LIMIT THEOREMS FOR MARKED HAWKES PROCESSES WITH APPLICATION TO A RISK MODEL

DMYTRO KARABASH AND LINGJIONG ZHU

Abstract. This paper focuses on limit theorems for linear Hawkes processes with random marks. We prove a large deviation principle, which answers the question raised by Bordenave and Torrisi. A central limit theorem is also obtained. We conclude with an example of application in finance.

1. Introduction and Main Results

1.1. **Introduction.** Let N be a simple point process on $\mathbb{R}$ and let $\mathcal{F}_t^{-\infty} := \sigma(N(C), C \in \mathcal{B}(\mathbb{R}), C \subset (-\infty, t])$ be an increasing family of σ-algebras. Any nonnegative $\mathcal{F}_t^{-\infty}$-progressively measurable process λ_t with

$$\mathbb{E}\left[N(a,b]|\mathcal{F}_a^{-\infty}\right] = \mathbb{E}\left[\int_a^b \lambda_s ds \big| \mathcal{F}_a^{-\infty}\right] \tag{1.1}$$

a.s. for all intervals $(a, b]$ is called an $\mathcal{F}_t^{-\infty}$-intensity of N. We use the notation $N_t := N(0, t]$ to denote the number of points in the interval $(0, t]$.

A linear Hawkes process with random marks is a simple point process with $\mathcal{F}_t^{\infty}$-predictable intensity

$$\lambda_t := \nu + Z_t, \quad Z_t := \sum_{\tau_i < t} h(t - \tau_i, a_i), \tag{1.2}$$

where $\nu > 0$, the $(\tau_i)_{i\geq1}$ are arrival times of the points, and the $(a_i)_{i\geq1}$ are i.i.d. random marks, a_i being independent of previous arrival times τ_j, $j \leq i$. We assume that $N(-\infty, 0] = 0$. If one considers the stationary version of the process, one should start from time $-\infty$ in (1.2).

We further assume that a_i has a common distribution $q(da)$ on a measurable space $\mathbb{X}$. Here, $h(\cdot, \cdot) : \mathbb{R}^+ \times \mathbb{X} \to \mathbb{R}^+$ is integrable, i.e. $\int_0^\infty \int_{\mathbb{X}} h(t, a)q(da)dt < \infty$. Let $H(a) := \int_0^\infty h(t, a)dt$ for any $a \in \mathbb{X}$. We also assume that

$$\int_{\mathbb{X}} H(a)q(da) < 1. \tag{1.3}$$

Let $\mathbb{P}^q$ denote the probability measure for the a_i's with the common law $q(da)$. Under assumption (1.3), it is well known that there exists a unique stationary version of the linear marked Hawkes process satisfying the dynamics (1.2) and that

Date: 16 August 2014. *Revised:* 25 February 2015.
2000 *Mathematics Subject Classification.* 60G55, 60F10, 60F05.
Key words and phrases. marked point processes, Hawkes processes, self-exciting processes, large deviations, central limit theorem.

by ergodic theorem, a law of large numbers holds,

$$\lim_{t\to\infty}\frac{N_t}{t}=\frac{\nu}{1-\mathbb{E}^q[H(a)]}. \tag{1.4}$$

This paper is organized as the following. In Section 1.2, we will review some results about the limit theorems for unmarked Hawkes processes. In Section 1.3, we will introduce the main results of this paper, i.e. the central limit theorem and the large deviation principle for linear marked Hawkes processes. The proof of the central limit theorem will be given in Section 2 and the proof of the large deviation principle will be given in Section 3. Finally, we will discuss an application of our results to a risk model in finance in Section 4.

1.2. **Limit Theorems for Unmarked Hawkes Processes.** Most of the literature about Hawkes processes considered the unmarked case, i.e. with intensity

$$\lambda_t:=\lambda\left(\sum_{\tau<t}h(t-\tau)\right), \tag{1.5}$$

where $h(\cdot):\mathbb{R}^+\to\mathbb{R}^+$ is integrable and $\|h\|_{L^1}<1$ and $\lambda(\cdot):\mathbb{R}^+\to\mathbb{R}^+$ is locally integrable.

When $\lambda(\cdot)$ is linear, the Hawkes process is said to be linear and it is nonlinear otherwise. The stability results for both linear and nonlinear Hawkes processes are known. For the linear case, we refer to Hawkes and Oakes [9]. For the nonlinear case, Brémaud and Massoulié [3] proved the stability results for α-Lipschitz $\lambda(\cdot)$ such that $\alpha\|h\|_{L^1}<1$. Karabash [11] obtained stability results for certain non-Lipschitz $\lambda(\cdot)$ and discontinuous $\lambda(\cdot)$.

The limit theorems for both linear and nonlinear Hawkes processes are well studied in the literature.

For the linear Hawkes process, assume $\lambda(z)=\nu+z$, for some $\nu>0$ and $\|h\|_{L^1}<1$, it has an immigration-birth representation, see for example Hawkes and Oakes [9]. For linear Hawkes process, limit theorems are very well understood. There is the law of large numbers (see for instance Daley and Vere-Jones [5]), i.e.

$$\frac{N_t}{t}\to\frac{\nu}{1-\|h\|_{L^1}},\quad \text{as } t\to\infty \text{ a.s.} \tag{1.6}$$

Moreover, Bordenave and Torrisi [2] proved a large deviation principle for $(\frac{N_t}{t}\in\cdot)$ with the rate function

$$I(x)=\begin{cases}x\log\left(\frac{x}{\nu+x\|h\|_{L^1}}\right)-x+x\|h\|_{L^1}+\nu & \text{if } x\in[0,\infty)\\ +\infty & \text{otherwise}\end{cases}. \tag{1.7}$$

By applying the techniques of large deviations, the asymptotics of the ruin probabilities for risk processes in insurance were studied in Stabile and Torrisi [13] for the light-tailed claims and in Zhu [21] for the heavy-tailed claims.

The limit theorems have also been studied for an extension of linear Hawkes processes and Cox-Ingersoll-Ross processes in Zhu [20], which has applications in short interest rate models in finance.

Recently, Bacry et al. [1] proved a functional central limit theorem for the linear multivariate Hawkes process under certain assumptions which includes the linear

Hawkes process as a special case and they proved that

$$\frac{N_{\cdot t} - \cdot \mu t}{\sqrt{t}} \to \sigma B(\cdot), \quad \text{as } t \to \infty, \tag{1.8}$$

weakly on $D[0,1]$ equipped with Skorokhod topology, where

$$\mu = \frac{\nu}{1 - \|h\|_{L^1}} \quad \text{and} \quad \sigma^2 = \frac{\nu}{(1 - \|h\|_{L^1})^3}. \tag{1.9}$$

Moderate deviation principle for linear Hawkes processes is obtained in Zhu [19], which fills in the gap between central limit theorem and large deviation principle.

