Stochastic Models, 23:593–625, 2007

ISSN: 1532-6349 print/1532-4214 online
DOI: 10.1080/15326340701645959

LARGE DEVIATIONS OF POISSON CLUSTER PROCESSES

Charles Bordenave □ *INRIA/ENS, Départment d'Informatique, Paris, France*

Giovanni Luca Torrisi □ *Istituto per le Applicazioni del Calcolo "Mauro Picone" (IAC), Consiglio Nazionale delle Ricerche (CNR), Roma, Italia*

□ *In this paper we prove scalar and sample path large deviation principles for a large class of Poisson cluster processes. As a consequence, we provide a large deviation principle for ergodic Hawkes point processes.*

Keywords Hawkes processes; Large deviations; Poisson cluster processes; Poisson processes.

Mathematics Subject Classification 60F10; 60G55.

1. INTRODUCTION

Poisson cluster processes are one of the most important classes of point process models (see Daley and Vere-Jones[7] and Møller; Waagepetersen[24]). They are natural models for the location of objects in the space, and are widely used in point process studies whether theoretical or applied. Very popular and versatile Poisson cluster processes are the so-called self-exciting or Hawkes processes (Hawkes[12,13]; Hawkes and Oakes[14]). From a theoretical point of view, Hawkes processes combine both a Poisson cluster process representation and a simple stochastic intensity representation.

Poisson cluster processes found applications in cosmology, ecology and epidemiology; see, respectively, Neyman and Scott[25], Brix and Chadoeuf[5], and Møller[19,20]. Hawkes processes are particularly appealing for seismological applications. Indeed, they are widely used as statistical models for the standard activity of earthquake series; see the papers by

Received October 2006; Accepted August 2007
Address correspondence to Charles Bordenave, INRIA/ENS, Départment d'Informatique, 45 rue d'Ulm, F-75230 Paris Cedex 05, France; E-mail: charles.bordenave@ens.fr

Ogata and Akaike[28], Vere-Jones and Ozaki[32], and Ogata[26,27]. Hawkes processes have also aspects appealing to neuroscience applications; see the paper by Johnson[16]. More recently, Hawkes processes found applications in finance; see Chavez-Demoulin et al.[6]; and in DNA modeling; see Gusto and Schbath[11].

In this paper we derive scalar and sample path large deviation principles for Poisson cluster processes. Our results are potentially useful to study risk processes with Poisson cluster arrivals and light-tailed claims. Particularly, they may lead to determine the asymptotic behavior of the ruin probability, the most likely path to ruin and an efficient Monte Carlo algorithm to estimate the ruin probability. Results in this direction can be found in Stabile and Torrisi[30], where risk processes with non-stationary Hawkes arrivals are studied.

The paper is organized as follows. In Section 2 we give some preliminaries on Poisson cluster processes, Hawkes processes and large deviations. In Section 3 we provide scalar large deviation principles for Poisson cluster processes, under a light-tailed assumption on the number of points per cluster. As consequence, we provide scalar large deviations for ergodic Hawkes processes. Section 4 is devoted to sample path large deviations of Poisson cluster processes. First, we prove a sample path large deviation principle on $D[0,1]$ equipped with the topology of point-wise convergence, under a light-tailed assumption on the number of points per cluster. Second, we give a sample path large deviation principle on $D[0,1]$ equipped with the topology of uniform convergence, under a super-exponential assumption on the number of points per cluster. In Section 5 we prove large deviations for spatial Poisson cluster processes, and we provide the asymptotic behavior of the void probability function and the empty space function. We conclude the paper with a short discussion.

2. PRELIMINARIES

In this section we recall the definition of Poisson cluster process, Hawkes process, and the notion of large deviation principle.

2.1. Poisson Cluster Processes

A Poisson cluster process $\mathbf{X} \subset \mathbb{R}$ is a point process. The clusters centers of $\mathbf{X}$ are given by particular points called immigrants; the other points of the process are called offspring. The formal definition of the process is the following:

(a) The immigrants are distributed according to a homogeneous Poisson process I with points $X_i \in \mathbb{R}$ and intensity $\nu > 0$.

(b) Each immigrant X_i generates a cluster $C_i = C_{X_i}$, which is a finite point process (i.e., it has a finite number of points almost surely) containing X_i.
(c) Given the immigrants, the centered clusters

$$C_i - X_i = \{Y - X_i : Y \in C_i\}, \quad X_i \in I$$

are independent, identically distributed (iid for short), and independent of I.
(d) $\mathbf{X}$ consists of the union of all clusters.

The number of points in a cluster is denoted by S. We will assume that $\mathrm{E}[S] < \infty$. Let $\mathbf{Y}$ be a point process on $\mathbb{R}$ and $N_{\mathbf{Y}}(0, t]$ the number of points of $\mathbf{Y}$ in the interval $(0, t]$. $\mathbf{Y}$ is said stationary if its law is translations invariant, is said ergodic if it is stationary, with a finite intensity $\mathrm{E}[N_{\mathbf{Y}}(0, 1]]$, and

$$\lim_{t\to\infty} \frac{N_{\mathbf{Y}}(0, t]}{t} = \mathrm{E}[N_{\mathbf{Y}}(0, 1]], \quad \text{a.s.}$$

By the above definition of Poisson cluster process it is clear that $\mathbf{X}$ is ergodic with finite intensity $\nu\mathrm{E}[S]$. In particular,

$$\lim_{t\to\infty} \frac{N_{\mathbf{X}}(0, t]}{t} = \nu\mathrm{E}[S], \quad \text{a.s.} \tag{1}$$

2.2. Hawkes Processes

We say that $\mathbf{X} \subset \mathbb{R}$ is a Hawkes process if it is a Poisson cluster process with (b) in the definition above replaced by:

(b) Each immigrant X_i generates a cluster $C_i = C_{X_i}$, which is the random set formed by the points of generations $n = 0, 1, \ldots$ with the following branching structure: the immigrant X_i is said to be of generation 0. Given generations $0, 1, \ldots, n$ in C_i, each point $Y \in C_i$ of generation n generates a Poisson process on (Y, ∞), say Φ, of offspring of generation $n + 1$ with intensity function $h(\cdot - Y)$. Here $h : (0, \infty) \to [0, \infty)$ is a non-negative Borel function called fertility rate.

We refer the reader to Section 2 in Møller and Rasmussen[21] for more insight into the branching structure and self-similarity property of clusters. Consider the mean number of points in any offspring process Φ:

$$\mu = \int_0^\infty h(t)\mathrm{d}t.$$

As usual in the literature on Hawkes processes, throughout this paper we assume

$$0 < \mu < 1. \tag{2}$$

Condition $\mu > 0$ excludes the trivial case in which there are almost surely no offspring. Recalling that the total number of points in a cluster is equivalent to the total progeny of the Galton-Watson process with one ancestor and number of offspring per individual following a Poisson distribution with mean μ (see p. 496 of Hawkes and Oakes[14]), the other condition $\mu < 1$ is equivalent to assuming that $\mathrm{E}[S] = 1/(1-\mu) < \infty$. For our purposes it is important to recall that for Hawkes processes the distribution of S is given by

$$P(S = k) = \frac{\mathrm{e}^{-k\mu}(k\mu)^{k-1}}{k!}, \quad k = 1, 2, \ldots \tag{3}$$

This follows by Theorem 2.11.2 in the book by Jagers[15]. Finally, since $\mathbf{X}$ is ergodic with a finite and positive intensity equal to $\nu/(1-\mu)$ it holds:

$$\lim_{t\to\infty} \frac{N_{\mathbf{X}}(0, t]}{t} = \frac{\nu}{1-\mu}, \quad \text{a.s.} \tag{4}$$

2.3. Large Deviation Principles

We recall here some basic definitions in large deviations theory (see, for instance, the book by Dembo and Zeitouni[9]). A family of probability measures $\{\mu_\alpha\}_{\alpha\in(0,\infty)}$ on a topological space $(M, \mathcal{T}_M)$ satisfies the large deviations principle (LDP for short) with rate function $J(\cdot)$ and speed $v(\cdot)$ if $J : M \to [0, \infty]$ is a lower semi-continuous function, $v : [0, \infty) \to [0, \infty)$ is a measurable function which increases to infinity, and the following inequalities hold for every Borel set B:

$$-\inf_{x\in B^\circ} J(x) \le \liminf_{\alpha\to\infty} \frac{1}{v(\alpha)} \log \mu_\alpha(B) \le \limsup_{\alpha\to\infty} \frac{1}{v(\alpha)} \log \mu_\alpha(B) \le -\inf_{x\in\overline{B}} J(x),$$

where B° denotes the interior of B and $\overline{B}$ denotes the closure of B. Similarly, we say that a family of M-valued random variables $\{V_\alpha\}_{\alpha\in(0,\infty)}$ satisfies the LDP if $\{\mu_\alpha\}_{\alpha\in(0,\infty)}$ satisfies the LDP and $\mu_\alpha(\cdot) = P(V_\alpha \in \cdot)$. We point out that the lower semi-continuity of $J(\cdot)$ means that its level sets:

$$\{x \in M : J(x) \le a\}, \quad a \ge 0,$$

are closed; when the level sets are compact the rate function $J(\cdot)$ is said to be good.

3. SCALAR LARGE DEVIATIONS

3.1. Scalar Large Deviations of Poisson Cluster Processes

Consider the ergodic Poisson cluster process $\mathbf{X}$ described above. In this section we prove that the process $\{N_{\mathbf{X}}(0,t]/t\}$ satisfies a LDP on $\mathbb{R}$. Define the set

$$\mathcal{D}_S = \left\{\theta \in \mathbb{R} : \mathrm{E}[\mathrm{e}^{\theta S}] < \infty\right\}.$$

With a little abuse of notation, denote by C_0 the cluster generated by an immigrant at 0 and let $L = \sup_{Y \in C_0} |Y|$ be the radius of C_0. We shall consider the following conditions:

$$\text{the function} \quad \theta \mapsto \mathrm{E}[\mathrm{e}^{\theta S}] \quad \text{is essentially smooth and } 0 \in \mathcal{D}_S^{\circ} \tag{5}$$

and

$$\mathrm{E}[L\mathrm{e}^{\theta S}] < \infty \quad \text{for all } \theta \in \mathcal{D}_S^{\circ}. \tag{6}$$

For the definition of essentially smooth function, we refer the reader to Definition 2.3.5. in Dembo and Zeitouni[9].

It is worthwhile to mention that conditions (5) and (6) are satisfied by many classes of Poisson cluster processes, which are used in the applications. For instance, we shall show in Subsection 3.2 that, under suitable assumptions on the fertility rate h, they are satisfied by ergodic Hawkes processes (see the proof of Theorem 3.2.1). We refer the reader to the comment after the statement of Theorem 5.1.1 for a similar remark in the spatial context.

Remark 3.1.1. Since $S \geq 1$ we have that the function $\varphi(\theta) = \mathrm{E}[\mathrm{e}^{\theta S}]$ is increasing. It follows that $\mathcal{D}_S^{\circ} = (-\infty, \theta_0)$ with $\theta_0 \in [0, \infty]$. By the dominated convergence theorem we have that $\varphi'(\theta) = \mathrm{E}[S\mathrm{e}^{\theta S}]$ and $\varphi''(\theta) = \mathrm{E}[S^2\mathrm{e}^{\theta S}]$, for all $\theta \in \mathcal{D}_S^{\circ}$. Hence, if $\theta_0 < \infty$, to prove that φ is essentially smooth it suffices to show that $\mathrm{E}[S\mathrm{e}^{\theta_0 S}] = \infty$. On the other hand, if $\theta_0 = +\infty$, the function φ is always essentially smooth.

It holds:

Theorem 3.1.1. *Assume* (5) *and* (6). *Then* $\{N_{\mathbf{X}}(0,t]/t\}$ *satisfies a LDP on* $\mathbb{R}$ *with speed* t *and good rate function*

$$\Lambda^*(x) = \sup_{\theta \in \mathbb{R}} (\theta x - \Lambda(\theta)). \tag{7}$$

where $\Lambda(\theta) = \nu(\mathrm{E}[\mathrm{e}^{\theta S}] - 1)$.

It is easily verified that $\Lambda^*(\nu \mathrm{E}[S]) = 0$. Moreover, this is the unique zero of $\Lambda^*(\cdot)$. Therefore the probability law of $N_{\mathbf{X}}(0,t]/t$ concentrates in arbitrarily small neighborhoods of $\nu\mathrm{E}[S]$ as $t \to \infty$, as stated by the law of large numbers (1). The LDP is a refinement of the law of large numbers in that it gives us the probability of fluctuations away the most probable value.

