1.17}$$

where y_c is 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

$$I(x) = \sup_{\theta\leq\theta_c}\left\{\theta x - bcy(\theta) - \alpha(e^{ay(\theta)+\theta}-1)\right\}, \tag{1.18}$$

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

$$-by(\theta) + \frac{1}{2}\sigma^2 y^2(\theta) + \beta(e^{ay(\theta)+\theta}-1) = 0, \tag{1.19}$$

and

$$\theta_c = \log\left(\frac{\sqrt{\sigma^4+a^2b^2+2a^2\sigma^2\beta}-\sigma^2}{a^2\beta}\right) - \frac{\sigma^2+ab-\sqrt{\sigma^4+a^2b^2+2a^2\sigma^2\beta}}{\sigma^2}. \tag{1.20}$$

Remark 6. *It is easy to see that when $c = 0$ and $\sigma = 0$, our results of Theorem 3 (ii), Theorem 4 (ii) and Theorem 5 (ii) are consistent with the law of large numbers and central limit theorem results for linear Hawkes process with exponential exciting function as in Bacry et al.* [1] *and the large deviation principle as in Bordenave and Torrisi* [3]*.*

Proposition 7. *Assume $b > a\beta$ and $2bc \geq \sigma^2$. Then, there exists a unique invariant probability measure for r_t.*

Proposition 8. *For any $\theta > 0$, the Laplace transform of r_t satisfies $\mathbb{E}[e^{-r_t}|r_0 = r] = e^{A(t)r+B(t)}$, where $A(t), B(t)$ satisfy the ordinary differential equations*

$$\begin{cases} A'(t) = -bA(t) + \frac{1}{2}\sigma^2 A(t)^2 + \beta(e^{aA(t)}-1), \\ B'(t) = bcA(t) + \alpha(e^{aA(t)}-1), \\ A(0) = -\theta, B(0) = 0. \end{cases} \tag{1.21}$$

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

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

$$P(t,T,r) := \mathbb{E}\left[e^{-\int_t^T r_s ds}\middle| r_t = r\right]. \tag{1.22}$$

Proposition 9. *(i)* $P(t,T,r) = e^{A(t)r+b(t)}$*, where* $A(t), B(t)$ *satisfy the following ordinary differential equations,*

$$\begin{cases} A'(t) - bA(t) + \frac{1}{2}\sigma^2 A(t)^2 + \beta(e^{aA(t)} - 1) - 1 = 0, \\ B'(t) + bcA(t) + \alpha(e^{aA(t)} - 1) = 0, \\ A(T) = B(T) = 0. \end{cases} \tag{1.23}$$

(ii) We have the following asymptotic result,

$$\lim_{T\to\infty} \frac{1}{T} \log P(t,T,r) = bcx_* + \alpha(e^{ax_*} - 1), \tag{1.24}$$

where x_* *is the unique negative solution to the following equation,*

$$-bx + \frac{1}{2}\sigma^2 x^2 + \beta(e^{ax} - 1) - 1 = 0. \tag{1.25}$$

Remark 10. *A natural way to generalize the Cox-Ingersoll-Ross process with Hawkes jumps is to allow the jump size to be random, i.e.*

$$dr_t = b(c - r_t)dt + \sigma\sqrt{r_t}dW_t + dJ_t, \tag{1.26}$$

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

We can write down the generator as

$$\mathcal{A}f(r) = bc\frac{\partial f}{\partial r} - br\frac{\partial f}{\partial r} + \frac{1}{2}\sigma^2 r\frac{\partial^2 f}{\partial r^2} + (\alpha + \beta r)\int_{\mathbb{R}^+}[f(r+a) - f(r)]Q(da). \tag{1.27}$$

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

Remark 11. *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* (1.6) *and* N_t *is a simple point process with intensity* $\lambda(r_t)$*, where* $\lambda(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ *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 expectations on both sides of (1.6), we have

$$d\mathbb{E}[r_t] = b(c - \mathbb{E}[r_t])dt + a(\alpha + \beta\mathbb{E}[r_t])dt, \tag{2.1}$$

