Renewal in Hawkes processes with self-excitation and inhibition

Manon Costa, Carl Graham, Laurence Marsalle, Viet-Chi Tran

Renewal in Hawkes processes with self-excitation and inhibition

Manon Costa[*], Carl Graham[†], Laurence Marsalle[‡], Viet Chi Tran[§]

January 15, 2018

Abstract

This paper investigates Hawkes processes on the positive real line exhibiting both self-excitation and inhibition. Each point of this point process impacts its future intensity by the addition of a signed reproduction function. The case of a nonnegative reproduction function corresponds to self-excitation, and has been widely investigated in the literature. In particular, there exists a cluster representation of the Hawkes process which allows to apply results known for Galton-Watson trees. In the present paper, we establish limit theorems for Hawkes process with signed reproduction functions by using renewal techniques. We notably prove exponential concentration inequalities, and thus extend results of Reynaud-Bouret and Roy [24] which were proved for nonnegative reproduction functions using this cluster representation which is no longer valid in our case. An important step for this is to establish the existence of exponential moments for renewal times of M/G/∞ queues that appear naturally in our problem. These results have their own interest, independently of the original problem for the Hawkes processes.

MSC 2010 subject classification: 60G55, 60F99, 60K05, 60K25, 44A10.

Key-words: Point processes, self-excitation, inhibition, renewal theory, ergodic limit theorems, concentration inequalities, Galton-Watson trees, M/G/∞ queues.

Acknowledgements: The authors thank Patricia Reynaud-Bouret for introducing them to Hawkes processes and for proposing the problem of inhibition in Hawkes processes. All the authors have been supported by the Chair "Modélisation Mathématique et Biodiversité" of Veolia Environnement-Ecole Polytechnique-Museum National d'Histoire Naturelle-Fondation X. L.M. and V.C.T. acknowledge support from Labex CEMPI (ANR-11-LABX-0007-01).

[*]Institut de Mathématiques de Toulouse, UMR 5219; Université de Toulouse, CNRS, UPS IMT, F-31062 Toulouse Cedex 9, France; E-mail: manon.costa@math.univ-toulouse.fr

[†]Centre de Mathématiques Appliquées, École Polytechnique, Route de Saclay, 91120 Palaiseau; E-mail: carl.graham@polytechnique.edu

[‡]Univ. Lille, CNRS, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille, France; E-mail: laurence.marsalle@math.univ-lille1.fr

[§]Univ. Lille, CNRS, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille, France; E-mail: chi.tran@math.univ-lille1.fr

1 Introduction and main results

Hawkes processes have been introduced by Hawkes [16], and are now widely used in many applications, for example: modelling of earthquake occurrences [17, 22], finance [1, 2, 3], genetics [25], neurosciences [8, 13, 23]. Hawkes processes are random point processes on the line (see [9, 10, 19] for an introduction) where each atom is associated with a (possibly signed) reproduction measure generating further atoms or adding repulsion. When the reproduction measure is nonnegative, Hawkes and Oakes [18] have provided a cluster representation of the Hawkes processes based on immigration of ancestors, each of which is at the head of the branching point process of its offspring. Exponential concentration inequalities for ergodic theorems and tools for statistical applications have been developed, *e.g.*, by Reynaud-Bouret and Roy [24] in this case by using a coupling *à la* Berbee [4].

For many applications however, it is important to allow the reproduction measure to be a signed measure. The positive part of the measure can be interpreted as self-excitation, and its negative part as self-inhibition. For instance, in neurosciences this can be used to model the existence of a latency period before the successive activations of a neuron.

A large part of the literature on Hawkes processes for neurosciences uses large systems approximations by mean-field limits (*e.g.*, [7, 12, 11, 13]) or stabilization properties (*e.g.*, [14] using Erlang kernels). Here, we consider a single Hawkes process for which the reproduction measure is a signed measure and aim to extend the ergodic theorem and deviation inequalities obtained in [24] for a nonnegative reproduction measure.

A main issue here is that when inhibition is present then the cluster representation of [18] is no longer valid. An important tool in our study is a coupling construction of the Hawkes process with signed reproduction measure and of a Hawkes process with a positive measure. The former is shown to be a thinning of the latter, for which the cluster representation is valid.

We then define renewal times for these general Hawkes processes. We introduce an auxiliary strong Markov process for this purpose. This allows to split the sample paths into i.i.d. distributed excursions, and use limit theorems for i.i.d. sequences.

In order to obtain concentration inequalities, a main difficulty is to obtain exponential bounds for the tail distribution of the renewal times. In the case in which the reproduction function is nonnegative, we associate to the Hawkes process a $M/G/\infty$ queue, and control the length of its excursions using Laplace transform techniques. These results have independent interest in themselves. We then extend our techniques to Hawkes processes with signed reproduction functions using the coupling.

1.1 Definitions and notations

Measure-theoretic and topological framework

Throughout this paper, an appropriate filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual assumptions is given. All processes will be assumed to be adapted.

Let $\mathcal{N}(\mathbb{R})$ denote the space of counting measures on the real line $\mathbb{R} = (-\infty, +\infty)$ which are boundedly finite; these are the Borel measures with values in $\mathbb{N}_0 \cup \{+\infty\}$ (where $\mathbb{N}_0 = \{0, 1, \ldots\}$) which are finite on any bounded set. The space $\mathcal{N}(\mathbb{R})$ is endowed with the weak topology $\sigma(\mathcal{N}(\mathbb{R}), \mathcal{C}_{bs}(\mathbb{R}))$ and the corresponding Borel σ-field, where $\mathcal{C}_{bs}$ denotes the space of continuous functions with bounded support.

If N is in $\mathcal{N}(\mathbb{R})$ and $I \subset \mathbb{R}$ is an interval then $N|_I$ denotes the restriction of N to I. Then $N|_I$ belongs to the space $\mathcal{N}(I)$ of boundedly finite counting measures on I. By abuse of notation, a point process on I is often identified with its extension which is null outside of I, and in particular $N|_I \in \mathcal{N}(I)$ with $\mathbb{1}_I N \in \mathcal{N}(\mathbb{R})$. Accordingly, $\mathcal{N}(I)$ is endowed with the trace topology and σ-field.

A random point process on $I \subset \mathbb{R}$ will be considered as a random variable taking values in the Polish space $\mathcal{N}(I)$. We shall also consider random processes with sample paths in the Skorohod space $\mathbb{D}(\mathbb{R}_+, \mathcal{N}(I))$.

All these spaces are Polish, see [9, Prop. A2.5.III, Prop. A2.6.III].

Hawkes processes

In this paper we study a random point process on the real line $\mathbb{R} = (-\infty, +\infty)$ specified by a stochastic evolution on the half-line $(0, +\infty)$ and an initial condition given by a point process on the complementary half-line $(-\infty, 0]$. This is much more general than considering a stationary version of the point process, does not require its existence, and can be used to prove the latter. The time origin 0 can be interpreted as the start of some sort of action with regards to the process (*e.g.*, computation of statistical estimators).

In the following definition of a Hawkes process with a signed reproduction measure, the initial condition N^0 is always assumed to be $\mathcal{F}_0$-measurable and $N^h|_{(0,+\infty)}$ to be adapted to $(\mathcal{F}_t)_{t\geq 0}$. We refer to [9, Sect. 7.2] for the definition of the conditional intensity measure and denote $x^+ = \max(x, 0)$ for $x \in \mathbb{R}$.

Definition 1.1. *Let $\lambda > 0$, a signed measurable function $h : (0, +\infty) \to \mathbb{R}$, and a boundedly finite point process N^0 on $(-\infty, 0]$ with law $\mathfrak{m}$ be given. The point process N^h on $\mathbb{R}$ is a Hawkes process on $(0, +\infty)$ with initial condition N^0 and reproduction measure $\mu(dt) \triangleq h(t)\,dt$ if $N^h|_{(-\infty,0]} = N^0$ and the conditional intensity measure of $N^h|_{(0,+\infty)}$ w.r.t. $(\mathcal{F}_t)_{t\geq 0}$ is absolutely continuous w.r.t. the Lebesgue measure and has density*

$$\Lambda^h : t \in (0, +\infty) \mapsto \Lambda^h(t) = \left(\lambda + \int_{(-\infty,t)} h(t-u)\, N^h(du)\right)^+ . \tag{1.1}$$

Hawkes processes can be defined for reproduction measures μ which are not absolutely continuous w.r.t. the Lebesgue measure, but we shall consider here this case only. This avoids in particular the issue of multiplicities of points in N^h. Since h is the density of μ, the support of h is naturally defined as the support of the measure μ:

$$\mathrm{supp}(h) \triangleq \mathrm{supp}(\mu) \triangleq (0, +\infty) - \bigcup_{G \text{ open},\ |\mu|(G)=0} G\,,$$

where $|\mu|(dt) = |h(t)|\,dt$ is the total variation measure of μ. We assume w.l.o.g. that $h = h\mathbb{1}_{\mathrm{supp}(h)}$ and define

$$L(h) \triangleq \sup(\mathrm{supp}(h)) \triangleq \sup\{t > 0, |h(t)| > 0\} \in [0, +\infty]\,.$$

The constant λ can be considered as the intensity of a Poisson immigration phenomenon on $(0, +\infty)$. The function h corresponds to self-excitation and self-repulsion phenomena: each point of N^h increases, or respectively decreases, the conditional intensity measure wherever the appropriately translated function h is positive (self-excitation), or respectively negative (self-inhibition).

In the sequel, the notation $\mathbb{P}_\mathfrak{m}$ and $\mathbb{E}_\mathfrak{m}$ is used to specify that N^0 has distribution $\mathfrak{m}$. In the case where $\mathfrak{m} = \delta_\nu$ some $\nu \in \mathcal{N}((-\infty, 0])$, we use the notation $\mathbb{E}_\nu$ and $\mathbb{P}_\nu$. We often consider the case when $\nu = \emptyset$, the null measure for which there is no point on $(-\infty, 0]$.

In Definition 1.1, the density Λ^h of the conditional intensity measure of N^h depends on N^h itself, hence existence and uniqueness results are needed. In Proposition 2.1, under the further assumptions that $\|h^+\|_1 < 1$ and that

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

we prove that the Hawkes processes can be constructed as the solution of the equation

$$\begin{cases} N^h = N^0 + \displaystyle\int_{(0,+\infty)\times(0,+\infty)} \delta_u \mathbb{1}_{\{\theta \le \Lambda^h(u)\}}\, Q(du, d\theta)\,, \\ \Lambda^h(u) = \Big(\lambda + \displaystyle\int_{(-\infty,u)} h(u-s)\, N^h(ds)\Big)^+\,, \qquad\qquad u > 0, \end{cases} \tag{1.2}$$

where Q is a $(\mathcal{F}_t)_{t\ge 0}$-Poisson point process on $(0, +\infty) \times (0, +\infty)$ with unit intensity, characterized by the fact that for every $t, h, a > 0$, the random variable $Q((t, t+h] \times (0, a])$ is $\mathcal{F}_{t+h}$-measurable, independent of $\mathcal{F}_t$, and Poisson of parameter ha. Such equations have been introduced and studied by Brémaud et al. [5, 6, 21]. Let us remark that for a given N^0, the counting process $(N^h_t)_{t\ge 0}$ with sample paths in $\mathbb{D}(\mathbb{R}_+, \mathbb{N})$ defined by $N^h_t = N^h((0, t])$ satisfies a pure jump time-inhomogeneous stochastic differential equation which is equivalent to the formulation (1.2).

If h is a nonnegative function satisfying $\|h\|_1 < 1$, then there exists an alternate existence and uniqueness proof based on a cluster representation involving subcritical continuous-time Galton-Watson trees, see [18], which we shall describe and use later.

1.2 Main Results

Our goal in this paper is to establish limit theorems for a Hawkes process N^h with general reproduction function h. We aim at studying the limiting behaviour of the process on a sliding finite time window of length A. We therefore introduce a time-shifted version of the Hawkes process.