Before proving Theorem 3.1.1 we show that the same LDP holds for the non-stationary Poisson cluster process $\mathbf{X}_{t,T}$ with immigrant process empty on $(-\infty, -T) \cup (t+T, \infty)$, where $T > 0$ is a fixed constant. Furthermore, the LDP for $\mathbf{X}_{t,T}$ holds under a weaker condition.

Theorem 3.1.2. *Assume* (5). *Then* $\{N_{\mathbf{X}_{t,T}}(0,t]/t\}$ *satisfies a LDP on* $\mathbb{R}$ *with speed* t *and good rate function* (7).

Proof. The proof is based on the Gärtner-Ellis theorem (see, for instance, Theorem 2.3.6 in Dembo and Zeitouni[9]). We start proving that

$$\lim_{t\to\infty} \frac{1}{t} \log \mathrm{E}\big[\mathrm{e}^{\theta N_{\mathbf{X}_{t,T}}(0,t]}\big] = \begin{cases} \nu(\mathrm{E}[\mathrm{e}^{\theta S}] - 1) & \text{if } \theta \in \mathcal{D}_S \\ +\infty & \text{if } \theta \notin \mathcal{D}_S \end{cases} \tag{8}$$

For a Borel set $A \subset \mathbb{R}$, let $I_{|A} = I \cap A$ be the point process of immigrants in A. Clearly $I_{|(0,t]}$, $I_{|[-T,0]}$ and $I_{|(t,t+T]}$ are independent Poisson processes with intensity ν, respectively on $(0,t]$, $[-T,0]$ and $(t, T+t]$. Since $I_{|(0,t]}$, $I_{|[-T,0]}$, and $I_{|(t,t+T]}$ are independent, by the definition of Poisson cluster process it follows that the random sets $\{C_i : X_i \in I_{|(0,t]}\}$, $\{C_i : X_i \in I_{|[-T,0]}\}$ and $\{C_i : X_i \in I_{|(t,t+T]}\}$ are independent. Therefore, for all $\theta \in \mathbb{R}$,

$$\begin{aligned} \mathrm{E}\big[\mathrm{e}^{\theta N_{\mathbf{X}_{t,T}}(0,t]}\big] &= \mathrm{E}\Big[\mathrm{e}^{\theta(\sum_{X_i\in I_{|(0,t]}} N_{C_i}(0,t] + \sum_{X_i \in I_{|[-T,0]}} N_{C_i}(0,t] + \sum_{X_i\in I_{|(t,t+T]}} N_{C_i}(0,t])}\Big] \\ &= \mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(0,t]}} N_{C_i}(0,t]}\Big] \mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|[-T,0]}} N_{C_i}(0,t]}\Big] \mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(t,t+T]}} N_{C_i}(0,t]}\Big]. \end{aligned}$$

We shall show

$$\lim_{t\to\infty} \frac{1}{t} \log \mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(0,t]}} N_{C_i}(0,t]}\Big] = \begin{cases} \nu(\mathrm{E}\big[\mathrm{e}^{\theta S}\big] - 1) & \text{if } \theta \in \mathcal{D}_S \\ +\infty & \text{if } \theta \notin \mathcal{D}_S \end{cases} \tag{9}$$

and

$$\begin{aligned} \lim_{t\to\infty} \frac{1}{t} \log \mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|[-T,0]}} N_{C_i}(0,t]}\Big] &= \lim_{t\to\infty} \frac{1}{t} \log \mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(t,t+T]}} N_{C_i}(0,t]}\Big] \\ &= 0, \quad \text{for } \theta \in \mathcal{D}_S. \end{aligned} \tag{10}$$

Note that (8) is a consequence of (9) and (10). We first prove (9). With a little abuse of notation, denote by C_0 the cluster generated by an immigrant at 0. Since $\{(X_i, C_i) : X_i \in I_{|(0,t]}\}$ is an independently marked Poisson process, by Lemma 6.4.VI in Daley and Vere-Jones[7] we have

$$\begin{aligned}
\mathrm{E}\Big[\mathrm{e}^{\theta \sum_{X_i \in I_{|(0,t]}} N_{C_i}(0,t]}\Big] &= \mathrm{E}\Big[\mathrm{e}^{\theta \sum_{X_i \in I_{|(0,t]}} N_{C_i - X_i}(-X_i, t-X_i]}\Big] \\
&= \exp\left(\nu \int_0^t \mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(-x,t-x]} - 1\big]\mathrm{d}x\right) \\
&= \exp\left(\nu t \int_0^1 \mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(-tz,(1-z)t]} - 1\big]\mathrm{d}z\right). \qquad (11)
\end{aligned}$$

Therefore if $\theta \in \mathcal{D}_S$, the expectation in (11) goes to $\mathrm{E}[\mathrm{e}^{\theta S} - 1]$ as $t \to \infty$ by the monotone convergence theorem. Hence, for $\theta \in \mathcal{D}_S$ the limit (9) follows from the dominated convergence theorem. For $\theta \notin \mathcal{D}_S$ the expectation in (11) goes to $+\infty$ as $t \to \infty$ by the monotone convergence theorem, and the limit (9) follows by Fatou's lemma. We now show (10). Here again, since $\{(X_i, C_i) : X_i \in I_{|[-T,0]}\}$ is an independently marked Poisson process, by Lemma 6.4.VI in Daley and Vere-Jones[7] we have

$$\begin{aligned}
\mathrm{E}\Big[\mathrm{e}^{\theta \sum_{X_i \in I_{|[-T,0]}} N_{C_i}(0,t]}\Big] &= \mathrm{E}\Big[\mathrm{e}^{\theta \sum_{X_i \in I_{|[-T,0]}} N_{C_i - X_i}(-X_i, t-X_i]}\Big] \\
&= \exp\left(\nu \int_0^T \mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(x,x+t]} - 1\big]\mathrm{d}x\right). \qquad (12)
\end{aligned}$$

Now note that, for $\theta \in \mathcal{D}_S \cap [0, \infty)$, we have

$$0 \le \frac{1}{t} \log \mathrm{E}\big[\mathrm{e}^{\theta \sum_{X_i \in I_{|[-T,0]}} N_{C_i}(0,t]}\big] \le \frac{\nu}{t} \int_0^T \mathrm{E}\big[\mathrm{e}^{\theta S} - 1\big]\mathrm{d}x < \infty$$

and, for each $\theta \le 0$,

$$\frac{\nu T}{t}\mathrm{E}\big[\mathrm{e}^{\theta S} - 1\big] \le \frac{\nu}{t} \int_0^T \mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(x,x+t]} - 1\big]\mathrm{d}x \le 0.$$

By passing to the limit as $t \to \infty$ we get that the first limit in (10) is equal to 0. The proof for the second limit in (10) is rigorously the same. Hence we proved (8). Using assumption (5), the conclusion is a consequence of the Gärtner–Ellis theorem.

Proof of Theorem 3.1.1. The proof is similar to that one of Theorem 3.1.2 and is again based on the Gärtner–Ellis theorem. We start showing that

$$\lim_{t\to\infty} \frac{1}{t} \log \mathrm{E}\big[\mathrm{e}^{\theta N_{\mathbf{X}}(0,t]}\big] = \begin{cases} \nu(\mathrm{E}[\mathrm{e}^{\theta S}] - 1) & \text{if } \theta \in \mathcal{D}_S^{\circ} \\ +\infty & \text{if } \theta \notin \mathcal{D}_S \end{cases} \qquad (13)$$

By similar arguments, as in the proof of Theorem 3.1.2, using the definition of $\mathbf{X}$, we have

$$\mathrm{E}\big[\mathrm{e}^{\theta N_{\mathbf{X}}(0,t]}\big] = \mathrm{E}\big[\mathrm{e}^{\theta N_{\mathbf{X}_{t,T}}(0,t]}\big]\mathrm{E}\big[\mathrm{e}^{\theta(N_{\mathbf{X}}(0,t]-N_{\mathbf{X}_{t,T}}(0,t])}\big], \quad \text{for all } \theta \in \mathbb{R}, \ \ t > 0.$$

By the computations in the proof of Theorem 3.1.2, in order to prove (13) we only need to check that

$$\lim_{t\to\infty} \frac{1}{t} \log \mathrm{E}\big[\mathrm{e}^{\theta(N_{\mathbf{X}}(0,t]-N_{\mathbf{X}_{t,T}}(0,t])}\big] = 0, \qquad \text{for all } \theta \in \mathscr{D}_S^\circ. \tag{14}$$

It is easily verified for $\theta \leq 0$ (the argument of the expectation is bounded below by $\mathrm{e}^{\theta S}$ and above by 1). We only check (14) for $\theta \in \mathscr{D}_S^\circ \cap (0,\infty)$. Here again, for a Borel set $A \subset \mathbb{R}$, let $I_{|A} = I \cap A$ denote the point process of immigrants in A. Note that

$$N_{\mathbf{X}}(0,t] - N_{\mathbf{X}_{t,T}}(0,t] = \sum_{X_i \in I_{|(-\infty,-T)}} N_{C_i}(0,t] + \sum_{X_i \in I_{|(t+T,\infty)}} N_{C_i}(0,t], \quad t > 0.$$

Clearly $I_{|(-\infty,-T)}$ and $I_{|(t+T,\infty)}$ are independent Poisson processes with intensity ν, respectively on $(-\infty,-T)$ and $(t+T,\infty)$. Thus, by the definition of Poisson cluster process it follows that the random sets $\{C_i : X_i \in I_{|(-\infty,-T)}\}$ and $\{C_i : X_i \in I_{|(t+T,\infty)}\}$ are independent. Therefore, for all $\theta \in \mathscr{D}_S^\circ \cap (0,\infty)$,

$$\mathrm{E}\big[\mathrm{e}^{\theta(N_{\mathbf{X}}(0,t]-N_{\mathbf{X}_{t,T}}(0,t])}\big] = \mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(-\infty,-T)}} N_{C_i}(0,t]}\Big]\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(t+T,\infty)}} N_{C_i}(0,t]}\Big].$$

Since $\{(X_i,C_i) : X_i \in I_{|(-\infty,-T)}\}$ and $\{(X_i,C_i) : X_i \in I_{|(t+T,\infty)}\}$ are independently marked Poisson processes, by Lemma 6.4.VI in Daley and Vere-Jones[7] we have

$$\begin{aligned}\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(-\infty,-T)}} N_{C_i}(0,t]}\Big] &= \mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(-\infty,-T)}} N_{C_i-X_i}(-X_i,t-X_i]}\Big] \\ &= \exp\left(\nu\int_T^\infty \mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(x,t+x]}-1\big]\mathrm{d}x\right)\end{aligned}$$

and

$$\begin{aligned}\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(t+T,\infty)}} N_{C_i}(0,t]}\Big] &= \exp\left(\nu\int_{t+T}^\infty \mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(-x,t-x]}-1\big]\mathrm{d}x\right) \\ &= \exp\left(\nu\int_T^\infty \mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(-t-z,-z]}-1\big]\mathrm{d}z\right).\end{aligned}$$

Now notice that since $\theta > 0$ we have

$$\mathrm{e}^{\theta N_{C_0}(x,x+t]} - 1 \leq \big(\mathrm{e}^{\theta N_{C_0}(\mathbb{R})} - 1\big)\mathbf{1}\{x \leq L\}, \qquad \text{for all } x \geq T$$

and

$$\mathrm{e}^{\theta N_{C_0}(-t-z,-z]}-1\leq\big(\mathrm{e}^{\theta N_{C_0}(\mathbb{R})}-1\big)\mathbf{1}\{z\leq L\},\quad\text{for all } z\geq T.$$

Relation (14) follows by assumption (6) noticing that the above relations yield

$$\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(-\infty,-T)}}N_{C_i}(0,t]}\Big]\\ \leq\exp(v\mathrm{E}[L(\mathrm{e}^{\theta S}-1)]),\quad\text{for all }\theta\in\mathscr{D}_S^\circ\cap(0,\infty),\ \ t>0.$$

and

$$\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|(t+T,\infty)}}N_{C_i}(0,t]}\Big]\\ \leq\exp(v\mathrm{E}[L(\mathrm{e}^{\theta S}-1)]),\quad\text{for all }\theta\in\mathscr{D}_S^\circ\cap(0,\infty),\ \ t>0.$$