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

$$\mathbb{E}[r_s|r_t] = \frac{bc + a\alpha}{b - a\beta} - e^{-(b-a\beta)(s-t)}\left[\frac{bc + a\alpha}{b - a\beta} - r_t\right]. \tag{2.2}$$

(ii) By Itô's formula, we have

$$d(r_t^2) = 2r_t[b(c-r_t)dt + \sigma\sqrt{r_t}dW_t] + \sigma^2 r_t dt + 2r_t a dN_t + a^2 dN_t. \tag{2.3}$$

Taking expectations on both sides, we get

$$\frac{d\mathbb{E}[r_t^2]}{dt} = 2bc\mathbb{E}[r_t] - 2b\mathbb{E}[r_t^2] + \sigma^2\mathbb{E}[r_t] + 2a(\alpha\mathbb{E}[r_t] + \beta\mathbb{E}[r_t^2]) + a^2\alpha + a^2\beta\mathbb{E}[r_t]. \tag{2.4}$$

This implies that

$$\begin{aligned}
&\mathbb{E}[r_s^2|r_t]e^{2(b-a\beta)s} - r_t^2 e^{2(b-a\beta)t} \\
&= (2bc+\sigma^2+2a\alpha+a^2\beta)\int_t^s e^{2(b-a\beta)u}\mathbb{E}[r_u|r_t]du + a^2\alpha\int_t^s e^{2(b-a\beta)u}du \\
&= \left[(2bc+\sigma^2+2a\alpha+a^2\beta)\frac{bc+a\alpha}{2(b-a\beta)^2} + \frac{a^2\alpha}{2(b-a\beta)}\right][e^{2(b-a\beta)s} - e^{2(b-a\beta)t}] \\
&\qquad - (2bc+\sigma^2+2a\alpha+a^2\beta)\frac{bc+a\alpha}{b-a\beta}\frac{e^{(b-a\beta)t}}{(b-a\beta)}[e^{(b-a\beta)s} - e^{(b-a\beta)t}] \\
&\qquad + (2bc+\sigma^2+2a\alpha+a^2\beta)r_t\frac{e^{(b-a\beta)t}}{(b-a\beta)}[e^{(b-a\beta)s} - e^{(b-a\beta)t}],
\end{aligned} \tag{2.5}$$

which yields (1.9). □

Proof of Theorem 3. (i) To prove the convergence in the $L^2(\mathbb{P})$ norm, we need to show that

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{t}\int_0^t r_s ds - \frac{bc+a\alpha}{b-a\beta}\right)^2 \\
&= \frac{1}{t^2}\mathbb{E}\left(\int_0^t r_s ds\right)^2 - \frac{2}{t}\int_0^t \mathbb{E}[r_s]ds\cdot\frac{bc+a\alpha}{b-a\beta} + \left(\frac{bc+a\alpha}{b-a\beta}\right)^2 \to 0,
\end{aligned} \tag{2.6}$$

as $t\to\infty$. From (1.8) of Proposition 1, it is clear that $\frac{1}{t}\int_0^t \mathbb{E}[r_s]ds \to \frac{bc+a\alpha}{b-a\beta}$ as $t\to\infty$. Therefore, it suffices to show that $\frac{1}{t^2}\mathbb{E}\left(\int_0^t r_s ds\right)^2 \to \frac{bc+a\alpha}{b-a\beta}$ as $t\to\infty$. Applying (1.8) of Proposition 1, we get

$$\begin{aligned}
&\frac{1}{t^2}\mathbb{E}\left(\int_0^t r_s ds\right)^2 \\
&= \frac{2}{t^2}\iint_{0<s_1<s_2<t}\mathbb{E}[r_{s_1}\mathbb{E}[r_{s_2}|r_{s_1}]]ds_1 ds_2 \\
&= \frac{2}{t^2}\iint_{0<s_1<s_2<t}\frac{bc+a\alpha}{b-a\beta}\mathbb{E}[r_{s_1}]ds_1 ds_2 \\
&\qquad - \frac{2}{t^2}\iint_{0<s_1<s_2<t} e^{-(b-a\beta)(s_2-s_1)}\left[\frac{bc+a\alpha}{b-a\beta}\mathbb{E}[r_{s_1}] - \mathbb{E}[r_{s_1}^2]\right]ds_1 ds_2.
\end{aligned} \tag{2.7}$$