Using classical notations for point processes, for $t \in \mathbb{R}$ we define

$$S_t : N \in \mathcal{N}(\mathbb{R}) \mapsto S_t N \triangleq N(\cdot + t) \in \mathcal{N}(\mathbb{R})\,. \tag{1.3}$$

Then $S_t N$ is the image measure of N by the shift by t units of time, and if $a < b$ then

$$\begin{aligned} S_t N((a, b]) &= N((t+a, t+b])\,, \\ (S_t N)|_{(a,b]} &= S_t(N|_{(t+a,t+b]}) = N|_{(t+a,t+b]}(\cdot + t) \end{aligned} \tag{1.4}$$

(with abuse of notation between $N|_{(t+a,t+b]}$ and $\mathbb{1}_{(t+a,t+b]} N$, etc.).

The quantities of interest will be of the form

$$\frac{1}{T}\int_0^T f((S_t N^h)|_{(-A,0]})\,dt = \frac{1}{T}\int_0^T f(N^h(\cdot+t)|_{(-A,0]})dt \tag{1.5}$$

in which $T > 0$ is a finite time horizon, $A > 0$ a finite window length, and f belongs to the set $\mathcal{B}_{lb}(\mathcal{N}((-A,0]))$ of real Borel functions on $\mathcal{N}((-A,0])$ which are locally bounded, *i.e.*, uniformly bounded on $\{\nu \in \mathcal{N}((-A,0]) : \nu((-A,0]) \le n\}$ for each $n \ge 1$. Such quantities appear commonly in the field of statistical inference of random processes; time is labelled so that observation has started by time $-A$.

Using renewal theory, we are able to obtain results without any non-negativity assumption on the reproduction function h. We first establish an ergodic theorem and a central limit theorem for such quantities. We then generalize the concentration inequalities which were obtained by Reynaud-Bouret and Roy [24] under the assumptions that h is a subcritical reproduction law.This leads us to make the following hypotheses on the reproduction function $h = h\mathbb{1}_{\mathrm{supp}(h)}$.

Assumption 1.2. *The signed measurable function $h : (0,+\infty) \to \mathbb{R}$ is such that*

$$L(h) < \infty\,, \qquad \|h^+\|_1 \triangleq \int_{(0,+\infty)} h^+(t)\,dt < 1\,.$$

The distribution $\mathfrak{m}$ of the initial condition N^0 is such that

$$\mathbb{E}_{\mathfrak{m}}(N^0(-L(h),0]) < \infty. \tag{1.6}$$

Under these assumptions, we may and will assume that the window $A < \infty$ is such that $A \ge L(h)$. Then the quantities (1.5) actually depend only on the restriction $N^0|_{(-A,0]}$ of the initial condition N^0 to $(-A,0]$. Thus, in the sequel we identify $\mathfrak{m}$ with its marginal on $\mathcal{N}((-A,0])$ with abuse of notation. Note that even though (1.6) does not imply that $\mathbb{E}_{\mathfrak{m}}(N^0((-A,0])) < \infty$, our results hold under (1.6) (see Remark 1.7 below).

The following important results for the Hawkes process N^h are obtained using its regeneration structure, which will be investigated using a Markov process we now introduce.

In Proposition 3.1 we prove that if $A \ge L(h)$ then the process $(X_t)_{t\ge0}$ defined by

$$X_t \triangleq (S_t N^h)|_{(-A,0]} \triangleq N^h|_{(t-A,t]}(\cdot+t)$$

is a strong Markov process which admits an unique invariant law denoted by π_A, see Theorem 3.5 below.

We introduce τ, the first return time to $\emptyset$ (the null point process) for this Markov process defined by

$$\tau \triangleq \inf\{t>0 : X_{t-} \neq \emptyset, X_t = \emptyset\} = \inf\{t>0 : N^h[t-A,t) \neq 0, N^h(t-A,t] = 0\}\,. \tag{1.7}$$

The probability measure π_A on $\mathcal{N}((-A,0])$ can be classically represented as the intensity of an occupation measure over an excursion: for any non-negative Borel function f,

$$\pi_A f \triangleq \frac{1}{\mathbb{E}_\emptyset(\tau)}\mathbb{E}_\emptyset\left(\int_0^\tau f((S_t N)|_{(-A,0]})\,dt\right). \tag{1.8}$$

Note that we may then construct a Markov process X_t in equilibrium on $\mathbb{R}_+$ and a time-reversed Markov process in equilibrium on $\mathbb{R}_+$, with identical initial conditions (drawn according to π_A) and independent transitions, and build from these a Markov process in equilibrium on $\mathbb{R}$. This construction would yield a stationary version of N^h on $\mathbb{R}$.

We now state our main results.

Theorem 1.3 (Ergodic theorems)**.** *Let N^h be a Hawkes process with immigration rate $\lambda > 0$, reproduction function $h : (0, +\infty) \to \mathbb{R}$, and initial condition N^0 with law $\mathfrak{m}$, satisfying Assumption 1.2. Let $A < \infty$ be such that $A \geq L(h)$, and π_A be the probability measure on $\mathcal{N}((-A, 0])$ defined by* (1.8)*.*

a) If $f \in \mathcal{B}_{lb}(\mathcal{N}((-A, 0]))$ is nonnegative or π_A-integrable, then

$$\frac{1}{T}\int_0^T f((S_t N^h)|_{(-A,0]})\, dt \xrightarrow[T\to\infty]{\mathbb{P}_{\mathfrak{m}}-\text{a.s.}} \pi_A f\,.$$

b) Convergence to equilibrium for large times holds in the following sense:

$$\mathbb{P}_{\mathfrak{m}}\big((S_t N^h)|_{[0,+\infty)} \in \cdot\big) \xrightarrow[t\to\infty]{\text{total variation}} \mathbb{P}_{\pi_A}\big(N^h|_{[0,+\infty)} \in \cdot\big)\,.$$

The following result provides the asymptotics of the fluctuations around the convergence result a), and yields asymptotically exact confidence intervals for it. We define the variance

$$\sigma^2(f) \triangleq \frac{1}{\mathbb{E}_\emptyset(\tau)}\mathbb{E}_\emptyset\bigg(\Big(\int_0^\tau \big(f((S_t N^h)|_{(-A,0]}) - \pi_A f\big)\, dt\Big)^2\bigg). \tag{1.9}$$

Theorem 1.4 (Central limit theorem)**.** *Let N^h be a Hawkes process with immigration rate $\lambda > 0$, reproduction function $h : (0, +\infty) \to \mathbb{R}$, and initial law $\mathfrak{m}$, satisfying Assumption 1.2. Let $A < \infty$ be such that $A \geq L(h)$, the hitting time τ be given by* (1.7)*, and the probability measure π_A on $\mathcal{N}((-A, 0])$ be given by* (1.8)*. If $f \in \mathcal{B}_{lb}(\mathcal{N}((-A, 0]))$ is π_A-integrable and satisfies $\sigma^2(f) < \infty$ then*

$$\sqrt{T}\Big(\frac{1}{T}\int_0^T f((S_t N^h)|_{(-A,0]})\, dt - \pi_A f\Big) \xrightarrow[T\to\infty]{\text{in law}} \mathcal{N}(0, \sigma^2(f))\,.$$

Now we provide non-asymptotic exponential concentration bounds for Theorem 1.3 a). The first entrance time at $\emptyset$ is defined by

$$\tau_0 \triangleq \inf\{t \geq 0 : N^h(t - A, t] = 0\}\,. \tag{1.10}$$

Recall that $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$ for $x \in \mathbb{R}$, and let set $(x)^k_\pm = (x^\pm)^k$.

Theorem 1.5 (Concentration inequalities)**.** *Let N^h be a Hawkes process with immigration rate $\lambda > 0$, reproduction function $h : (0, +\infty) \to \mathbb{R}$, and initial law $\mathfrak{m}$, satisfying Assumption 1.2. Let $A < \infty$ be such that $A \geq L(h)$. Consider the hitting time τ given by* (1.7)*, the entrance time τ_0 given by* (1.10)*, and the probability measure on $\mathcal{N}((-A, 0])$ defined* *in* (1.8). *Consider* $f \in \mathcal{B}_{lb}(\mathcal{N}((-A,0]))$ *taking its values in a bounded interval* $[a,b]$*, and define* $\sigma^2(f)$ *as in* (1.9) *and*

$$c^{\pm}(f) \triangleq \sup_{k\geq 3}\left(\frac{2}{k!}\frac{\mathbb{E}_\emptyset\big((\int_0^\tau (f((S_tN^h)|_{(-A,0]}) - \pi_A f)\,dt)_\pm^k\big)}{\mathbb{E}_\emptyset(\tau)\sigma^2(f)}\right)^{\frac{1}{k-2}},$$

$$c^{\pm}(\tau) \triangleq \sup_{k\geq 3}\left(\frac{2}{k!}\frac{\mathbb{E}_\emptyset\big((\tau - \mathbb{E}_\emptyset(\tau))_\pm^k\big)}{\mathrm{Var}_\emptyset(\tau)}\right)^{\frac{1}{k-2}},$$

$$c^{+}(\tau_0) \triangleq \sup_{k\geq 3}\left(\frac{2}{k!}\frac{\mathbb{E}_\mathfrak{m}\big((\tau_0 - \mathbb{E}_\mathfrak{m}(\tau_0))_+^k\big)}{\mathrm{Var}_\mathfrak{m}(\tau_0)}\right)^{\frac{1}{k-2}}.$$

Then, for all $\varepsilon > 0$ *and* T *sufficiently large*

$$\begin{aligned}
&\mathbb{P}_\mathfrak{m}\left(\left|\frac{1}{T}\int_0^T f((S_tN^h)|_{(-A,0]}) - \pi_A f\right| \geq \varepsilon\right)\\
&\leq \exp\left(-\frac{((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))^2}{8T\sigma^2(f) + 4c^+(f)((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))}\right)\\
&\quad + \exp\left(-\frac{((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))^2}{8T\sigma^2(f) + 4c^-(f)((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))}\right)\\
&\quad + \exp\left(-\frac{((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))^2}{8T|b-a|^2\frac{\mathrm{Var}_\emptyset(\tau)}{\mathbb{E}_\emptyset(\tau)} + 4|b-a|c^+(\tau)((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))}\right)\\
&\quad + \exp\left(-\frac{((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))^2}{8T|b-a|^2\frac{\mathrm{Var}_\emptyset(\tau)}{\mathbb{E}_\emptyset(\tau)} + 4|b-a|c^-(\tau)((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))}\right)\\
&\quad + \exp\left(-\frac{(\sqrt{T}\varepsilon - 2|b-a|\mathbb{E}_\mathfrak{m}(\tau_0))^2}{8|b-a|^2\,\mathrm{Var}_\mathfrak{m}(\tau_0) + 4|b-a|c^+(\tau_0)(\sqrt{T}\varepsilon - 2|b-a|\mathbb{E}_\mathfrak{m}(\tau_0))}\right).
\end{aligned} \tag{1.11}$$

If $N|_{(-A,0]} = \emptyset$ *then the last term of the r.h.s. is null and the upper bound is true with* T *instead of* $T - \sqrt{T}$ *in the other terms.*

This concentration inequality can be simplified, using upper bounds for the constants $c^{\pm}(f)$ and $c^{\pm}(\tau)$. In the following corollary, we use explicitly the fact that the hitting time τ admits an exponential moment (see Proposition 3.4).