Therefore, (13) is proved. Now, if $\mathscr{D}_S=\mathscr{D}_S^\circ$ then the claim is a consequence of the Gärtner-Ellis theorem and assumption (5). It remains to deal with the case $\mathscr{D}_S\neq\mathscr{D}_S^\circ$. We shall show the large deviations upper and lower bounds proving that for any sequence $\{t_n\}_{n\geq1}\subset(0,\infty)$ diverging to $+\infty$, as $n\to\infty$, there exists a subsequence $\{s_n\}\subseteq\{t_n\}$ such that

$$\limsup_{n\to\infty}\frac{1}{s_n}\log P(N_{\mathbf{X}}(0,s_n]/s_n\in F)\leq-\inf_{x\in F}\Lambda^*(x),\qquad\text{for all closed sets } F \tag{15}$$

and

$$\liminf_{n\to\infty}\frac{1}{s_n}\log P(N_{\mathbf{X}}(0,s_n]/s_n\in G)\geq-\inf_{x\in G}\Lambda^*(x),\quad\text{for all open sets } G, \tag{16}$$

where Λ^* is defined by (7) (then the large deviations upper and lower bounds hold for any sequence $\{t_n\}$ and the claim follows). By assumption (5), there exists $\theta_0>0$ such that $\mathscr{D}_S=(\infty,\theta_0]$. Let $\{t_n\}_{n\geq1}\subset(0,\infty)$ be a sequence diverging to $+\infty$, as $n\to\infty$, and define the extended non-negative real number $l\in[0,\infty]$ by

$$l\equiv\limsup_{n\to\infty}\frac{1}{t_n}\log\mathrm{E}\big[\mathrm{e}^{\theta_0N_{\mathbf{X}}(0,t_n]}\big].$$

Clearly, there exists a subsequence $\{s_n\}\subseteq\{t_n\}$ that realizes this lim sup, i.e.,

$$\lim_{n\to\infty}\frac{1}{s_n}\log\mathrm{E}\big[\mathrm{e}^{\theta_0N_{\mathbf{X}}(0,s_n]}\big]=l.$$

By (13) it follows that

$$\lim_{n\to\infty} \frac{1}{s_n} \log \mathrm{E}\left[\mathrm{e}^{\theta N_{\mathbf{X}}(0,s_n]}\right] = \widetilde{\Lambda}(\theta), \quad \theta \in \mathbb{R}$$

where

$$\widetilde{\Lambda}(\theta) = \begin{cases} \Lambda(\theta) & \text{if } \theta < \theta_0 \\ l & \text{if } \theta = \theta_0 \\ +\infty & \text{if } \theta > \theta_0. \end{cases}$$

Note that, irrespective to the value of l, $\widetilde{\Lambda}$ is essentially smooth (however, it may be not lower semi-continuous). We now show that the Legendre transform of Λ and $\widetilde{\Lambda}$ coincide, i.e.,

$$\widetilde{\Lambda}^*(x) = \Lambda^*(x), \quad x \in \mathbb{R}. \tag{17}$$

A straightforward computation gives $\widetilde{\Lambda}^*(x) = \Lambda^*(x) = +\infty$, for $x < 0$, and $\widetilde{\Lambda}^*(0) = \Lambda^*(0) = \nu$. Now, note that since $\theta_0 < \infty$, $\widetilde{\Lambda}^*(x)$ and $\Lambda^*(x)$ are both finite, for $x > 0$. Moreover, since Λ and $\widetilde{\Lambda}$ are essentially smooth, if $x > 0$ we have that $\Lambda^*(x) = \theta_x x - \Lambda(\theta_x)$ and $\widetilde{\Lambda}^*(x) = \tilde{\theta}_x x - \widetilde{\Lambda}(\tilde{\theta}_x)$, where θ_x (respectively $\tilde{\theta}_x$) is the unique solution of $\Lambda'(\theta) = x$ (respectively $\widetilde{\Lambda}(\theta) = x$) on $(-\infty, \theta_0)$. The claim (17) follows recalling that $\widetilde{\Lambda}(\theta) = \Lambda(\theta) = \nu(\mathrm{E}[\mathrm{e}^{\theta S}] - 1)$ on $\mathscr{D}_S^\circ$. Now, applying part (a) of Theorem 2.3.6 in Dembo and Zeitouni[9] we have (15). Applying part (b) of Theorem 2.3.6 in Dembo and Zeitouni[9] we get

$$\liminf_{n\to\infty} \frac{1}{s_n} \log P(N_{\mathbf{X}}(0, s_n]/s_n \in G) \geq - \inf_{x \in G\cap\mathscr{F}} \Lambda^*(x), \quad \text{for any open set } G, \tag{18}$$

where $\mathscr{F}$ is the set of exposed points of Λ^* whose exposing hyperplane belongs to $(-\infty, \theta_0)$, i.e.,

$$\mathscr{F} = \left\{y \in \mathbb{R} : \exists\, \theta \in \mathscr{D}_S^\circ \text{ such that for all } x \neq y, \theta y - \Lambda^*(y) > \theta x - \Lambda^*(x)\right\}.$$

We now prove that $\mathscr{F} = (0, +\infty)$. For $y < 0$, $\Lambda^*(y) = \infty$, therefore an exposing hyperplane satisfying the corresponding inequality does not exist. For $y > 0$ consider the exposing hyperplane $\theta = \theta_y$, where θ_y is the unique positive solution on $(-\infty, \theta_0)$ of $\mathrm{E}[S\mathrm{e}^{\theta S}] = y/\nu$. Note that $\Lambda'(\theta) = \mathrm{E}[S\mathrm{e}^{\theta S}]$ and $\Lambda''(\theta) = \mathrm{E}[S^2\mathrm{e}^{\theta S}]$ for all $\theta < \theta_0$. In particular, since $S \geq 1$, we have that Λ is strictly convex on $(-\infty, \theta_0)$. Therefore, for all $x \neq y$, it follows

$$\theta_y y - \Lambda^*(y) = \Lambda(\theta_y) > \Lambda(\theta_x) + \Lambda'(\theta_x)(\theta_y - \theta_x) = \theta_y x - \Lambda^*(x).$$

It remains to check that $0 \notin \mathcal{F}$. Notice that since $\mathrm{E}[Se^{\theta_x S}] = x/v$, $\lim_{x\downarrow 0} \theta_x = -\infty$. Also, by the implicit function theorem, $x \mapsto \theta_x$ is a continuous mapping on $(0, \infty)$. Now assume that $0 \in \mathcal{F}$, then there would exist $\theta < \theta_0$, such that for all $x > 0$, $-\Lambda^*(0) > \theta x - \Lambda^*(x)$. However, by the intermediate values theorem, there exists $y > 0$ such that $\theta = \theta_y$, and we obtain a contradiction. This implies $\mathcal{F} = (0, +\infty)$ as claimed. Now recall that $\Lambda^*(x) = +\infty$ for $x < 0$; moreover, $\lim_{x\downarrow 0} \Lambda^*(x) = \Lambda^*(0) = v$ (indeed, $\lim_{x\downarrow 0} \theta_x = -\infty$). Therefore,

$$\inf_{x \in G \cap \mathcal{F}} \Lambda^*(x) \leq \inf_{x \in G} \Lambda^*(x), \quad \text{for any open set } G.$$

Finally, by (18) and the above inequality we obtain (16).

3.2. Scalar Large Deviations of Hawkes Processes

Consider the ergodic Hawkes process $\mathbf{X}$ described before. In this section we prove that the process $\{N_{\mathbf{X}}(0, t]/t\}$ satisfies a LDP, and we give the explicit expression of the rate function. Our result is a refinement of the law of large numbers (4). The following theorem holds:

Theorem 3.2.1. *Assume* (2) *and*

$$\int_0^\infty th(t)\mathrm{d}t < \infty. \tag{19}$$

Then $\{N_{\mathbf{X}}(0, t]/t\}$ *satisfies a LDP on* $\mathbb{R}$ *with speed* t *and good rate function*

$$\Lambda^*(x) = \begin{cases} x\theta_x + v - \dfrac{vx}{v + \mu x} & \textit{if } x \in (0, \infty) \\ v & \textit{if } x = 0 \\ +\infty & \textit{if } x \in (-\infty, 0) \end{cases}, \tag{20}$$

where $\theta = \theta_x$ *is the unique solution in* $(-\infty, \mu - 1 - \log \mu)$ *of*

$$\mathrm{E}[Se^{\theta S}] = x/v, \quad x > 0, \tag{21}$$

or equivalently of

$$\mathrm{E}[e^{\theta S}] = \frac{x}{v + x\mu}, \quad x > 0.$$

Proof. Since a Hawkes process is a Poisson cluster process, the proof is a consequence of Theorem 3.1.1. We need to check assumptions (5) and (6).

We start noticing that by (3) we have

$$\mathrm{E}[\mathrm{e}^{\theta S}] = \sum_{k\geq 1} \frac{(\mathrm{e}^{\theta-\mu})^k (k\mu)^{k-1}}{k!},$$

and this sum is infinity for $\theta > \mu - 1 - \log\mu$ and finite for $\theta < \mu - 1 - \log\mu$ (apply, for instance, the ratio criterion). If $\theta = \mu - 1 - \log\mu$ the sum above is finite. Indeed, in this case

$$\mathrm{E}[\mathrm{e}^{\theta S}] = (1/\mu) \sum_{k\geq 1} \frac{\mathrm{e}^{-k} k^{k-1}}{k!} = 1/\mu.$$

Therefore, $\mathscr{D}_S = (-\infty, \mu - 1 - \log\mu]$. The origin belongs to $\mathscr{D}_S^\circ$ in that by (2) and the inequality $\mathrm{e}^x > x + 1$, $x \neq 0$, we have $\frac{\mathrm{e}^{\mu-1}}{\mu} > 1$. The function $\theta \mapsto \mathrm{E}[\mathrm{e}^{\theta S}]$ is essentially smooth. Indeed, it is differentiable in the interior of $\mathscr{D}_S$ and

$$\mathrm{E}\big[S\mathrm{e}^{(\mu-1-\log\mu)S}\big] = \infty$$

because

$$\mathrm{E}\big[S\mathrm{e}^{(\mu-1-\log\mu)S}\big] = (1/\mu) \sum_{k\geq 1} \frac{\mathrm{e}^{-k} k^{k}}{k!}$$

and this sum is infinity since by Stirling's formula $\frac{\mathrm{e}^{-k}k^k}{k!} \sim 1/\sqrt{2\pi k}$. We now check assumption (6). By the structure of the clusters, it follows that there exists a sequence of independent non-negative random variables $\{V_n\}_{n\geq 1}$, independent of S, such that V_1 has probability density $h(\cdot)/\mu$ and the following stochastic domination holds:

$$L \leq \sum_{n=1}^{S} V_n, \quad \text{a.s.}$$

(see Reynaud-Bouret and Roy[29]). Therefore, for all $\theta < \mu - 1 - \log\mu$, we have

$$\mathrm{E}[L\mathrm{e}^{\theta S}] \leq \mathrm{E}\left[\mathrm{e}^{\theta S} \sum_{n=1}^{S} V_n\right] = \mathrm{E}[V_1]\mathrm{E}[S\mathrm{e}^{\theta S}].$$

Since $\theta < \mu - 1 - \log\mu$, we have $\mathrm{E}[S\mathrm{e}^{\theta S}] < \infty$; moreover, assumption (19) yields

$$\mathrm{E}[V_1] = \frac{1}{\mu} \int_0^\infty t h(t)\mathrm{d}t < \infty.$$

Hence, condition (6) holds, and by Theorem 3.1.1, $\{N_{\mathbf{X}}(0,t]/t\}$ satisfies a LDP on $\mathbb{R}$ with speed t and good rate function

$$\Lambda^*(x) = \sup_{\theta\in\mathbb{R}}(\theta x - \Lambda(\theta)) = \sup_{\theta\leq\mu-1-\log\mu}(\theta x - \Lambda(\theta)).$$

Now $\Lambda^*(x) = \infty$ if $x < 0$, in that in such a case $\lim_{\theta\to-\infty}(\theta x - \Lambda(\theta)) = \infty$. If $x > 0$, letting $\theta_x \in (-\infty, \mu - 1 - \log\mu)$ denote the unique solution of the equation (21) easily follows that

$$\Lambda^*(x) = x\theta_x - \Lambda(\theta_x). \tag{22}$$

It is well-known (see, for instance, p. 39 in Jagers[15]) that, for all $\theta \in (-\infty, \mu - 1 - \log\mu)$, $\mathrm{E}[\mathrm{e}^{\theta S}]$ satisfies

$$\mathrm{E}[\mathrm{e}^{\theta S}] = \mathrm{e}^{\theta}\exp\big\{\mu(\mathrm{E}[\mathrm{e}^{\theta S}] - 1)\big\},$$

therefore differentiating with respect to θ we get

$$\mathrm{E}[S\mathrm{e}^{\theta S}] = \frac{\mathrm{e}^{\theta}\exp\big\{\mu(\mathrm{E}[\mathrm{e}^{\theta S}] - 1)\big\}}{1 - \mu\mathrm{e}^{\theta}\exp\big\{\mu(\mathrm{E}[\mathrm{e}^{\theta S}] - 1)\big\}} = \frac{\mathrm{E}[\mathrm{e}^{\theta S}]}{1 - \mu\mathrm{E}[\mathrm{e}^{\theta S}]}. \tag{23}$$

Setting $\theta = \theta_x$ in the above equality and using (21) we have

$$\frac{x}{v} = \frac{\mathrm{E}[\mathrm{e}^{\theta_x S}]}{1 - \mu\mathrm{E}[\mathrm{e}^{\theta_x S}]},$$

which yields

$$\mathrm{E}\big[\mathrm{e}^{\theta_x S}\big] = \frac{x}{v + x\mu}.$$

Thus, by (22) we have for $x > 0$

$$\Lambda^*(x) = x\theta_x + v - \frac{vx}{v + \mu x}.$$

The conclusion follows noticing that a direct computation gives $\Lambda^*(0) = v$.