For nonlinear Hawkes processes, a central limit theorem is obtained in Zhu [18]. In Bordenave and Torrisi [2], they raised two questions about large deviations for Hawkes processes. One question is about large deviations for nonlinear Hawkes process and the other is about large deviations for linear marked Hawkes processes. Recently, Zhu [16] considered a special case for nonlinear Hawkes processes when $h(\cdot)$ is exponential or sums of exponentials and proved the large deviations. In another paper, Zhu [17] proved a process-level, i.e. level-3 large deviation principle for nonlinear Hawkes processes for general $h(\cdot)$ and hence by contraction principle, the level-1 large deviation principle for $(N_t/t \in \cdot)$. In this paper, we will prove the large deviations for linear marked Hawkes processes and thus both questions raised in Bordenave and Torrisi [2] have been answered. The large deviation theory studies the small probability of rare events. Unlike the unmarked case, the rare events in marked Hawkes processes can also be due to the presence of random marks. It is mixture of atypical behavior of unmarked Hawkes processes and atypical behavior of random marks. The role of random marks in the large deviations of Hawkes processes is what we need to understand.

1.3. **Main Results.** 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 (LDP) with rate function $I : X \to \mathbb{R}$ if I is non-negative, lower semicontinuous and for any measurable set A, we have

$$-\inf_{x\in A^o} I(x) \leq \liminf_{n\to\infty} \frac{1}{n} \log P_n(A) \leq \limsup_{n\to\infty} \frac{1}{n} \log P_n(A) \leq -\inf_{x\in\overline{A}} I(x). \tag{1.10}$$

Here, A^o is the interior of A and $\overline{A}$ is its closure. See Dembo and Zeitouni [6] or Varadhan [15] for general background regarding large deviations and their applications. Also Varadhan [14] has an excellent survey article on this subject.

For a linear marked Hawkes process satisfying the dynamics (1.2), we prove the following large deviation principle in this article.

Theorem 1 (Large Deviation Principle)**.** *Assume the conditions* (1.3) *and* $H(a) > 0$ *with positive probability. Also assume that there exists some* $\theta > 0$*, so that* $\int_{\mathbb{X}} e^{\theta H(a)} q(da) < \infty$*. Then,* $\mathbb{P}(N_t/t \in \cdot)$ *satisfies a large deviation principle with rate function,*

$$\begin{aligned}\Lambda(x) &:= \begin{cases} \inf_{\hat{q}} \left\{ x\mathbb{E}^{\hat{q}}[H(a)] + \nu - x + x\log\left(\frac{x}{x\mathbb{E}^{\hat{q}}[H(a)]+\nu}\right) + x\mathbb{E}^{\hat{q}}\left[\log\frac{d\hat{q}}{dq}\right]\right\} & x \geq 0 \\ +\infty & x < 0 \end{cases} \\ &= \begin{cases} \theta_* x - \nu(x_* - 1) & x \geq 0 \\ +\infty & x < 0 \end{cases},\end{aligned}$$

where the infimum of $\hat{q}$ is taken over $\mathcal{M}(\mathbb{X})$, the space of probability measures on $\mathbb{X}$ such that $\hat{q}$ is absolutely continuous w.r.t. q. Here, θ_* and x_* satisfy the following equations

$$\begin{cases} x_* = \mathbb{E}^q\left[e^{\theta_*+(x_*-1)H(a)}\right] \\ \frac{x}{\nu} = x_* + \frac{x}{\nu}\mathbb{E}^q\left[H(a)e^{\theta_*+(x_*-1)H(a)}\right] \end{cases}. \tag{1.11}$$

Theorem 2 (Central Limit Theorem)**.** *Assume* $\lim_{t\to\infty} t^{1/2}\int_t^\infty \mathbb{E}^q[h(s,a)]ds = 0$ *and that* (1.3) *holds. Then,*

$$\frac{N_t - \frac{\nu t}{1-\mathbb{E}^q[H(a)]}}{\sqrt{t}} \to N\left(0, \frac{\nu(1+Var^q[H(a)])}{(1-\mathbb{E}^q[H(a)])^3}\right), \tag{1.12}$$

in distribution as $t\to\infty$.

Remark 3. *Comparing Theorem 1 and Theorem 2 with* (1.7) *and* (1.8)*, it is easy to see that our results are consistent with the limit theorems for unmarked Hawkes process.*

2. Proof of Central Limit Theorem

Proof of Theorem 2. First, let us observe that

$$\begin{aligned}\int_0^t \lambda_s ds &= \nu t + \sum_{\tau_i<t}\int_{\tau_i}^t h(s-\tau_i,a_i)ds \\ &= \nu t + \sum_{\tau_i<t} H(a_i) - \mathcal{E}_t,\end{aligned} \tag{2.1}$$

where the error term $\mathcal{E}_t$ is given by

$$\mathcal{E}_t := \sum_{\tau_i<t}\int_t^\infty h(s-\tau_i,a_i)ds. \tag{2.2}$$

Therefore,

$$\begin{aligned}\frac{N_t-\int_0^t\lambda_s ds}{\sqrt{t}} &= \frac{N_t-\nu t-\sum_{\tau_i<t}H(a_i)}{\sqrt{t}} + \frac{\mathcal{E}_t}{\sqrt{t}} \\ &= (1-\mathbb{E}^q[H(a)])\frac{N_t-\mu t}{\sqrt{t}} + \frac{\mathbb{E}^q[H(a)]N_t-\sum_{\tau_i<t}H(a_i)}{\sqrt{t}} + \frac{\mathcal{E}_t}{\sqrt{t}},\end{aligned} \tag{2.3}$$

where $\mu := \frac{\nu}{1-\mathbb{E}^q[H(a)]}$. Rearranging the terms in (2.3), we get

$$\frac{N_t-\mu t}{\sqrt{t}} = \frac{1}{1-\mathbb{E}^q[H(a)]}\left[\frac{N_t-\int_0^t\lambda_s ds}{\sqrt{t}} + \frac{\sum_{\tau_i<t}(H(a_i)-\mathbb{E}^q[H(a)])}{\sqrt{t}} - \frac{\mathcal{E}_t}{\sqrt{t}}\right]. \tag{2.4}$$

It is easy to check that $\frac{\mathcal{E}_t}{\sqrt{t}} \to 0$ in probability as $t\to\infty$. To see this, first notice that $\mathbb{E}[\lambda_t] \leq \frac{\nu}{1-\mathbb{E}^q[H(a)]}$ uniformly in t. Let $g(t,a) := \int_t^\infty h(s,a)ds$. We have

