

# Limit theorems for Hawkes processes including inhibition

Patrick Cattiaux<sup>\*</sup>, Laetitia Colombani, Manon Costa

*Institut de Mathématiques de Toulouse. CNRS UMR 5219. Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse cedex 09, France*

Received 17 September 2021; received in revised form 11 March 2022; accepted 6 April 2022

Available online 12 April 2022

## Abstract

In this paper we consider some non linear Hawkes processes with signed reproduction function (or memory kernel) thus exhibiting both self-excitation and inhibition. We provide a Law of Large Numbers, a Central Limit Theorem and large deviation results, as time grows to infinity. The proofs lie on a renewal structure for these processes introduced in Costa et al. (2020) which leads to a comparison with cumulative processes. Explicit computations are made on some examples. Similar results have been obtained in the literature for self-exciting Hawkes processes only.

© 2022 Elsevier B.V. All rights reserved.

MSC: 60G55; 60F05; 60K15

Keywords: Hawkes processes; Inhibition; Renewal theory; Limit theorems

## 1. Introduction

Hawkes processes have been introduced by Hawkes [16] and are widely used for modeling purposes: originally as models for the appearances of earthquakes [16,18], but now in finance [3,17] and econometrics or in neuroscience as models of spike trains of neurons [15,22]. We refer to the bibliography of our references for more details.

A Hawkes process  $t \mapsto N_t^h = N^h([0, t])$  is a point process on the real line  $\mathbb{R}$  characterized by its initial condition on  $]-\infty, 0]$  and its intensity process  $t \mapsto \Lambda(t)$  through the infinitesimal relation

$$\mathbb{P}(N_t^h \text{ has a jump in } ]t, t + dt[ | \mathcal{F}_t) = \Lambda(t) dt,$$

<sup>\*</sup> Corresponding author.

E-mail addresses: [patrick.cattiaux@math.univ-toulouse.fr](mailto:patrick.cattiaux@math.univ-toulouse.fr) (P. Cattiaux),  
[laetitia.colombani@math.univ-toulouse.fr](mailto:laetitia.colombani@math.univ-toulouse.fr) (L. Colombani), [manon.costa@math.univ-toulouse.fr](mailto:manon.costa@math.univ-toulouse.fr) (M. Costa).

where  $\mathcal{F}_t = \sigma(N^h(\cdot - \infty, s]; s \leq t)$  is the natural filtration of the process and

$$\Lambda(t) = f \left( \lambda + \int_{-\infty, t[} h(t-s) N^h(ds) \right). \quad (1.1)$$

Here  $\lambda \in \mathbb{R}$ ,  $f : \mathbb{R} \rightarrow \mathbb{R}^+$  is the *jump rate function* and  $h : \mathbb{R}^+ \rightarrow \mathbb{R}$  is the *reproduction function* (or *memory kernel*). We shall give a more precise definition in the next section (in particular on what happens before time 0) as well as results on existence and stability.

When  $f$  is linear or affine, the process is said to be linear. In this case one has to assume that  $\lambda \geq 0$  and  $h \geq 0$  too. Note that when  $h$  vanishes identically we recover a standard Poisson process. Otherwise the Hawkes process is called non linear. Actually, except for the behavior of the shifted process [8,23], very few papers are dealing with possibly negative or signed  $h$ . The negative part of  $h$  can be interpreted as self-inhibition.

It is very natural to look at the large time behavior of  $N^h$ , in particular the Law of Large Numbers (LLN) the Central Limit Theorem (CLT) and the deviations from the asymptotic mean or more generally the large deviations (LD).

In the linear case (recall that  $h$  is thus assumed to be non-negative) and assuming that  $\|h\|_{L^1(du)} < 1$ , both the LLN

$$\frac{N_t^h}{t} \rightarrow \frac{\lambda}{1 - \|h\|_{L^1(du)}} := \mu \quad a.s. \text{ as } t \rightarrow +\infty, \quad (1.2)$$

and the CLT

$$\frac{N_t^h - \mu t}{\sqrt{t}} \Rightarrow \mathcal{N}^h(0, \sigma^2) \quad \text{with } \sigma^2 = \frac{\lambda}{(1 - \|h\|_{L^1(du)})^3},$$

where the convergence holds in distribution, have been shown (see e.g. [9]). Actually Bacry et al. [2] have obtained the functional version of the CLT (convergence to some Brownian motion) in the multivariate case. In a different direction, [12] have shown a CLT for fixed  $t$  as  $\lambda \rightarrow +\infty$ .

The easiest way to derive LLN and CLT in the linear case is presumably to use the immigration-birth representation also called the cluster process representation in [18], connecting  $N^h$  to subcritical Galton–Watson processes. This representation was used in [4] in order to get the Large Deviation (LD) principle for  $N_t^h/t$  with rate function

$$I(x) = x \ln \left( \frac{x}{\lambda + x \|h\|_{L^1(du)}} \right) - x(1 - \|h\|_{L^1(du)}) + \lambda.$$

For this explicit expression of the rate function see [27] p.761. The LD principle is obtained in [4] under the additional assumption  $\int_0^{+\infty} t h(t) dt < +\infty$ . It is claimed in the introduction of [13] that this assumption is not necessary. Under more restrictive assumptions, [13] contains precise deviations (see e.g. Theorem 2 therein).

The non linear case is of course more difficult. According to the general seminal paper by Brémaud and Massoulié [5], if  $f$  is  $L$ -Lipschitz and  $L \|h\|_{L^1(du)} < 1$ , there exists a unique stationary version of the Hawkes process. Rate of convergence to equilibrium is studied in [6] in two specific cases. As a consequence of Brémaud and Massoulié result, we get that

$$\frac{N_t^h}{t} \rightarrow \mu = \mathbb{E}_s[N^h([0, 1])] \quad a.s. \text{ as } t \rightarrow +\infty, \quad (1.3)$$

where  $\mathbb{E}_s$  denotes the expectation w.r.t. the stationary ergodic distribution.

In the particular situation where  $h$  is an exponential, the Hawkes process becomes Markovian and some results of large deviation have been obtained [29]. In [27], Zhu proved a functional CLT *at equilibrium* from which the following follows

**Theorem 1.1.** *Assume that*

- (1)  $f$  is  $L$ -Lipschitz,
- (2)  $h$  is non-negative, decreasing and such that  $\int_0^{+\infty} t h(t) dt < +\infty$ ,
- (3)  $L \int_0^{+\infty} h(t) dt < 1$ ,
- (4)  $\lambda \geq 0$ .

*Then the stationary Hawkes process satisfies  $\frac{N_t^h - \mu t}{\sqrt{t}} \Rightarrow \mathcal{N}^h(0, \sigma^2)$  as  $t \rightarrow +\infty$  in distribution, with*

$$\sigma^2 := \text{Var}_s(N^h([0, 1])) + 2 \sum_{j \geq 1} \text{Cov}_s(N^h([0, 1]), N^h([j, j + 1]))$$

*where  $\text{Var}_s$  and  $\text{Cov}_s$  denote the variance and covariance w.r.t. the stationary distribution.*

The proof is based on martingales techniques for the functional CLT. As the author himself is saying, to obtain an explicit expression for  $\mu$  and  $\sigma^2$  can rapidly become a difficult task. In the same work, Zhu also obtained a Strassen iterated logarithm law. One can also mention [28] where a large deviation result is obtained by contracting the level-3 LDP, i.e. by considering the shifted occupation measure. Theorem 2 in [28] then furnishes a LDP for  $N_t^h/t$ , provided  $h$  is non decreasing and non-negative and  $f$  is sub-linear at infinity. The expression of the rate function, as the infimum of the entropy on some set of measures satisfying a linear constraint is however not really tractable.

Since we are interested in neurosciences, our goal in this work is to understand the role of self-inhibition in the asymptotic behavior of Hawkes processes. Since inhibition will slow down the neuronal activity, we thus have to consider signed functions  $h$  (the positive part modeling the self-excitation), but also jump rate functions  $f$  satisfying  $f(u) = 0$  if  $u \leq 0$ . In the present paper, we will study the case of a general, signed, reproduction function with compact support and the specific jump rate function  $f(u) = u^+ = \max(u, 0)$ . This choice is of course the simplest one allowing us to introduce inhibition, and to compare this situation with linear models.

We will obtain a LLN, a CLT and deviation inequalities, where the parameters are characterized by the renewal structure of the process introduced in [8] replacing the classical cluster representation of the self-exciting case established in [18] which is no more valid. This renewal structure allows us to write the Hawkes process almost as a cumulative process.

The main tools are then limit theorems for cumulative processes and actually, the technical work consists in showing that one can apply these theorems in the present situation. An important tool is a comparison between the considered Hawkes process, the self excited process associated to the positive part of the reproduction function, furnishing an upper bound, and a purely inhibited process corresponding to the (negative) lower bound of the reproduction function (see Proposition 2.4), furnishing a lower bound.

For simplicity we restrict ourselves to an empty initial condition (see below). Some explicit computations are done in simple particular cases of pure inhibition ( $h$  non-positive). Precise statements will require some definitions, so that they are postponed to the next section. We emphasize, that the inhibition part introduces new intricacies.

As we said, very few papers are dealing with inhibition. In [11] some specific kernels are considered, but the addressed problem is not the one we are considering here. Looking at possibly negative reproduction functions is not only of mathematical interest. As shown in [15,20,25] a multivalued version of the model we are studying is particularly well suited for modeling spike train of neurons, at least in an almost stationary regime. To extend our results to the multivalued framework should thus be an interesting question.