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 only depending on r, say $M(r)$. Therefore,

$$\begin{aligned}
&\left|\frac{2}{t^2}\iint_{0<s_1<s_2<t} e^{-(b-a\beta)(s_2-s_1)}\left[\frac{bc+a\alpha}{b-a\beta}\mathbb{E}[r_{s_1}] - \mathbb{E}[r_{s_1}^2]\right]ds_1 ds_2\right| \\
&\le \frac{2}{t^2}M(r)\left[\frac{bc+a\alpha}{b-a\beta}+1\right]\iint_{0<s_1<s_2<t} e^{-(b-a\beta)(s_2-s_1)}ds_1 ds_2 \to 0,
\end{aligned} \tag{2.8}$$

as $t \to \infty$. Again by (1.8), it is easy to check that

$$\frac{2}{t^2}\iint_{0<s_1<s_2<t}\frac{bc+a\alpha}{b-a\beta}\mathbb{E}[r_{s_1}]ds_1ds_2 \to \left(\frac{bc+a\alpha}{b-a\beta}\right)^2, \tag{2.9}$$

as $t \to \infty$. Hence, we 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 martinagle and

$$\mathbb{E}\left[\left(\frac{N_t - \int_0^t \lambda_s}{t}\right)^2\right] = \frac{1}{t^2}\mathbb{E}\left[\int_0^t \lambda_s ds\right] = \frac{\alpha}{t} + \frac{\beta}{t^2}\int_0^t \mathbb{E}[r_s]ds \to 0, \tag{2.10}$$

as $t \to \infty$ by Proposition 1. Therefore, we have

$$\frac{N_t}{t} - \alpha - \frac{\beta}{t}\int_0^t r_s ds \to 0, \tag{2.11}$$

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

Remark 12. *The L^2 convergence in Theorem 3 implies the convergence in probability. Indeed, the convergence in Theorem 3 also holds almost surely by using Proposition 7 and ergodic theorem. For example, by ergodic theorem, $\frac{1}{t}\int_0^t r_s ds \to \mathbb{E}[r_\infty]$ 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 $\mathbb{E}[r_\infty] = \frac{bc+a\alpha}{b-a\beta}$. Similarly, we can show that $\frac{N_t}{t} \to \frac{b(\alpha+\beta c)}{b-a\beta}$ as $t \to \infty$ almost surely. Indeed, the a.s. 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 4. (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

$$\mathcal{A}f(r) = K\left[(a\beta - b)r + (\alpha a + bc)\right]. \tag{2.12}$$

Let us choose $K = \frac{1}{b-a\beta} > 0$. Then, we have

$$\int_0^t r_s ds - \frac{bc+a\alpha}{b-a\beta}t = \left[f(r_t) - f(r_0) - \int_0^t \mathcal{A}f(r_s)ds\right] - f(r_t) + f(r_0). \tag{2.13}$$

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

$$\frac{\mathbb{E}[f(r_t)]}{\sqrt{t}} = \frac{K\mathbb{E}[r_t]}{\sqrt{t}} = \frac{K}{\sqrt{t}}\left\{\frac{bc+a\alpha}{b-a\beta} - e^{-(b-a\beta)t}\left[\frac{bc+a\alpha}{b-a\beta} - r_0\right]\right\} \to 0, \tag{2.14}$$