$\mathcal{E}_t = \sum_{\tau_i<t} g(t-\tau_i, a_i)$ and thus

$$\begin{aligned}\mathbb{E}[\mathcal{E}_t] &= \int_0^t \int_{\mathbb{X}} g(t-s,a)q(da)\mathbb{E}[\lambda_s]ds \\ &\leq \frac{\nu}{1-\mathbb{E}^q[H(a)]}\int_0^t\int_{\mathbb{X}} g(t-s,a)q(da)ds \\ &= \frac{\nu}{1-\mathbb{E}^q[H(a)]}\int_0^t \mathbb{E}^q[g(s,a)]ds.\end{aligned} \tag{2.5}$$

Hence, by L'Hôpital's rule,

$$\begin{aligned}\lim_{t\to\infty}\frac{1}{t^{1/2}}\int_0^t \mathbb{E}^q[g(s,a)]ds &= \lim_{t\to\infty}\frac{\mathbb{E}^q[g(t,a)]}{\frac{1}{2}t^{-1/2}} \\ &= \lim_{t\to\infty} 2t^{1/2}\int_t^\infty \mathbb{E}^q[h(s,a)]ds = 0.\end{aligned} \tag{2.6}$$

Hence, $\frac{\mathcal{E}_t}{\sqrt{t}} \to 0$ in probability as $t\to\infty$.

Furthermore, $M_1(t) := N_t - \int_0^t \lambda_s ds$ and $M_2(t) := \sum_{\tau_i<t}(H(a_i) - \mathbb{E}^q[H(a)])$ are both martingales.

Moreover, since $\int_0^t \lambda_s ds$ is of finite variation, the quadratic variation of $M_1(t) + M_2(t)$ is the same as the quadratic variation of $N_t + M_2(t)$. And notice that $N_t + M_2(t) = \sum_{\tau_i<t}(1 + H(a_i) - \mathbb{E}^q[H(a)])$ which has quadratic variation

$$\sum_{\tau_i<t}(1 + H(a_i) - \mathbb{E}^q[H(a)])^2. \tag{2.7}$$

By the standard law of large numbers, we have

$$\begin{aligned}\frac{1}{t}\sum_{\tau_i<t}(1 + H(a_i) - \mathbb{E}^q[H(a)])^2 &= \frac{N_t}{t}\cdot\frac{1}{N_t}\sum_{\tau_i<t}(1 + H(a_i) - \mathbb{E}^q[H(a)])^2 \\ &\to \frac{\nu}{1-\mathbb{E}^q[H(a)]}\cdot\mathbb{E}^q\left[(1+H(a)-\mathbb{E}^q[H(a)])^2\right] \\ &= \frac{\nu(1+\mathrm{Var}^q[H(a)])}{1-\mathbb{E}^q[H(a)]},\end{aligned} \tag{2.8}$$

a.s. as $t\to\infty$. By a standard martingale central limit theorem (see. e.g. Theorem VIII-3.11 of Jacod and Shiryaev [10]), we conclude that

$$\frac{N_t - \frac{\nu t}{1-\mathbb{E}^q[H(a)]}}{\sqrt{t}} \to N\left(0, \frac{\nu(1+\mathrm{Var}^q[H(a)])}{(1-\mathbb{E}^q[H(a)])^3}\right), \tag{2.9}$$

in distribution as $t\to\infty$. □

3. Proof of Large Deviation Principle

3.1. **Limit of a Logarithmic Moment Generating Function.** In this subsection, we prove the existence of the limit of the logarithmic moment generating function $\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}[e^{\theta N_t}]$ and give a variational formula and a more explicit formula for this limit.

Theorem 4. *Assume* (1.3) *and that there exists some* $\theta > 0$*, so that* $\int_{\mathbb{X}} e^{\theta H(a)} q(da) < \infty$ *and* $q(da)$ *has a continuous density. The limit* $\Gamma(\theta)$ *of the logarithmic moment generating function is*

$$\Gamma(\theta) = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}[e^{\theta N_t}] = \begin{cases} \nu(f(\theta)-1) & \text{if } \theta \in (-\infty, \theta_c] \\ +\infty & \text{otherwise} \end{cases}, \tag{3.1}$$

where $f(\theta)$ *is the minimal solution to* $x = \int_{\mathbb{X}} e^{\theta + H(a)(x-1)} q(da)$ *and*

$$\theta_c = -\log \int_{\mathbb{X}} H(a) e^{H(a)(x_c-1)} q(da) > 0, \tag{3.2}$$

where $x_c > 1$ *satisfies the equation* $x \int_{\mathbb{X}} H(a) e^{H(a)(x-1)} q(da) = \int_{\mathbb{X}} e^{H(a)(x-1)} q(da)$*.*

Now, let us prove Theorem 4.

Proof of Theorem 4. It is well known that a linear Hawkes process has an immigration-birth representation, see e.g. Hawkes and Oakes [9]. The immigrants (roots) arrive via a standard Poisson process with constant intensity $\nu > 0$. Each immigrant generates children according to a Galton-Watson tree. Consider a random, rooted tree (with root, i.e. immigrant, at time 0) associated to the Hawkes process via the Galton-Watson interpretation. Note the root is unmarked at the start of the process so the marking goes into the expectation calculation later. Let K be the number of children of the root node, and let $S_t^{(1)}, S_t^{(2)}, \ldots, S_t^{(K)}$ be the number of descendants of root's k-th child that were born before time t (including k-th child if an only if it was born before time t). Let S_t be the total number of children in tree before time t including root node. Then

$$\begin{aligned}
F_S(t) &:= \mathbb{E}[\exp(\theta S_t)] && (3.3)\\
&= \sum_{k=0}^{\infty} \mathbb{E}[\exp(\theta S_t)|K=k]\mathbb{P}(K=k) && (3.4)\\
&= \exp(\theta) \sum_{k=0}^{\infty} \mathbb{P}(K=k) \prod_{i=1}^{k} \mathbb{E}\left[\exp\left(\theta S_t^{(i)}\right)\right] && (3.5)\\
&= \exp(\theta) \sum_{k=0}^{\infty} \mathbb{E}\left[\exp\left(\theta S_t^{(1)}\right)\right]^k \mathbb{P}(K=k) && (3.6)\\
&= \exp(\theta) \sum_{k=0}^{\infty} \int_{\mathbb{X}} \left[\left(\int_0^t \frac{h(s,a)}{H(a)} F_S(t-s) ds\right)^k e^{-H(a)} \frac{H(a)^k}{k!}\right] q(da) && (3.7)\\
&= \int_{\mathbb{X}} \exp\left(\theta + \int_0^t h(s,a)(F_S(t-s)-1) ds\right) q(da). && (3.8)
\end{aligned}$$

Now observe that by definition $F_S(t) = \mathbb{E}[\exp(\theta S_t)]$. When $\theta \leq 0$, $F_S(t)$ is decreasing in t and also $0 \leq F_S(t) \leq 1$. Thus, $F_S(t)$ converges to a finite limit x_* as $t \to \infty$, which satisfies

$$x = \int_{\mathbb{X}} \exp\left[\theta + H(a)(x-1)\right] q(da). \tag{3.9}$$

Similarly, for $\theta > 0$, $F_S(t)$ is increasing in t and either $F_S(t) \to \infty$ as $t \to \infty$ or it converges to some finite limit x_* that satisfies (3.9).