Corollary 1.6. *Under assumptions and notation of Theorem 1.5, there exists* $\alpha > 0$ *such that* $\mathbb{E}_\emptyset(e^{\alpha\tau}) < \infty$*. We set*

$$v = \frac{2(b-a)^2}{\alpha^2}\left\lfloor\frac{T}{\mathbb{E}_\emptyset(\tau)}\right\rfloor \mathbb{E}_\emptyset(e^{\alpha\tau})e^{\alpha\mathbb{E}_\emptyset(\tau)}, \quad \textit{and} \quad c = \frac{|b-a|}{\alpha}.$$

Then for all $\varepsilon > 0$

$$\mathbb{P}_\emptyset\left(\left|\frac{1}{T}\int_0^T f((S_tN^h)|_{(-A,0]}) - \pi_A f\right| \geq \varepsilon\right) \leq 4\exp\left(-\frac{\big(T\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau)\big)^2}{4\,(2v + c(T\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau)))}\right),$$

or equivalently for all $1 \geq \eta > 0$

$$P_\emptyset\Big(\Big|\frac{1}{T}\int_0^T f((S_t N^h)|_{(-A,0]})dt - \pi_A f\Big| \geq \varepsilon_\eta\Big) \leq \eta\,, \tag{1.12}$$

where

$$\varepsilon_\eta = \frac{1}{T}\left(|b-a|\mathbb{E}_\emptyset(\tau) - 2c\log\big(\frac{\eta}{4}\big) + \sqrt{4c^2\log^2\big(\frac{\eta}{4}\big) - 8v\log\big(\frac{\eta}{4}\big)}\right).$$

Remark 1.7. *All these results hold under* (1.6) *even if* $\mathbb{E}_\mathfrak{m}(N^0((-A,0])) = +\infty$*. Indeed,*

$$\begin{aligned}\frac{1}{T}\int_0^T f(N^h(\cdot+t)|_{(-A,0]})dt = \frac{1}{T}\int_0^{A-L(h)} & f(N^h(\cdot+t)|_{(-A,0]})dt \\ & + \frac{1}{T}\int_{A-L(h)}^T f(N^h(.+t)|_{(-A,0]})dt\,.\end{aligned}$$

The first r.h.s. term converges $\mathcal{P}_\mathfrak{m}$*-a.s. to zero, even when multiplied by* $\sqrt{T}$*. For the second r.h.s. term, we can apply the Markov property at time* $A - L(h)$ *(which will be justified when proving that* $(S_. N^h)|_{(-A,0]}$ *is a Markov process) and show that*

$$\mathbb{E}_{(S_{A-L(h)}N^h)|_{(-A,0]}}(N^0((-A,0])) < +\infty.$$

2 Hawkes processes

In this Section, we first give a constructive proof of Eq. (1.2), which yields a coupling between N^h and N^{h^+} satisfying $N^h \leq N^{h^+}$. The renewal times on which are based the proofs of our main results are the instants at which the intensity Λ^h has returned and then stayed at λ for a duration long enough to be sure that the dependence on the past has vanished, in order to be able to write the process in terms of i.i.d. excursions. The coupling will allow us to control the renewal times for N^h by the renewal times for N^{h^+}.

When dealing with h^+, we use the well-known cluster representation for a Hawkes process with nonnegative reproduction function. This representation allows us to interpret the renewal times as times at which an $M/G/\infty$ queue is empty, and we use this interpretation in order to obtain tail estimates for the interval between these times.

2.1 Solving the equation for the Hawkes process

In this paragraph, we give an algorithmic proof of the existence and uniqueness of strong solutions of Equation (1.2). This algorithmic proof can be used for simulations, which are shown in Fig. 2.1.

Proposition 2.1. *Let* Q *be a* $(\mathcal{F}_t)_{t\geq 0}$*-Poisson point process on* $(0,+\infty)\times(0,+\infty)$ *with unit intensity.Consider Equation* (1.2), i.e.,

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

in which $h : (0, +\infty) \to \mathbb{R}$ is a signed measurable reproduction function, $\lambda > 0$ an immigration rate, and N^0 an initial condition in $\mathcal{N}((-\infty, 0])$ and law $\mathfrak{m}$. Consider the similar equation for N^{h^+} in which h is replaced by h^+. Assume that

$$\|h^+\|_1 < 1 \tag{2.1}$$

and that the distribution $\mathfrak{m}$ of the initial condition N^0 satisfies

$$\forall t > 0, \quad \int_0^t \mathbb{E}_{\mathfrak{m}}\Big(\int_{(-\infty,0]} h^+(u-s)\, N^0(ds) \Big)\, du < +\infty. \tag{2.2}$$

a) Then there exists a pathwise unique strong solution N^h of Equation (1.2), and this solution is a Hawkes process in the sense of Definition 1.1.

b) The same holds for N^{h^+}, and moreover $N^h \le N^{h^+}$ a.s. (in the sense of measures).

Remark 2.2. *In order to prove the strong existence and pathwise uniqueness of the solution of Eq. (1.2), we propose a proof based on an algorithmic construction similar to the Poisson embedding of [5], also referred in [10] as thinning. A similar result is also proved in these references using Picard iteration techniques, with Assumption (2.2) replaced by the stronger hypothesis that there exists $D_{\mathfrak{m}} > 0$ such that*

$$\forall t > 0, \quad \mathbb{E}_{\mathfrak{m}}\Big(\int_{(-\infty,0]} |h(t-s)|\, N^0(ds) \Big) < D_{\mathfrak{m}}\,. \tag{2.3}$$

When h is nonnegative, the result can be deduced from the cluster representation of the self-exciting Hawkes process, since $N^h([0, t])$ is upper bounded by the sum of the sizes of a Poisson number of sub-critical Galton-Watson trees, see [18, 24].

Remark 2.3. *Proposition 2.1 does not require that $L(h)$ be finite. When $L(h) < \infty$ then the assumption (2.2) can be rewritten as*

$$\int_0^{L(h)} \mathbb{E}_{\mathfrak{m}}\Big(\int_{(-L(h),0]} h^+(u-s)\, N^0(ds) \Big)\, du < +\infty\,. \tag{2.4}$$

A sufficient condition for (2.4) to hold is that $\mathbb{E}_{\mathfrak{m}}(N^0(-L(h), 0]) < +\infty$. Indeed, using the Fubini-Tonelli theorem, the l.h.s. of (2.4) can be bounded by $\|h^+\|_1\, \mathbb{E}_{\mathfrak{m}}(N^0(-L(h), 0])$. Therefore, the results of Proposition 2.1 hold under Assumptions 1.2.

Before proving Proposition 2.1, we start with a lemma showing that Assumption (2.2) implies a milder condition which will be used repeatedly in the proof of the proposition.

Lemma 2.4. *Suppose that Assumption (2.2) is satisfied. Then for any nonnegative random variable U and $r > 0$,*

$$\mathbb{P}_{\mathfrak{m}}\Big(\int_U^{U+r} \int_{(-\infty,0]} h^+(t-s)\, N^0(ds)\, dt < +\infty, \quad U < +\infty \Big) = \mathbb{P}_{\mathfrak{m}}(U < +\infty)\,.$$

Proof. First note that, for every integer n,

$$\int_0^n \int_{(-\infty,0]} h^+(t-s)\, N^0(ds) dt < +\infty\,,\ \mathbb{P}_{\mathfrak{m}} - \text{a.s.},$$

using condition (2.2) and the Fubini-Tonelli theorem. This leads easily to

$$\mathbb{P}_{\mathfrak{m}}\Big(\forall n \geq 0,\ \int_0^n \int_{(-\infty,0]} h^+(t-s)\, N^0(ds)dt < +\infty\Big) = 1\,,$$

and, for a positive real number r, to

$$\mathbb{P}_{\mathfrak{m}}\Big(\forall u \geq 0,\ \int_u^{u+r} \int_{(-\infty,0]} h^+(t-s)\, N^0(ds)dt < +\infty\Big) = 1\,,$$

which gives the announced result. □

Proof of Proposition 2.1. Proofs of both a) and b) will be obtained by induction on the successive atoms of N^h.

Proof of a): initialization. Let

$$\Lambda_0^h(t) = \Big(\lambda + \int_{(-\infty,0]} h(t-s)\, N^0(ds)\Big)^+, \qquad t > 0\,, \tag{2.5}$$

$$U_1^h = \inf\Big\{u > 0 : \int_{(0,u]} \int_{(0,\Lambda_0^h(v)]} Q(dv, d\theta) > 0\Big\}\,, \tag{2.6}$$

with the usual convention that $\inf \emptyset = +\infty$. First note that conditionally on N^0,

$$Q(\{(v,\theta) \in (0,\varepsilon] \times (0,+\infty) : \theta \leq \Lambda_0^h(v)\})$$

follows a Poisson law with parameter $\int_0^\varepsilon \Lambda_0^h(t)dt$. Using Assumption (2.2) and Lemma 2.4, we can find $\varepsilon_0 > 0$ such that $\int_0^{\varepsilon_0} \int_{(-\infty,0]} h^+(t-s)\, N^0(ds)dt < +\infty$. We thus have, $\mathbb{P}_{\mathfrak{m}}$-a.s.,

$$\begin{aligned}\int_0^{\varepsilon_0} \Lambda_0^h(t)dt &= \int_0^{\varepsilon_0} \Big(\lambda + \int_{(-\infty,0]} h(t-s)\, N^0(ds)\Big)^+ dt \\ &\leq \lambda\varepsilon_0 + \int_0^{\varepsilon_0} \int_{(-\infty,0]} h^+(t-s)\, N^0(ds)dt < +\infty\,.\end{aligned}$$

Consequently, $Q(\{(v,\theta) \in (0,\varepsilon_0] \times (0,+\infty) : \theta \leq \Lambda_0^h(v)\})$ is finite $\mathbb{P}_{\mathfrak{m}}$-a.s. If it is null then $U_1^h = +\infty$ and $N^h = N^0$. Else, U_1^h is the first atom on $(0,+\infty)$ of the point process of conditional intensity Λ_0^h. Since $\Lambda_0^h(t) = \Lambda^h(t)$ for $t \in (0, U_1^h]$, thus U_1^h is also the first atom of N^h on $(0,+\infty)$.
On $\{U_1^h = +\infty\}$, we define $U_k^h = +\infty$ for all $k \geq 2$.

Proof of a): recursion. Assume that we have built $U_1^h, \dots, U_k^h$ such that on the event $\{U_k^h < +\infty\}$ these are the first k atoms of N^h in increasing order. We are going to construct U_{k+1}^h, which will be when finite an atom of N^h greater than U_k^h.