## 2. Notation, definitions and results

### 2.1. Hawkes processes

We consider an appropriate filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  satisfying the usual assumptions.

**Definition 2.1.** Let  $\lambda > 0$  and  $h : (0, +\infty) \rightarrow \mathbb{R}$  a signed measurable function. Let  $N^0$  a locally finite point process on  $(-\infty, 0]$  with law  $\mathbf{m}$ .

The point process  $N^h$  on  $\mathbb{R}$  is a Hawkes process on  $(0, +\infty)$ , with initial condition  $N^0$  and reproduction measure  $\mu(dt) = h(t)dt$  if:

- $N^h|_{(-\infty, 0]} = N^0$ ,
- the conditional intensity measure of  $N^h|_{(0, +\infty)}$  with respect to  $(\mathcal{F}_t)_{t \geq 0}$  is absolutely continuous w.r.t the Lebesgue measure and has stochastic intensity:

$$\Lambda^h : t \in (0, +\infty) \mapsto \left( \lambda + \int_{(-\infty, t)} h(t-u) N^h(du) \right)^+. \quad (2.1)$$

where  $x^+ = \max(x, 0)$ .

The next proposition gives an explicit representation of the Hawkes process as solution of an SDE driven by a Poisson point process and states an important coupling property.

**Proposition 2.2** (Proposition 2.1 in [8]). Let  $Q$  be a  $(\mathcal{F}_t)_{t \geq 0}$  — two-dimensional Poisson point process on  $(0, +\infty) \times (0, +\infty)$  with unit intensity. We consider the equation

$$\begin{cases} N^h = N^0 + \int_{(0, +\infty) \times (0, +\infty)} \delta_u \mathbb{1}_{\theta \leq \Lambda^h(u)} Q(du, d\theta) \\ \Lambda^h(u) = \left( \lambda + \int_{(-\infty, u)} h(u-s) N^h(ds) \right)^+, \quad u > 0, \end{cases} \quad (2.2)$$

where  $\lambda > 0$  is an immigration rate,  $h : (0, +\infty) \rightarrow \mathbb{R}$  is a signed measurable function and  $N^0$  is an initial condition of law  $\mathbf{m}$  on  $(-\infty, 0]$ .

We consider the similar equation for  $N^{h^+}$  in which  $h$  is replaced by  $h^+(\cdot) = \max(h(\cdot), 0)$ . We assume that  $\|h^+\|_1 := \|h^+\|_{L^1(du)} < 1$  and that the distribution  $\mathbf{m}$  satisfies:

$$\forall t > 0, \int_0^t \mathbb{E}_{\mathbf{m}} \left( \int_{(-\infty, 0]} h^+(u-s) N^0(ds) \right) du < +\infty. \quad (2.3)$$

Then:

- There exists a pathwise strong solution  $N^h$  of Eq. (2.2), and this solution is a Hawkes process.
- This property is true for  $N^{h^+}$ . Moreover, in the sense of measures,  $N^h \leq N^{h^+}$ , meaning that for all  $0 \leq s \leq t < +\infty$ ,  $N^h([s, t]) \leq N^{h^+}([s, t])$ .

## 2.2. Definitions and assumptions

In this paper we consider a Hawkes process  $N^h$  according to [Definition 2.1](#). We focus on the case of a signed reproduction function  $h$  which represents a possible inhibition on the appearance of future points.

### Assumption 2.3.

In all the paper, we will make the following assumptions:

- (i)  $h : (0, +\infty) \rightarrow \mathbb{R}$  is a compactly supported signed measurable function. We define  $L(h)$  as the supremum of the support of  $h$ :  $L(h) := \sup\{t > 0, |h(t)| > 0\} < \infty$ .

(ii)

$$\|h^+\|_1 := \int_0^{+\infty} h^+(u) du < 1,$$

where  $h^+(x) = \max(h(x), 0)$ .

(iii)  $\lambda > 0$ ,

(iv) the initial condition on  $]-\infty, 0[$  does not contain any point i.e.  $\mathbf{m} = \delta_\emptyset$ .

We are interested in the asymptotic behavior of the number of jumps of the process  $N^h$  on the interval  $[0, t]$ , and we denote:

$$N_t^h = N^h([0, t]), \quad \forall t \geq 0$$

In particular we aim at quantifying precisely the loss of points due to inhibition. We will prove asymptotic results for  $\frac{N_t^h}{t}$  and give exact computations on specific examples.

First we show another comparison result, this time furnishing a lower bound for  $N_t^h$ . This result motivates the detailed study of the canceling of intensity example.

**Proposition 2.4** (Minoration of Hawkes Process). *Let  $h$  be a function satisfying [Assumption 2.3](#). Let  $\lambda > 0$  and define  $g = -\lambda \mathbb{1}_{[0, L(h)]}$ .*

*One can find a coupling of two Hawkes processes  $N^h$  and  $N^g$ , respectively associated with the reproduction functions  $h$  and  $g$  and with basal intensity  $\lambda$ , such that for any  $t \geq 0$ :*

$$N_t^h \geq N_t^g \quad a.s.$$

Note that this comparison result is weaker than the majoration via  $h^+$ , since we do not have  $N^h([s, t]) \geq N^g([s, t])$  for all  $s$ , but only for  $s = 0$ .

**Proof.** The main idea is to construct these two processes with the same Poisson point process  $Q$  on  $(0, +\infty)^2$ . We consider the successive jumps of  $N^h$ :  $U_1^h, U_2^h, U_3^h, \dots$ ; and the ones of  $N^g$ :  $U_1^g, U_2^g, U_3^g, \dots$

We will prove by induction, that

$$\forall j \geq 1, N_{U_j^g}^h \geq N_{U_j^g}^g = j \quad a.s.$$

by studying the intervals associated with the  $[U_j^g, U_{j+1}^g)$  for  $j \in \mathbb{N}$ . We stress out that considering the definition of the function  $g$ , the intensity  $\Lambda^g$  of the Hawkes process  $N^g$  can only take the two values 0 and  $\lambda$ .

*First interval:* First remark that  $\forall t < \min(U_1^h, U_1^g)$ ,

$$\Lambda^h(t) = \lambda = \Lambda^g(t),$$

thus we have  $U_1^h = U_1^g$  and consequently  $N_{U_1^g}^h = N_{U_1^g}^g$ .

*Second interval:* For  $j = 2$ : by definition, there is only one jump for  $N^g$  on  $[U_1^g, U_2^g)$ . There are two possibilities for  $N^h$ :

- Assume that there is no other jump that  $U_1^h$  in this interval. Since  $U_2^g \geq U_1^g + L(h)$ , we have  $\Lambda^g(U_2^g-) = \lambda = \Lambda^h(U_2^g-)$ . Accordingly,  $U_2^h = U_2^g$  and in particular,  $N_{U_2^g}^h = N_{U_2^g}^g$  a.s.
- Otherwise, there is at least one other jump of  $N^h$  in  $(U_1^g, U_1^g + L(h))$ . In this case,  $N_{U_2^g}^h \geq 2 = N_{U_2^g}^g$  a.s.

*Recursion step:* We fix  $j$  and we suppose that the statement holds for  $i \leq j$ . Let  $k = N_{U_j^g}^h \geq j$  by assumption. Then consider the two following cases:

- If  $U_j^g$  is a jump of  $N^h$ , there is either at least one other jump of  $N^h$  in  $(U_j^g, U_{j+1}^g)$  or no other jump. If there is no other jump, then  $\Lambda^h(U_{j+1}^g-) = \Lambda^g(U_{j+1}^g-) = \lambda$ , since  $U_{j+1}^g > U_j^g + L(h)$ . So,  $U_{k+1}^h = U_{j+1}^g$ . In both situations,

$$N_{U_{j+1}^g}^h \geq 1 + N_{U_j^g}^h \geq 1 + N_{U_j^g}^g = N_{U_{j+1}^g}^g.$$

- If  $U_j^g$  is not a jump of  $N^h$ , then

$$\Lambda^h(U_j^g-) < \lambda = \Lambda^g(U_j^g-).$$

Therefore since the support of  $h$  is of length  $L(h)$  we deduce that

$$U_k^h < U_j^g < U_k^h + L(h).$$

By the induction hypothesis, we know that  $k \geq j$ . Then, there is either at least one jump of  $N^h$  in  $(U_j^g, U_j^g + L(h))$ , or the next jump is  $U_{j+1}^g$ , i.e.  $U_{k+1}^h = U_{j+1}^g$ . In both cases, we have  $N_{U_{j+1}^g}^h \geq 1 + k \geq 1 + j = N_{U_{j+1}^g}^g$ .

This concludes the induction. Let us come back to a general  $t \in \mathbb{R}^+$ . For any fixed  $\omega$ , there exists  $j = j(\omega) \in \mathbb{N}$ , such that:  $U_j^g(\omega) \leq t < U_{j+1}^g(\omega)$ . Then using the monotonicity of  $N^h$  we have

$$N_t^h \geq N_{U_j^g}^h \geq N_{U_j^g}^g = N_t^g. \quad \square$$

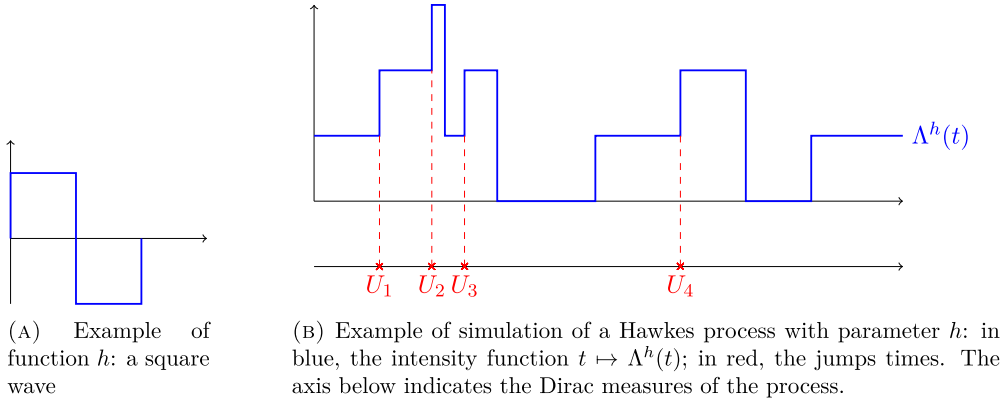
Both comparison results may be used in the sequel.