Next, we need to determine for what values of θ the solution of (3.9) exists. Let

$$G(x) = e^{\theta} \int_{\mathbb{X}} e^{H(a)(x-1)} q(da) - x. \tag{3.10}$$

If $\theta = 0$, then $G(x) = \int_{\mathbb{X}} e^{H(a)(x-1)} q(da) - x$ satisfies $G(1) = 0$, $G(\infty) = \infty$ and $G'(1) = \mathbb{E}^q[H(a)] - 1 < 0$ which implies $\min_{x>1} G(x) < 0$. Hence, there exists some critical $\theta_c > 0$ such that $\min_{x>1} G(x) = 0$. The critical values x_c and θ_c satisfy $G(x_c) = G'(x_c) = 0$, which implies

$$\theta_c = -\log \int_{\mathbb{X}} H(a) e^{H(a)(x_c - 1)} q(da), \tag{3.11}$$

where $x_c > 1$ satisfies the equation $x \int_{\mathbb{X}} H(a) e^{H(a)(x-1)} q(da) = \int_{\mathbb{X}} e^{H(a)(x-1)} q(da)$. Hence (3.9) has finite solutions if and only if $\theta \leq \theta_c$. Moreover, it is easy to see that $G(x)$ is strictly convex in x and hence there can be at most two (positive) solutions to (3.9). When $\theta < 0$,

$$G(e^{\theta}) = e^{\theta} \left[\int_{\mathbb{X}} e^{H(a)(e^{\theta}-1)} q(da) - 1 \right] = e^{\theta} \left[\mathbb{E}^q \left[e^{H(a)(e^{\theta}-1)} \right] - 1 \right] < 0, \tag{3.12}$$

and $F_S(0) = e^{\theta}$ and $F_S(t)$ is decreasing in t and therefore it converges to the smaller solution of (3.9). Similary, when $\theta > 0$,

$$G(e^{\theta}) = e^{\theta} \left[\int_{\mathbb{X}} e^{H(a)(e^{\theta}-1)} q(da) - 1 \right] = e^{\theta} \left[\mathbb{E}^q \left[e^{H(a)(e^{\theta}-1)} \right] - 1 \right] > 0, \tag{3.13}$$

and $F_S(0) = e^{\theta}$ and $F_S(t)$ is increasing in t and for $\theta \leq \theta_c$ we know that $F_S(t)$ converges to a finite limit and therefore it must converge to the smaller solution of (3.9) as $t \to \infty$.

Finally, since random roots arrive according to a Poisson process with constant intensity $\nu > 0$, we have

$$F_N(t) := \mathbb{E}[\exp(\theta N_t)] = \exp \left[\nu \int_0^t (F_S(t-s) - 1) ds \right]. \tag{3.14}$$

But since $F_S(s) \uparrow x_*$ as $s \to \infty$ we obtain the main result

$$\frac{1}{t} \log F_N(t) = \nu \frac{1}{t} \left[\int_0^t (F_S(s) - 1)\, ds \right] \underset{t\to\infty}{\longrightarrow} \nu(x_* - 1), \tag{3.15}$$

if $\theta \leq \theta_c$, which proves the desired formula. □

3.2. **Large Deviation Principle.** In this section, we prove the main result, i.e. Theorem 1 by using the Gärtner-Ellis theorem for the upper bound and tilting method for the lower bound. Before we proceed, let us first state a lemma that will be used in the proof the lower bound in Theorem 1. This result can be found in Brémaud et al. [4].

Lemma 5. *[Brémaud et al. [4]] Consider a linear marked Hawkes process with intensity*

$$\lambda_t := \alpha + \beta Z_t := \alpha + \beta \sum_{\tau_i < t} h(t - \tau_i, a_i), \tag{3.16}$$

and $\beta \mathbb{E}^q[H(a)] < 1$, where the a_i are i.i.d. random marks with the common law $q(da)$ independent of the previous arrival times, then there exists a unique invariant *measure π for Z_t such that*

$$\int_0^\infty \lambda(z)\pi(dz) = \frac{\alpha}{1-\beta\mathbb{E}^q[H(a)]}. \tag{3.17}$$

Now we are ready to prove Theorem 1.

Proof of Theorem 1. For the upper bound, since we have Theorem 4, we can simply apply Gärtner-Ellis theorem (see e.g. Dembo and Zeitouni [6]). To prove the lower bound, it suffices to show that for any $x > 0$, $\epsilon > 0$, we have

$$\liminf_{t\to\infty} \frac{1}{t}\log\mathbb{P}\left(\frac{N_t}{t}\in B_\epsilon(x)\right) \geq -\sup_{\theta\in\mathbb{R}}\{\theta x - \Gamma(\theta)\}, \tag{3.18}$$

where $B_\epsilon(x)$ denotes the open ball centered at x with radius ϵ.

The intensity at time t is $\lambda_t := \lambda(Z_t)$ where $\lambda(z) = \nu + z$ and $Z_t = \sum_{\tau_i<t} h(t-\tau_i, a_i)$. We tilt λ to $\hat{\lambda}$ and q to $\hat{q}$ such that by Girsanov formula the tilted probability measure $\hat{\mathbb{P}}$ is given by

$$\left.\frac{d\hat{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left\{\int_0^t (\lambda(Z_s) - \hat{\lambda}(Z_s))ds + \int_0^t \left[\log\left(\frac{\hat{\lambda}(Z_s)}{\lambda(Z_s)}\right) + \log\left(\frac{d\hat{q}}{dq}\right)\right] dN_s\right\}. \tag{3.19}$$

Let $\mathcal{Q}_e$ be the set of $(\hat{\lambda}, \hat{q}, \hat{\pi})$ such that the marked Hawkes process with intensity $\hat{\lambda}(Z_t)$ and random marks distributed as $\hat{q}$ is ergodic with $\hat{\pi}$ as the invariant measure of Z_t.