as $t \to \infty$ by Proposition 1. Therefore, $\frac{f(r_t)}{\sqrt{t}} \to 0$ as $t \to \infty$ in probability. The quadratic variation of the martingale $f(r_t) - f(r_0) - \int_0^t \mathcal{A}f(r_s)ds$ is the same as the quadratic variation of $f(r_t) = \frac{1}{b-a\beta}r_t$, which is the same as the quadratic variation of $\frac{1}{b-a\beta}(aN_t + \int_0^t \sigma\sqrt{r_s}dW_s)$, which is $\frac{1}{(b-a\beta)^2}\left[a^2N_t + \sigma^2\int_0^t r_s ds\right]$. By the law of large numbers in Theorem 3, we have

$$\frac{1}{t}\frac{1}{(b-a\beta)^2}\left[a^2N_t + \sigma^2\int_0^t r_s ds\right] \to \frac{1}{(b-a\beta)^2}\left[a^2\alpha + \frac{(a^2\beta+\sigma^2)(bc+a\alpha)}{b-a\beta}\right], \tag{2.15}$$

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

$$\frac{\int_0^t r_s ds - \frac{bc+a\alpha}{b-a\beta}t}{\sqrt{t}} \to N\left(0, \frac{a^2\alpha(b-a\beta)+(a^2\beta+\sigma^2)(bc+a\alpha)}{(b-a\beta)^3}\right), \tag{2.16}$$

in distribution as $t \to \infty$.

(ii) From (1.6), 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

(2.17)

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

where $f(r) = -\frac{b}{a(b-a\beta)}$ 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 as in (i). Now,

$$\left[f(r_t) - f(r_0) - \int_0^t \mathcal{A}f(r_s)ds\right] - \frac{\sigma}{a}\int_0^t \sqrt{r_s}dW_s \tag{2.18}$$

is a martingale and it has the same quadratic variation as

$$-\frac{b}{b-a\beta}N_t - \frac{b\sigma}{a(b-a\beta)}\int_0^t \sqrt{r_s}dW_s - \frac{\sigma}{a}\int_0^t \sqrt{r_s}dW_s, \tag{2.19}$$

which has quadratic variation $\frac{b^2}{(b-a\beta)^2}N_t + \frac{4\sigma^2 b^2}{a^2(b-a\beta)^2}\int_0^t r_s ds$. By law of large numbers, i.e. Theorem 3, we have

$$\begin{aligned} &\frac{1}{t}\left[\frac{b^2}{(b-a\beta)^2}N_t + \frac{4\sigma^2 b^2}{a^2(b-a\beta)^2}\int_0^t r_s ds\right] \\ &\to \frac{b^2}{(b-a\beta)^2}\frac{b(a\alpha+\beta c)}{b-a\beta} + \frac{4\sigma^2 b^2}{a^2(b-a\beta)^2}\frac{bc+a\alpha}{b-a\beta}, \end{aligned} \tag{2.20}$$

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

$$\frac{N_t - \frac{b(\alpha+\beta c)}{b-a\beta}t}{\sqrt{t}} \to N\left(0, \frac{b^3a^2(\alpha+\beta c)+4\sigma^2 b^2(bc+a\alpha)}{a^2(b-a\beta)^3}\right), \tag{2.21}$$

in distribution as $t \to \infty$. □

Proof of Theorem 5. (i) Let $u(\theta,t,r) := \mathbb{E}[e^{\theta\int_0^t r_s ds}]$. Then, by Feynman-Kac formula, we have

(2.22)

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

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

$$\begin{cases} A'(t)=-bA(t)+\frac{1}{2}\sigma^2A(t)^2+\beta(e^{aA(t)}-1)+\theta, \\ B'(t)=bcA(t)+\alpha(e^{aA(t)}-1), \\ A(0)=B(0)=0. \end{cases} \tag{2.23}$$

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

$$-by+\frac{1}{2}\sigma^2y^2+\beta(e^{ay}-1)+\theta=0, \tag{2.24}$$

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 the equation (2.24) for $\theta\leq\theta_c$, where

$$\begin{aligned}\theta_c&=\max_{y\in\mathbb{R}^+}\left\{by-\frac{1}{2}\sigma^2y^2-\beta(e^{ay}-1)\right\} \\ &=by_c-\frac{1}{2}\sigma^2y_c^2-\beta(e^{ay_c}-1),\end{aligned} \tag{2.25}$$

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

Let us give more explanations here. The function $F(y):=-by+\frac{1}{2}\sigma^2y^2+\beta(e^{ay}-1)+\theta$ is convex and have two distinct solutions of $F(y)=0$ when $\theta<\theta_c$ and has 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 non-negative solution of $F(y)=0$.