### 2.3. Hawkes processes as cumulative processes

Our study of the large time behavior of Hawkes processes lies on a renewal structure for Hawkes processes first introduced in [8] we shall partly recall below. Notice that this structure is used in [8] for a completely different purpose.

Let  $N^h$  be a Hawkes process according to Definition 2.1, with initial condition  $N^0 = \emptyset$ . We denote by  $U_1, U_2, U_3, \dots$  its successive jumps (see Figs. 1 and 3).

Let us introduce the renewal times of the process which splits the time line into independent and identically distributed time windows of length  $\tau_1, \tau_2, \dots$



**Fig. 1.** Example of Hawkes process.

Define the stopping time

$$\tau_1 = \inf\{t > U_1, N^h((t - L(h), t]) = 0\},$$

that is the first time after  $U_1$  such that there has been no jump during a time  $L(h)$ . We also set

$$S_0 = 0 \quad \text{and} \quad S_1 = \tau_1.$$

Let us now define

$$W_1 = N^h([U_1^1, S_1]) = N^h([0, S_1]),$$

the number of jumps of the process in this first time window and rename the jump times in the first time window as:

$$U_j^1 = U_j, \quad \forall j \in \{1, \dots, W_1\}.$$

We shall see below that  $\tau_1$  and  $W_1$  are almost surely finite. Recursively let  $i \in \mathbb{N}^*$  such that  $(\tau_1, W_1), \dots, (\tau_i, W_i)$  are well defined (and a.s. finite). Let  $S_i = \sum_{k=1}^i \tau_k$  and define

$$U_1^{i+1} = U_{W_1 + \dots + W_i + 1},$$

and

$$\tau_{i+1} = \inf\{t > U_1^{i+1}, N^h((t - L(h), t]) = 0\} - S_i, \quad (2.4)$$

Notice that there is at least one jump in  $[S_i, S_i + \tau_{i+1}]$ . We finally introduce the number of jumps in the  $(i + 1)$ th window as

$$W_{i+1} = N^h(U_1^{i+1}, S_i + \tau_{i+1}) = N^h([S_i, S_i + \tau_{i+1}]), \quad (2.5)$$

and rename the associated jump times as:

$$U_j^{i+1} = U_{W_1 + \dots + W_i + j}, \quad \forall j \in \{1, \dots, W_{i+1}\}.$$

Fig. 2 is an example of this splitting of the time and the renumbering of the jumps, in the case where  $h(t) = -\lambda \mathbb{1}_{(1,2)}(t)$ , so that  $L(h) = 2$ .

The next Proposition gathers important properties on the law of  $(\tau_i, W_i)$  defined above. However more explicit information are difficult to obtain except in specific cases (see Section 3).





By translation,  $U_1^2 - \tau_1$  is the first jump of a Poisson point process  $\mathcal{Q}'$  on  $(0, \infty) \times [0, \infty]$ , independent of  $\mathcal{Q}$  on  $(0, \tau_1) \times [0, \infty]$ , and  $U_1^2 - \tau_1$  is independent of  $\tau_1 = S_1$ .

Since the jumps of  $N^h$  before  $\tau_1$  do not influence  $\Lambda^h(t)$  for  $t > \tau_1$  (by definition of  $L(h)$  and  $\tau_1$ ),

$$\tau_2 = \inf\{t > U_1^2 - S_1, N^h((t + S_1 - L(h), t + S_1]) = 0\},$$

only depends on  $\mathcal{Q}$  on  $(U_1^2, +\infty) \times [0, \infty]$ . Moreover,  $(0, S_1) \times [0, \infty]$  and  $(U_1^2, +\infty) \times [0, \infty]$  are almost surely disjoint. Hence  $\mathcal{Q}$  on  $(U_1^2, +\infty) \times [0, \infty]$  is independent of  $\mathcal{Q}$  on  $(0, S_1) \times [0, \infty]$  so that  $\tau_2$  is independent of  $(\tau_1, W_1)$ .

The number of points in the second time window  $W_2 = N^h([U_1^2, S_2]) = N^h([S_1, S_2])$  only depends on  $\mathcal{Q}$  on  $(U_1^2, +\infty) \times [0, \infty]$ .  $W_1$  depends on  $\mathcal{Q}$  on  $(0, S_1) \times [0, \infty]$ . For the same reason as before,  $W_2$  is independent of  $(\tau_1, W_1)$ . The same argument can be used for each  $k$ : as the  $(S_k)_k$  split  $\mathbb{R}^+$  in disjoint intervals, then  $\mathcal{Q}$  on each of these intervals is independent of  $\mathcal{Q}$  of another interval.

In particular,  $U_1^1 = U_1^1 - S_0$ ,  $U_1^2 - S_1$  (and all the following) are independent and can be defined as the first jump of a Poisson point process on  $(0, +\infty) \times [0, \infty]$ . Then they follow an exponential law of parameter  $\lambda$ .

Using time translation, we see that  $\tau_1, \tau_2$  (and so on) are defined the same way and follow the same law. Then  $W_1, W_2$  (and so on) are defined the same way and follow the same law.  $\square$

This construction indicates the renewal structure generated by the Hawkes process. We shall use this structure to prove limit theorems.

To this end remark that

$$N_t^h = \sum_{i=1}^{\infty} \mathbb{1}_{U_i \leq t} = \sum_{i=1}^{\infty} \sum_{j=1}^{W_i} \mathbb{1}_{U_i^j \leq t}. \quad (2.6)$$

Introduce the renewal process associated to the  $S_i$ 's

$$M_t^h := \sum_{i=1}^{\infty} \mathbb{1}_{S_i \leq t}. \quad (2.7)$$

Since  $S_i = \sum_{k=1}^i \tau_k$  we may introduce

$$\hat{N}_t^h := \sum_{i=1}^{\infty} W_i \mathbb{1}_{S_i \leq t} = \sum_{i=1}^{M_t^h} W_i. \quad (2.8)$$

For any  $t \in \mathbb{R}^+$ , the current window is the  $M_t^h + 1$ th.  $\hat{N}_t^h$  includes only the jumps up to the  $M_t^h$ th window, while  $N_t^h$  can have more jumps. In particular,

$$\hat{N}_t^h \leq N_t^h \leq \hat{N}_t^h + W_{M_t^h+1} \quad \text{a.s.} \quad (2.9)$$

We thus have

$$N_t^h = \sum_{i=1}^{M_t^h} W_i + R_t^h$$

for some renewal process  $M_t^h$  and a remaining term  $R_t^h \leq W_{M_t^h+1}$ , the  $W_i$ 's being i.i.d. Such processes are known as *cumulative processes* in the literature. A LLN and a CLT for  $\hat{N}_t^h$  can

be found in [1] theorems 3.1 and 3.2. The LD principle for cumulative processes is studied in [21] in the special case

$$W_i = F(\tau_i)$$

for some non-negative, bounded and continuous function  $F$  (see the references in [21] for some previous results in still more specific cases). These results do not apply for Hawkes processes, and we had to establish a more general LD principle in the companion paper [7]. In order to get similar results for  $N_t^h/t$  it will remain to study the remaining  $R_t^h$ .

#### 2.4. LLN, CLT and asymptotic deviations for signed reproduction function $h$

We now state the main results of the paper. The key is to get enough moments for the  $(\tau_i, W_i)$ . The first result deals with this problem.

**Proposition 2.6.** *Let  $h$  be a signed function satisfying Assumption 2.3. Let us consider the Hawkes process  $N^h$  and the i.i.d. couples of random variables  $(\tau_i, W_i)$  defined in (2.4)–(2.5).*

- (i) *For  $\alpha < \alpha_0 := \min\left(\lambda, \frac{\|h^+\|_1 - \ln(\|h^+\|_1) - 1}{L(h)}\right)$  we have  $\mathbb{E}(e^{\alpha\tau_1}) < +\infty$ .*
- (ii) *There exists  $\theta_0 > 0$  such that for  $\theta < \theta_0$ ,  $\mathbb{E}(e^{\theta W_1}) < +\infty$ .*

*In particular  $\tau_1$  and  $W_1$  have polynomial moments of any order.*

The proof of this proposition is given in Section 4.1. Actually, one can give a lower bound on  $\theta_0$ . This lower bound differs whether  $h \leq 0$  or not.

In the general case, the upper bound for  $\theta_0$  depends on a random variable  $S$  with distribution

$$\mathbb{P}(S = k) = \frac{e^{-k\|h^+\|_1} (k\|h^+\|_1)^{k-1}}{k!}.$$

Using a comparison with a queue process that will be detailed in the proof of Proposition 2.6 one proves that  $\theta_0$  can be chosen as

$$\theta_0 < \|h^+\|_1 - \ln(\|h^+\|_1) - 1 \quad \text{and} \quad \lambda(\mathbb{E}(e^{2\theta_0 S}) - 1) < \alpha_0.$$

In the case of pure inhibition, i.e.  $h \leq 0$ , the quantity  $\|h^+\|_1 - \ln(\|h^+\|_1)$  becomes infinite. However using a comparison with a Poisson Process one can get another explicit bound for  $\theta_0$ , whose proof will also be given in Section 4.1.