4. SAMPLE PATH LARGE DEVIATIONS

Let $\mathbf{X}$ be the ergodic Poisson cluster process described at the beginning. The results proved in this section are sample path LDP for $\mathbf{X}$.

4.1. Sample Path Large Deviations in the Topology of Point-Wise Convergence

Let $D[0,1]$ be the space of càdlàg functions on the interval $[0,1]$. Here we prove that $\{\frac{N_{\mathbf{X}}(0,\alpha\cdot]}{\alpha}\}$ satisfies a LDP on $D[0,1]$ equipped with the topology of point-wise convergence on $D[0,1]$. The LDP we give is a refinement of the following functional law of large numbers:

$$\lim_{\alpha\to\infty}\frac{N_{\mathbf{X}}(0,\alpha\cdot]}{\alpha}=\chi(\cdot)\quad \text{a.s.,} \tag{24}$$

where $\chi(t)=\nu\mathrm{E}[S]t$. As this is a corollary of the LDP we establish, we do not include a separate proof of this result. Letting $\Lambda^*(\cdot)$ denote the rate function of the scalar LDP, we have:

Theorem 4.1.1. *Assume* (5) *and* (6). *If moreover* $\mathscr{D}_S$ *is open, then* $\{\frac{N_{\mathbf{X}}(0,\alpha\cdot]}{\alpha}\}$ *satisfies a LDP on* $D[0,1]$, *equipped with the topology of point-wise convergence, with speed* α *and good rate function*

$$J(f)=\begin{cases}\int_0^1 \Lambda^*(\dot{f}(t))\,dt & \mathit{if}\ f\in AC_0[0,1]\\ \infty & \mathit{otherwise}\end{cases}, \tag{25}$$

where $AC_0[0,1]$ *is the family of absolutely continuous functions* $f(\cdot)$ *defined on* $[0,1]$, *with* $f(0)=0$.

While it is tempting to conjecture that the result above holds even if the effective domain of S is not open, we do not have a proof of this claim. If we take $\chi(t)=\nu\mathrm{E}[S]t$, then $J(\chi)=0$. Moreover this is the unique zero of $J(\cdot)$. Thus, the law of $N_{\mathbf{X}}(0,\alpha\cdot]/\alpha$ concentrates in arbitrarily small neighborhoods of $\chi(\cdot)$ as $\alpha\to\infty$, as ensured by the functional law of large numbers (24). The sample path LDP is a refinement of the functional law of large numbers in that it gives the probability of fluctuations away the most likely path.

As in Section 3.1, denote by $\mathbf{X}_{t,T}$ the non-stationary Poisson cluster process with immigrant process empty on $(-\infty,-T)\cup(t+T,\infty)$, where $T>0$ is a fixed constant. Before proving Theorem 4.1.1 we show that the same LDP holds for $\mathbf{X}_{t,T}$. Furthermore, the LDP for $\mathbf{X}_{t,T}$ holds under a weaker condition.

Theorem 4.1.2. *Assume* (5). *Then* $\{N_{\mathbf{X}_{\alpha\cdot,T}}(0,\alpha\cdot]/\alpha\}$ *satisfies a LDP on* $D[0,1]$, *equipped with the topology of point-wise convergence, with speed* α *and good rate function* (25).

To prove this theorem we need Lemma 4.1.1 below, whose proof can be found in Ganesh et al.[10] (see Lemma 2.3 therein).

Lemma 4.1.1. *Let $(\theta_1, \ldots, \theta_n) \in \mathbb{R}^n$ and let $w_1, \ldots, w_n \geq 0$ be such that $w_1 \leq \cdots \leq w_n$. Then $\sum_{i=k}^n \theta_i w_i \leq \theta^* w^*$ for all $k \in \{1, \ldots, n\}$, for any $\theta^* \geq \max\big\{\max\{\sum_{i=k}^n \theta_i : k \in \{1, \ldots, n\}\}, 0\big\}$ and any $w^* \geq w_n$.*

Proof of Theorem 4.1.2. With a little abuse of notation denote by C_0 the cluster generated by an immigrant at 0. We first show the theorem under the additional condition

$$N_{C_0}((-\infty, 0)) = 0, \quad \text{a.s.} \tag{26}$$

The idea in proving Theorem 4.1.2 is to apply the Dawson-Gärtner theorem to "lift" a LDP for the finite-dimensional distributions of $\{N_{\mathbf{X}_{\alpha t,T}}(0, \alpha t]/\alpha\}$ to a LDP for the process. Therefore, we first show the following claim:

(**C**) For all $n \geq 1$ and $0 \leq t_1 < \cdots < t_n \leq 1$, $\big(N_{\mathbf{X}_{\alpha t_1,T}}(0, \alpha t_1]/\alpha, \ldots, N_{\mathbf{X}_{\alpha t_n,T}}(0, \alpha t_n]/\alpha\big)$ satisfies the LDP in $\mathbb{R}^n$ with speed α and good rate function

$$J_{t_1,\ldots,t_n}(x_1, \ldots, x_n) = \sum_{j=1}^n (t_j - t_{j-1}) \Lambda^*\left(\frac{x_j - x_{j-1}}{t_j - t_{j-1}}\right), \tag{27}$$

where $x_0 = 0$ and $t_0 = 0$.

Claim (**C**) is a consequence of the Gärtner-Ellis theorem in $\mathbb{R}^n$, and will be shown in three steps:

(a) For each $(\theta_1, \ldots, \theta_n) \in \mathbb{R}^n$, we prove that

$$\begin{aligned}\Lambda_{t_1,\ldots,t_n}(\theta_1, \ldots, \theta_n) &\equiv \lim_{\alpha \to \infty} \frac{1}{\alpha} \log \mathrm{E}\left[\exp\left(\sum_{i=1}^n \theta_i N_{\mathbf{X}_{\alpha t_i,T}}(0, \alpha t_i]\right)\right] \\ &= \sum_{j=1}^n (t_j - t_{j-1}) \Lambda\left(\sum_{i=j}^n \theta_i\right),\end{aligned} \tag{28}$$

where the existence of the limit (as an extended real number) is part of the claim, and $\Lambda(\cdot)$ is defined in the statement of Theorem 3.1.1.

(b) The function $\Lambda_{t_1,\ldots,t_n}(\cdot)$ satisfies the hypotheses of the Gärtner-Ellis theorem.

(c) The rate function

$$J_{t_1,\ldots,t_n}(x_1, \ldots, x_n) \equiv \sup_{(\theta_1,\ldots,\theta_n) \in \mathbb{R}^n} \left[\sum_{i=1}^n \theta_i x_i - \Lambda_{t_1,\ldots,t_n}(\theta_1, \ldots, \theta_n)\right]$$

coincides with the rate function defined in (27).

Proof of (a). For a Borel set $A \subset \mathbb{R}$, denote by $I_{|A} = I \cap A$ the Poisson process of immigrants in A. Since, for each t, $I_{|(0,t]}$ and $I_{|[-T,0]}$ are independent, it follows from the definition of Poisson cluster process that, for each i, the random sets $\{C_k : X_k \in I_{|(0,\alpha t_i]}\}$ and $\{C_k : X_k \in I_{|[-T,0]}\}$ are independent. Therefore,

$$\mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i N_{\mathbf{X}_{\alpha t_i,T}}(0,\alpha t_i]\right] = \mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i \sum_{X_k\in I_{|(0,\alpha t_i]}} N_{C_k}(0,\alpha t_i]\right] \times \mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i \sum_{X_k\in I_{|[-T,0]}} N_{C_k}(0,\alpha t_i]\right], \quad (29)$$

where we used the independence and the assumption that $N_{C_0}(-\infty,0) = 0$ a.s. In order to prove (28), we treat successively the two terms in (29). Viewing $I_{|(0,\alpha t_i]}$ as the superposition of the i independent Poisson processes: $I_{|(\alpha t_{j-1},\alpha t_j]}$ on $(\alpha t_{j-1}, \alpha t_j]$ $(j = 1, \ldots, i)$ with intensity ν we get

$$\begin{aligned}\mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i \sum_{X_k\in I_{|(0,\alpha t_i]}} N_{C_k}(0,\alpha t_i]\right] &= \mathrm{E}\left[\exp\sum_{i=1}^{n}\sum_{j=1}^{i}\sum_{X_k\in I_{|(\alpha t_{j-1},\alpha t_j]}}\theta_i N_{C_k}(0,\alpha t_i]\right]\\ &= \mathrm{E}\left[\exp\sum_{j=1}^{n}\sum_{i=j}^{n}\sum_{X_k\in I_{|(\alpha t_{j-1},\alpha t_j]}}\theta_i N_{C_k}(0,\alpha t_i]\right]\\ &= \prod_{j=1}^{n}\mathrm{E}\left[\exp\sum_{i=j}^{n}\sum_{X_k\in I_{|(\alpha t_{j-1},\alpha t_j]}}\theta_i N_{C_k}(0,\alpha t_i]\right],\end{aligned} \quad (30)$$

where in the latter equality we used the independence of $\{C_k : X_k \in I_{|(\alpha t_{j-1},\alpha t_j]}\}$ $(j = 1, \ldots, n)$. Since, for each j, $\{(X_k, C_k) : X_k \in I_{|(\alpha t_{j-1},\alpha t_j]}\}$ is an independently marked Poisson process, by Lemma 6.4.VI in Daley and Vere-Jones[7] we have

$$\begin{aligned}&\mathrm{E}\left[\exp\sum_{i=j}^{n}\sum_{X_k\in I_{|(\alpha t_{j-1},\alpha t_j]}}\theta_i N_{C_k}(0,\alpha t_i]\right]\\ &= \exp\left(\nu\int_0^{\alpha(t_j-t_{j-1})}\mathrm{E}\left[\exp\left(\sum_{i=j}^{n}\theta_i N_{C_0}(-\alpha t_{j-1}-s, \alpha(t_i-t_{j-1})-s]\right)-1\right]\mathrm{d}s\right).\end{aligned} \quad (31)$$

We now show

$$\lim_{\alpha\to\infty}\frac{1}{\alpha}\log \mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{|(0,\alpha t_i]}}N_{C_k}(0,\alpha t_i]\right]=\sum_{j=1}^{n}(t_j-t_{j-1})\Lambda\left(\sum_{i=j}^{n}\theta_i\right) \quad (32)$$

for each $(\theta_1,\ldots,\theta_n)\in\mathbb{R}^n$. We first notice that by (30) and (31) we have

$$\frac{1}{\alpha}\log \mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{|(0,\alpha t_i]}}N_{C_k}(0,\alpha t_i]\right]=\sum_{j=1}^{n}(t_j-t_{j-1})J_j(\alpha),$$

where

$$J_j(\alpha)=\frac{v}{\alpha(t_j-t_{j-1})}\times\int_0^{\alpha(t_j-t_{j-1})}E\left[\exp\sum_{i=j}^{n}\theta_i N_{C_0}\big((-\alpha t_{j-1}-s,\alpha(t_i-t_{j-1})-s]\big)-1\right]\mathrm{d}s. \quad (33)$$

Now suppose that $(\theta_1,\ldots,\theta_n)\in\mathbb{R}^n$ is such that $\sum_{i=j}^{n}\theta_i\in\mathcal{D}_S$ for each $j\in\{1,\ldots,n\}$. Then by Lemma 4.1.1 it follows that there exists $\theta^*\in\mathcal{D}_S$ such that $\theta^*\geq 0$, $\sum_{i=j}^{n}\theta_i\leq\theta^*$ for all $j\in\{1,\ldots,n\}$, and

$$\sum_{i=j}^{n}\theta_i N_{C_0}(-\alpha t_{j-1}-s,\alpha(t_i-t_{j-1})-s]\leq\theta^* N_{C_0}(\mathbb{R}),\quad \text{a.s.}$$