On $\{U_k^h = +\infty\}$ we set $U_{k+1}^h = +\infty$. Henceforth, we work on $\{U_k^h < +\infty\}$. Let

$$\Lambda_k^h(t) = \Big(\lambda + \int_{(-\infty,0]} h(t-s)\, N^0(ds) + \int_{(0,U_k^h]} h(t-s)\, N^h(ds)\Big)^+, \qquad t > 0\,, \tag{2.7}$$

$$U_{k+1}^h = \inf\Big\{u > U_k^h : \int_{(U_k^h,u]} \int_{(0,\Lambda_k^h(v)]} Q(dv, d\theta) > 0\Big\}\,.$$

As in Step 1, we first prove that there exists $\varepsilon > 0$ such that $Q(\mathcal{R}_\varepsilon)$ is a.s. finite, where

$$\mathcal{R}_\varepsilon = \{(v,\theta) : v \in (U_k^h, U_k^h + \varepsilon],\, \theta \in (0, \Lambda_k^h(v)]\}\,.$$

Since the random function Λ_k^h is measurable with respect to $\mathcal{F}_{U_k^h}$, conditionally on $\mathcal{F}_{U_k^h}$, $Q(\mathcal{R}_\varepsilon)$ follows a Poisson law with parameter $\int_{U_k^h}^{U_k^h+\varepsilon} \Lambda_k^h(t)dt$ (see Lemma A.3) so that

$$\mathbb{P}(Q(\mathcal{R}_\varepsilon) < +\infty) = \mathbb{E}\Big(\mathbb{P}(Q(\mathcal{R}_\varepsilon) < +\infty \,|\, \mathcal{F}_{U_k^h})\Big) = \mathbb{E}\left(\mathbb{P}\left(\int_{U_k^h}^{U_k^h+\varepsilon} \Lambda_k^h(t)dt < +\infty \,\middle|\, \mathcal{F}_{U_k^h}\right)\right).$$

Using the fact that $x \le x^+$ and the monoticity of $x \mapsto x^+$, we obtain from (2.7) that

$$\begin{aligned}\int_{U_k^h}^{U_k^h+\varepsilon} \Lambda_k^h(t)dt \le \lambda\varepsilon &+ \int_{U_k^h}^{U_k^h+\varepsilon} \int_{(-\infty,0]} h^+(t-s)\, N^0(ds)dt \\ &+ \int_{U_k^h}^{U_k^h+\varepsilon} \int_{(0,U_k^h]} h^+(t-s)\, N^h(ds)dt\,.\end{aligned}$$

On $\{U_k^h < +\infty\}$ the second term in the r.h.s. is finite thanks to Assumption (2.2) and Lemma 2.4. It is thus also finite, a.s., on $\{U_k^h < +\infty\}$, conditionally on $\mathcal{F}_{U_k^h}$. Now, using the Fubini-Tonelli Theorem and Assumption (2.1), we obtain that

$$\begin{aligned}\int_{U_k^h}^{U_k^h+\varepsilon} \int_{(0,U_k^h]} h^+(t-s)\, N^h(ds)dt &= \int_{(0,U_k^h]} \left(\int_{U_k^h}^{U_k^h+\varepsilon} h^+(t-s)dt\right) N^h(ds) \\ &\le \|h^+\|_1\, N^h((0,U_k^h]) = k\|h^+\|_1 < +\infty.\end{aligned}$$

This concludes the proof of the finiteness of $\int_{U_k^h}^{U_k^h+\varepsilon} \Lambda_k^h(t)dt$, so that $Q(\mathcal{R}_\varepsilon) < +\infty$, $\mathbb{P}_{\mathfrak{m}}$-a.s.

If $Q(\mathcal{R}_\varepsilon)$ is null then $U_{k+1}^h = +\infty$ and thus $N^h = N^0 + \sum_{i=1}^k \delta_{U_i^h}$. Else, U_{k+1}^h is actually a minimum, implying that $U_k^h < U_{k+1}^h$ and, since Λ^h and Λ_k^h coincide on $(0, U_{k+1}^h)$, that U_{k+1}^h is the $(k+1)$-th atom of N^h.

We have finally proved by induction the existence of a random nondecreasing sequence $(U_k^h)_{k\ge1}$, which is either stationary equal to infinity, or strictly increasing. On this last event, the U_k^h are exactly the atoms of the random point process N^h on $(0,+\infty)$.

To complete the proof, it is enough to prove that $\lim_{k\to+\infty} U_k^h = +\infty$, $\mathbb{P}_{\mathfrak{m}}$-a.s. For this, we compute $\mathbb{E}_{\mathfrak{m}}(N^h(0,t))$ for $t > 0$. For all $k \ge 1$,

$$\begin{aligned}\mathbb{E}_{\mathfrak{m}}(N^h(0, t\wedge U_k^h)) &= \mathbb{E}_{\mathfrak{m}}\left(\int_0^{t\wedge U_k^h} \Lambda^h(u)du\right) \\ &= \mathbb{E}_{\mathfrak{m}}\left(\int_0^{t\wedge U_k^h} \left(\lambda + \int_{(-\infty,u)} h(u-s)\, N^h(ds)\right)^+ du\right) \\ &\le \lambda t + \mathbb{E}_{\mathfrak{m}}\left(\int_0^t \int_{(-\infty,0]} h^+(u-s)\, N^0(ds)du\right) \\ &\quad + \mathbb{E}_{\mathfrak{m}}\left(\int_0^{t\wedge U_k^h} \int_{(0,u)} h^+(u-s)\, N^h(ds)du\right).\end{aligned}$$

Using the nonnegativity of h^+ and Assumption (2.2),

$$\mathbb{E}_{\mathfrak{m}}\Big(\int_0^t \int_{(-\infty,0]} h^+(u-s)\, N^0(ds)du\Big) \le \int_0^t \mathbb{E}_{\mathfrak{m}}\Big(\int_{(-\infty,0]} h^+(u-s)\, N^0(ds)\Big)du < +\infty\,.$$

For the last term, we use again the Fubini-Tonelli theorem and obtain

$$\begin{aligned}\mathbb{E}_{\mathfrak{m}}\Big(\int_0^{t\wedge U_k^h} \int_{(0,u)} h^+(u-s)\, N^h(ds)\; du\Big) &= \mathbb{E}_{\mathfrak{m}}\Big(\int_{(0,t\wedge U_k^h)} \int_s^{t\wedge U_k^h} h^+(u-s)du\, N^h(ds)\Big)\\ &\le \|h^+\|_1\, \mathbb{E}_{\mathfrak{m}}\Big(N^h(0,t\wedge U_k^h)\Big).\end{aligned}$$

These three inequalities and the fact that $\|h^+\|_1 < 1$, see Assumption (2.1), yield that

$$0 \le \mathbb{E}_{\mathfrak{m}}(N^h(0,t\wedge U_k^h)) \le \frac{1}{1-\|h^+\|_1}\Big(\lambda t + \int_0^t \mathbb{E}_{\mathfrak{m}}\Big(\int_{(-\infty,0]} h^+(u-s)\, N^0(ds)\Big)du\Big) \quad (2.8)$$

where the upper bound is finite and independent of k.

As a consequence, we necessarily have that $\lim_{k\to+\infty} U_k^h = +\infty$, a.s. Otherwise, there would exist $T > 0$ and Ω_0 such that $\mathbb{P}(\Omega_0) > 0$ and $\lim_{k\to+\infty} U_k^h \le T$ on Ω_0. But this would entail that $\mathbb{E}_{\mathfrak{m}}(N^h(0,T\wedge U_k^h)) \ge (k-1)\mathbb{P}_{\mathfrak{m}}(\Omega_0)$ which converges to $+\infty$ with k and cannot be upper bounded by (2.8).
Note additionally that once we know that $\lim_{k\to+\infty} U_k^h = +\infty$, a.s., we can use the Beppo-Levi theorem, which leads to $\mathbb{E}_{\mathfrak{m}}\big(N^h(0,t)\big) < +\infty$ for all $t > 0$.

Note that uniqueness comes from the algorithmic construction of the sequence $(U_k^h)_{k\ge1}$.

Proof of b). The assumptions of the theorem are valid both for h and for h^+, and the result a) which we have just proved allows to construct strong solutions N^h and N^{h^+} of Eq. (1.2) driven by the same Poisson point process Q. Proving b) is equivalent to showing that the atoms of N^h are also atoms of N^{h^+}, which we do using the following recursion.

If $U_1^h = +\infty$ then N^h as no atom on $(0,+\infty)$ and there is nothing to prove.

Else, we first show that the first atom U_1^h of N^h is also an atom of N^{h^+}. The key point is to establish that

$$\forall t \in (0, U_1^h), \quad \Lambda^h(t) \le \Lambda^{h^+}(t). \quad (2.9)$$

Indeed, from the definition of U_1^h, there exists an atom of the Poisson measure Q at some (U_1^h, θ) with $\theta \le \Lambda^h((U_1^h)_-)$. If (2.9) is true we may deduce that (U_1^h, θ) is also an atom of Q satisfying $\theta \le \Lambda^{h^+}((U_1^h)_-)$, and thus that U_1^h is also an atom of N^{h^+}.

We now turn to the proof of (2.9). For every $t \in (0, U_1^h)$, we clearly have

$$\Lambda^h(t) = \Lambda_0^h(t) \triangleq \Big(\lambda + \int_{(-\infty,0]} h(t-s)\, N^0(ds)\Big)^+,$$

we use the fact that $x \mapsto x^+$ is nondecreasing on $\mathbb{R}$ to obtain that

$$\Lambda^h(t) \le \lambda + \int_{(-\infty,t)} h^+(t-s)\, N^{h^+}(ds) \triangleq \Lambda^{h^+}(t)\,.$$

We now prove that if $U_1^h, \ldots, U_k^h$ are atoms of N^{h^+} and $U_{k+1}^h < +\infty$ then U_{k+1}^h is also an atom of N^{h^+}. By construction, $\Lambda^h(t) = \Lambda_k^h(t)$ for all $t \in (0, U_{k+1}^h)$, and there exists $\theta > 0$

such that (U_{k+1}^h, θ) is an atom of Q satisfying $\theta \leq \Lambda^h((U_{k+1}^h)_-)$. To obtain that U_{k+1}^h is also an atom of N^{h^+}, it is thus enough to prove that

$$\forall t \in [U_k^h, U_{k+1}^h), \quad \Lambda^h(t) \leq \Lambda^{h^+}(t).$$

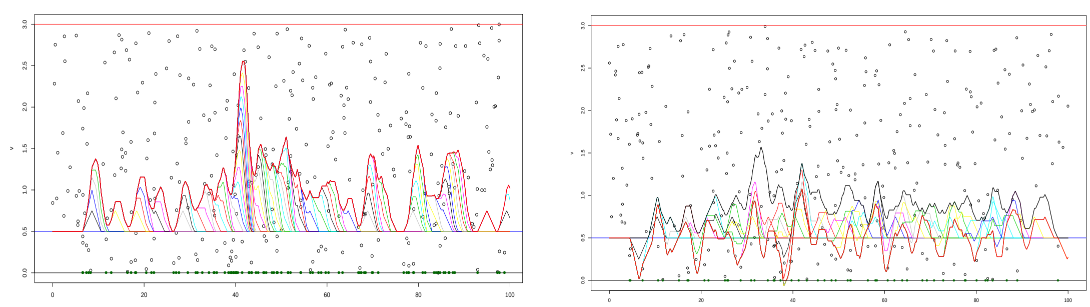
Using that $h \leq h^+$ and the induction hypothesis that the first k atoms $U_1^h, \ldots, U_k^h$ of N^h are also atoms of N^{h^+}, we obtain for all $t \in (U_k^h, U_{k+1}^h)$ that

$$\int_{(0,U_k^h]} h(t-s)\, N^h(ds) \leq \int_{(0,U_k^h]} h^+(t-s)\, N^h(ds) \leq \int_{(0,t)} h^+(t-s)\, N^{h^+}(ds)\,.$$

This upper bound and the definition (2.7) of Λ_k^h yield that, for all $t \in (U_k^h, U_{k+1}^h)$,

$$\Lambda_k^h(t) \leq \Lambda^{h^+}(t)\,,$$

and since Λ_k^h and Λ^h coincide on $(0, U_{k+1}^h)$, we have finally proved that U_{k+1}^h is an atom of N^{h^+}. This concludes the proof of the proposition. $\square$



(a) (b)

Figure 2.1: *(a) Hawkes process with a positive reproduction function h. (b) Hawkes process with a general reproduction function h. The dots in the plane represent the atoms of the Poisson point process Q used for the construction. The atoms of the Hawkes processes are the green dots on the abscissa axis. The bold red curve corresponds to the intensity Λ^h and the colored curves represent the partial cumulative contributions of the successive atoms of the Hawkes process. In (b), the bold blue curve corresponds to the intensity of the dominating Hawkes process with reproduction function h^+.*

2.2 The cluster representation for nonnegative reproduction functions

In this subsection, we consider the case in which the reproduction function h is nonnegative. The intensity process of a corresponding Hawkes process can be written, for $t > 0$,

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

The first term can be interpreted as an immigration rate of *ancestors*. Let $(V_k)_{k\geq 1}$ be the corresponding sequence of arrival times, forming a Poisson process of intensity λ.