By Jensen's inequality,

$$\begin{aligned}
&\frac{1}{t}\log\mathbb{P}\left(\frac{N_t}{t}\in B_\epsilon(x)\right) \\
&= \frac{1}{t}\log\int_{\frac{N_t}{t}\in B_\epsilon(x)} \frac{d\mathbb{P}}{d\hat{\mathbb{P}}} d\hat{\mathbb{P}} \\
&= \frac{1}{t}\log\hat{\mathbb{P}}\left(\frac{N_t}{t}\in B_\epsilon(x)\right) - \frac{1}{t}\log\left[\frac{1}{\hat{\mathbb{P}}\left(\frac{N_t}{t}\in B_\epsilon(x)\right)}\int_{\frac{N_t}{t}\in B_\epsilon(x)} \frac{d\hat{\mathbb{P}}}{d\mathbb{P}} d\hat{\mathbb{P}}\right] \\
&\geq \frac{1}{t}\log\hat{\mathbb{P}}\left(\frac{N_t}{t}\in B_\epsilon(x)\right) - \frac{1}{\hat{\mathbb{P}}\left(\frac{N_t}{t}\in B_\epsilon(x)\right)}\cdot\frac{1}{t}\cdot\hat{\mathbb{E}}\left[1_{\frac{N_t}{t}\in B_\epsilon(x)}\log\frac{d\hat{\mathbb{P}}}{d\mathbb{P}}\right].
\end{aligned} \tag{3.20}$$

By the ergodic theorem (see e.g. Chapter 16.4. of Koralov and Sinai [12]),

$$\liminf_{t\to\infty}\frac{1}{t}\log\mathbb{P}\left(\frac{N_t}{t}\in B_\epsilon(x)\right) \geq -\inf_{\substack{0<K<\mathbb{E}^{\hat{q}}[H(a)]^{-1}\\ (\hat{\lambda},\hat{q},\hat{\pi})\in\mathcal{Q}_e^x, \hat{\lambda}=K\lambda}} \mathcal{H}(\hat{\lambda},\hat{q},\hat{\pi}). \tag{3.21}$$

where $\mathcal{Q}_e^x$ is defined by

$$\mathcal{Q}_e^x = \left\{(\hat{\lambda},\hat{q},\hat{\pi})\in\mathcal{Q}_e : \int\hat{\lambda}(z)\hat{\pi}(dz) = x\right\}. \tag{3.22}$$

and the relative entropy $\mathcal{H}$ is

$$\mathcal{H}(\hat{\lambda},\hat{q},\hat{\pi}) = \int(\lambda-\hat{\lambda})\hat{\pi} + \int\log(\hat{\lambda}/\lambda)\hat{\lambda}\hat{\pi} + \iint\log(d\hat{q}/dq)\hat{q}\hat{\lambda}\hat{\pi}. \tag{3.23}$$

By Lemma 5,

$$\begin{aligned}
&\inf_{0<K<\mathbb{E}^{\hat{q}}[H(a)]^{-1},x=\frac{\nu K}{1-K\mathbb{E}^{\hat{q}}[H(a)]},(\hat{\lambda},\hat{q},\hat{\pi})\in\mathcal{Q}_e,\hat{\lambda}=K\lambda}\mathcal{H}(\hat{\lambda},\hat{q},\hat{\pi}) \qquad (3.24)\\
&=\inf_{K=\frac{x}{x\mathbb{E}^{\hat{q}}[H(a)]+\nu},(\hat{\lambda},\hat{q},\hat{\pi})\in\mathcal{Q}_e,\hat{\lambda}=K\lambda}\left\{\frac{1}{K}-1+\log K+\mathbb{E}^{\hat{q}}\left[\log\frac{d\hat{q}}{dq}\right]\right\}\int\hat{\lambda}\hat{\pi}\\
&=\inf_{\hat{q}}\left\{\mathbb{E}^{\hat{q}}[H(a)]+\frac{\nu}{x}-1+\log\left(\frac{x}{x\mathbb{E}^{\hat{q}}[H(a)]+\nu}\right)+\mathbb{E}^{\hat{q}}\left[\log\frac{d\hat{q}}{dq}\right]\right\}x\\
&=\inf_{\hat{q}}\left\{x\mathbb{E}^{\hat{q}}[H(a)]+\nu-x+x\log\left(\frac{x}{x\mathbb{E}^{\hat{q}}[H(a)]+\nu}\right)+x\mathbb{E}^{\hat{q}}\left[\log\frac{d\hat{q}}{dq}\right]\right\}.
\end{aligned}$$

Next, let us find a more explict form for the Legendre-Fenchel transform of $\Gamma(\theta)$.

$$\sup_{\theta\in\mathbb{R}}\{\theta x-\Gamma(\theta)\}=\sup_{\theta\in\mathbb{R}}\{\theta x-\nu(f(\theta)-1)\}, \tag{3.25}$$

where $f(\theta)=\mathbb{E}^q[e^{\theta+(f(\theta)-1)H(a)}]$. Here,

$$f'(\theta)=\mathbb{E}^q\left[(1+f'(\theta)H(a))e^{\theta+(f(\theta)-1)H(a)}\right]. \tag{3.26}$$

So the optimal θ_* for (3.25) would satisfy $f'(\theta_*)=\frac{x}{\nu}$ and θ_* and $x_*=f(\theta_*)$ satisfy the following equations

$$\begin{cases}x_*=\mathbb{E}^q\left[e^{\theta_*+(x_*-1)H(a)}\right]\\ \frac{x}{\nu}=x_*+\frac{x}{\nu}\mathbb{E}^q\left[H(a)e^{\theta_*+(x_*-1)H(a)}\right]\end{cases}, \tag{3.27}$$

and $\sup_{\theta\in\mathbb{R}}\{\theta x-\Gamma(\theta)\}=\theta_*x-\nu(x_*-1)$.

On the other hand, letting $dq_*=\frac{e^{(x_*-1)H(a)}}{\mathbb{E}^q[e^{(x_*-1)H(a)}]}dq$, we have

$$\mathbb{E}^{q_*}[H(a)]=\frac{\mathbb{E}^q\left[e^{\theta_*+(x_*-1)H(a)}\right]}{\mathbb{E}^q\left[e^{(x_*-1)H(a)}\right]}=\frac{1}{x_*}-\frac{\nu}{x}, \tag{3.28}$$

and $\mathbb{E}^{q_*}[\log\frac{dq_*}{dq}]=(x_*-1)\mathbb{E}^{q_*}[H(a)]-\log\mathbb{E}^q[e^{(x_*-1)H(a)}]$, which imply

$$\begin{aligned}
&\liminf_{t\to\infty}\frac{1}{t}\log\mathbb{P}\left(\frac{N_t}{t}\in B_\epsilon(x)\right) \qquad (3.29)\\
&\geq-\inf_{\hat{q}}\left\{x\mathbb{E}^{\hat{q}}[H(a)]+\nu-x+x\log\left(\frac{x}{x\mathbb{E}^{\hat{q}}[H(a)]+\nu}\right)+x\mathbb{E}^{\hat{q}}\left[\log\frac{d\hat{q}}{dq}\right]\right\}\\
&\geq-\left\{x\mathbb{E}^{q_*}[H(a)]+\nu-x+x\log\left(\frac{x}{x\mathbb{E}^{q_*}[H(a)]+\nu}\right)+x\mathbb{E}^{q_*}\left[\log\frac{dq_*}{dq}\right]\right\}\\
&=\theta_*x-\nu(x_*-1)=\sup_{\theta\in\mathbb{R}}\{\theta x-\Gamma(\theta)\}.
\end{aligned}$$