By (33) and the dominated convergence theorem, we have

$$\lim_{\alpha\to\infty}J_j(\alpha)=v\big(\mathrm{E}\big[\mathrm{e}^{\sum_{i=j}^{n}\theta_i S}\big]-1\big).$$

Hence we proved (32) whenever $(\theta_1,\ldots,\theta_n)\in\mathbb{R}^n$ satisfies $\sum_{i=j}^{n}\theta_i\in\mathcal{D}_S$ for every $j\in\{1,\ldots,n\}$. Now suppose that $(\theta_1,\ldots,\theta_n)\in\mathbb{R}^n$ is such that $\sum_{i=j}^{n}\theta_i\notin\mathcal{D}_S$ for some $j\in\{1,\ldots,n\}$. We have that $J_j(\alpha)$ is bigger than or equal to

$$\frac{v}{\alpha(t_j-t_{j-1})}\int_0^{\alpha(t_j-t_{j-1})}\mathrm{E}\left[\exp\left(\sum_{i=j}^{n}\mathbf{1}\{\theta_i<0\}\theta_i S+\sum_{i=j}^{n}\mathbf{1}\{\theta_i>0\}\theta_i N_{C_0}[0,\alpha(t_j-t_{j-1})-s]\right)-1\right]\mathrm{d}s$$

$$= v\int_0^1 \mathrm{E}\Bigg[\exp\Bigg(\sum_{i=j}^{n}\mathbf{1}\{\theta_i < 0\}\theta_i S + \sum_{i=j}^{n}\mathbf{1}\{\theta_i > 0\}\theta_i N_{C_0}[0,\alpha(t_j - t_{j-1})(1-z)]\Bigg) - 1\Bigg]\mathrm{d}z.$$

The expectation in the latter formula goes to $\mathrm{E}\big[\exp\big(\sum_{i=j}^{n}\theta_i S\big) - 1\big]$ as $\alpha \to \infty$ by the monotone convergence theorem. Therefore, by Fatou's lemma we have

$$\lim_{\alpha\to\infty} J_j(\alpha) \geq v\mathrm{E}\Bigg[\exp\Bigg(\sum_{i=j}^{n}\theta_i S\Bigg) - 1\Bigg] = \infty.$$

Thus, since the quantities $J_1(\alpha),\ldots,J_n(\alpha)$ are bounded below by $-v$, we get (32) also in this case. We now show

$$\lim_{\alpha\to\infty}\frac{1}{\alpha}\log\mathrm{E}\Bigg[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{[-T,0]}} N_{C_k}(0,\alpha t_i]\Bigg] = 0 \tag{34}$$

for all $(\theta_1,\ldots,\theta_n)\in\mathbb{R}^n$ such that $\sum_{i=j}^{n}\theta_i \in \mathscr{D}_S$ for each $j\in\{1,\ldots,n\}$. By Lemma 4.1.1 we have that there exists $\theta^*\in\mathscr{D}_S$ such that $\theta^*\geq 0$, $\sum_{i=j}^{n}\theta_i \leq \theta^*$ for all $j\in\{1,\ldots,n\}$ and

$$\theta_- \sum_{X_k\in I_{[-T,0]}} N_{C_k}(\mathbb{R}) \leq \sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{[-T,0]}} N_{C_k}(0,\alpha t_i] \leq \theta^* \sum_{X_k\in I_{[-T,0]}} N_{C_k}(\mathbb{R}), \quad \text{a.s.},$$

where $\theta_- \equiv \sum_{i:\theta_i<0}\theta_i$ and $\theta_- \equiv 0$ if $\{i:\theta_i<0\}=\emptyset$. Therefore, using again Lemma 6.4 VI in Daley and Vere-Jones[7], we have

$$\exp\big(vT(\mathrm{E}[\mathrm{e}^{\theta_- S}]-1)\big) \leq \mathrm{E}\Bigg[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{[-T,0]}} N_{C_k}(0,\alpha t_i]\Bigg] \leq \exp\big(vT(\mathrm{E}[\mathrm{e}^{\theta^* S}]-1)\big).$$

Equation (34) follows taking the logarithms in the above inequalities and passing to the limit. The conclusion follows putting together (29), (32) and (34).

Proof of (b) *and Proof of* (c). Part (b) can be shown using assumption (5) and following the lines of the proof of part (b) of Proposition 2.2 in Ganesh et al.[10]. The proof of part (c) is identical to the proof of part (c) of Proposition 2.2 in Ganesh et al.[10].

End of the proof under condition (26). By claim (**C**) and the Dawson-Gärtner theorem, $\{N_{\mathbf{X}_{\alpha\cdot,T}}(0,\alpha\cdot]/\alpha\}$ satisfies the LDP on $D[0,1]$, equipped with the topology of point-wise convergence, with speed α and good rate function

$$\widetilde{J}(f) = \sup\left\{\sum_{k=1}^{n}(t_k - t_{k-1})\Lambda^*\left(\frac{f(t_k) - f(t_{k-1})}{t_k - t_{k-1}}\right) : n \geq 1, 0 \leq t_1 < \cdots < t_n \leq 1\right\}.$$

The conclusion follows noticing that $\widetilde{J}(\cdot)$ coincides with $J(\cdot)$ in (25), as can be checked following the same lines as in the proof of Lemma 5.1.6 in Dembo and Zeitouni[9].

Removing the additional condition (26). The general case is solved as follows. Since C_k is almost surely finite, there exists a left-most extremal point $Y_k \in C_k$ such that $N_{C_k}(-\infty, Y_k) = 0$ a.s. Note that, given the immigrants, $Y_k - X_k$ is an iid sequence. Therefore, by a classical result on Poisson processes we have that $\{Y_k\}$ is a Poisson process with intensity ν. Viewing $\mathbf{X}_{t,T}$ as a Poisson cluster process with cluster centers Y_k and clusters C_k, the conclusion follows by the first part of the proof.

Proof of Theorem 4.1.1. The proof uses similar steps as in the proof of Theorem 4.1.2. Here we sketch the main difference. Assume the additional condition $N_{C_0}((-\infty, 0)) = 0$ a.s. (the general case can be treated as in the proof of Theorem 4.1.2). Define the following subsets of $\mathbb{R}^n$:

$$A_1 \equiv \left\{(\theta_1, \ldots, \theta_n) \in \mathbb{R}^n : \sum_{i=j}^{n} \theta_i \in \mathcal{D}_S \text{ for all } j \in \{1, \ldots, n\}\right\}$$

and

$$A_2 \equiv \left\{(\theta_1, \ldots, \theta_n) \in \mathbb{R}^n : \sum_{i=j}^{n} \theta_i \notin \mathcal{D}_S \text{ for some } j \in \{1, \ldots, n\}\right\}$$

We start showing that for all $n \geq 1$ and $0 \leq t_1 < \cdots < t_n \leq 1$

$$\Lambda_{t_1,\ldots,t_n}(\theta_1, \ldots, \theta_n) = \begin{cases} \sum_{j=1}^{n}(t_j - t_{j-1})\Lambda\left(\sum_{i=j}^{n}\theta_i\right) & \text{for } (\theta_1, \ldots, \theta_n) \in A_1 \\ +\infty & \text{for } (\theta_1, \ldots, \theta_n) \in A_2, \end{cases} \tag{35}$$

where

$$\Lambda_{t_1,\ldots,t_n}(\theta_1, \ldots, \theta_n) \equiv \lim_{\alpha\to\infty} \frac{1}{\alpha} \log \mathrm{E}\left[\exp\left(\sum_{i=1}^{n} \theta_i N_{\mathbf{X}}(0, \alpha t_i]\right)\right]$$

and $\Lambda(\cdot)$ is defined in the statement of Theorem 3.1.1. Using the definition of $\mathbf{X}$ and the assumption $N_{C_0}((-\infty,0))=0$ a.s., we have

$$\begin{aligned}
&\mathrm{E}\left[\exp\left(\sum_{i=1}^{n}\theta_i N_{\mathbf{X}}(0,\alpha t_i]\right)\right]\\
&\quad=\mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{|(0,\alpha t_i]}}N_{C_k}(0,\alpha t_i]\right]\mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{|[-T,0]}}N_{C_k}(0,\alpha t_i]\right]\\
&\qquad\times\mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{|(-\infty,-T)}}N_{C_k}(0,\alpha t_i]\right].
\end{aligned}$$

As noticed in the proof of Theorem 4.1.2 we have

$$\begin{aligned}
&\mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i N_{\mathbf{X}_{\alpha t_i,T}}(0,\alpha t_i]\right]\\
&\quad=\mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{|(0,\alpha t_i]}}N_{C_k}(0,\alpha t_i]\right]\mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{|[-T,0]}}N_{C_k}(0,\alpha t_i]\right].
\end{aligned}$$

Therefore, by the computations in the proof of Theorem 4.1.2, to prove (35) we only need to check that

$$\lim_{\alpha\to\infty}\frac{1}{\alpha}\log\mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{|(-\infty,-T)}}N_{C_k}(0,\alpha t_i]\right]=0,\quad\text{for all }(\theta_1,\dots,\theta_n)\in A_1. \tag{36}$$

Since $\{(X_i,C_i): X_i\in I_{|(-\infty,-T)}\}$ is an independently marked Poisson process, by Lemma 6.4. VI in Daley and Vere-Jones[7] we have

$$\begin{aligned}
&\mathrm{E}\left[\exp\sum_{i=1}^{n}\theta_i\sum_{X_k\in I_{|(-\infty,-T)}}N_{C_k}(0,\alpha t_i]\right]\\
&\quad=\mathrm{E}\left[\exp\sum_{X_k\in I_{|(-\infty,-T)}}\sum_{i=1}^{n}\theta_i N_{C_k-X_k}(-X_k,\alpha t_i-X_k]\right]\\
&\quad=\exp\left(\nu\int_T^{\infty}\mathrm{E}\left[e^{\sum_{i=1}^{n}\theta_i N_{C_0}(x,\alpha t_i+x]}-1\right]\mathrm{d}x\right)
\end{aligned}$$

Take $(\theta_1,\dots,\theta_n)\in A_1$. By Lemma 4.1.1 we have that there exists $\theta^*\in\mathcal{D}_S$ such that $\theta^*\geq 0$, $\sum_{i=j}^{n}\theta_i\leq\theta^*$ for all $j\in\{1,\dots,n\}$ and

$$\theta_- N_{C_0}(\mathbb{R})\leq\sum_{i=1}^{n}\theta_i N_{C_0}(x,\alpha t_i+x]\leq\theta^* N_{C_0}(\mathbb{R}),\quad\text{a.s.}$$

where $\theta_- \equiv \sum_{i:\theta_i<0} \theta_i$ and $\theta_- \equiv 0$ if $\{i : \theta_i < 0\} = \emptyset$. Thus,

$$e^{\sum_{i=1}^{n} \theta_i N_{C_0}(x, \alpha t_i + x]} - 1 \le \left(e^{\theta^* N_{C_0}(\mathbb{R})} - 1\right)\mathbf{1}\{x \le L\}, \quad \text{for all } x \ge T$$

and

$$e^{\sum_{i=1}^{n} \theta_i N_{C_0}(x, \alpha t_i + x]} - 1 \ge \left(e^{\theta_- N_{C_0}(\mathbb{R})} - 1\right)\mathbf{1}\{x \le L\}, \quad \text{for all } x \ge T$$

The limit (36) follows by assumption (6) noticing that the above relations yield, for all $(\theta_1, \dots, \theta_n) \in A_1$:

$$\mathrm{E}\left[\exp \sum_{i=1}^{n} \theta_i \sum_{X_k \in I_{|(-\infty,-T)}} N_{C_k}(0, \alpha t_i]\right] \le \exp\left(\nu \mathrm{E}\left[L\left(e^{\theta^* S} - 1\right)\right]\right)$$

and

$$\mathrm{E}\left[\exp \sum_{i=1}^{n} \theta_i \sum_{X_k \in I_{|(-\infty,-T)}} N_{C_k}(0, \alpha t_i]\right] \ge \exp\left(\nu \mathrm{E}\left[\left(e^{\theta_- S} - 1\right)(L - T)\mathbf{1}\{L \ge T\}\right]\right).$$

Now since $\mathscr{D}_S$ is open, the claim follows by applying first the Gärtner-Ellis theorem in $\mathbb{R}^n$ to get the LDP for the finite-dimensional distributions, and then the Dawson-Gärtner theorem to have the LDP for the process (argue as in the proof of Theorem 4.1.2 for the remaining steps).