**Proposition 2.7.** *If  $h \leq 0$ , one can choose  $\theta_0 < -\ln(1 - e^{-\lambda L(h)})$  in Proposition 2.6.*

**Remark 2.8.** Exact computations for moments of  $\tau$  and  $W$  are difficult. Let us consider here and in Section 3 some specific cases.

Notice that for  $h = 0$  (i.e. in the case of a Poisson process),  $W_1 = 1$  has exponential moments of any order and  $\tau_1$  whose distribution is exponential with parameter  $\lambda$ , has exponential moments up to order  $\lambda$ .

Another basic case is the *canceling of intensity* case, i.e. choosing the reproduction function as  $g = -\lambda \mathbb{1}_{[0,A]}$  for some positive  $\lambda$  and  $A$ . We have seen in Proposition 2.4 that the corresponding  $N_t^g$  is smaller than any  $N_t^h$  with  $L(h) = A$ . Since for  $t \in (U_1^1, U_1^1 + A)$  it holds  $\Lambda^h(t) = 0$ , it immediately follows that  $\tau_1 = U_1^1 + A$  and  $W_1 = 1$ , so that

$$(W_1, \tau_1) \sim (1, A + \mathcal{E}(\lambda)),$$

so that  $\mathbb{E}(\tau_1) = A + \lambda^{-1}$ ,  $\text{Var}(\tau_1) = \lambda^{-2}$ ,  $\alpha_0 = \lambda$  and  $\theta_0 = +\infty$ . ◇

From these moments properties and the renewal structure of the Hawkes process, we will derive the following asymptotic results:

**Theorem 2.9** (Law of Large Numbers). *Let  $h$  be a signed function satisfying [Assumption 2.3](#) and consider the Hawkes process  $N^h$  given by (2.2). Then we have the following:*

$$\frac{N_t^h}{t} \xrightarrow[t \rightarrow \infty]{a.s.} \frac{\mathbb{E}[W_1]}{\mathbb{E}(\tau_1)}.$$

Thanks to our comparison results and to (1.2) we have

$$\frac{\lambda}{1 + \lambda L(h)} \leq \frac{\mathbb{E}[W_1]}{\mathbb{E}(\tau_1)} \leq \frac{\lambda}{1 - \|h^+\|_1}.$$

Our method will also provide us with a CLT.

**Theorem 2.10** (Central Limit Theorem). *Let  $h$  be a signed function satisfying [Assumption 2.3](#) and consider the Hawkes process  $N^h$  given by (2.2). Then*

$$\sqrt{t} \left( \frac{N_t^h}{t} - \frac{\mathbb{E}[W_1]}{\mathbb{E}(\tau_1)} \right) \xRightarrow[t \rightarrow \infty]{} \mathcal{N}(0, \sigma^2)$$

with

$$\sigma^2 = \frac{\text{Var} \left( W_1 - \tau_1 \frac{\mathbb{E}[W_1]}{\mathbb{E}(\tau_1)} \right)}{\mathbb{E}(\tau_1)}.$$

We finally state deviation results based on the results in the companion paper [7], in which we obtain large deviation for general cumulative processes. To this end we need to introduce some notations. We also refer to [10] for a more general introduction on the topic.

**Definition 2.11.** We introduce the Cramer transform for  $(a, b) \in \mathbb{R}^2$ ,

$$\Lambda^*(a, b) = \sup_{x, y} \{ ax + by - \ln (\mathbb{E} [e^{x\tau_1 + yW_1}]) \}.$$

We also define for  $z \in \mathbb{R}^+$ ,

$$J(z) = \inf_{\beta > 0} \left( \beta \Lambda^* \left( \frac{1}{\beta}, \frac{z}{\beta} \right) \right).$$

Similarly we define  $\Lambda_n^*$  and  $J_n$  replacing  $W_1$  by  $\min(W_1, n)$ . Finally define

$$\tilde{J}(z) = \sup_{\delta > 0} \liminf_{n \rightarrow \infty} \inf_{|y-z| < \delta} J_n(y).$$

Thanks to [Proposition 2.6](#) we may apply Theorem 2.4 in [7], telling us that the distributions of  $\hat{N}_t^h/t$  satisfy asymptotic deviation inequalities.

**Theorem 2.12.** *Recall that  $\theta_0$  is defined in [Proposition 2.6\(ii\)](#).*

- If  $\theta_0 = +\infty$ , the laws of the family  $\hat{N}_t^h/t$  satisfy a large deviation principle with rate function  $\tilde{J}$ , i.e

(1) for any closed set  $C \in \mathbb{R}$ ,

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \ln \mathbb{P} \left( \hat{N}_t^h/t \in C \right) \leq - \inf_{m \in C} \tilde{J}(m)$$

(2) for any open set  $\mathcal{O} \in \mathbb{R}$ ,

$$\liminf_{t \rightarrow \infty} \frac{1}{t} \ln \mathbb{P} \left( \hat{N}_t^h / t \in \mathcal{O} \right) \geq - \inf_{m \in \mathcal{O}} \tilde{J}(m),$$

- If  $\theta_0 < +\infty$ , denoting  $m = \mathbb{E}(W_1)/\mathbb{E}(\tau_1)$  we have for all  $a > 0$  and all  $\kappa \in (0, 1)$

$$\limsup_{t \rightarrow +\infty} \frac{1}{t} \ln \mathbb{P} \left( \frac{\hat{N}_t^h}{t} \geq m + a \right) \leq - \min \left[ \inf_{z \geq m + \kappa a} J(z), (1 - \kappa)\theta_0 a \right],$$

and

$$\limsup_{t \rightarrow +\infty} \frac{1}{t} \ln \mathbb{P} \left( \frac{\hat{N}_t^h}{t} \leq m - a \right) \leq - \min \left[ \inf_{z \leq m - \kappa a} J(z), (1 - \kappa)\theta_0 a \right].$$

The latter deviation inequalities are obtained using that  $J \leq \tilde{J}$  (see [7]). The proof of Theorem 2.4 in [7] is inspired by the one in [21] and uses a contraction principle applied to the Large Deviation Principle (LDP for short) at the level of the empirical process. The full LDP is obtained when all exponential moments exist. Otherwise we mainly obtain the upper bound, which is enough to get asymptotic deviations results, as explained in the Theorem below.

**Corollary 2.13.** Recall that  $\theta_0$  is defined in Proposition 2.6(ii).

- (1) If  $\theta_0 = +\infty$ ,  $N_t^h/t$  satisfies the same LDP as  $\hat{N}_t^h/t$ .
- (2) If  $\theta_0 < +\infty$ , we have for all  $a > 0$  and  $\kappa \in (0, 1)$

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \ln \mathbb{P} \left( \frac{N_t^h}{t} \geq m + a \right) \leq - \min \left[ \inf_{z - m \geq \kappa a} J(z), \frac{1 - \kappa}{2} \theta_0 a \right], \quad (2.10)$$

Similarly

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \ln \mathbb{P} \left( \frac{N_t^h}{t} \leq m - a \right) \leq - \min \left[ \inf_{m - z \leq \kappa a} J(z), \frac{(1 - \kappa)}{2} \theta_0 a \right]. \quad (2.11)$$

for  $\kappa \in (0, 1)$ .

**Remark 2.14.** Once again we may get an explicit expression for the rate function in the canceling intensity case  $h = -\lambda \mathbb{1}_{[0, A]}$ . Since  $W_1 = 1$  and  $\tau_1 - A$  is an exponential variable with parameter  $\lambda$ , we have

$$\beta \Lambda^* \left( \frac{1}{\beta}, \frac{m}{\beta} \right) = \sup_{x, y} \left( x + (m - \beta)y + \beta \ln \left( 1 - \frac{x}{\lambda} \right) - \beta A x \right).$$

Notice that for a given  $x$ ,  $\sup_y (x + (m - \beta)y + \beta \ln (1 - \frac{x}{\lambda}) - \beta A x) < +\infty$  if and only if  $\beta = m$  due to the linear term in  $y$ . We deduce

$$J(m) = m \Lambda^* \left( \frac{1}{m}, 1 \right).$$

It easily follows

$$J(m) = \lambda(1 - mA) - m + m \ln \left( \frac{m}{\lambda(1 - mA)} \right).$$

◇

**Remark 2.15** (*Added to the Proof*). After completing the present paper and [7], another proof of a LDP for cumulative processes was proposed in the preprint [26]. Some of the results in [26] should possibly complete the picture when  $\theta_0 < +\infty$ .  $\diamond$

### 3. One more example with explicit calculations: canceling intensity with delay

We already discussed in Remark 2.8 the *canceling of intensity* case  $h = -\lambda \mathbb{1}_{[0,A]}$ .

In our second example we add a delay to the previous case: the inhibition only occurs after a lag period of length  $r > 0$ . Let  $\lambda > 0$ ,  $r > 0$  and  $A > r$  we consider  $h = -\lambda \mathbb{1}_{[r,r+A]}$ . Then  $L(h) = r + A$ . We can again explicitly compute the law of  $W_i$  and  $\tau_i$ .