The second term is the sum of all the contributions of the atoms of N^h before time t and can be seen as self-excitation. If U is an atom of N^h, it contributes to the intensity by the addition of the function $t \mapsto h(t-U)$, hence generating new points regarded as its *descendants* or *offspring*. Each individual has a *lifelength* $L(h) = \sup(\mathrm{supp}(h))$, the number of its descendants follows a Poisson distribution with mean $\|h\|_1$, and the ages at which it gives birth to them have density $h/\|h\|_1$, all this independently. This induces a Galton-Watson process in continuous time, see [18, 24], and Fig. 2.2.

To each ancestor arrival time V_k we can associate a cluster of times, composed of the times of birth of its descendents. The condition $\|h\|_1 < 1$ is a necessary and sufficient condition for the corresponding Galton-Watson process to be sub-critical, which implies that the cluster sizes are finite almost surely. More precisely, if we define H_k by saying that $V_k + H_k$ is the largest time of the cluster associated with V_k, then the $(H_k)_{k\geq 1}$ are i.i.d random variables independent from the sequence $(V_k)_{k\geq 1}$.

Reynaud-Bouret and Roy [24] proved the following tail estimate for H_1.

Proposition 2.5 ([24, Prop. 1.2])**.** *Under Assumption 1.2, we have*

$$\forall x \geq 0, \quad \mathbb{P}(H_1 > x) \leq \exp\Big(-\frac{x}{L(h)}(\|h\|_1 - \log\|h\|_1 - 1) + 1 - \|h\|_1\Big). \tag{2.10}$$

If we define

$$\gamma \triangleq \frac{\|h\|_1 - \log(\|h\|_1) - 1}{L(h)} \tag{2.11}$$

then $\mathbb{P}(H_1 > x) \leq \exp(1 - \|h\|_1)\exp(-\gamma x)$, and γ is an upper bound of the rate of decay of the Galton-Watson cluster length.

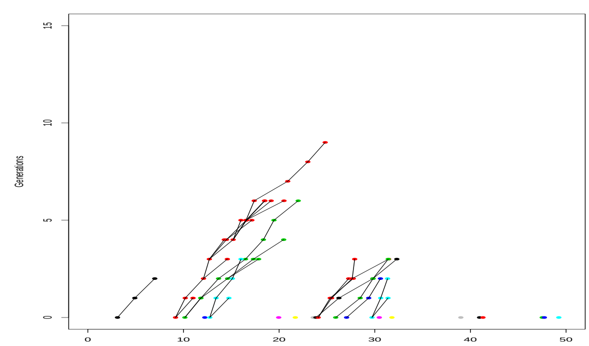


Figure 2.2: *Cluster representation of a Hawkes process with positive reproduction function. The abscissa of the dots give its atoms. Offspring are colored according to their ancestor, and their ordinates correspond to their generation in this age-structured Galton-Watson tree.*

When h is nonnegative, it is possible to associate to the Hawkes process a $M/G/\infty$ queue. For $A \geq L(h)$, we consider that the arrival times of ancestors $(V_k)_{k\geq 1}$ correspond to the arrival of customers in the queue and associate to the k-th customer a service time

$\widetilde{H}_k(A) \triangleq H_k + A$. We assume that the queue is empty at time 0, and then the number Y_t of customers in the queue at time $t \geq 0$ is given by

$$Y_t = \sum_{k:V_k \leq t} \mathbb{1}_{\{V_k + \widetilde{H}_k(A) > t\}} \,. \tag{2.12}$$

Let $\mathcal{T}_0 = 0$, and the successive hitting times of 0 by the process $(Y_t)_{t\geq 0}$ be given by

$$\forall k \geq 1, \quad \mathcal{T}_k = \inf\{t \geq \mathcal{T}_{k-1},\ Y_{t-} \neq 0,\ Y_t = 0\}. \tag{2.13}$$

The time interval $[V_1, \mathcal{T}_1)$ is called the first busy period, and is the first time interval during which the queue is never empty. Note that the $\mathcal{T}_k$ are times at which the conditional intensity of the underlying Hawkes process has returned to λ and there is no remaining influence of its previous atoms, since $\widetilde{H}_k(A) \triangleq H_k + A \geq H_k + L(h)$.

Thus the Hawkes process after $\mathcal{T}_k$ has the same law as the Hawkes process with initial condition the null point process $\emptyset \in \mathcal{N}((-A, 0])$, translated by $\mathcal{T}_k$. This allows us to split the random measure N^h into i.i.d. parts. We will prove all this in the next section.

We end this part by giving tail estimates for the $\mathcal{T}_k$, which depend on λ and on γ given in (2.11) which respectively control the exponential decays of $\mathbb{P}(V_1 > x)$ and $\mathbb{P}(H_1 > x)$.

Proposition 2.6. *In this situation, for all $x > 0$, if $\lambda < \gamma$, then $\mathbb{P}(\mathcal{T}_1 > x) = O(\mathrm{e}^{-\lambda x})$, and if $\gamma \leq \lambda$, for any $\alpha < \gamma$ then $\mathbb{P}(\mathcal{T}_1 > x) = O(\mathrm{e}^{-\alpha x})$. Notably, if $\alpha < \min(\lambda, \gamma)$, $\mathbb{E}(\mathrm{e}^{\alpha \mathcal{T}_1})$ is finite.*

Proof of Proposition 2.6. The proof follows from Proposition 2.5, from which we deduce that the service time $\widetilde{H}_1 = H_1 + A$ satisfies:

$$\mathbb{P}(\widetilde{H}_1 > x) = \mathbb{P}(H_1 > x - A) \leq \exp(-(x - A)\gamma + 1 - \|h\|_1) = O(\mathrm{e}^{-\gamma x}) \,. \tag{2.14}$$

We then conclude by applying Theorem A.1 to the queue $(Y_t)_{t\geq 0}$ defined by (2.12). $\square$

Theorem A.1 in Appendix establishes the decay rates for the tail distributions of $\mathcal{T}_1$ and of the length of the busy period $[V_1, \mathcal{T}_1)$. It has an interest in itself, independently of the results for Hawkes processes considered here.

3 An auxiliary Markov Process

When the reproduction function h has a bounded support, $N^h|_{(t,+\infty)}$ depends on $N^h|_{(-\infty,t]}$ only through $N^h|_{(t-L(h),t]}$. The process $t \mapsto N^h|_{(t-L(h),t]}$ will then be seen to be Markovian, which yields regenerative properties for N^h. It is the purpose of this section to formalize that idea by introducing an auxiliary Markov process.

3.1 Definition of a strong Markov process

We suppose that Assumption 1.2 holds and consider the Hawkes process N^h solution of the corresponding Equation (1.2) constructed in Proposition 2.1. We recall that $L(h) < \infty$. Then, for any $t > 0$ and $u \in (-\infty, -L(h)]$, $h(t-u) = 0$, and thus

$$\Lambda^h(t) = \left(\lambda + \int_{(-\infty,t)} h(t-u)\, N^h(du)\right)^+ = \left(\lambda + \int_{(-L(h),t)} h(t-u)\, N^h(du)\right)^+ . \tag{3.1}$$

In particular $N^h|_{(0,+\infty)}$ depends only on the restriction $N^0|_{(-L(h),0]}$ of the initial condition.

Recall that the shift operator S_t is defined in (1.3) and (1.4). Note that if $t, s \geq 0$ then $S_{s+t}N^h = S_t S_s N^h = S_s S_t N^h$. Let $A < \infty$ be such that $A \geq L(h)$. Consider the $(\mathcal{F}_t)$-adapted process $X = (X_t)_{t\geq 0}$ defined by

$$X_t = (S_t N^h)|_{(-A,0]} = N^h|_{(t-A,t]}(\cdot + t)) \,, \tag{3.2}$$

i.e.,

$$\begin{array}{rcl} X_t \; : \; \mathcal{B}((-A,0]) & \to & \mathbb{R}_+ \\ B & \mapsto & X_t(B) = N^h|_{(t-A,t]}(B+t), \end{array}$$

The measure X_t is the point process N^h in the time window $(t-A,t]$, shifted back to $(-A,0]$. This is a function of $N^h|_{(-A,+\infty)}$. Using Equation (3.1) and the remark below it, we see that the law of $N^h|_{(-A,+\infty)}$ depends on the initial condition N^0 only through $N^0|_{(-A,0]}$. Therefore, with abuse of notation, when dealing with the process $(X_t)_{t\geq 0}$ we shall use the notations $\mathbb{P}_{\mathfrak{m}}$ and $\mathbb{E}_{\mathfrak{m}}$ even when $\mathfrak{m}$ is a law on $\mathcal{N}((-A,0])$, and $\mathbb{P}_\nu$ and $\mathbb{E}_\nu$ even when ν is an element of $\mathcal{N}((-A,0])$.

Note that X depends on A, and that we omit this in the notation.

Proposition 3.1. *Under Assumption 1.2. Let $A < \infty$ be such that $A \geq L(h)$. Then $(X_t)_{t\geq 0}$ defined in* (3.2) *is a strong $(\mathcal{F}_t)_{t\geq 0}$-Markov process with initial condition $X_0 = N^0|_{(-A,0]}$ and sample paths in the Skorohod space $\mathbb{D}(\mathbb{R}_+, \mathcal{N}((-A,0]))$.*

Proof. This follows from the fact that N^h is the unique solution of Eq. (1.2). Indeed, let T be a stopping time. On $\{T < \infty\}$, by definition

$$X_{T+t} = (S_{T+t}N^h)|_{(-A,0]} = (S_t S_T N^h)|_{(-A,0]} \,.$$

Using that N^h satisfies Eq. (1.2) driven by the process Q, we have

$$\begin{aligned} S_T N^h &= S_T(N^h|_{(-\infty,T]}) + S_T(N^h|_{(T,+\infty)}) \\ &= (S_T N^h)|_{(-\infty,0]} + \int_{(T,+\infty)\times(0,+\infty)} \delta_{u-T} \mathbb{1}_{\{\theta \leq \Lambda^h(u)\}} Q(du, d\theta) \\ &= (S_T N^h)|_{(-\infty,0]} + \int_{(0,+\infty)\times(0,+\infty)} \delta_v \mathbb{1}_{\{\theta \leq \widetilde{\Lambda}^h(v)\}} \, S_T Q(dv, d\theta), \end{aligned}$$

where $S_T Q$ is the (randomly) shifted process with bivariate cumulative distribution function given by

$$S_T Q((0,t] \times (0,a]) = Q((T, T+t] \times (0,a]) \,, \qquad t, a > 0, \tag{3.3}$$

and where for $v > 0$,

$$\widetilde{\Lambda}^h(v) = \Lambda^h(v+T) = \Big(\lambda + \int_{(-\infty,v)} h(v-s) S_T N^h(ds)\Big)^+.$$

This shows that $S_T N^h$ satisfies Eq. (1.2) driven by $S_T Q$ with initial condition $(S_T N^h)|_{(-\infty,0]}$. Since $A \geq L(h)$, moreover $S_T N^h|_{(0,+\infty)}$ actually depends only on $(S_T N^h)|_{(-A,0]} \triangleq X_T$.

Let us now condition on $\{T < \infty\}$ and on $\mathcal{F}_T$. Since Q is a $(\mathcal{F}_t)_{t\geq 0}$-Poisson point process with unit intensity, $S_T Q$ is a $(\mathcal{F}_{T+t})_{t\geq 0}$-Poisson point process with unit intensity, see Lemma A.3 for this classic fact. In particular it is independent of the $\mathcal{F}_T$-measurable random variable X_T. Additionally, X_T satisfies Assumption (2.2), which becomes in this case: for all $r > 0$

$$\int_0^r \int_{(-A,0]} h^+(u-s)(S_T N^h)(ds)\ du < +\infty \qquad \mathbb{P}_{\mathfrak{m}}\text{-a.s.}$$

We have indeed that:

$$\begin{aligned}
&\int_0^r \int_{(-A,0]} h^+(u-s)(S_T N^h)(ds) du \\
&\qquad = \int_0^r \int_{(-A+T,T]} h^+(T+u-s) N^h(ds)\ du \\
&\qquad = \int_T^{T+r} \int_{(-A+T,T]} h^+(v-s) N^h(ds)\ dv \\
&\qquad \leq \int_T^{T+r} \int_{(-\infty,0]} h^+(v-s) N^0(ds)\ dv + \int_T^{T+r} \int_{(0,T]} h^+(v-s) N^h(ds)\ dv \\
&\qquad \leq \int_T^{T+r} \int_{(-\infty,0]} h^+(v-s) N^0(ds)\ dv + \|h^+\|_1 N^h(0,T] \\
&\qquad < +\infty \qquad \mathbb{P}_{\mathfrak{m}}\text{-a.s.},
\end{aligned}$$

since the distribution $\mathfrak{m}$ of N^0 satisfies (2.2), and since we have shown at the end of the proof of Proposition 2.1 that $\mathbb{E}_{\mathfrak{m}}(N^h(0,t]) < +\infty$ for all $t > 0$.