4.2. Sample Path Large Deviations in the Topology of Uniform Convergence

In the applications, one usually derives LDPs for continuous functions of sample paths of stochastic processes by using the contraction principle. Since the topology of uniform convergence is finer than the topology of point-wise convergence, it has a larger class of continuous functions. Thus, it is of interest to understand if $\{N_{\mathbf{X}}(0, \alpha\cdot]/\alpha\}$ satisfies a LDP on $D[0,1]$ equipped with the topology of uniform convergence. In this section we give an answer to this question assuming that the tails of S decay super-exponentially. We do not know the answer when the distribution of S is light-tailed. However, we notice that if there exists $\theta_0 \in (0, \infty)$ such that $\mathrm{E}\left[e^{\theta S}\right] < \infty$ for $\theta < \theta_0$ and $\lim_{\theta \to \theta_0^-} \mathrm{E}\left[S e^{\theta S}\right] = \infty$ (as it happens for Hawkes processes) the same argument used in Section 4 of Ganesh et al.[10] shows that the rate function $J(\cdot)$ is not good. This means that, even if Theorem 4.1.1 holds with respect to the topology of uniform convergence, the contraction principle is not applicable, as it requires goodness of the rate function.

Theorem 4.2.1. *Assume*

$$E\left[e^{\theta S}\right] < \infty \quad \textit{for each } \theta \in \mathbb{R} \tag{37}$$

and

$$E\left[Le^{\theta S}\right] < \infty \quad \textit{for each } \theta \in \mathbb{R}. \tag{38}$$

Then $\{\frac{N_{\mathbf{X}}(0,\alpha\cdot]}{\alpha}\}$ *satisfies a LDP on* $D[0,1]$, *equipped with the topology of uniform convergence, with speed* α *and good rate function* (25).

In this section, without loss of generality we assume that the points of I are $\{X_i\}_{i\in\mathbb{Z}^*}$, where $\mathbb{Z}^* = \mathbb{Z}\backslash\{0\}$, $X_i < X_{i+1}$, and we set $X_0 = 0$. As usual, we denote by $\mathbf{X}_{t,T}$ the non-stationary Poisson cluster process with immigrant process empty on $(-\infty, -T) \cup (t+T, \infty)$, where $T > 0$ is a fixed constant, and by C_0 the cluster generated by an immigrant at 0.

Before proving Theorem 4.2.1 we show that the same LDP holds for $\mathbf{X}_{t,T}$, under a weaker condition.

Theorem 4.2.2. *Assume* (37). *Then* $\{N_{\mathbf{X}_{\alpha\cdot,T}}(0,\alpha\cdot]/\alpha\}$ *satisfies a LDP on* $D[0,1]$, *equipped with the topology of uniform convergence, with speed* α *and good rate function* (25).

To prove Theorem 4.2.2 above we use the following Lemma 4.2.1, whose proof is omitted since it is similar to the proof of Lemma 3.3 in Ganesh et al.[10]. Let $\{S_k\}_{k\in\mathbb{Z}}$ be the iid sequence of random variables (distributed as S) defined by $S_k = N_{C_k}(\mathbb{R})$.

Lemma 4.2.1. *Assume* (37), $N_{C_0}(-\infty, 0) = 0$ *a.s., and define*

$$A_n = \sum_{k=0}^{n-1}(S_k - N_{C_k - X_k}(0, X_k]), \quad n \geq 1.$$

It holds

$$\lim_{n\to\infty} \frac{1}{n} \log P(A_n \geq n\delta) = -\infty \quad \textit{for each } \delta > 0.$$

Proof of Theorem 4.2.2. We prove the theorem assuming that $N_{C_0}(-\infty, 0) = 0$ a.s. The general case is solved as in the proof of Theorem 4.1.2. As usual, denote by $I_{|A}$ the restriction of I on the Borel set $A \subset \mathbb{R}$. Define

$$C(t) = \sum_{X_k \in I_{|(0,t]}} S_k, \quad t > 0.$$

We prove that $\{N_{\mathbf{X}_{\alpha\cdot,T}}(0,\alpha\cdot]/\alpha\}$ and $\{C(\alpha\cdot)/\alpha\}$ are exponentially equivalent (see, for instance, Definition 4.2.10 in the book of Dembo and Zeitouni[9]) with respect to the topology of uniform convergence. Therefore the conclusion follows by a well-known result on sample path large deviations, with respect to the uniform topology, of compound Poisson processes (see, for instance, Borovkov[1]; see also de Acosta[8] and the references cited therein) and Theorem 4.2.13 in Dembo and Zeitouni[9]. Define

$$C_T(t) = \sum_{X_k \in I_{|[-T,t]}} S_k, \quad t > 0.$$

Using Chernoff bound and condition (37) can be easily realized that the processes $\{C(\alpha\cdot)/\alpha\}$ and $\{C_T(\alpha\cdot)/\alpha\}$ are exponentially equivalent with respect to the topology of uniform convergence. Therefore, it suffices to show that $\{C_T(\alpha\cdot)/\alpha\}$ and $\{N_{\mathbf{X}_{\alpha\cdot,T}}(0,\alpha\cdot]/\alpha\}$ are exponentially equivalent with respect to the topology of uniform convergence. Note that the assumption $N_{C_0}(-\infty,0) = 0$ a.s. gives

$$N_{\mathbf{X}_{t,T}}(0,t] = \sum_{X_k \in I_{|(0,t]}} N_{C_k}(0,t] + \sum_{X_k \in I_{|[-T,0]}} N_{C_k}(0,t] \quad t > 0, \text{ a.s.}$$

Therefore, we need to show that

$$\lim_{\alpha\to\infty} \frac{1}{\alpha} \log P(M_\alpha > \delta) = -\infty, \quad \text{for any } \delta > 0, \tag{39}$$

where

$$M_\alpha = \frac{1}{\alpha} \sup_{t\in[0,1]} \left| C_T(\alpha t) - \sum_{X_k \in I_{|(0,\alpha t]}} N_{C_k}(0,\alpha t] - \sum_{X_k \in I_{|[-T,0]}} N_{C_k}(0,\alpha t] \right|.$$

Since $S_k \geq N_{C_k}(0,\alpha t]$, we have:

$$\begin{aligned} &\left| C_T(\alpha t) - \sum_{X_k \in I_{|(0,\alpha t]}} N_{C_k}(0,\alpha t] - \sum_{X_k \in I_{|[-T,0]}} N_{C_k}(0,\alpha t] \right| \\ &\quad = \sum_{X_k \in I_{|[-T,0]}} (S_k - N_{C_k}(0,\alpha t]) + \sum_{X_k \in I_{|(0,\alpha t]}} (S_k - N_{C_k}(0,\alpha t]). \end{aligned}$$

Hence

$$M_\alpha \leq M_\alpha^{(1)} + M_\alpha^{(2)}, \quad \text{a.s.,}$$

where

$$M_\alpha^{(1)} = \frac{1}{\alpha} \sum_{X_k \in I_{|[-T,0]}} S_k \quad \text{and} \quad M_\alpha^{(2)} = \frac{1}{\alpha} \sup_{t\in[0,1]} \sum_{X_k \in I_{|(0,\alpha t]}} (S_k - N_{C_k}(0, \alpha t]),$$

the limit (39) follows if we prove

$$\lim_{\alpha\to\infty} \frac{1}{\alpha} \log P\big(M_\alpha^{(1)} > \delta/2\big) = -\infty, \quad \text{for any } \delta > 0 \tag{40}$$

and

$$\lim_{\alpha\to\infty} \frac{1}{\alpha} \log P\big(M_\alpha^{(2)} > \delta/2\big) = -\infty, \quad \text{for any } \delta > 0. \tag{41}$$

The limit (40) easily follows by the Chernoff bound and condition (37). It remains to show (41). Since the random function $t \mapsto N_{C_k}(0, \alpha t]$ is non-decreasing, it is clear that the supremum over t is attained at one of the points X_n, $n \geq 1$. Thus,

$$M_\alpha^{(2)} = \frac{1}{\alpha} \max_{n\geq 1: X_n \leq \alpha} \sum_{k=1}^{n} (S_k - N_{C_k}(0, X_n]).$$

Note that

$$M_\alpha^{(2)} \leq \widetilde{M}_\alpha \quad \text{where } \widetilde{M}_\alpha = \frac{1}{\alpha} \max_{n\geq 1: X_n \leq \alpha} \sum_{k=1}^{n} (S_k - N_{C_k - X_k}(0, X_n - X_k]) \text{ a.s.}$$

Therefore, (41) follows if we show

$$\lim_{\alpha\to\infty} \frac{1}{\alpha} \log P\big(\widetilde{M}_\alpha > \delta/2\big) = -\infty, \quad \text{for any } \delta > 0. \tag{42}$$

Since $X_n, n \geq 1$, is the sum of n exponential random variables with mean $1/v$, using Chernoff bound and taking the logarithm, we have that, for all $\eta > 0$ and all integers $K > v$,

$$\frac{1}{\alpha} \log P(X_{K[\alpha]} < \alpha) \leq \eta + \frac{K[\alpha]}{\alpha} \log\left(\frac{v}{v+\eta}\right). \tag{43}$$

Here the symbol $[\alpha]$ denotes the integer part of α. Next, observe that using the union bound we get

$$\begin{aligned} &P\big(\widetilde{M}_\alpha > \delta/2, X_{K[\alpha]} \geq \alpha\big) \\ &\quad \leq K[\alpha] \max_{1\leq n\leq K[\alpha]} P\left(\sum_{k=1}^{n} (S_k - N_{C_k - X_k}(0, X_n - X_k]) \geq \alpha\delta/2\right), \end{aligned}$$

Now we remark that, for $n \geq 1$, $(X_n - X_1, \ldots, X_n - X_{n-1})$ and $(X_{n-1}, \ldots, X_1)$ have the same joint distribution. Moreover, given I, the centered processes $C_k - X_k$ are iid and independent of the $\{X_k\}$. Hence, letting A_n denote the random variable defined in the statement of Lemma 4.2.1, we have

$$P\big(\widetilde{M}_\alpha > \delta/2, X_{K[\alpha]} \geq \alpha\big) \leq K[\alpha] \max_{1 \leq n \leq K[\alpha]} P(A_n \geq \alpha\delta/2).$$

The random variables A_n are increasing in n, therefore,

$$P\big(\widetilde{M}_\alpha > \delta, X_{K[\alpha]} \geq \alpha\big) \leq K[\alpha] P(A_{K[\alpha]} \geq \alpha\delta/2),$$

and by Lemma 4.2.1 we have

$$\lim_{\alpha\to\infty} \frac{1}{\alpha} \log P\big(\widetilde{M}_\alpha > \delta/2, X_{K[\alpha]} \geq \alpha\big) = -\infty. \tag{44}$$

Now note that

$$P\big(\widetilde{M}_\alpha > \delta/2\big) \leq P\big(\widetilde{M}_\alpha > \delta/2, X_{K[\alpha]} \geq \alpha\big) + P(X_{K[\alpha]} < \alpha),$$

for arbitrary $K > \nu$. Hence by (43) and(44) we have

$$\limsup_{\alpha\to\infty} \frac{1}{\alpha} \log P\big(\widetilde{M}_\alpha > \delta/2\big) \leq \inf_{\eta>0} \left(\eta + K \log\Big(\frac{\nu}{\nu+\eta}\Big)\right) = K - \nu - K \log \frac{K}{\nu}.$$

Then we obtain (42) by letting K tend to ∞.