We can summarize the results of this two cases and apply Theorem 2.9 to obtain

**Proposition 3.1.** *Let us consider  $A > 0$  and  $r \geq 0$ . The Hawkes process associated with  $h = -\lambda \mathbb{1}_{[r,r+A]}$  satisfies*

$$\lim_{t \rightarrow \infty} \frac{N_t^h}{t} = \frac{\lambda(1 + \lambda r)}{\lambda A + 2\lambda t + e^{-r}} \quad a.s.$$

#### Remark 3.2.

This result naturally leads to some comments on the issues brought by inhibition.

- Let us first remark that as  $r \rightarrow 0$  we recover the result of the canceling intensity case given in Remark 2.8.
- Secondly we wonder whether one of both examples admits more points asymptotically. Therefore we are lead to study the ratio

$$\frac{\frac{\lambda}{\lambda A + 1}}{\frac{\lambda(1 + \lambda r)}{\lambda A + 2\lambda r + e^{-\lambda r}}} = \frac{\lambda A + 2\lambda r + e^{-\lambda r}}{(\lambda A + 1)(1 + \lambda r)},$$

or equivalently the sign of

$$e^{-\lambda r} - 1 + \lambda r - \lambda^2 A r = \lambda^2 r \left( \frac{r}{2} - A \right) + \sum_{k=3}^{\infty} \frac{(-\lambda r)^k}{k!}.$$

using the series expansion of the exponential. We therefore deduce that since  $A > r$ , the right hand side is negative, and thus the ratio is less than 1. Consequently, this proves that the lag induces asymptotically more points in the inhibited Hawkes process. Note that even if our proof only holds in the case of  $A > r$  since in this case there is a single time interval with canceled intensity between successive windows (see the proof), we expect that delay in inhibition will also induce more points in more general framework.

Notice finally that  $\|h\|_1$  is the same in the delayed and the non delayed case, therefore we deduce that the asymptotic proportion of points is not fully characterized by  $\|h\|_1$  and that formula (1.2) cannot be extended to the inhibited setting.

$\diamond$

**Proof of Proposition 3.1.** Let us study  $\Lambda^h$  on the time interval  $[U_1^1, U_1^1 + r + A]$ :

- for  $t \in [U_1^1, U_1^1 + r)$ , then for any  $u \in (0, t)$ ,  $t - u$  belongs to  $(0, t)$  and thus

$$\Lambda^h(t) = \left( \lambda + \int_0^t -\lambda \mathbb{1}_{(t-u) \in [r,r+A]} N^h(du) \right)^+ = \lambda,$$

- for  $t \in [U_1^1 + r, U_1^1 + r + A]$ , then

$$\Lambda^h(t) = \left( \lambda + \int_{(0,t)} -\lambda \mathbb{1}_{(t-u) \in [r, r+A]} N^h(du) \right)^+ \leq \left( \lambda - \lambda \mathbb{1}_{(t-U_1^1) \in [r, r+A]} \right)^+ = 0.$$

From this, we deduce that all the points of  $N^h$  in  $]U_1^1, U_1^1 + r + A]$  actually belong to the interval  $]U_1^1, U_1^1 + r[$ . In particular, if  $N^h$  has no points in  $]U_1^1, U_1^1 + r + A]$ , then  $W_1 = 1$  and  $\tau_1 = U_1^1 + r + A$ .

Let us now remark that  $N^h([U_1^1, U_1^1 + r])$  follows a Poisson law of parameter  $\lambda r$  since the intensity of Hawkes process is constant on this interval. In particular  $N^h([U_1^1, U_1^1 + r])$  is finite almost surely. More generally for any  $1 < k \leq N^h([U_1^1, U_1^1 + r])$ , then  $U_k^1 \leq U_1^1 + r$  and

$$\forall t \in [U_k^1 + r, U_k^1 + r + A], \quad \Lambda^h(t) = 0.$$

Finally, since  $A > r$  we have that  $U_k^1 + r \leq U_1^1 + r + r \leq U_1^1 + r + A$ , and thus the intensity remains null on the interval  $[U_1^1 + r, U_k^1 + r + A]$ .

We can conclude that

$$W_1 = N^h([U_1^1, U_1^1 + r]), \quad (3.1)$$

$$\tau_1 = U_{W_1}^1 + r + A. \quad (3.2)$$

Since the first point in  $N^h$  in the interval  $[U_1^1, U_1^1 + r]$  is in  $U_1^1$  we actually have

$$W_1 = 1 + N^h((U_1^1, U_1^1 + r)).$$

It follows that

$$W_1 - 1 \sim \mathcal{P}(\lambda r)$$

and

$$\mathbb{E}(W_1) = 1 + \lambda r, \quad \text{Var}(W_1) = \lambda r \quad \text{and} \quad \theta_0 = +\infty.$$

We finally study the law of  $\tau_1$ . From Eq. (3.2), we can write

$$\begin{aligned} \tau_1 &= U_{W_1}^1 + r + A \\ &= r + A + U_1^1 + (U_{W_1}^1 - U_1^1), \end{aligned}$$

where  $U_1^1 \sim \mathcal{E}(\lambda)$  by Proposition 2.5 and  $U_1^1$  and  $(U_{W_1}^1 - U_1^1)$  are independent.

It remains to study the law of  $(U_{W_1}^1 - U_1^1)$ .

Thanks to (3.1),  $0 \leq U_{W_1}^1 - U_1^1 \leq r$ .

Let  $t \in [0, r]$ , we have:

$$\mathbb{P}\left(0 \leq U_{W_1}^1 - U_1^1 \leq t\right) = \sum_{k=1}^{\infty} \mathbb{P}\left(\{W_1 = k\} \cap \{U_{W_1}^1 - U_1^1 \leq t\}\right).$$

For  $k = 1$ :  $\mathbb{P}\left(\{W_1 = 1\} \cap \{U_{W_1}^1 - U_1^1 \leq t\}\right) = \mathbb{P}(W_1 = 1) = e^{-\lambda r}$ .

For  $k > 1$  since the intensity of the Hawkes process remains constant equal to  $\lambda$  on  $[U_1^1, U_{W_1}^1]$  we can write

$$U_{W_1}^1 - U_1^1 \stackrel{(law)}{=} \sum_{k=1}^{W_1-1} T_k$$

where  $(T_k)_{k \in \mathbb{N}}$  is a sequence of i.i.d  $\mathcal{E}(\lambda)$ . We can consider  $(T_k)_{k \in \mathbb{N}}$  as the interarrival times of a Poisson process of parameter  $\lambda$  coupled with our Hawkes process, as in [Proposition 2.2](#). Then,  $T_0 = U_1^1$ , and  $(T_k)_{k \geq W_1}$  are defined. Then:

$$\begin{aligned} \mathbb{P}(\{W_1 = k\} \cap \{U_{W_1}^1 - U_1^1 \leq t\}) &= \mathbb{P}\left(\left\{0 \leq \sum_{i=1}^{k-1} T_i \leq t\right\} \cap \left\{T_k + \sum_{i=1}^{k-1} T_i > r\right\}\right) \\ &= \mathbb{E}\left[\mathbb{1}_{0 \leq \sum_{i=1}^{k-1} T_i \leq t} \mathbb{P}\left(T_k + \sum_{i=1}^{k-1} T_i > r \mid (T_1, \dots, T_{k-1})\right)\right] \\ &= \mathbb{E}\left[\mathbb{1}_{0 \leq \sum_{i=1}^{k-1} T_i \leq t} e^{-\lambda(r - \sum_{i=1}^{k-1} T_i)}\right] \\ &= \int_{(\mathbb{R}^+)^{k-1}} \mathbb{1}_{0 \leq \sum_{i=1}^{k-1} s_i \leq t} \lambda^{k-1} e^{-\lambda \sum_{i=1}^{k-1} s_i} \times e^{-\lambda(r - \sum_{i=1}^{k-1} s_i)} ds_2 \dots ds_k \\ &= e^{-\lambda r} \lambda^{k-1} I_{k-1}(t) \\ &= e^{-\lambda r} \frac{(\lambda t)^{k-1}}{(k-1)!}. \end{aligned}$$

with

$$I_k(t) := \int_{(\mathbb{R}^+)^k} \mathbb{1}_{0 \leq \sum_{i=1}^k s_i \leq t} ds_1 \dots ds_k = \frac{t^k}{k!}.$$

Thus

$$\begin{aligned} \mathbb{P}\left(0 \leq U_{W_1}^1 - U_1^1 \leq t\right) &= e^{-\lambda r} + \sum_{k \geq 2} e^{-\lambda r} \frac{(\lambda t)^{k-1}}{(k-1)!} \\ &= e^{-\lambda(r-t)}. \end{aligned}$$

Hence the distribution of  $U_{W_1}^1 - U_1^1$  is given by  $e^{-\lambda r} \delta_0 + \lambda e^{-\lambda(r-t)} \mathbb{1}_{(0,r]}(t) dt$ .

An easy computation gives  $\mathbb{E}(U_{W_1}^1 - U_1^1) = r - \frac{1}{\lambda}(1 - e^{-\lambda r})$ . Finally we obtain that

$$\begin{aligned} \mathbb{E}(\tau_1) &= r + A + \mathbb{E}(U_1^1) + \mathbb{E}(U_{W_1}^1 - U_1^1) \\ &= r + A + \lambda^{-1} + r - \lambda^{-1}(1 - e^{-\lambda r}) \\ &= 2r + A + \lambda^{-1} e^{-\lambda r} \end{aligned}$$

From [Theorem 2.9](#) we obtain the following LLN

$$\frac{N_t^h}{t} \xrightarrow[t \rightarrow \infty]{a.s.} \frac{1 + \lambda r}{2r + A + \lambda^{-1} e^{-\lambda r}} = \frac{\lambda(1 + \lambda r)}{\lambda A + 1 + (e^{-\lambda r} - 1) + 2\lambda r}. \quad \square$$

## 4. Proofs

### 4.1. Proofs of [Propositions 2.6](#) and [2.7](#)

We start by proving that the random variables  $\tau$  and  $W$  admit exponential moments.