Thus the assumptions of Proposition 2.1 are satisfied, which yields that $(X_{T+t})_{t\geq 0}$ is the pathwise unique, and hence weakly unique, strong solution of Eq. (1.2) started at X_T and driven by the $(\mathcal{F}_{T+t})_{t\geq 0}$-Poisson point process $S_T Q$. Hence, it is a process started at X_T which is a $(\mathcal{F}_{T+t})_{t\geq 0}$-Markov process with same transition semi-group as $(X_t)_{t\geq 0}$. If we wish to be more specific, for every bounded Borel function F on $\mathbb{D}(\mathbb{R}_+, \mathcal{N}((-A,0]))$ we set

$$\Pi F(x) \triangleq \mathbb{E}_x(F((X_t)_{t\geq 0}))$$

and note that existence and uniqueness in law for (1.2) yield that

$$\mathbb{E}_x(F((X_t)_{t\geq 0}) \,|\, T < \infty, \mathcal{F}_T) = \Pi F(X_T)\,.$$

This is the strong Markov property we were set to prove. □

3.2 Renewal of X at $\emptyset$

Using $(X_t)_{t\geq 0}$ and Proposition 3.1, we obtain that if T is a stopping time such that $N^h|_{(T-A,T]} = \emptyset$, then $N^h|_{(T,+\infty)}$ is independent of $N^h|_{(-\infty,T]}$ and behaves as N^h started from $\emptyset$ and translated by T. Such renewal times lead to an interesting decomposition of N^h, enlightening its dependence structure.

The successive hitting times of $\emptyset \in \mathcal{N}((-A,0])$ for the Markov process X are such renewal times. This subsection is devoted to the study of their properties. Recall that we have introduced in (1.7) the first hitting time of $\emptyset \in \mathcal{N}((-A,0])$ for X, given by

$$\tau \triangleq \inf\{t > 0 : X_{t-} \neq \emptyset, X_t = \emptyset\} = \inf\{t > 0 : N^h[t-A,t) \neq 0, N^h(t-A,t] = 0\}\,.$$

It depends on A, but this is omitted in the notation. It is natural to study whether τ is finite or not. When the reproduction function h is nonnegative, we introduce the queue $(Y_t)_{t\geq 0}$ defined by (2.12), and its return time to zero $\mathcal{T}_1$ defined in (2.13). The following result will yield the finiteness of τ.

Lemma 3.2. *Under Assumption 1.2. Let $A < \infty$ be such that $A \geq L(h)$. Let τ and $\mathcal{T}_1$ be defined in* (1.7) *and* (2.13)*. If h is nonnegative then $\mathbb{P}_\emptyset(\tau = \mathcal{T}_1) = 1$.*

Proof. We use the notations defined in Subsection 2.2. To begin with, we remark that $\tau > V_1$. First, let us consider t such that $V_1 < t < \mathcal{T}_1$. By definition, there exists $i \geq 1$, such that

$$V_i \leq t \leq V_i + \widetilde{H}_i(A) = V_i + H_i + A.$$

Since the interval $[V_i, V_i + H_i]$ corresponds to the cluster of descendants of V_i, there exists a sequence of points of N^h in $[V_i, V_i + H_i]$ which are distant by less than $L(h)$ and thus less than A. Therefore, if $t \in [V_i, V_i + H_i]$, then $N^h(t-A,t] > 0$

If $t \in [V_i + H_i, V_i + H_i + A]$, then $N^h(t-A,t] > 0$ as well since $V_i + H_i \in N^h$ (it is the last birth time in the Galton-Watson tree steming from V_i, by definition of H_i). Since this reasoning holds for any $t \leq \mathcal{T}_1$, thus $\tau \geq \mathcal{T}_1$.

Conversely, for any $t \in [V_1, \tau)$, by definition of τ necessarily $N^h(t-A,t] > 0$. Thus there exists an atom of N^h in $(t-A,t]$, and from the cluster representation, there exists $i \geq 1$ such that this atom belongs to the cluster of V_i, hence to $[V_i, V_i + H_i]$. We easily deduce that

$$V_i \leq t \leq V_i + H_i + A$$

and thus $Y_t \geq 1$, for all $t \in [V_1, \tau)$. This proves that $\tau \leq \mathcal{T}_1$ and concludes the proof. $\square$

To extend the result of finiteness of τ when no assumption is made on the sign of h, we use the coupling between N^h and N^{h^+} stated in Proposition 2.1, b).

Proposition 3.3. *Under Assumptions 1.2. Let $A < \infty$ be such that $A \geq L(h)$. Let τ be defined in* (1.7)*, and τ^+ be defined similarly with h^+ instead of h. Then*

$$\mathbb{P}_\mathfrak{m}(\tau \leq \tau^+) = 1\,.$$

Proof. We use the coupling (N^h, N^{h^+}) of Proposition 2.1, b), which satisfies $N^h \leq N^{h^+}$. If $\tau = +\infty$, since the immigration rate λ is positive, for any $t \geq 0$ necessarily $N^h(t-A,t] > 0$ and thus $N^{h^+}(t-A,t] > 0$, which implies that $\tau^+ = +\infty$ also, a.s.

Now, it is enough to prove that $\tau \leq \tau^+$ when both times are finite. In this case, since N^{h^+} is locally finite a.s., $\tau^+ - A$ is an atom of N^{h^+} such that $N^{h^+}(\tau^+ - A, \tau^+] = 0$. This implies that $N^h(\tau^+ - A, \tau^+] = 0$. If $\tau^+ - A$ is also an atom of N^h, then $\tau \leq \tau^+$.

Else, first prove that $N^h(-A, \tau^+ - A) > 0$. The result is obviously true if $N^0 \neq \emptyset$. When $N^0 = \emptyset$, the first atoms of N^h and N^{h^+} coincide because $\Lambda_0^h = \Lambda_0^{h^+}$, where these functions are defined in (2.5). This first atom is necessarily before $\tau^+ - A$, and hence $N^h(-A, \tau^+ - A) > 0$. The last atom U of N^h before $\tau^+ - A$ is thus well defined, and necessarily satisfies $N^h(U, U+A] = 0$ and $N^h[U, U+A) \neq 0$ so that $\tau \leq U + A \leq \tau^+$. We have thus proved that $\tau \leq \tau^+$, $\mathbb{P}_{\mathfrak{m}}$-a.s. □

We now prove that the regeneration time τ admits an exponential moment which ensures that it is finite almost surely. The results will rely on the coupling between N^h and N^{h^+} and on the results obtained in Section 2.2. We define

$$\gamma^+ \triangleq \frac{\|h^+\|_1 - \log(\|h^+\|_1) - 1}{L(h^+)} .$$

Proposition 3.4. *Under Assumption 1.2. Let $A < \infty$ be such that $A \geq L(h)$, and assume that $\mathbb{E}_{\mathfrak{m}}(N^0(-A, 0]) < +\infty$. Then τ given by (1.7) satisfies*

$$\forall \alpha < \min(\lambda, \gamma^+) , \quad \mathbb{E}_{\mathfrak{m}}(\mathrm{e}^{\alpha\tau}) < +\infty .$$

In particular τ is finite, $\mathbb{P}_{\mathfrak{m}}$-a.s., and $\mathbb{E}_{\mathfrak{m}}(\tau) < +\infty$.

Proof. Using Proposition 3.3, it is sufficient to prove this for τ^+. When $\mathfrak{m}$ is the Dirac measure at $\emptyset$, the result is a direct consequence of Lemma 3.2 and Proposition 2.6. We now turn to the case when $\mathfrak{m}$ is different from $\delta_\emptyset$. The proof is separated in several steps.

Step 1: Analysis of the problem. To control τ^+, we distinguish the points of N^h coming from the initial condition from the points coming from ancestors arrived after zero. We thus introduce $K = N^0((-A, 0])$, the number of atoms of N^0, $(V_i^0)_{1\leq i \leq K}$, these atoms, and $(\widetilde{H}_i^0(A))_{1\leq i\leq K}$ the durations such that $V_i^0 + \widetilde{H}_i^0(A) - A$ is the time of birth of the last descendant of V_i^0. Note that V_i^0 has no offspring before time 0, so that the reproduction function of V_i^0 is a truncation of h. We finally define the time when the influence of the past before 0 has vanished, given by

$$E = \max_{1\leq i\leq K} (V_i^0 + \widetilde{H}_i^0(A)),$$

with the convention that $E = 0$ if $K = 0$. If $K > 0$, since $V_i^0 \in (-A, 0]$ and $\widetilde{H}_i^0(A) \geq A$, we have $E > 0$. Note that $\tau^+ \geq E$.
We now consider the sequence $(V_i)_{i\geq 1}$ of ancestors arriving after time 0 at rate λ. We recall that they can be viewed as the arrival of customers in a $M/G/\infty$ queue with time service of law that of $\widetilde{H}_1(A)$. In our case, the queue may not be empty at time 0, when $E > 0$. In that case, the queue returns to 0 when all the customers arrived before time 0 have left the system (which is the case at time E) and when all the busy periods containing the customers arrived at time between 0 and E are over. The first hitting time of 0 for the queue is thus equal to

$$\tau^+ = \begin{cases} E & \text{if} \quad Y_E = 0 , \\ \inf\{t \geq E : Y_t = 0\} & \text{if} \quad Y_E > 0 , \end{cases} \tag{3.4}$$

where Y_t is given in (2.12) by $Y_t = \sum_{k:0\leq V_k\leq t} \mathbb{1}_{\{V_k + \widetilde{H}_k(A) > t\}}$.