Hence, we have

$$\Gamma(\theta):=\lim_{t\to\infty}\frac{1}{t}\log u(\theta,t,r)=\begin{cases} bcy(\theta)+\alpha(e^{ay(\theta)}-1) & \text{if } \theta\leq\theta_c \\ +\infty & \text{otherwise} \end{cases}. \tag{2.26}$$

Since $b>a\beta$, for y being positive and sufficiently small in (2.25), we have $by-\frac{1}{2}\sigma^2y^2-\beta(e^{ay}-1)\sim by-\beta ay>0$ and thus $\theta_c>0$. Also $\Gamma(\theta)$ is differentiable for $\theta<\theta_c$ and differentiating with respect to θ to (2.24), we get

$$\frac{\partial y}{\partial\theta}=\frac{1}{b-\sigma^2y-\beta ae^{ay}}\rightarrow+\infty, \tag{2.27}$$

as $\theta\uparrow\theta_c$, since $y\uparrow y_c$ as $\theta\uparrow\theta_c$. Therefore, we have the essential smoothness and by Gärtner-Ellis theorem (for the definition of essential smoothness and statement of Gärtner-Ellis theorem, we refer to Dembo and Zeitouni [7]), $(\frac{1}{t}\int_0^t r_sds\in\cdot)$ satisfies a large deviation principle with rate function

$$I(x)=\sup_{\theta\in\mathbb{R}}\left\{\theta x-bcy(\theta)-\alpha(e^{ay(\theta)}-1)\right\}. \tag{2.28}$$

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

$$\mathcal{A}f(r,n)=bc\frac{\partial f}{\partial r}-br\frac{\partial f}{\partial r}+\frac{1}{2}\sigma^2r\frac{\partial^2f}{\partial r^2}+(\alpha+\beta r)[f(r+a,n+1)-f(r,n)]. \tag{2.29}$$

Let $u(t,r):=u(\theta,t,r):=\mathbb{E}[e^{\theta N_t}|r_0=r]$. Consider $f(t,r_t,N_t)=\mathbb{E}[e^{\theta N_T}|r_t,N_t]$ and $f(t,r_t,N_t)_{t\leq T}$ is a martingale only if $\frac{\partial f}{\partial t}+\mathcal{A}f=0$ and $f(T,r_T,N_T)=e^{\theta N_T}$. Let

$f(t,r,n)=u(t,r)e^{\theta n}$ and make the time change $t\mapsto T-t$ to change the backward equation to the forward equation, we have

$$\begin{cases}\frac{\partial u}{\partial t}=bc\frac{\partial u}{\partial r}-br\frac{\partial u}{\partial r}+\frac{1}{2}\sigma^2 r\frac{\partial^2 u}{\partial r^2}+(\alpha+\beta r)[u(t,r+a)e^{\theta}-u(t,r)],\\ u(0,r)\equiv 1.\end{cases}\tag{2.30}$$

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

$$\begin{cases}A'(t)=-bA(t)+\frac{1}{2}\sigma^2A^2(t)+\beta(e^{aA(t)+\theta}-1),\\ B'(t)=bcA(t)+\alpha(e^{aA(t)+\theta}-1),\\ A(0)=B(0)=0.\end{cases}\tag{2.31}$$

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

$$-by(\theta)+\frac{1}{2}\sigma^2y^2(\theta)+\beta(e^{ay(\theta)+\theta}-1)=0,\tag{2.32}$$

if the above equation (2.32) has a solution and $+\infty$ otherwise. Similar to the arguments in (i), $y(\theta)$ is the smaller solution of (2.32) when $\theta\leq\theta_c$ and $+\infty$ otherwise. θ_c is to be determined as the following. We can rewrite the equation (2.32) as

$$e^{\theta}=\left(by-\frac{1}{2}\sigma^2y^2+\beta\right)\frac{1}{\beta}e^{-ay}.\tag{2.33}$$