**Proof of [Proposition 2.6](#).** Let  $h$  be a signed measurable function and  $h^+$  its positive part. We generate  $N^h$  and  $N^{h^+}$  by coupling as in [Proposition 2.2](#). Recall that  $\|h^+\|_1 < 1$ .

We denote by  $W_i$ ,  $\tau_i$ ,  $S_i$ , ... (respectively  $W_i^+$ ,  $\tau_i^+$ ,  $S_i^+$ , ...) the renewal quantities associated to  $N^h$  (resp.  $N^{h^+}$ ). Be careful that the previous construction of  $W_i^+$ ,  $\tau_i^+$ ,  $S_i^+$  is done by

using intervals of length  $L(h)$  not  $L(h^+)$ . Notice that since  $L(h) \geq L(h^+)$ , then the renewal structure is well defined for  $N^{h^+}$ . Moreover, if  $h \leq 0$ ,  $L(h^+) = 0$ , one can replace  $h^+$  by  $h_\varepsilon^+ = h^+ + \varepsilon \mathbb{1}_{[0, L(h)]}$  and then let  $\varepsilon$  go to 0 in order to compare with [8].

Thanks to Proposition 2.2, we have  $N^h \leq N^{h^+}$  a.s. We also know that  $U_1^1 = U_1^{+,1}$ .

Moreover,  $\tau_1 \leq \tau_1^+$  a.s. because the jumps of  $N^h$  are included in those of  $N^{h^+}$ . We also have  $W_1 = N^h([0, \tau_1]) \leq N^h([0, \tau_1^+]) \leq N^{h^+}([0, \tau_1^+]) = W_1^+$  a.s. So  $W_1 \leq W_1^+$  a.s.

Study of  $N^{h^+}$ :

First, we focus on  $N^{h^+}$ . According to [8], we can associate a  $M/G/\infty$  queue to  $N^{h^+}$ . To do this, we consider:

$$\Lambda^{h^+}(t) = \lambda + \int_{(-L(h), t]} h^+(t-u) N^{h^+}(du).$$

We can consider the Hawkes process as the sum of:

- the arrivals of ancestors  $V_k$  at rate  $\lambda$  and
- a continuous time Galton–Watson process beginning at each  $V_k$  with the following characteristics: the number of descendants follows a Poisson distribution with mean  $\|h^+\|_1$  and the times of births have the density  $h^+/\|h^+\|_1$ .

In fact, to each arrival of an ancestor  $V_k$ , we can associate a time  $H_k$  corresponding to the life time of the cluster of  $V_k$ .  $V_k$  is independent of  $H_k$  and the  $(H_k)_k$ 's are independent.

We can associate to this process a queue in the following way:

- the customers are the ancestors and arrive at rate  $\lambda$ ,
- the service time for each customer is  $H_k + L(h)$ .

We denote by  $Y_t$  the number of customers in the queue at time  $t$ :

$$Y_t = \sum_k \mathbb{1}_{V_k \leq t < V_k + H_k + L(h)}.$$

Let  $\mathcal{T}_1^+ = \inf\{t \geq 0, Y_{t-} \neq 0, Y_t = 0\}$ , be the first time the queue is empty. By proposition 2.6 of [8], we have:

$$\forall \alpha < \alpha_0 := \min\left(\lambda, \frac{\|h^+\|_1 - \log(\|h^+\|_1) - 1}{L(h)}\right), \quad \text{it holds} \quad \mathbb{E}[e^{\alpha \mathcal{T}_1^+}] < \infty.$$

Of course  $\lambda > 0$  and  $\|h^+\|_1 - \log(\|h^+\|_1) - 1 > 0$ , and so  $\mathcal{T}_1^+$  admits an exponential moment.

Since  $\tau_1^+$  is the first time after  $U_1^{+,1}$  such that there were no jump during a time  $L(h)$ . Thus  $\tau_1^+ = \mathcal{T}_1^+$  and since  $\tau_1 \leq \tau_1^+$ , part (i) of the proposition is proved.

In order to prove (ii) it is enough to show that the distribution of  $W_1^+$  admits exponential moments. Recall that

$$W_1^+ = N^{h^+}([0, \tau_1^+]).$$

According to [4] (see proof of Theorem 3.2 and proof of Theorem 3.4 therein),

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \ln \mathbb{E}\left(e^{\theta N^{h^+}([0, t])}\right) = \lambda(\mathbb{E}(e^{\theta S}) - 1) := \mu(\theta) < +\infty$$

as soon as  $\theta < \|h^+\|_1 - \log(\|h^+\|_1) - 1$ . Here  $S$  is distributed according to (see (3) in [4])

$$\mathbb{P}(S = k) = \frac{e^{-k\|h^+\|_1} (k\|h^+\|_1)^{k-1}}{k!}.$$

It is thus immediate that  $\mu(\theta)$  goes to 0 as  $\theta$  goes to 0.



For  $\varepsilon > 0$  we may thus choose  $\theta$  small enough such that

$$\alpha_0 - 2\varepsilon \geq \mu(2\theta) + \varepsilon.$$

For this  $\theta$ , one can find  $t_\theta$  such that for  $t \geq t_\theta$ ,

$$\mathbb{E}\left(e^{2\theta N^{h^+}([0,t])}\right) \leq e^{t(\mu(2\theta)+\varepsilon)}.$$

It follows

$$\begin{aligned} \mathbb{E}(e^{\theta W_1^+}) &= \mathbb{E}\left(e^{\theta N^{h^+}([0,\tau_1^+])}\right) \\ &\leq \sum_{k=1}^{\infty} \mathbb{E}\left(e^{\theta N^{h^+}([0,k])} \mathbb{1}_{k-1 \leq \tau_1^+ < k}\right) \\ &\leq \sum_{k=1}^{\infty} \left(\beta_k \mathbb{E}\left(e^{2\theta N^{h^+}([0,k])}\right) + \frac{1}{\beta_k} \mathbb{P}(k-1 \leq \tau_1^+)\right) \\ &\leq A(t_\theta) + \sum_{k=[t_\theta]+1}^{\infty} \left(\beta_k e^{k(\mu(2\theta)+\varepsilon)} + \frac{\mathbb{E}(e^{(\alpha_0-\varepsilon)\tau_1^+})}{\beta_k} e^{-(k-1)(\alpha_0-\varepsilon)}\right) \end{aligned}$$

where  $A(t_\theta)$  denotes the finite sum up to  $k = [t_\theta]$ . Choosing  $\beta_k = k^{-2} e^{-k(\mu(2\theta)+\varepsilon)}$  the  $k$ 'th term of the remaining sum is smaller than  $1/k^2 + c k^2 e^{-\varepsilon(k-1)}$  and the series is thus convergent. Since  $\varepsilon$  is arbitrary, (ii) follows.  $\square$

**Proof of Proposition 2.7.** We consider a process  $N^h$ , generated by the Poisson point process  $Q$ , as in Proposition 2.2. Since  $h \leq 0$ , we will couple (and upper-bound) this time the Hawkes process with the Poisson point process  $\mathcal{R}$  on  $\mathbb{R}^+$ , with intensity  $\lambda$ , generated by the same Poisson point process  $Q$  on  $(0, \infty)^2$ . Since  $\forall t \geq 0$

$$\lambda \geq \Lambda^h(t) \quad \text{a.s.}$$

we deduce that

$$\mathcal{R} \geq N^h.$$

We can now upper bound the length of the first time window  $\tau_1$  by a similar quantity associated with  $\mathcal{R}$ . Recall that  $U_1^1$  is the first jump time of  $N^h$  and define:

$$\tau = \inf\{t > U_1^1, \mathcal{R}[t - L(h), t] \neq 0, \mathcal{R}(t - L(h), t] = 0\}. \quad (4.1)$$

$\tau$  indicates the first moment such that there were no jump for the process  $\mathcal{R}$  during an interval of length  $L(h)$ . In particular, there were not jump for  $N^h$  either. Therefore  $\tau_1 \leq \tau$  a.s. and

$$W_1 = N^h([U_1^1, \tau_1]) \leq \mathcal{R}([U_1^1, \tau_1]) \leq \mathcal{R}([0, \tau]).$$

It is thus enough to get an upper bound for  $\mathbb{E}\left(e^{\theta \mathcal{R}([0,\tau])}\right)$ .