□

Remark 6. *An alternative proof of the lower bound of the large deviation principle in Theorem 1 is by checking the essential smoothness condition in Gärtner-Ellis theorem and applying it to prove the lower bound directly. Nevertheless, our tilting approach has the advantage of pin-pointing to the most likely path when a rare event occurs.*

4. Risk Model with Marked Hawkes Claims Arrivals

We consider the following risk model for the surplus process R_t of an insurance portfolio,

$$R_t = u + \rho t - \sum_{i=1}^{N_t} C_i, \tag{4.1}$$

where $u > 0$ is the initial reserve, $\rho > 0$ is the constant premium and the C_i's are i.i.d. positive random variables with the common distribution $\mu(dC)$. C_i represents the claim size at the ith arrival time, these being independent of N_t, a marked Hawkes process.

For $u > 0$, let

$$\tau_u = \inf\{t > 0 : R_t \leq 0\}, \tag{4.2}$$

and denote the infinite and finite horizon ruin probabilities by

$$\psi(u) = \mathbb{P}(\tau_u < \infty), \quad \psi(u, uz) = \mathbb{P}(\tau_u \leq uz), \quad u, z > 0. \tag{4.3}$$

We first consider the case when the claim sizes have light-tails, i.e. there exists some $\theta > 0$ so that $\int e^{\theta C}\mu(dC) < \infty$.

By the law of large numbers,

$$\lim_{t\to\infty} \frac{1}{t}\sum_{i=1}^{N_t} C_i = \frac{\mathbb{E}^{\mu}[C]\nu}{1 - \mathbb{E}^{q}[H(a)]}. \tag{4.4}$$

Therefore, to exclude the trivial case, we need to assume that

$$\frac{\mathbb{E}^{\mu}[C]\nu}{1 - \mathbb{E}^{q}[H(a)]} < \rho < \frac{\nu(x_c - 1)}{\theta_c}, \tag{4.5}$$

where the critical values θ_c and $x_c = f(\theta_c)$ satisfy

$$\begin{cases} x_c = \int_{\mathbb{R}^+}\int_{\mathbb{X}} e^{\theta_c C + H(a)(x_c-1)} q(da)\mu(dC) \\ 1 = \int_{\mathbb{R}^+}\int_{\mathbb{X}} H(a) e^{H(a)(x_c-1)+\theta_c C} q(da)\mu(dC) \end{cases}. \tag{4.6}$$

Following the proofs of large deviation results in Section 3, we have

$$\Gamma_C(\theta) := \lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}\left[e^{\theta\sum_{i=1}^{N_t} C_i}\right] = \begin{cases} \nu(x-1) & \text{if } \theta \in (-\infty, \theta_c] \\ +\infty & \text{otherwise} \end{cases}, \tag{4.7}$$

where x is the minimal solution to the equation

$$x = \int_{\mathbb{R}^+}\int_{\mathbb{X}} e^{\theta C + (x-1)H(a)} q(da)\mu(dC). \tag{4.8}$$

Before we proceed, let us quote a result from Glynn and Whitt [7], which will be used in our proof of Theorem 9.

Theorem 7 (Glynn and Whitt [7]). *Let S_n be random variables. $\tau_u = \inf\{n : S_n > u\}$ and $\psi(u) = \mathbb{P}(\tau_u < \infty)$. Assume that there exist $\gamma, \epsilon > 0$ such that*

(i) $\kappa_n(\theta) = \log\mathbb{E}[e^{\theta S_n}]$ is well defined and finite for $\gamma - \epsilon < \theta < \gamma + \epsilon$.

(ii) $\limsup_{n\to\infty}\mathbb{E}[e^{\theta(S_n - S_{n-1})}] < \infty$ for $-\epsilon < \theta < \epsilon$.

(iii) $\kappa(\theta) = \lim_{n\to\infty}\frac{1}{n}\kappa_n(\theta)$ exists and is finite for $\gamma - \epsilon < \theta < \gamma + \epsilon$.

(iv) $\kappa(\gamma) = 0$ and κ is differentiable at γ with $0 < \kappa'(\gamma) < \infty$.

Then, $\lim_{u\to\infty}\frac{1}{u}\log\psi(u) = -\gamma$.

Remark 8. *We claim that $\Gamma_C(\theta) = \rho\theta$ has a unique positive solution $\theta^\dagger < \theta_c$. Let $G(\theta) = \Gamma_C(\theta) - \rho\theta$. Notice that $G(0) = 0$, $G(\infty) = \infty$, and that G is convex. We also have $G'(0) = \frac{\mathbb{E}^\mu[C]\nu}{1-\mathbb{E}^q[H(a)]} - \rho < 0$ and $\Gamma_C(\theta_c) - \rho\theta_c > 0$ since we assume that $\rho < \frac{\nu(f(\theta_c)-1)}{\theta_c}$. Therefore, there exists only one solution $\theta^\dagger \in (0, \theta_c)$ of $\Gamma_C(\theta^\dagger) = \rho\theta^\dagger$.*

Theorem 9 (Infinite Horizon)**.** *Assume all the assumptions in Theorem 1 and in addition (4.5), we have $\lim_{u\to\infty} \frac{1}{u} \log \psi(u) = -\theta^\dagger$, where $\theta^\dagger \in (0, \theta_c)$ is the unique positive solution of $\Gamma_C(\theta) = \rho\theta$.*

Proof. Take $S_t = \sum_{i=1}^{N_t} C_i - \rho t$ and $\kappa_t(\theta) = \log \mathbb{E}[e^{\theta S_t}]$. Then $\lim_{t\to\infty} \frac{1}{t}\kappa_t(\theta) = \Gamma_C(\theta) - \rho\theta$. Consider $\{S_{nh}\}_{n\in\mathbb{N}}$. We have $\lim_{n\to\infty} \frac{1}{n}\kappa_{nh}(\theta) = h\Gamma_C(\theta) - h\rho\theta$. Checking the conditions in Theorem 7 and applying it, we get

$$\lim_{u\to\infty} \frac{1}{u} \log \mathbb{P}\left(\sup_{n\in\mathbb{N}} S_{nh} > u\right) = -\theta^\dagger. \tag{4.9}$$