Let

(2.34)

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

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

$$\Gamma(\theta):=\lim_{t\to\infty}\frac{1}{t}\log u(\theta,t,r)=\begin{cases}bcy(\theta)+\alpha(e^{ay(\theta)+\theta}-1)&\text{if }\theta\leq\theta_c\\+\infty&\text{otherwise}\end{cases}.\tag{2.35}$$

Since $b>\beta a$, for y being positive and sufficiently small in (2.34), we have $(by-\frac{1}{2}\sigma^2y^2+\beta)\frac{1}{\beta}e^{-ay}\sim(\frac{b}{\beta}y+1)(1-ay)\sim 1+(\frac{b}{\beta}-1)y>1$ and thus $\theta_c>0$. Also $\Gamma(\theta)$ is differentiable for $\theta<\theta_c$ and differentiating with respect to θ to (2.32), we get

$$\frac{\partial y}{\partial\theta}=\frac{\beta e^{ay+\theta}}{b-\sigma^2y-\beta ae^{ay+\theta}}\to+\infty,\tag{2.36}$$

as $\theta\uparrow\theta_c$ since $y\uparrow y_c$ as $\theta\uparrow\theta_c$ and by (2.34), we have $e^{\theta_c}=\frac{b-\sigma^2y_c}{a\beta}e^{-ay_c}$. Therefore, we have the essential smoothness and by Gärtner-Ellis theorem (for the definition of essential smoothness and statement of Gärtner-Ellis theorem, we refer to Dembo and Zeitouni [7]), $(N_t/t \in \cdot)$ satisfies a large deviation principle with rate function

$$I(x) = \sup_{\theta\in\mathbb{R}} \left\{ \theta x - bcy(\theta) - \alpha(e^{ay(\theta)+\theta} - 1) \right\}. \tag{2.37}$$

$\square$

Proof of Proposition 7. The lecture notes [11] by Hairer gives the criterion for the existence and uniqueness of the invariant probability measure for Markov processes. Suppose we have a jump diffusion process with generator $\mathcal{A}$. If we can find u such that $u \geq 0$, $\mathcal{A}u \leq C_1 - C_2 u$ for some constants $C_1, C_2 > 0$, then, there exists an invariant probability measure. In our problem, recall that

$$\mathcal{A}u(r) = bc\frac{\partial u}{\partial r} - br\frac{\partial u}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 u}{\partial r^2} + (\alpha + \beta r)[u(r+a) - u(r)]. \tag{2.38}$$

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

$$\begin{aligned} \mathcal{A}u + C_2 u &= bc - br + \alpha a + \beta a r + C_2 r \\ &= (bc + \alpha a) + (\beta a - b + C_2) r \\ &\leq bc + \alpha a \leq C_1. \end{aligned} \tag{2.39}$$

Next, we will prove the uniqueness of the invariant probability measure. To get 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 law 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. $\square$

Proof of Proposition 8. By Kolmogorov equation, $u(t, r) = \mathbb{E}[e^{-\theta r_t} | r_0 = r]$ satisfies

$$\begin{cases} \frac{\partial u}{\partial t} = bc\frac{\partial u}{\partial r} - br\frac{\partial u}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 u}{\partial r^2} + (\alpha + \beta r)[u(t, r+a) - u(t, r)], \\ u(0, r) = e^{-\theta r}. \end{cases} \tag{2.40}$$

Now, try $u(t, r) = e^{A(t)r + B(t)}$, we get the desired results. $\square$

Proof of Proposition 9. (i) By Feynman-Kac formula, $P(t, T, r)$ satisfies the following integro-partial differential equation,

$$\begin{cases} \frac{\partial P}{\partial t} + bc\frac{\partial P}{\partial r} - br\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 P}{\partial r^2} \\ \qquad\qquad + (\alpha + \beta r)[P(t, T, r+a) - P(t, T, r)] - rP(t, T, r) = 0, \\ P(T, T, r) = 1. \end{cases} \tag{2.41}$$