Proof of Theorem 4.2.1. Throughout the proof we assume $N_{C_0}((-\infty, 0)) = 0$ a.s. The general case is solved as in the proof of Theorem 4.1.2. Let $\{C_T(t)\}$ be the process defined in the proof of Theorem 4.2.2. The claim follows if we show that $\{C_T(\alpha\cdot)/\alpha\}$ and $\{N_{\mathbf{X}}(0, \alpha\cdot]/\alpha\}$ are exponentially equivalent with respect to the topology of uniform convergence. Note that the assumption $N_{C_0}(-\infty, 0) = 0$ a.s. implies

$$N_{\mathbf{X}}(0, t] = N_{\mathbf{X}_{t,T}}(0, t] + \sum_{X_k \in I_{|(-\infty,-T)}} N_{C_k}(0, t], \quad t > 0 \text{ a.s.}$$

Therefore, since we already proved that $\{C_T(\alpha\cdot)/\alpha\}$ and $\{N_{\mathbf{X}_{\alpha\cdot,T}}(0, \alpha\cdot]/\alpha\}$ are exponentially equivalent with respect to the uniform topology (see the proof of Theorem 4.2.2), the claim follows if we prove that

$$\lim_{\alpha\to\infty} \frac{1}{\alpha} \log P\Bigg(\sum_{X_k \in I_{|(-\infty,-T)}} N_{C_k}(0, \alpha] > \alpha\delta\Bigg) = -\infty, \quad \text{for any } \delta > 0. \tag{45}$$

Using the Chernoff bound we have, for all $\theta > 0$,

$$P\Bigg(\sum_{X_k\in I|(-\infty,-T)} N_{C_k}(0,\alpha] > \alpha\delta\Bigg) \le \mathrm{e}^{-\alpha\theta\delta}\mathrm{E}\Bigg[\exp\sum_{X_k\in I|(-\infty,-T)} \theta N_{C_k}(0,\alpha]\Bigg]$$
$$= \mathrm{e}^{-\alpha\theta\delta}\exp\left(\nu\int_T^\infty \mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(x,\alpha+x]}-1\big]\mathrm{d}x\right)$$
$$\le \mathrm{e}^{-\alpha\theta\delta}\exp\Big(\nu\mathrm{E}\big[\big(\mathrm{e}^{\theta S}-1\big)L\big]\Big).$$

Taking the logarithm, dividing by α, letting α tend to ∞ and using assumption (38) we get

$$\limsup_{\alpha\to\infty}\frac{1}{\alpha}\log P\Bigg(\sum_{X_k\in I|(-\infty,-T)} N_{C_k}(0,\alpha] > \alpha\delta\Bigg) \le -\theta\delta, \quad \text{for all } \theta > 0.$$

Relation (45) follows letting θ tend to infinity in the above inequality.

5. LARGE DEVIATIONS OF SPATIAL POISSON CLUSTER PROCESSES

5.1. The Large Deviations Principle

A spatial Poisson cluster process $\mathbf{X}$ is a Poisson cluster process in $\mathbb{R}^d$, where $d \ge 1$ is an integer. The clusters centers are the points $\{X_i\}$ of a homogeneous Poisson process $I \subset \mathbb{R}^d$ with intensity $\nu \in (0,\infty)$. Each immigrant $X_i \in I$ generates a cluster $C_i = C_{X_i}$, which is a finite point process. Given I, the centered clusters $\{C_{X_i} - X_i\}$ are iid and independent of I. $\mathbf{X}$ is the union of all clusters. As in dimension 1, we denote by S the number of points in a cluster, with a little abuse of notation by C_0 the cluster generated by a point at 0, and by L the radius of C_0. Moreover, we denote by $N_{\mathbf{X}}(b(0,r))$ the number of points of $\mathbf{X}$ in the ball $b(0,r)$, and by

$$\omega_d(r) = \frac{r^d\pi^{d/2}}{\Gamma(1+d/2)}$$

the volume of $b(0,r)$. The following LDP holds:

Theorem 5.1.1. *Assume* (5) *and*

$$E\big[L^d e^{\theta S}\big] < \infty, \quad \textit{for all } \theta \in \mathscr{D}_S^\circ. \tag{46}$$

Then $\{N_{\mathbf{X}}(b(0,r))/\omega_d(r)\}$ *satisfies a LDP on* $\mathbb{R}$ *with speed* $\omega_d(r)$ *and good rate function* (7).

Assumptions (5) and (46) hold for some classes of spatial Poisson cluster processes that are of interest in the applications. In particular, they hold for ergodic Hawkes processes, under certain integrability conditions on the fertility rate (see Subsection 5.3). One can also show that conditions (5) and (46) hold for the class of stationary shot noise Cox processes (see Møller[19,20]; Møller and Torrisi[22] for more insight into this class of stochastic processes), under some integrability conditions on the shot shape function that describes the stochastic intensity of the process.

Before proving Theorem 5.1.1, we show that the same LDP holds for the non-stationary Poisson cluster process $\mathbf{X}_{r,R}$ with immigrant process empty in $\mathbb{R}^d\backslash b(0,R+r)$. As usual this LDP holds under a weaker condition.

Theorem 5.1.2. *Assume* (5). *Then* $\{N_{\mathbf{X}_{r,R}}(b(0,r))/\omega_d(r)\}$ *satisfies a LDP on* $\mathbb{R}$ *with speed* $\omega_d(r)$ *and good rate function* (7).

Proof of Theorem 5.1.2. The proof is similar to that one for the non-stationary Poisson cluster process on the line. Here we just sketch the main differences. As in the proof of Theorem 3.1.2, the claim follows by the Gärtner-Ellis theorem. Indeed, letting $I_{|\,b(0,r)}$ denote the point process of immigrants in $b(0,r)$, and $I_{|\,b(0,R+r)\backslash b(0,r)}$ the point process of immigrants in $b(0,R+r)\backslash b(0,r)$ we have, for each $\theta\in\mathbb{R}$,

$$\mathrm{E}\Big[\mathrm{e}^{\theta N_{\mathbf{X}_{r,R}}(b(0,r))}\Big]=\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|\,b(0,r)}}N_{C_i}(b(0,r))}\Big]\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|\,b(0,R+r)\backslash b(0,r)}}N_{C_i}(b(0,r))}\Big].$$

As usual, with a little abuse of notation denote by C_0 the cluster generated by an immigrant at 0. It holds:

$$\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|b(0,r)}}N_{C_i}(b(0,r))}\Big]=\exp\left(\nu\int_{b(0,r)}\mathrm{E}\Big[\mathrm{e}^{\theta N_{C_0}(b(-x,r))}-1\Big]\mathrm{d}x\right),$$

for each $\theta\in\mathbb{R}$;

$$\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|\,b(0,R+r)\backslash b(0,r)}}N_{C_i}(b(0,r))}\Big]\le\exp\big(\nu(\omega_d(R+r)-\omega_d(r))\mathrm{E}\big[\mathrm{e}^{\theta S}-1\big]\big),$$

for $\theta\in[0,\infty)\cap\mathcal{D}_S$;

$$1\ge\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|\,b(0,R+r)\backslash b(0,r)}}N_{C_i}(b(0,r))}\Big]\ge\exp\big(\nu(\omega_d(R+r)-\omega_d(r))\mathrm{E}\big[\mathrm{e}^{\theta S}-1\big]\big),$$

for $\theta\le 0$. Therefore,

$$\lim_{r\to\infty}\frac{1}{\omega_d(r)}\log\mathrm{E}\Big[\mathrm{e}^{\theta\sum_{X_i\in I_{|\,b(0,r)}}N_{C_i}(b(0,r))}\Big]=\lim_{r\to\infty}\frac{\nu}{\omega_d(r)}\int_{b(0,r)}\mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(b(-x,r))}-1\big]\mathrm{d}x$$

$$= \frac{\nu}{\omega_d(1)} \lim_{r\to\infty} \int_{b(0,1)} \mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(b(-ry,r))} - 1\big]\mathrm{d}y$$
$$= \nu \mathrm{E}\big[\mathrm{e}^{\theta S} - 1\big], \quad \text{for each } \theta \in \mathbb{R},$$

and, since $\lim_{r\to\infty} \omega_d(R+r)/\omega_d(r) = 1$, for each $\theta \in \mathscr{D}_S$,

$$\lim_{r\to\infty} \frac{1}{\omega_d(r)} \log \mathrm{E}\Big[\mathrm{e}^{\theta \sum_{X_i \in I_{|\, b(0,R+r)\setminus b(0,r)}} N_{C_i}(b(0,r))}\Big] = 0.$$

The rest of the proof is exactly as in the one-dimensional case.

Proof of Theorem 5.1.1. The proof is similar to that one of Theorem 5.1.2 and is again based on the Gärtner-Ellis theorem. We start showing that

$$\lim_{r\to\infty} \frac{1}{\omega_d(r)} \log \mathrm{E}\big[\mathrm{e}^{\theta(N_{\mathbf{X}}(b(0,r)) - N_{\mathbf{X}_{r,R}}(b(0,r)))}\big] = 0, \quad \text{for all } \theta \in \mathscr{D}_S^\circ. \tag{47}$$

This relation is easily verified for $\theta \le 0$. Thus we only check (47) for $\theta \in \mathscr{D}_S^\circ \cap (0,\infty)$. We have, for all $\theta \in \mathscr{D}_S^\circ \cap (0,\infty)$,

$$\mathrm{E}\big[\mathrm{e}^{\theta(N_{\mathbf{X}}(b(0,r)) - N_{\mathbf{X}_{r,R}}(b(0,r)))}\big] = \mathrm{E}\Big[\mathrm{e}^{\theta \sum_{X_i \in I_{|\, \mathbb{R}^d \setminus b(0,r+R)}} N_{C_i}(b(0,r))}\Big]$$
$$= \exp\left(\nu \int_{\mathbb{R}^d \setminus b(0,r+R)} \mathrm{E}\big[\mathrm{e}^{\theta N_{C_0}(b(-x,r))} - 1\big]\, \mathrm{d}x\right).$$

Now notice that since $\theta > 0$ we have

$$\mathrm{e}^{\theta N_{C_0}(b(-x,r))} - 1 \le \big(\mathrm{e}^{\theta N_{C_0}(\mathbb{R}^d)} - 1\big)\mathbf{1}\{\|x\| \le L + r\}, \quad \text{for all } x \in \mathbb{R}^d.$$

The limit (47) follows by assumption (46) noticing that the above relations yield, for all $\theta \in \mathscr{D}_S^\circ \cap (0,\infty)$, $r > 0$,

$$\mathrm{E}\Big[\mathrm{e}^{\theta \sum_{X_i \in I_{|\, \mathbb{R}^d \setminus b(0,r+R)}} N_{C_i}(b(0,r))}\Big] \le \exp((\nu\pi^{d/2}/\Gamma(1+d/2))\mathrm{E}[(L+r)^d(\mathrm{e}^{\theta S} - 1)]).$$

Now notice that

$$\mathrm{E}\big[\mathrm{e}^{\theta N_{\mathbf{X}}(b(0,r))}\big] = \mathrm{E}\big[\mathrm{e}^{\theta N_{\mathbf{X}_{r,R}}(b(0,r))}\big]\mathrm{E}\big[\mathrm{e}^{\theta(N_{\mathbf{X}}(b(0,r)) - N_{\mathbf{X}_{r,R}}(b(0,r)))}\big], \quad \text{for all } \theta \in \mathbb{R}, \ \ r > 0.$$

Therefore, if $\mathscr{D}_S = \mathscr{D}_S^\circ$ then the claim is a consequence of the computation of the log-Laplace limit of $\{N_{\mathbf{X}_{r,R}}(b(0,r))\}$ in the proof of Theorem 5.1.2, the Gärtner-Ellis theorem and assumption (5). It remains to deal with the case $\mathscr{D}_S \neq \mathscr{D}_S^\circ$. Arguing exactly as in the proof of Theorem 3.1.1 it can be proved that for any sequence $\{r_n\}_{n\geq 1} \subset (0, \infty)$ diverging to $+\infty$, as $n \to \infty$, there exists a subsequence $\{q_n\} \subseteq \{r_n\}$ such that

$$\limsup_{n\to\infty} \frac{1}{q_n} \log P(N_{\mathbf{X}}(b(0, q_n))/\omega_d(q_n) \in F)$$
$$\leq -\inf_{x\in F} \Lambda^*(x), \quad \text{for all closed sets } F$$

and

$$\liminf_{n\to\infty} \frac{1}{q_n} \log P(N_{\mathbf{X}}(b(0, q_n))/\omega_d(q_n) \in G)$$
$$\geq -\inf_{x\in G} \Lambda^*(x), \quad \text{for all open sets } G,$$

where Λ^* is defined by (7). Then the large deviations upper and lower bounds hold for any sequence $\{r_n\}$ and the claim follows.