Finally, notice that

$$\sup_{t\in\mathbb{R}^+} S_t \geq \sup_{n\in\mathbb{N}} S_{nh} \geq \sup_{t\in\mathbb{R}^+} S_t - \rho h. \tag{4.10}$$

Hence, $\lim_{u\to\infty} \frac{1}{u} \log \psi(u) = -\theta^\dagger$. □

Theorem 10 (Finite Horizon)**.** *Under the same assumptions as in Theorem 9, we have*

$$\lim_{u\to\infty} \frac{1}{u} \log \psi(u, uz) = -w(z), \quad \textit{for any } z > 0. \tag{4.11}$$

Here

$$w(z) = \begin{cases} z\Lambda_C\left(\frac{1}{z} + \rho\right) & \textit{if } 0 < z < \frac{1}{\Gamma_C'(\theta^\dagger)-\rho} \\ \theta^\dagger & \textit{if } z \geq \frac{1}{\Gamma_C'(\theta^\dagger)-\rho} \end{cases}, \tag{4.12}$$

$\Lambda_C(x) = \sup_{\theta\in\mathbb{R}}\{\theta x - \Gamma_C(\theta)\}$ and $\theta^\dagger \in (0, \theta_c)$ is the unique positive solution of $\Gamma_C(\theta) = \rho\theta$, as before.

Proof. The proof is similar to that in Stabile and Torrisi [13] and we omit it here. □

Next, we are interested to study the case when the claim sizes have heavy tails, i.e. $\int_{\mathbb{R}^+} e^{\theta C}\mu(dC) = +\infty$ for any $\theta > 0$.

A distribution function B is subexponential, i.e. $B \in \mathcal{S}$ if

$$\lim_{x\to\infty} \frac{\mathbb{P}(C_1 + C_2 > x)}{\mathbb{P}(C_1 > x)} = 2, \tag{4.13}$$

where C_1, C_2 are i.i.d. random variables with distribution function B. Let us denote $B(x) := \mathbb{P}(C_1 \geq x)$ and let us assume that $\mathbb{E}[C_1] < \infty$ and define $B_0(x) := \frac{1}{\mathbb{E}[C]}\int_0^x \overline{B}(y)dy$, where $\overline{F}(x) = 1 - F(x)$ is the complement of any distribution function $F(x)$.

Goldie and Resnick [8] showed that if $B \in \mathcal{S}$ and satisfies some smoothness conditions, then B belongs to the maximum domain of attraction of either the Frechet distribution or the Gumbel distribution. In the former case, $\overline{B}$ is regularly varying, i.e. $\overline{B}(x) = L(x)/x^{\alpha+1}$, for some $\alpha > 0$ and we write it as $\overline{B} \in \mathcal{R}(-\alpha - 1)$, $\alpha > 0$.

We assume that $B_0 \in \mathcal{S}$ and either $\overline{B} \in \mathcal{R}(-\alpha - 1)$ or $B \in \mathcal{G}$, i.e. the maximum domain of attraction of Gumbel distribution, that is, there exist sequences $a_n > 0$,

$b_n \in \mathbb{R}$, such that $\lim_{n\to\infty} n\overline{B}(a_n x + b_n) = e^{-x}$, $x \in \mathbb{R}$. $\mathcal{G}$ includes Weibull and lognormal distributions.

When the arrival process N_t satisfies a large deviation result, the probability that it deviates away from its mean is exponentially small, which is dominated by subexonential distributions. The results in Zhu [21] for the asymptotics of ruin probabilities for risk processes with non-stationary, non-renewal arrivals and subexponential claims can be applied in the context of marked Hawkes arrivals. We have the following infinite-horizon and finite-horizon ruin probability estimates when the claim sizes are subexponential.

Theorem 11. *Assume the net profit condition* $\rho > \mathbb{E}[C_1]\frac{\nu}{1-\mathbb{E}^q[H(a)]}$.

(i) (Infinite-Horizon)

$$\lim_{u\to\infty} \frac{\psi(u)}{\overline{B}_0(u)} = \frac{\nu\mathbb{E}[C_1]}{\rho(1-\mathbb{E}^q[H(a)]) - \nu\mathbb{E}[C_1]}. \tag{4.14}$$

(ii) (Finite-Horizon) For any $T > 0$,

$$\lim_{u\to\infty} \frac{\psi(u, uz)}{\overline{B}_0(u)} \tag{4.15}$$

$$= \begin{cases} \frac{\nu\mathbb{E}[C_1]}{\rho(1-\mathbb{E}^q[H(a)])-\nu\mathbb{E}[C_1]} \left[1 - \left(1 + \left(\frac{\rho(1-\mathbb{E}^q[H(a)])-\nu\mathbb{E}[C_1]}{\rho(1-\mathbb{E}^q[H(a)])}\right)\frac{T}{\alpha}\right)^{-\alpha}\right] & \text{if } \overline{B} \in \mathcal{R}(-\alpha-1) \\ \frac{\nu\mathbb{E}[C_1]}{\rho(1-\mathbb{E}^q[H(a)])-\nu\mathbb{E}[C_1]} \left[1 - e^{-\frac{\rho(1-\mathbb{E}^q[H(a)])-\nu\mathbb{E}[C_1]}{\rho(1-\mathbb{E}^q[H(a)])}T}\right] & \text{if } B \in \mathcal{G} \end{cases}.$$

5. Examples with Explicit Formulas

In this section, we discuss two examples where an explicit formula exists.

Example 12 is about the exponential asymptotics of the infinite-horizon ruin probability when $H(a)$ and the claim size C are exponentially distributed. Example 13 gives an explicit expression for the rate function of the large deviation principle when $H(a)$ is exponentially distributed.

Example 12. *Recall that x is the minimal solution of*

$$x = \int_{\mathbb{R}^+}\int_{\mathbb{X}} e^{\theta C + (x-1)H(a)} q(da)\mu(dC). \tag{5.1}$$

Now, assume that $H(a)$ is exponentially distributed with parameter $\lambda > 0$, then, we have

$$x = \mathbb{E}^{\mu}[e^{\theta C}]\frac{\lambda}{\lambda - (x-1)}, \tag{5.2}$$

which implies that

$$x = \frac{1}{2}\left\{\lambda + 1 - \sqrt{(\lambda+1)^2 - 4\lambda\mathbb{E}^{\mu}[e^{\theta C}]}\right\}. \tag{5.3}$$

Now, assume that C is exponentially distributed with parameter $\gamma > 0$. Then,

$$x = \frac{1}{2}\left\{\lambda + 1 - \sqrt{(\lambda+1)^2 - 4\lambda\frac{\gamma}{\gamma-\theta}}\right\}. \tag{5.4}$$