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

$$\begin{cases} A'(t) - bA(t) + \frac{1}{2}\sigma^2 A(t)^2 + \beta(e^{aA(t)} - 1) - 1 = 0, \\ B'(t) + bcA(t) + \alpha(e^{aA(t)} - 1) = 0, \\ A(T) = B(T) = 0. \end{cases} \tag{2.42}$$

(ii) By using the same arguments as in the proof of Theorem 5, we have the following asymptotic result,

$$\lim_{T\to\infty} \frac{1}{T} \log P(t, T, r) = bcx_* + \alpha(e^{ax_*} - 1), \tag{2.43}$$

where x_* is the unique negative solution to the following equation,

$$-bx+\frac{1}{2}\sigma^2x^2+\beta(e^{ax}-1)-1=0. \tag{2.44}$$

$\square$

Acknowledgements

The author is supported by NSF grant DMS-0904701, DARPA grant and MacCracken Fellowship at New York University. The author is very grateful to an anonymous referee for the helpful comments and suggestions.

References

[1] Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J. F. Scaling limits for Hawkes processes and application to financial statistics. Preprint. arXiv:1202.0842.
[2] Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J. F. Modeling microstructure noise with mutually exciting point processes. To appear in *Quantitative Finance*. arXiv:1101.3422.
[3] Bordenave, C. and Torrisi, G. L. (2007). Large deviations of Poisson cluster processes. *Stochastic Models*. **23**, 593-625.
[4] Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Probab.*. **24**, 1563-1588.
[5] Cox, J. C., Ingersoll, J. E. 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: Mathematics and Economics*. **35**, 113-136.
[7] Dembo, A. and Zeitouni, O. *Large Deviations Techniques and Applications*. 2nd Edition, Springer, New York, 1998
[8] Duffie, D. (2005). Credit risk modeling with affine processes. *Journal of Banking & Finance*. **29**, 2751-2802.
[9] Errais, E., Giesecke, K. and Goldberg, L. (2010). Affine point processes and portfolio credit risk. *SIAM J. Financial Math*. **1**, 642-665.
[10] Feller, W. (1951). Two singular diffusion problems. *Annals of Mathematics*. **54**, 173-182.
[11] Hairer, M., *Convergence of Markov Processes*, Lecture Notes, University of Warwick, available at `http://www.hairer.org/notes/Convergence.pdf`, August 2010
[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 bound and currency options. *Review of Financial Studies*. **6**, 327-343.
[15] Kou, S., and Peng, X. *Default Clustering and Valuation of Collateralized Debt Obligations*. Working Paper, Columbia University, January 2009
[16] Stabile, G. and Torrisi, G. L. (2010). Risk processes with non-stationary Hawkes arrivals. *Methodol. Comput. Appl. Prob.* **12**, 415-429.
[17] Varadhan, S. R. S. *Large Deviations and Applications*, SIAM, Philadelphia, 1984.
[18] Zhu, L. Large deviations for Markovian nonlinear Hawkes processes. Preprint. arXiv:1108.2432.
[19] Zhu, L. Process-level large deviations for nonlinear Hawkes point processes. To appear in *Annales de l'Institut Henri Poincaré*. arXiv:1108.2431.
[20] Zhu, L. Central limit theorem for nonlinear Hawkes processes. To appear in *Journal of Applied Probability*. arXiv:1204.1067.
[21] Zhu, L. (2013). Moderate deviations for Hawkes processes. *Statistics & Probability Letters* **83**, 885-890.
[22] Zhu, L. (2013). Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. *Insurance: Mathematics and Economics*. **53**, 544-550.

Courant Institute of Mathematical Sciences
New York University
251 Mercer Street
New York, NY-10012
United States of America
E-mail address: ling@cims.nyu.edu