5.2. The Asymptotic Behavior of the Void Probability Function and the Empty Space Function

Apart some specific cases, the void probability function $v(r) = P(N_{\mathbf{X}}(b(0, r)) = 0), r > 0$, of a spatial Poisson cluster process is not known in closed form. Comparing $\mathbf{X}$ with the immigrant process I we easily obtain

$$v(r) \leq P(N_I(b(0, r)) = 0) = \mathrm{e}^{-\nu\omega_d(r)}, \quad r > 0. \tag{48}$$

A more precise information on the asymptotic behavior of $v(\cdot)$, as $r \to \infty$, is provided by the following proposition:

Proposition 5.2.1. *Assume* $\mathrm{E}[L^d] < \infty$. *Then*

$$\lim_{r\to\infty} \frac{1}{\omega_d(r)} \log v(r) = -\nu.$$

Proof. Note that

$$\begin{aligned} v(r) &= P(N_{C_i}(b(0, r)) = 0, \text{ for all } X_i \in I) \\ &= \mathrm{E}\Bigg[\mathbf{1}\{N_I(b(0, r)) = 0\} \prod_{X_i\in I_{|\mathbb{R}^d\setminus b(0,r)}} \mathbf{1}\{N_{C_i}(b(0, r)) = 0\}\Bigg] \\ &= \mathrm{e}^{-\nu\omega_d(r)} \mathrm{E}\Bigg[\prod_{X_i\in I_{|\mathbb{R}^d\setminus b(0,r)}} \mathbf{1}\{N_{C_i}(b(0, r)) = 0\}\Bigg] \\ &= \mathrm{e}^{-\nu\omega_d(r)} \exp\left(-\nu \int_{\mathbb{R}^d\setminus b(0,r)} P(N_{C_0}(b(-x, r)) > 0)\, \mathrm{d}x\right) \end{aligned} \tag{49}$$

where in (49) we used Lemma 6.4.VI in Daley and Vere-Jones[7]. Thus the claim follows if we prove

$$\lim_{r\to\infty} \frac{1}{\omega_d(r)} \int_{\mathbb{R}^d \setminus b(0,r)} P(N_{C_0}(b(-x,r)) > 0)\mathrm{d}x = 0.$$

For this note that

$$\mathbf{1}\{N_{C_0}(b(0,r)) > 0\} \leq \mathbf{1}\{\|x\| - r \leq L\}, \quad \text{for all } x \in \mathbb{R}^d, \ \ r > 0.$$

Therefore,

$$\frac{1}{\omega_d(r)} \int_{\mathbb{R}^d \setminus b(0,r)} P(N_{C_0}(b(-x,r)) > 0)\mathrm{d}x \leq \mathrm{E}[(1 + L/r)^d - 1],$$

and the right-hand side in the above inequality goes to 0 as $r \to \infty$ by the dominated convergence theorem (note that $\mathrm{E}[L^d] < \infty$ by assumption).

In spatial statistics, a widely used summary statistic is the so-called empty space function, which is the distribution function of the distance from the origin to the nearest point in $\mathbf{X}$ (see, for instance, Møller and Waagepetersen[24]), that is

$$e(r) = 1 - v(r), \quad r > 0.$$

Apart from some specific cases, the empty space function of Poisson cluster processes seems to be intractable. Next Corollary 5.2.1 concerns the asymptotic behavior of $e(r)$, as $r \to \infty$.

Corollary 5.2.1. *Under the assumption of Proposition* 5.2.1 *it holds*

$$\lim_{r\to\infty} \frac{1}{\omega_d(r)} \log\log e(r)^{-1} = -\nu.$$

Proof. The proof is an easy consequence of Proposition 5.2.1. By the upper bound (48) we obtain

$$\limsup_{r\to\infty} \frac{1}{\omega_d(r)} \log\log e(r)^{-1} \leq \lim_{r\to\infty} \frac{1}{\omega_d(r)} \log\log(1 - \mathrm{e}^{-\nu\omega_d(r)})^{-1} = -\nu.$$

To get the matching lower bound we note that the inequality $\log(1 - x) \leq -x, x \in [0, 1)$, gives $\log e(r)^{-1} \geq v(r), r > 0$, and therefore by Proposition 5.2.1 we get

$$\liminf_{r\to\infty} \frac{1}{\omega_d(r)} \log\log e(r)^{-1} \geq \liminf_{r\to\infty} \frac{1}{\omega_d(r)} \log v(r) = -\nu.$$

5.3. Spatial Hawkes Processes

Spatial Hawkes processes have been introduced in Daley and Vere-Jones[7]. Brémaud et al.[3] considered spatial Hawkes processes with random fertility rate and not necessarily Poisson immigrants, and computed the Bartlett spectrum; the reader is directed to Daley and Vere-Jones[7] for the notion of Bartlett spectrum. Møller and Torrisi[22] derived the pair correlation function of spatial Hawkes processes; we refer the reader to Møller and Waagepetersen[24] for the notion of pair correlation function.

For the sake of completeness, we briefly recall the definition of spatial Hawkes process. A spatial Hawkes process is a Poisson cluster process $\mathbf{X} \subset \mathbb{R}^d$ where $d \geq 1$ is an integer. The clusters centers are the points $\{X_i\}$ of a homogeneous Poisson process $I \subset \mathbb{R}^d$ with intensity $v \in (0, \infty)$. Each immigrant $X_i \in I$ generates a cluster $C_i = C_{X_i}$ which is formed by the points of generations $n = 0, 1, \ldots$ with the following branching structure: the immigrant $X_i \in I$ is said to be of zeroth generation. Given generations $0, 1, \ldots, n$ in C_i, each point $Y \in C_i$ of generation n generates a Poisson process on $\mathbb{R}^d$ of offspring of generation $n+1$ with intensity function $h(\cdot - Y)$. Here $h : \mathbb{R}^d \to [0, \infty)$ is a non-negative Borel function. In the model it is assumed that, given the immigrants, the centered clusters $\{C_i - X_i\}$ are iid, and independent of I. By definition the spatial Hawkes process is $\mathbf{X} \equiv \bigcup_i C_i$. As in the one-dimensional case, it is assumed

$$0 < \mu \equiv \int_{\mathbb{R}^d} h(\xi) \mathrm{d}\xi < 1. \tag{50}$$

This assumption guarantees that the number of points in a cluster has a finite mean equal to $1/(1-\mu)$, excludes the trivial case where there are no offspring, and ensures that $\mathbf{X}$ is ergodic, with a finite and positive intensity given by $v/(1-\mu)$. Due to the branching structure, the number S of offspring in a cluster follows the distribution (3). Finally, we note that the classical Hawkes process considered in the previous sections corresponds to the special case where $d = 1$ and $h(t) = 0$ for $t \leq 0$.

A LDP for spatial Hawkes processes can be obtained by Theorem 5.1.1. The precise statement is as Theorem 5.1.1 with (50) and

$$\int_{\mathbb{R}^d} \|\xi\| h(\xi) \mathrm{d}\xi < \infty,$$

in place of (5) and (46), moreover the rate function is $\Lambda^*(\cdot)$ defined by (20). Recall that the symbol $\|\cdot\|$ denotes the Euclidean norm.

Similarly, the asymptotic behavior of the void probability function and the empty space function of spatial Hawkes processes can be obtained as immediate consequences of Proposition 5.2.1 and Corollary 5.2.1, respectively. The precise statements are as Proposition 5.2.1 and Corollary 5.2.1, with conditions (50) and

$$\int_{\mathbb{R}^d} \|\xi\|^d h(\xi) \mathrm{d}\xi < \infty$$

in place of $\mathrm{E}[L^d] < \infty$.

6. EXTENSIONS AND OPEN PROBLEMS

In this paper we studied large deviations of Poisson cluster processes. Applications in insurance of some results in this work can be found in Stabile and Torrisi[30].

The definition of Hawkes process extends immediately to the case of random fertility rate $h(\cdot, Z)$, where Z_k's are iid unpredictable marks associated to the points X_k (see Daley and Vere-Jones[7] for the definition of unpredictable marks, and Brémaud et al.[3] for the construction of Hawkes processes with random fertility rate specified by an unpredictable mark). Due to the form of the distribution of S in this case (see formula (6) in Møller and Rasmussen[21]) it is not clear if the LDPs for Hawkes processes proved in this paper are still valid for Hawkes processes with random fertility rate.

The generalization of our results to non-linear Hawkes processes (Brémaud and Massoulié[2]; Brémaud et al.[4]; Kerstan[17]; Massoulié[18]; Torrisi[31]) would be interesting. However, since a non-linear Hawkes process is not even a Poisson cluster process, a different approach is needed.

ACKNOWLEDGMENT

We thank Kamil Szczegot for reporting two mistakes and for a careful reading of a first draft of the paper.

REFERENCES

1. Borovkov, A.A. Boundary values problems for random walks and large deviations for function spaces. Theory Probab. Appl. **1967**, *12*, 575–595.
2. Brémaud, P.; Massoulié, L. Stability of nonlinear Hawkes processes. Ann. Prob. **1996**, *24*, 1563–1588.
3. Brémaud, P.; Massoulié, L.; Ridolfi, A. Power spectra of random spike fields and related processes. Adv. Appl. Prob. **2005**, *37*, 1116–1146.
4. Brémaud, P.; Nappo, G.; Torrisi, G.L. Rate of convergence to equilibrium of marked Hawkes processes. J. Appl. Prob. **2002**, *39*, 123–136.
5. Brix, A.; Chadoeuf, J. Spatio-temporal modeling of weeds by shot-noise G Cox processes. Biometrical J. **2002**, *44*, 83–99.
6. Chavez-Demoulin, V.; Davison, A.C.; Mc Neil, A.J. Estimating value-at-risk: a point process approach. Quantitative Finance. **2005**, *5*, 227–234.

7. Daley, D.J.; Vere-Jones, D. *An Introduction to the Theory of Point Processes*, 2nd Ed.; Springer: New York, 2003.
8. de Acosta, A. Large deviations for vector valued Lévy processes. Stochastic Process. Appl. **1994**, *51*, 75–115.
9. Dembo, A.; Zeitouni, O. *Large Deviations Techniques and Applications*, 2nd Ed.; Springer: New York, 1998.
10. Ganesh, A.; Macci, C.; Torrisi, G.L. Sample path large deviations principles for Poisson shot noise processes and applications. Electron. J. Probab. **2005**, *10*, 1026–1043.
11. Gusto, G.; Schbath, S. F.A.D.O.: A statistical method to detect favored or avoided distances between occurrences of motifs using the Hawkes model. Stat. Appl. Genet. Mol. Biol. **2005**, *4*, Article 24.
12. Hawkes, A.G. Spectra of some self-exciting and mutually exciting point processes. Biometrika **1971**, *58*, 83–90.
13. Hawkes, A.G. Point spectra of some mutually exciting point processes. J. Roy. Statist. Soc. Ser. B **1971**, *33*, 438–443.
14. Hawkes, A.G.; Oakes, D. A cluster representation of a self-exciting process. J. Appl. Prob. **1974**, *11*, 493–503.
15. Jagers, P. *Branching Processes with Biological Applications*; John Wiley: London, 1975.
16. Jonnson, D.H. Point process models of single-neuron discharges. J. Computational Neuroscience **1996**, *3*, 275–299.
17. Kerstan, J. Teilprozesse poissonscher prozesse. Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, Random Processes 1964; 377–403.
18. Massoulié, L. Stability results for a general class of interacting point processes dynamics, and applications. Stoch. Proc. Appl. **1998**, *75*, 1–30.
19. Møller, J. A comparison of spatial point process models in epidemiological applications. In *Highly Structured Stochastic Systems*; Green, P.J.; Hjort, N.L.; Richardson, S., Eds; Oxford University Press, 2003; 264–268.
20. Møller, J. Shot noise Cox processes. Adv. Appl. Prob. **2003**, *35*, 614–640.
21. Møller, J.; Rasmussen, J.G. Perfect simulation of Hawkes processes. Adv. Appl. Prob. **2005**, *37*, 629–646.
22. Møller, J.; Torrisi, G.L. Generalised shot noise Cox processes. Adv. Appl. Prob. **2005**, *37*, 48–74.
23. Møller, J.; Torrisi, G.L. The pair correlation function of spatial Hawkes processes. Stat. Probab. Lett. **2007**, *77*, 995–1003.
24. Møller, J.; Waagepetersen, R.S. *Statistical Inference and Simulation for Spatial Point Processes*; Chapman and Hall: Boca Raton, 2004.
25. Neyman, J.; Scott, E.L. Statistical approach to problems of cosmology. J. R. Statist. Soc. B **1958**, *20*, 1–43.
26. Ogata, Y. Statistical models for earthquake occurrences and residual analysis for point processes. J. Amer. Statist. Assoc. **1988**, *83*, 9–27.
27. Ogata, Y. Space-time point process model for earthquake occurrences. Ann. Inst. Statist. Math. **1998**, *50*, 379–402.
28. Ogata, Y.; Akaike, H. On linear intensity models for mixed doubly stochastic poisson and self-exciting point processes. J. R. Statist. Soc. B **1982**, *44*, 102–107.
29. Reynaud-Bouret, P.; Roy, E. Some non asymptotic tail estimates for Hawkes processes. Bull. Belg. Math. Soc. Simon Stevin **2007**, *13*, 883–896.
30. Stabile, G.; Torrisi, G.L. Risk processes with non-stationary Hawkes arrivals. Submitted. 2007.
31. Torrisi, G.L. A class of interacting marked point processes: rate of convergence to equilibrium. J. Appl. Prob. **2002**, *39*, 137–161.
32. Vere-Jones, D.; Ozaki, T. Some examples of statistical estimation applied to earthquake data. Ann. Inst. Statist. Math. **1982**, *34*, 189–207.