To this end we shall study the random variable  $\tau$ . Denote by  $V_i$  the jumps of the Poisson point process  $\mathcal{R}$ . From the definition there exists a random integer  $K$  such that

$$\tau = V_K + L(h).$$

The definition of  $K$  leads to

$$K = \mathcal{R}[0, \tau].$$

From the independence of the times between jumps of  $\mathcal{R}$  we deduce that

$$\begin{aligned} \mathbb{P}(K = 1) &= \mathbb{P}[\tau = V_1 + L(h)] = \mathbb{P}[V_2 - V_1 \geq L(h)] \\ &= e^{-\lambda L(h)}, \\ \mathbb{P}(K = 2) &= \mathbb{P}[\tau = V_2 + L(h)] = \mathbb{P}[\{V_2 - V_1 < L(h)\} \cup \{V_3 - V_2 \geq L(h)\}] \\ &= \mathbb{P}[V_2 - V_1 < L(h)] \mathbb{P}[V_3 - V_2 \geq L(h)] \\ &= (1 - e^{-\lambda L(h)}) e^{-\lambda L(h)}, \\ \forall k \geq 2, \quad \mathbb{P}(K = k) &= \mathbb{P}[\tau = V_k + L(h)] = (1 - e^{-\lambda L(h)})^{k-1} e^{-\lambda L(h)}. \end{aligned}$$

$K$  is a geometric random variable with parameter  $e^{-\lambda L(h)}$  and thus admits exponential moments provided  $e^\theta (1 - e^{-\lambda L(h)}) < 1$  which concludes the proof.  $\square$

#### 4.2. Proof of the LLN and CLT

**Proof of Theorems 2.9 and 2.10.** Recall that

$$\hat{N}_t^h \leq N_t^h \leq \hat{N}_t^h + W_{M_t^h+1} \quad \text{a.s.}$$

where

$$\hat{N}_t^h := \sum_{i=1}^{\infty} W_i \mathbb{1}_{S_i \leq t} = \sum_{i=1}^{M_t^h} W_i$$

and

$$M_t^h := \sum_{i=1}^{\infty} \mathbb{1}_{S_i \leq t},$$

as explained in (2.6), (2.7), (2.9).

As we previously said Theorem 3.1 and Theorem 3.2 Chapter 6 in [1] furnish

$$\frac{\hat{N}_t^h}{t} \xrightarrow[t \rightarrow \infty]{a.s.} \frac{\mathbb{E}[W_1]}{\mathbb{E}(\tau_1)}$$

and

$$\sqrt{t} \left( \frac{\hat{N}_t^h}{t} - \frac{\mathbb{E}[W_1]}{\mathbb{E}(\tau_1)} \right) \xrightarrow[t \rightarrow \infty]{} \mathcal{N}(0, \sigma^2)$$

with

$$\sigma^2 = \frac{\text{Var} \left( W_1 - \tau_1 \frac{\mathbb{E}[W_1]}{\mathbb{E}(\tau_1)} \right)}{\mathbb{E}(\tau_1)}.$$

It is thus enough to control the remaining (or error) term  $W_{M_t^h+1}$  i.e to prove

$$\lim_{t \rightarrow \infty} \frac{W_{M_t^h+1}}{t} = 0 \quad \text{a.s.}, \quad (4.2)$$

and

$$\lim_{t \rightarrow \infty} \frac{W_{M_t^h+1}}{\sqrt{t}} = 0 \quad \text{in Probability}. \quad (4.3)$$

(4.2) will conclude the LLN and (4.3) the CLT.

Actually we will prove stronger results. Let  $\beta(n)$  an increasing sequence going to infinity and  $\varepsilon > 0$ . Introduce the independent events  $A_n = \{W_n > \varepsilon\beta(n)\}$ . Then  $\limsup_n A_n = \{\limsup_n \frac{W_n}{\beta(n)} > \varepsilon\}$ . Since the  $(W_i)_i$  are finite i.i.d random variables

$$\sum_n \mathbb{P}(A_n) = \sum_n \mathbb{P}(W_n > \varepsilon\beta(n)) = \sum_n \mathbb{P}(W_1 > \varepsilon\beta(n)).$$

Thanks to Proposition 2.6 and to Markov inequality, we know that for

$$\mathbb{P}(W_1 > \varepsilon\beta(n)) \leq \mathbb{E}[e^{\theta_0 W_1}] e^{-\theta_0 \varepsilon \beta(n)}.$$

We may now apply Borel–Cantelli, telling that provided  $\sum_n e^{-\theta_0 \varepsilon \beta(n)} < +\infty$ ,

$$\mathbb{P}(\limsup_n A_n) = 0.$$

The previous holds with  $\beta(n) = n^\alpha$  for any  $\alpha > 0$ . We have proved in particular that

$$\frac{W_n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{a.s.} 0.$$

Since  $M_t^h$  is a non-decreasing family of integers going to infinity almost surely,

$$\frac{W_{M_t^h+1}}{\sqrt{M_t^h+1}} \xrightarrow[t \rightarrow \infty]{a.s.} 0.$$

It remains to recall that

$$\frac{M_t^h+1}{t} \xrightarrow[t \rightarrow \infty]{} \frac{1}{\mathbb{E}(\tau_1)} \quad \text{a.s.}, \quad (4.4)$$

to conclude that

$$\frac{W_{M_t^h+1}}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{a.s.} 0. \quad \square$$

#### 4.3. Proof of Corollary 2.13

**Proof of Corollary 2.13.** In order to prove the first part of Corollary 2.13 it is enough to show that  $\hat{N}_t^h/t$  and  $N_t^h/t$  are exponentially equivalent, i.e. that for each  $\delta > 0$ ,

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \ln \mathbb{P} \left( \left| \frac{N_t^h}{t} - \frac{\hat{N}_t^h}{t} \right| > \delta \right) = -\infty.$$

To this end it is enough to show that for all  $\delta > 0$

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \ln \mathbb{P} \left( W_{M_t^h+1} > \delta t \right) = -\infty. \quad (4.5)$$

We will decompose the probability space into two events:  $M_t^h \leq t^2$  and  $M_t^h > t^2$ . It holds

$$\begin{aligned} \mathbb{P}(W_{M_t^h+1} > \delta t) &\leq \mathbb{P}(M_t^h > t^2) + \mathbb{P}(\{W_{M_t^h+1} > \delta t\} \cap \{M_t^h \leq t^2\}) \\ &\leq \mathbb{P}(M_t^h > t^2) + \mathbb{P}(\{\exists k \in \{1, \dots, \lfloor t^2 + 1 \rfloor\}, W_k > \delta t\} \cap \{M_t^h \leq t^2 + 1\}) \\ &\leq \mathbb{P}(M_t^h > t^2) + \mathbb{P}(\{\exists k \in \{1, \dots, \lfloor t^2 + 1 \rfloor\}, W_k > \delta t\}) \\ &\leq \mathbb{P}(M_t^h > t^2) + \sum_{j=1}^{\lfloor t^2+1 \rfloor} \mathbb{P}(W_j > \delta t) \\ &\leq \mathbb{P}(M_t^h > t^2) + (t^2 + 1)\mathbb{P}(W_1 > \delta t). \end{aligned}$$

On one hand, we have, by Markov's inequality, for all  $\theta_0 > \theta > 0$ ,

$$\mathbb{P}(W_1 > \delta t) \leq \mathbb{E}[e^{\theta W_1}] e^{-\theta \delta t}$$

so that

$$\limsup_{t \rightarrow +\infty} \frac{1}{t} \ln(1 + t^2) \mathbb{P}(W_1 > \delta t) \leq -\theta_0 \delta.$$

On the other hand, according to [19] Theorem 2.3, for all  $x > 0$ :

$$\limsup_{t \rightarrow +\infty} \frac{1}{t} \ln \mathbb{P}\left(\frac{M_t^h}{t} \geq x\right) \leq -J_{\tau_1}(x),$$

where  $J_{\tau_1}(x) = \sup_{\eta} \{\eta - x \ln \mathbb{E}[e^{\eta \tau_1}]\}$ .

Since  $\eta \mapsto \mathbb{E}(e^{\eta \tau_1})$  is continuous on  $\mathbb{R}^-$  there exists some  $\eta_0$  such that  $\mathbb{E}(e^{\eta_0 \tau_1}) = e^{-1}$ . It follows  $J_{\tau_1}(x) \geq \eta_0 + x$ .

Choose  $t_1, t_2, \dots$  an increasing sequence of times such that  $t_i \xrightarrow{i \rightarrow +\infty} +\infty$ . For a fixed  $i$ , we have for  $t$  large enough

$$\mathbb{P}(M_t^h > t^2) = \mathbb{P}\left(\frac{M_t^h}{t} > t\right) \leq \mathbb{P}\left(\frac{M_t^h}{t} > t_i\right)$$

Since

$$\limsup_{t \rightarrow +\infty} \frac{1}{t} \ln \mathbb{P}\left(\frac{M_t^h}{t} > t_i\right) \leq -J_{\tau_1}(t_i) \leq -t_i - \eta_0.$$

It follows,

$$\limsup_{t \rightarrow +\infty} \frac{1}{t} \ln \mathbb{P}\left(\frac{M_t^h}{t} > t\right) = -\infty.$$

Eventually,

$$\begin{aligned} \limsup_{t \rightarrow +\infty} \frac{1}{t} \ln \mathbb{P}\left(\frac{W_{M_t^h+1}}{t} > \delta\right) &\leq \limsup_{t \rightarrow +\infty} \frac{1}{t} \ln [\mathbb{P}(M_t^h > t^2) + (t^2 + 1)\mathbb{P}(W_1 > \delta t)] \\ &\leq \limsup_{t \rightarrow +\infty} \left( \frac{\ln 2}{t} + \max \left[ \frac{1}{t} \ln \mathbb{P}\left(\frac{M_t^h}{t} > t\right), \frac{1}{t} \ln ((t^2 + 1)\mathbb{P}(W_1 > \delta t)) \right] \right) \end{aligned}$$

$$\begin{aligned} &\leq \max \left[ \limsup_{t \rightarrow +\infty} \frac{1}{t} \ln \mathbb{P} \left( \frac{M_t^h}{t} > t \right), \limsup_{t \rightarrow +\infty} \left( \frac{\ln(t^2 + 1)}{t} + \frac{1}{t} \ln \mathbb{P}(W_1 > \delta t) \right) \right] \\ &\leq -\theta_0 \delta. \end{aligned}$$

This completes the proof for  $\theta_0 = +\infty$ .