The infinite horizon probability satisfies $\lim_{u\to\infty}\frac{1}{u}\log\psi(u) = -\theta^{\dagger}$, *where* $\theta^{\dagger}$ *satisfies*

$$\rho\theta^{\dagger} = \nu\left(\frac{1}{2}\left\{\lambda + 1 - \sqrt{(\lambda+1)^2 - 4\lambda\frac{\gamma}{\gamma-\theta^{\dagger}}}\right\} - 1\right), \tag{5.5}$$

which implies

$$\frac{2\rho\theta^\dagger}{\nu}+1-\lambda=-\sqrt{(\lambda+1)^2-\frac{4\lambda\gamma}{\gamma-\theta^\dagger}}, \tag{5.6}$$

and thus

$$\frac{\rho^2}{\nu^2}(\theta^\dagger)^2+\frac{\rho\theta^\dagger}{\nu}(1-\lambda)=\lambda-\frac{\lambda\gamma}{\gamma-\theta^\dagger}=\frac{-\lambda\theta^\dagger}{\gamma-\theta^\dagger}. \tag{5.7}$$

Since we are looking for positive $\theta^\dagger$, we get the quadratic equation,

$$\rho^2(\theta^\dagger)^2-(\rho^2\gamma-\rho\nu(1-\lambda))\theta^\dagger-(\rho\nu\gamma(1-\lambda)+\lambda\nu^2)=0. \tag{5.8}$$

Since $\rho>\frac{\mathbb{E}^\mu[C]\nu}{1-\mathbb{E}^q[H(a)]}=\frac{\nu\lambda}{\gamma(\lambda-1)}$, we have $\rho\nu\gamma(1-\lambda)+\lambda\nu^2>0$. Therefore,

$$\theta^\dagger=\frac{(\rho^2\gamma-\rho\nu(1-\lambda))+\sqrt{(\rho^2\gamma-\rho\nu(1-\lambda))^2+4\rho^2(\rho\nu\gamma(1-\lambda)+\lambda\nu^2)}}{2\rho^2}. \tag{5.9}$$

Example 13. *Now, let $H(a)$ be exponentially distributed with parameter $\lambda>0$. We want an explicit expression for the rate function of the large deviation principle for $(N_t/t\in\cdot)$. Notice that,*

$$\Gamma(\theta)=\begin{cases}\nu\left(\frac{1}{2}\left\{\lambda+1-\sqrt{(\lambda+1)^2-4\lambda e^\theta}\right\}-1\right) & \text{for } \theta\leq\log\left(\frac{(\lambda+1)^2}{4\lambda}\right)\\ +\infty & \text{otherwise}\end{cases}. \tag{5.10}$$

To get $I(x)=\sup_{\theta\in\mathbb{R}}\{\theta x-\Gamma(\theta)\}$, we optimize over θ and consider $x=\Gamma'(\theta)$. Evidently,

$$x+\frac{1}{2}\nu(-4\lambda)e^\theta\frac{1}{2\sqrt{(\lambda+1)^2-4\lambda e^\theta}}=0, \tag{5.11}$$

which gives us

$$\theta=\log\left(\frac{-2x^2+x\sqrt{4x^2+\nu^2(\lambda+1)^2}}{\lambda\nu^2}\right), \tag{5.12}$$

whence,

$$I(x)=\begin{cases}x\log\left(\frac{-2x^2+x\sqrt{4x^2+\nu^2(\lambda+1)^2}}{\lambda\nu^2}\right)\\ \qquad -\nu\left(\frac{1}{2}\left\{\lambda+1-\frac{-2x+\sqrt{4x^2+\nu^2(\lambda+1)^2}}{\nu}\right\}-1\right) & \text{if } x\geq 0\\ +\infty & \text{otherwise}\end{cases}. \tag{5.13}$$

Acknowledgements

The authors are both supported by NSF grant DMS-0904701, DARPA grant and MacCracken Fellowship at New York University. The authors are grateful to the editor and the referees for a very careful reading of the manuscript and also for the helpful suggestions.

References

[1] Bacry, E., Delattre, S., Hoffmann, M., and J. F. Muzy. (2013). Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Processes and their Applications*. **123**, 2475-2499.

[2] Bordenave, C. and G. L. Torrisi. (2007). Large deviations of Poisson cluster processes. *Stochastic Models*. **23**, 593-625.

[3] Brémaud, P., and L. Massoulié. (1996). Stability of nonlinear Hawkes processes. *Annals of Probability*. **24**, 1563-1588.

[4] Brémaud, P., Nappo, G. and G. L. Torrisi. (2002). Rate of convergence to equilibrium of marked Hawkes processes. *J. Appl. Prob.* **39**, 123-136.

[5] Daley, D. J. and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Volume I and II, Springer, Second Edition, 2003

[6] Dembo, A. and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd Edition, Springer, 1998

[7] Glynn, P. W. and W. Whitt. (1994). Logarithmic asymptotics for steady-state tail probabilities in a single-server queue. *J. Appl. Probab.* **31A**, 131-156.

[8] Goldie, C. M. and S. Resnick. (1988). Distributions that are both subexponential and in the domain of attraction of an extreme value distribution. *Adv. Appl. Probab.* **20**, 706-718.

[9] Hawkes, A. G. and D. Oakes. (1974). A cluster process representation of a self-exciting process. *J. Appl. Prob.* **11**, 493-503.

[10] Jacod, J. and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin, 1987.

[11] Karabash, D. On stability of Hawkes process. *arXiv:1201.1573*.

[12] Koralov, L. B. and Ya. G. Sinai. *Theory of Probability and Random Processes*. Springer, 2nd edition, 2012.

[13] Stabile, G. and G. L. Torrisi. (2010). Risk processes with non-stationary Hawkes arrivals. *Methodol. Comput. Appl. Prob.* **12**, 415-429.

[14] Varadhan, S. R. S. (2008). Large deviations. *Annals of Probability*. **36**, 397-419.

[15] Varadhan, S. R. S., *Large Deviations and Applications*, SIAM, 1984

[16] Zhu, L. (2015). Large deviations for Markovian nonlinear Hawkes processes. *Annals of Applied Probability*. **25**, 548-581.

[17] Zhu, L. (2014). Process-level large deviations for nonlinear Hawkes point processes. *Annales de l'Institut Henri Poincaré-Probabilités et Statistiques*. **50**, 845-871.

[18] Zhu, L. (2013). Central limit theorem for nonlinear Hawkes processes. *Journal of Applied Probability*. **50**, 760-771.

[19] Zhu, L. (2013). Moderate deviations for Hawkes processes. *Statistics & Probability Letters*. **83**, 885-890.

[20] Zhu, L. (2014). Limit theorems for a Cox-Ingersoll-Ross process with Hawkes jumps. *Journal of Applied Probability*. **51**, 699-712.

[21] 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: karabash@cims.nyu.edu

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