Step 2: Exponential moments of E. In (3.4), E depends only on N^0 and $(Y_t)_{t\geq 0}$ only on the arrivals and service times of customers entering the queue after time 0. A natural idea is then to condition with respect to E, and for this it is important to gather estimates on the moments of E. Since $V_i^0 \leq 0$, we have that

$$0 \leq E \leq \max_{1\leq i\leq K} \widetilde{H}_i^0(A).$$

The truncation mentioned in Step 1 implies that the $\widetilde{H}_i^0(A)$ are stochastically dominated by independent random variables distributed as $\widetilde{H}_1$, which we denote by $\bar{H}_i^0(A)$. Thus for $t > 0$, using (2.14),

$$\begin{aligned}\mathbb{P}_{\mathfrak{m}}(E > t) &\leq \mathbb{P}_{\mathfrak{m}}\big(\max_{1\leq i\leq K} \bar{H}_i^0(A) > t\big)\\ &= 1 - \mathbb{E}_{\mathfrak{m}}\Big(\big(1 - \mathbb{P}(\widetilde{H}_1(A) > t)\big)^K\Big)\\ &\leq 1 - \mathbb{E}_{\mathfrak{m}}\big((1 - C\mathrm{e}^{-\gamma^+ t})^K\big)\,.\end{aligned}$$

Thus there exists $t_0 > 0$ such that for any $t > t_0$,

$$\mathbb{P}_{\mathfrak{m}}(E > t) \leq C\mathbb{E}_{\mathfrak{m}}(N^0(-A,0])\mathrm{e}^{-\gamma^+ t}. \tag{3.5}$$

As a corollary, we have for any $\beta \in (0,\gamma^+)$ that

$$\mathbb{E}_{\mathfrak{m}}(\mathrm{e}^{\beta E}) < +\infty\,. \tag{3.6}$$

Step 3: Estimate of the tail distribution of τ^+. For $t > 0$, we have

$$\begin{aligned}\mathbb{P}_{\mathfrak{m}}(\tau^+ > t) &= \mathbb{P}_{\mathfrak{m}}(\tau^+ > t,\, E > t) + \mathbb{P}_{\mathfrak{m}}(\tau^+ > t,\, E \leq t)\\ &\leq \mathbb{P}_{\mathfrak{m}}(E > t) + \mathbb{E}_{\mathfrak{m}}\Big(\mathbb{1}_{\{E\leq t\}}\, \mathbb{P}_{\mathfrak{m}}(\tau^+ > t \,|\, E)\Big).\end{aligned}$$

The first term is controlled by (3.5). For the second term, we use Proposition A.2 which is a consequence of Theorem A.1. For this, let us introduce a constant κ such that $\kappa < \gamma^+$ if $\gamma^+ \leq \lambda$ and $\kappa = \lambda$ if $\lambda < \gamma^+$. We have:

$$\mathbb{E}_{\mathfrak{m}}\Big(\mathbb{1}_{\{E\leq t\}}\, \mathbb{P}(\tau^+ > t \,|\, E)\Big) \leq \mathbb{E}_{\mathfrak{m}}\big(\mathbb{1}_{\{E\leq t\}}\, \lambda C E\, \mathrm{e}^{-\kappa(t-E)}\big) = \lambda C \mathrm{e}^{-\kappa t}\mathbb{E}_{\mathfrak{m}}\big(\mathbb{1}_{\{E\leq t\}}\, E\, \mathrm{e}^{\kappa E}\big).$$

Since $\kappa < \gamma^+$, it is always possible to choose $\beta \in (\kappa,\gamma^+)$ such that (3.6) holds, which entails that $\mathbb{E}_{\mathfrak{m}}(\mathbb{1}_{\{E\leq t\}}\, E\, \mathrm{e}^{\kappa E})$ can be bounded by a finite constant independent of t.

Gathering all the results,

$$\mathbb{P}_{\mathfrak{m}}(\tau^+ > t) \leq C\mathbb{E}_{\mathfrak{m}}(N^0(-A,0])\mathrm{e}^{-\gamma^+ t} + \lambda C' \mathrm{e}^{-\kappa t} = O\big(\mathrm{e}^{-\kappa t}\big).$$

This yields that $\mathbb{E}_{\mathfrak{m}}(\mathrm{e}^{\alpha\tau^+}) < +\infty$ for any $\alpha < \kappa$, i.e. $\alpha < \min(\lambda,\gamma^+)$. □

Theorem 3.5. *Under Assumptions 1.2. The strong Markov process $X = (X_t)_{t\geq 0}$ defined by* (3.2) *admits a unique invariant law, π_A, defined on $\mathcal{N}((-A,0])$ by* (1.8)*: for every Borel nonnegative function f,*

$$\pi_A f = \frac{1}{\mathbb{E}_{\emptyset}(\tau)}\mathbb{E}_{\emptyset}\Big(\int_0^\tau f(X_t)\,dt\Big)\,.$$

Moreover, $\pi_A\{\emptyset\} = 1/(\lambda\mathbb{E}_{\emptyset}(\tau))$ and thus the null measure $\emptyset$ is a positive recurrent state in the classical sense for X.

Proof. We recall the classic proof. Let $(P_s)_{s\geq 0}$ denote the semi-group of X and f be a Borel nonnegative function. Then

$$\pi_A P_s f = \frac{1}{\mathbb{E}_\emptyset(\tau)}\mathbb{E}_\emptyset\Big(\int_0^\tau P_s f(X_t)\,dt\Big) = \frac{1}{\mathbb{E}_\emptyset(\tau)}\int_0^\infty \mathbb{E}_\emptyset\big(\mathbb{1}_{\{\tau>t\}}P_s f(X_t)\big)\,dt\,.$$

Using the Markov property at time t and since $\{\tau > t\} \in \mathcal{F}_t$,

$$\mathbb{E}_\emptyset\big(\mathbb{1}_{\{\tau>t\}}P_s f(X_t)\big) = \mathbb{E}_\emptyset\big(\mathbb{1}_{\{\tau>t\}}\mathbb{E}_\emptyset(f(X_{t+s}) \mid \mathcal{F}_t)\big) = \mathbb{E}_\emptyset\big(\mathbb{1}_{\{\tau>t\}}f(X_{t+s})\big)$$

and thus

$$\pi_A P_s f = \frac{1}{\mathbb{E}_\emptyset(\tau)}\int_0^\infty \mathbb{E}_\emptyset\big(\mathbb{1}_{\{\tau>t\}}f(X_{t+s})\big)\,dt = \frac{1}{\mathbb{E}_\emptyset(\tau)}\mathbb{E}_\emptyset\Big(\int_0^\tau f(X_{t+s})\,dt\Big)\,.$$

Using the strong Markov property at time τ,

$$\begin{aligned}
\mathbb{E}_\emptyset\Big(\int_0^\tau f(X_{t+s})\,dt\Big) &= \mathbb{E}_\emptyset\Big(\int_s^{\tau+s} f(X_t)\,dt\Big)\\
&= \mathbb{E}_\emptyset\Big(\int_s^{\tau} f(X_t)\,dt\Big) + \mathbb{E}_\emptyset\Big(\int_\tau^{\tau+s} f(X_t)\,dt\Big)\\
&= \mathbb{E}_\emptyset\Big(\int_s^{\tau} f(X_t)\,dt\Big) + \mathbb{E}_\emptyset\Big(\int_0^{s} f(X_t)\,dt\Big)\\
&= \mathbb{E}_\emptyset\Big(\int_0^{\tau} f(X_t)\,dt\Big).
\end{aligned}$$

Thus $\pi_A P_s f = \pi_A f$. We conclude that π_A is an invariant law for $(P_s)_{s\geq 0}$.

The proof of its uniqueness is an immediate consequence of Theorem 1.3 b), which will be proved later. Indeed, for any invariant law π of X it holds that

$$\pi = \mathbb{P}_\pi(X_t \in \cdot) \xrightarrow[t\to\infty]{\text{total variation}} \mathbb{P}_{\pi_A}(X_0 \in \cdot) = \pi_A.$$

From the definition of π_A, we obtain that

$$\pi_A\{\emptyset\} = \frac{1}{\mathbb{E}_\emptyset(\tau)}\mathbb{E}_\emptyset\Big(\int_0^\tau \mathbb{1}_{\{\emptyset\}}(X_t)\,dt\Big)\,.$$

Under $\mathbb{P}_\emptyset$, an excursion $(X_t)_{t\in(0,\tau]}$ proceeds as follows. First, $X_t = \emptyset$ for $t \in (0, U_1^h)$ with U_1^h the first atom of N^h defined in (2.6). Under $\mathbb{P}_\emptyset$, U_1^h follows an exponential distribution with expectation $1/\lambda$. Then, $X_t \neq \emptyset$ for $t \in [U_1^h, \tau)$ by definition of τ. We deduce from this that

$$\pi_A\{\emptyset\} = \frac{\mathbb{E}_\emptyset(U_1^h)}{\mathbb{E}_\emptyset(\tau)} = \frac{1}{\lambda\mathbb{E}_\emptyset(\tau)}\,.$$

This concludes the proof. □

The strong Markov property of X yields a sequence of regeneration times $(\tau_k)_{k\geq 0}$, which are the successive visits of X to the null measure $\emptyset$, defined as follows (the time τ_0 has already been introduced in (1.10)):

$$\begin{aligned}
&\tau_0 = \inf\{t \geq 0 : X_t = \emptyset\} && \text{(First entrance time of } \emptyset)\\
&\tau_k = \inf\{t > \tau_{k-1} : X_{t-} \neq \emptyset, X_t = \emptyset\}\,, \quad k \geq 1\,. && \text{(Successive return times at } \emptyset)
\end{aligned}$$

They provide a useful decomposition of the path of X in i.i.d. excursions:

Theorem 3.6. *Let N^h be a Hawkes process satisfying Assumption 1.2, and $A \geq L(h)$. Consider the Markov process X defined in* (3.2)*. Under $\mathbb{P}_{\mathfrak{m}}$ the following holds:*

a) The τ_k for $k \geq 0$ are finite stopping times, a.s.

b) The delay $(X_t)_{t\in[0,\tau_0)}$ is independent of the cycles $(X_{\tau_{k-1}+t})_{t\in[0,\tau_k-\tau_{k-1})}$ for $k \geq 1$.

c) These cycles are i.i.d. and distributed as $(X_t)_{t\in[0,\tau)}$ under $\mathbb{P}_\emptyset$. In particular their durations $(\tau_k - \tau_{k-1})_{k\geq 1}$ are distributed as τ under $\mathbb{P}_\emptyset$, so that $\lim_{k\to+\infty} \tau_k = +\infty$, $\mathbb{P}_{\mathfrak{m}}$-a.s.

Proof. The above items follow classically from the strong Markov property of X. Let us first prove the finiteness of the return times τ_k. For any $\mathfrak{m}$, from the definition of τ_0 and τ, we have that $\tau_0 \leq \tau$, $\mathbb{P}_{\mathfrak{m}}$-a.s. Then, $\mathbb{P}_{\mathfrak{m}}(\tau_0 < +\infty) = 1$ follows from Proposition 3.4 (i). For $k \geq 1$, using the strong Markov property of X, we have for any $\mathfrak{m}$:

$$\begin{aligned}
\mathbb{P}_{\mathfrak{m}}(\tau_k < +\infty) &= \mathbb{E}_{\mathfrak{m}}\big(\mathbb{1}_{\{\tau_{k-1}<+\infty\}}\,\mathbb{P}_{X_{\tau_{k-1}}}(\tau < +\infty)\big) \\
&= \mathbb{E}_{\mathfrak{m}}\big(\mathbb{1}_{\{\tau_{k-1}<+\infty\}}\,\mathbb{P}_{\emptyset}(\tau < +\infty)\big) \\
&= \mathbb{P}_{\mathfrak{m}}(\tau_{k-1} < +\infty) = \cdots = \mathbb{P}_{\mathfrak{m}}(\tau_0 < +\infty) = 1.
\end{aligned}$$

Let us now prove a) and b). It is sufficient to consider $(X_t)_{t\in[0,\tau_0)}$, $(X_{\tau_0+t})_{t\in[0,\tau_1-\tau_0)}$ and $(X_{\tau_1+t})_{t\in[0,\tau_2-\tau_1)}$. Let F_0, F_1, and F_2 be three measurable bounded real functions on $\mathbb{D}(\mathbb{R}_+, \mathcal{N}(-A,0])$. Then, using the strong Markov property successively at τ_1 and τ_0, we obtain:

$$\begin{aligned}
&\mathbb{E}_{\mathfrak{m}}\Big(F_0\big((X_t)_{t\in[0,\tau_0)}\big)\,F_1\big((X_{\tau_0+t})_{t\in[0,\tau_1-\tau_0)}\big)\,F_2\big((X_{\tau_1+t})_{t\in[0,\tau_2-\tau_1)}\big)\Big) \\
&\quad = \mathbb{E}_{\mathfrak{m}}\Big(F_0\big((X_t)_{t\in[0,\tau_0)}\big)\Big)\,\mathbb{E}_{\emptyset}\Big(F_1\big((X_t)_{t\in[0,\tau)}\big)\Big)\,\mathbb{E}_{\emptyset}\Big(F_2\big((X_t)_{t\in[0,\tau)}\big)\Big).
\end{aligned}$$

This concludes the proof. □

4 Proof of the main results

We translate the statements of the main results in terms of the Markov process X. Let $T > 0$ be fixed, and define

$$K_T \triangleq \max\{k \geq 0 : \tau_k \leq T\}\,, \tag{4.1}$$

which goes to infinity with T since the sequence $(\tau_k)_{k\geq 0}$ increases to infinity. For a locally bounded Borel function f on $\mathcal{N}((-A,0])$ we define the random variables

$$I_k f \triangleq \int_{\tau_{k-1}}^{\tau_k} f(X_t)\,dt\,, \quad k \geq 1\,, \tag{4.2}$$

which are finite a.s., i.i.d., and of the same law as $\int_0^\tau f(X_t)\,dt$ under $\mathbb{P}_\emptyset$, see Theorem 3.6.