Let us now assume  $\theta_0 < \infty$ . Recall that  $m = \frac{\mathbb{E}(W_1)}{\mathbb{E}(\tau_1)}$ , then (2.10) is a consequence of the following line of reasoning:

$$\begin{aligned} \mathbb{P} \left( \frac{N_t^h}{t} \geq m + a \right) &\leq \mathbb{P} \left( \frac{\hat{N}_t^h}{t} + \frac{W_{M_t^h+1}}{t} \geq m + a \right) \\ &\leq \mathbb{P} \left( \frac{\hat{N}_t^h}{t} \geq m + \kappa_1 a \right) + \mathbb{P} \left( \frac{W_{M_t^h+1}}{t} \geq (1 - \kappa_1) a \right) \end{aligned}$$

where  $\kappa_1 \in (0, 1)$ , yielding

$$\begin{aligned} &\limsup_t \frac{1}{t} \ln \mathbb{P} \left( \frac{N_t^h}{t} \geq m + a \right) \\ &\leq \max \left( \limsup_t \frac{1}{t} \ln \mathbb{P} \left( \frac{\hat{N}_t^h}{t} \geq m + \kappa_1 a \right), \limsup_t \frac{1}{t} \ln \mathbb{P} \left( \frac{W_{M_t^h+1}}{t} \geq (1 - \kappa_1) a \right) \right) \end{aligned}$$

Now applying Theorem 2.12 with  $\kappa_2$  and  $(1 - \kappa_2)$ , we deduce that

$$\begin{aligned} &\limsup_t \frac{1}{t} \ln \mathbb{P} \left( \frac{N_t^h}{t} \geq m + a \right) \\ &\leq \max \left( - \inf_{z-m \geq \kappa_2 \kappa_1 a} J(z), -(1 - \kappa_2) \kappa_1 \theta_0 a, -(1 - \kappa_1) a \theta_0 \right) \end{aligned}$$

yielding the result with  $\kappa = \kappa_1 \kappa_2$  and  $\kappa' = 1 - \kappa_1$ . The condition  $\kappa + 2\kappa' = 1$  arises from the equality of the last two terms.

Finally, (2.11) is a consequence of the same reasoning on  $\hat{N}_t^h \leq N_t^h$ :

$$\mathbb{P} \left( \frac{N_t^h}{t} \leq m - a \right) \leq \mathbb{P} \left( \frac{\hat{N}_t^h}{t} \leq m - a \right)$$

yielding the result.  $\square$

## 5. Final comments

As we recalled in the introduction, in the linear case the LLN, the CLT and the LDP are completely characterized by  $\|h\|_1$ . As we have shown in Section 3, the “almost linear case with inhibition” we are looking at is dramatically different, since the limiting behavior is not fully determined by  $\|h\|_1$  nor even by some moments of  $h$ . The renewal description of the Hawkes process we have used allows us to characterize all these limit theorems in terms of the joint law of  $(\tau_1, W_1)$ . It should be very interesting to link this distribution with  $h$ . As for the non linear self-excited case such a goal seems difficult to reach.

Another interesting direction should be to obtain non asymptotic deviation bounds (or concentration properties). Since the Large Deviation Principle for cumulative processes we have proved in [7] is based on the contraction of a higher level LDP, new methods are necessary for non asymptotic results.

The methods of the paper can be used for more general jump rate functions  $f$ , provided one can generalize the construction of the sequence  $(\tau_i, W_i)$ . This generalization is partly done in [14] in which a regenerative structure is exhibited without the assumption of bounded support for the reproduction function  $h$  and in [24] which exhibit renewal points for non linear Hawkes processes and age-dependent Hawkes processes.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

We warmly acknowledge two referees for their careful reading, their constructive comments and their criticism.

### References

- [1] Søren Asmussen, Applied Probability and Queues, second ed., in: Applications of Mathematics (New York), vol. 51, Springer-Verlag, New York, 2003, Stochastic Modelling and Applied Probability.
- [2] Emmanuel Bacry, Sylvain Delattre, Marc Hoffmann, Jean François Muzy, Some limit theorems for Hawkes processes and application to financial statistics, *Stochastic Process. Appl.* 123 (7) (2013) 2475–2499.
- [3] Emmanuel Bacry, Jean-François Muzy, Hawkes model for price and trades high-frequency dynamics, *Quant. Finance* 14 (7) (2014) 1147–1166.
- [4] Charles Bordenave, Giovanni Luca Torrisi, Large deviations of poisson cluster processes, *Stoch. Models* 23 (4) (2007) 593–625.
- [5] Pierre Brémaud, Laurent Massoulié, Stability of nonlinear hawkes processes, *Ann. Probab.* 24 (3) (1996) 1563–1588.
- [6] Pierre Brémaud, Giovanna Nappo, Giovanni Luca Torrisi, Rate of convergence to equilibrium of marked Hawkes processes, *J. Appl. Probab.* 39 (1) (2002) 123–136.
- [7] Patrick Cattiaux, Laetitia Colombani, Manon Costa, Large deviations principles for cumulative processes and applications, 2021, Available on [arXiv:2109.07800](https://arxiv.org/abs/2109.07800) [math.PR].
- [8] Manon Costa, Carl Graham, Laurence Marsalle, Viet Chi Tran, Renewal in Hawkes processes with self-excitation and inhibition, *Adv. Appl. Probab.* 52 (3) (2020) 879–915.
- [9] Daryl J. Daley, David Vere-Jones, An Introduction to the Theory of Point Processes, second ed., in: Probability and its Applications (New York), vol. I, Springer-Verlag, New York, 2003, Elementary theory and methods.
- [10] Amir Dembo, Ofer Zeitouni, Large Deviations Techniques and Applications, second ed., in: Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin Heidelberg, 2010.
- [11] Aline Duarte, Eva Löcherbach, Guilherme Ost, Stability, convergence to equilibrium and simulation of non-linear Hawkes processes with memory kernels given by the sum of Erlang kernels, *ESAIM Probab. Stat.* 23 (2019) 770–796.
- [12] Xuefeng Gao, Lingjiong Zhu, Functional central limit theorems for stationary Hawkes processes and application to infinite-server queues, *Queueing Syst.* 90 (1–2) (2018) 161–206.
- [13] Fuqing Gao, Lingjiong Zhu, Precise deviations for Hawkes processes, *Bernoulli* 27 (1) (2021) 221–248.
- [14] Carl Graham, Regenerative properties of the linear Hawkes process with unbounded memory, *Ann. Appl. Probab.* 31 (6) (2021) 2844–2863.
- [15] Niels Richard Hansen, Patricia Reynaud-Bouret, Vincent Rivoirard, Lasso and probabilistic inequalities for multivariate point processes, *Bernoulli* 21 (1) (2015) 83–143.
- [16] Alan G. Hawkes, Spectra of some self-exciting and mutually exciting point processes, *Biometrika* 58 (1971) 83–90.
- [17] Alan G. Hawkes, Hawkes processes and their applications to finance: a review, *Quant. Finance* 18 (2) (2018) 193–198.
- [18] Alan G. Hawkes, David Oakes, A cluster process representation of a self-exciting process, *J. Appl. Probab.* 11 (1974) 493–503.
- [19] Tie Feng Jiang, Large deviations for renewal processes, *Stochastic Process. Appl.* 50 (1) (1994) 57–71.

- [20] Régis Lambert, Christine Tuleau-Malot, Thomas Bessaih, Vincent Rivoirard, Yann Bouret, Nathalie Leresche, Patricia Reynaud-Bouret, Reconstructing the functional connectivity of multiple spike trains using Hawkes models, *J. Neurosci. Methods* 297 (2018) 9–21.
- [21] Raphaël Lefevre, Mauro Mariani, Lorenzo Zambotti, Large deviations for renewal processes, *Stochastic Process. Appl.* 121 (10) (2011) 2243–2271.
- [22] Eva Löcherbach, Spiking neurons: interacting hawkes processes, mean field limits and oscillations, in: Journées MAS 2016 de la SMAI—Phénomènes Complexes Et Hétérogènes, in: ESAIM Proc. Surveys, vol. 60, EDP Sci., Les Ulis, 2017, pp. 90–103.
- [23] Laurent Massoulié, Stability results for a general class of interacting point processes dynamics, and applications, *Stochastic Process. Appl.* 75 (1) (1998) 1–30.
- [24] Mads Bonde Raad, Renewal time points for hawkes processes, 2019, Available on [arXiv:1906.02036](https://arxiv.org/abs/1906.02036) [math.PR].
- [25] Patricia Reynaud-Bouret, Vincent Rivoirard, Franck Grammont, Christine Tuleau-Malot, Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis, *J. Math. Neurosci.* 4 (2014) 3, 41.
- [26] Marco Zamparo, Large deviation principles for renewal–reward processes, 2021, Available on [arXiv:2111.01679](https://arxiv.org/abs/2111.01679) [math.PR].
- [27] Lingjiong Zhu, Central limit theorem for nonlinear Hawkes processes, *J. Appl. Probab.* 50 (3) (2013) 760–771.
- [28] Lingjiong Zhu, Process-level large deviations for nonlinear Hawkes point processes, *Ann. Inst. H. Poincaré Probab. Statist.* 50 (2014) 845–871.
- [29] Lingjiong Zhu, Large deviations for Markovian nonlinear Hawkes processes, *Ann. Appl. Probab.* 25 (2) (2015) 548–581.