Proof of Theorem 1.3 a)

Assume first that $f \geq 0$. Then, with the notation (4.1) and (4.2),

$$\frac{1}{K_T}\sum_{k=1}^{K_T} I_k f \leq \frac{1}{K_T}\int_0^T f(X_t)\,dt \leq \frac{1}{K_T}\int_0^{\tau_0} f(X_t)\,dt + \frac{1}{K_T}\sum_{k=1}^{K_T+1} I_k f\,.$$

Since f is locally bounded, $\int_0^{\tau_0} f(X_t)\,dt$ is finite, $\mathbb{P}_{\mathfrak{m}}$-a.s., thus

$$\mathbb{P}_{\mathfrak{m}}\left(\frac{1}{K_T}\int_0^{\tau_0} f(X_t)\,dt \xrightarrow[T\to\infty]{} 0\right) = 1.$$

The strong law of large numbers applied to the i.i.d. sequence $(I_k f)_{k\geq 1}$ yields that

$$\frac{1}{K_T}\sum_{k=1}^{K_T+1} I_k f \xrightarrow[T\to\infty]{\mathbb{P}_{\mathfrak{m}}-\text{a.s.}} \mathbb{E}_{\mathfrak{m}}(I_1 f) = \mathbb{E}_{\emptyset}\left(\int_0^{\tau} f(X_t)\ dt\right), \quad \mathbb{P}_{\mathfrak{m}}\text{-a.s.}$$

Gathering the two last limits,

$$\frac{1}{K_T}\int_0^T f(X_t)\,dt \xrightarrow[T\to\infty]{\mathbb{P}_{\mathfrak{m}}-\text{a.s.}} \mathbb{E}_{\emptyset}\left(\int_0^{\tau} f(X_t)\,dt\right) = \mathbb{E}_{\emptyset}(\tau)\,\pi_A f\ .$$

Choosing $f = 1$ yields that

$$\frac{T}{K_T} \xrightarrow[T\to\infty]{\mathbb{P}_{\mathfrak{m}}-\text{a.s.}} \mathbb{E}_{\emptyset}(\tau) < \infty\,. \tag{4.3}$$

Dividing the first limit by the second concludes the proof for $f \geq 0$.

The case of π_A-integrable signed f follows by the decomposition $f = f^+ - f^-$.

Proof of Theorem 1.3 b)

This follows from a general result in Thorisson [30, Theorem 10.3.3 p.351], which yields that if the distribution of τ under $\mathbb{P}_{\emptyset}$ has a density with respect to the Lebesgue measure and if $\mathbb{E}_{\emptyset}(\tau) < +\infty$, then there exists a probability measure $\mathbb{Q}$ on $\mathbb{D}(\mathbb{R}_+, \mathcal{N}(-A,0])$ such that, for any initial law $\mathfrak{m}$,

$$\mathbb{P}_{\mathfrak{m}}\big((X_{t+u})_{u\geq 0} \in \cdot\big) \xrightarrow[t\to\infty]{\text{total variation}} \mathbb{Q}\,.$$

Since π_A is an invariant law, $\mathbb{P}_{\pi_A}\big((X_{t+u})_{u\geq 0} \in \cdot\big) = \mathbb{P}_{\pi_A}(X \in \cdot)$ for every $t \geq 0$. Hence, taking $\mathfrak{m} = \pi_A$ in the above convergence yields that $\mathbb{Q} = \mathbb{P}_{\pi_A}(X \in \cdot)$.

It remains to check the above assumptions of the theorem. Proposition 3.4 yields that $\mathbb{E}_{\emptyset}(\tau) < +\infty$. Moreover, under $\mathbb{P}_{\emptyset}$ we can rewrite τ as

$$\tau = U_1^h + \inf\big\{t > 0 :\ X_{(t+U_1^h)_-} \neq \emptyset \text{ and } X_{t+U_1^h} = \emptyset\big\}.$$

Using the strong Markov property, we easily prove independence of the two terms in the r.h.s. Since U_1^h has an exponential distribution under $\mathbb{P}_{\emptyset}$, τ has a density under $\mathbb{P}_{\emptyset}$.

Proof of Theorem 1.4

Let $\tilde{f} \triangleq f - \pi_A f$. With the notation (4.1) and (4.2), we have the decomposition

$$\int_0^T \tilde{f}(X_t)\,dt = \int_0^{\tau_0} \tilde{f}(X_t)\,dt + \sum_{k=1}^{K_T} I_k\tilde{f} + \int_{\tau_{K_T}}^T \tilde{f}(X_t)\,dt\,. \tag{4.4}$$

The $I_k\tilde{f}$ are i.i.d. and are distributed as $\int_0^\tau \tilde{f}(X_t)\,dt$ under $\mathbb{P}_\emptyset$, with expectation 0 and variance $\mathbb{E}_\emptyset(\tau)\sigma^2(f)$, see Theorem 3.6. Since f is locally bounded, so is $\tilde{f}$ and

$$\frac{1}{\sqrt{T}}\int_0^{\tau_0} \tilde{f}(X_t)\,dt \xrightarrow[T\to\infty]{\mathbb{P}_\mathfrak{m}-\text{a.s.}} 0\,.$$

Now, let $\varepsilon > 0$. For arbitrary $a > 0$ and $0 < u \leq T$,

$$\mathbb{P}_\mathfrak{m}\left(\left|\int_{\tau_{K_T}}^T \tilde{f}(X_t)\,dt\right| > a\right) \leq \mathbb{P}_\mathfrak{m}(T - \tau_{K_T} > u) + \mathbb{P}_\mathfrak{m}\left(\sup_{0\leq s\leq u}\left|\int_{T-s}^T \tilde{f}(X_t)\,dt\right| > a\right).$$

But

$$\mathbb{P}_\mathfrak{m}(T - \tau_{K_T} > u) = 1 - \mathbb{P}_\mathfrak{m}(\exists t \in [T-u, T] : X_{t-} \neq \emptyset, X_t = \emptyset)$$

and Theorem 1.3 b) yields that

$$\lim_{T\to\infty} \mathbb{P}_\mathfrak{m}(T - \tau_{K_T} > u) = 1 - \mathbb{P}_{\pi_A}(\exists t \in [0, u] : X_{t-} \neq \emptyset, X_t = \emptyset)\ ,$$

so that there exists u_0 large enough such that

$$\lim_{T\to\infty} \mathbb{P}_\mathfrak{m}(T - \tau_{K_T} > u_0) < \frac{\varepsilon}{2}\,.$$

Moreover Theorem 1.3 b) yields that

$$\lim_{T\to\infty} \mathbb{P}_\mathfrak{m}\left(\sup_{0\leq s\leq u_0}\left|\int_{T-s}^T \tilde{f}(X_t)\,dt\right| > a\right) = \mathbb{P}_{\pi_A}\left(\sup_{0\leq s\leq u_0}\left|\int_0^s \tilde{f}(X_t)\,dt\right| > a\right)$$

and thus there exists a_0 large enough such that

$$\lim_{T\to\infty} \mathbb{P}_\mathfrak{m}\left(\sup_{0\leq s\leq u_0}\left|\int_{T-s}^T \tilde{f}(X_t)\,dt\right| > a_0\right) < \frac{\varepsilon}{2}$$

and hence

$$\limsup_{T\to\infty} \mathbb{P}_\mathfrak{m}\left(\left|\int_{\tau_{K_T}}^T \tilde{f}(X_t)\,dt\right| > a_0\right) < \varepsilon\,.$$

This implies in particular that

$$\frac{1}{\sqrt{T}}\int_{\tau_{K_T}}^T \tilde{f}(X_t)\,dt \xrightarrow[T\to\infty]{\text{probab.}} 0\,.$$

It now remains to treat the second term in the r.h.s. of (4.4). The classic central limit theorem yields that

$$\frac{1}{\sqrt{T}}\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} I_k\tilde{f} \xrightarrow[T\to\infty]{\text{in law}} \frac{1}{\sqrt{\mathbb{E}_\emptyset(\tau)}}\mathcal{N}(0, \mathbb{E}_\emptyset(\tau)\sigma^2(f)) = \mathcal{N}(0, \sigma^2(f))$$

and we are left to control

$$\Delta_T \triangleq \frac{1}{\sqrt{T}} \sum_{k=1}^{K_T} I_k \tilde{f} - \frac{1}{\sqrt{T}} \sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} I_k \tilde{f} \,.$$

Let $\varepsilon > 0$ and define

$$v(T,\varepsilon) \triangleq \{\lfloor (1-\varepsilon^3)T/\mathbb{E}_\emptyset(\tau)\rfloor, \dots, \lfloor (1+\varepsilon^3)T/\mathbb{E}_\emptyset(\tau)\rfloor\} \,.$$

Note that

$$(1-\varepsilon^3)\frac{T}{\mathbb{E}_\emptyset(\tau)} < \frac{T}{\mathbb{E}_\emptyset(\tau)} < (1+\varepsilon^3)\frac{T}{\mathbb{E}_\emptyset(\tau)} \;,$$

which implies that $\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor \in v(T,\varepsilon)$. In view of (4.3), there exists t_ε such that if $T \geq t_\varepsilon$

$$\mathbb{P}_{\mathfrak{m}}(K_T \in v(T,\varepsilon)) > 1 - \varepsilon \,.$$

For $T \geq t_\varepsilon$, we thus have on $\{K_T \in v(T,\varepsilon)\}$ that

$$\begin{aligned} |\Delta_T| &\leq \left| \frac{1}{\sqrt{T}} \sum_{k=\lfloor (1-\varepsilon^3)T/\mathbb{E}_\emptyset(\tau)\rfloor}^{K_T} I_k \tilde{f} \right| + \left| \frac{1}{\sqrt{T}} \sum_{k=\lfloor (1-\varepsilon^3)T/\mathbb{E}_\emptyset(\tau)\rfloor}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} I_k \tilde{f} \right| \\ &\leq \frac{2}{\sqrt{T}} \max_{n \in v(T,\varepsilon)} \left| \sum_{k=\lfloor (1-\varepsilon^3)T/\mathbb{E}_\emptyset(\tau)\rfloor}^{n} I_k \tilde{f} \right| . \end{aligned}$$

Using now Kolmogorov's maximal inequality [15, Sect. IX.7 p.234] we obtain

$$\mathbb{P}_{\mathfrak{m}}(|\Delta_T| \geq \varepsilon) \leq \frac{\lfloor (1+\varepsilon^3)T/\mathbb{E}_\emptyset(\tau)\rfloor - \lfloor (1-\varepsilon^3)T/\mathbb{E}_\emptyset(\tau)\rfloor}{\varepsilon^2 T/4} \mathbb{E}_\emptyset(\tau)\sigma^2(f) \leq 8\sigma^2(f)\varepsilon \,.$$

Since $\varepsilon > 0$ is arbitrary, we conclude that

$$\left| \frac{1}{\sqrt{T}} \sum_{k=1}^{K_T} I_k \tilde{f} - \frac{1}{\sqrt{T}} \sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} I_k \tilde{f} \right| \xrightarrow[T\to\infty]{\text{probab.}} 0 \,.$$

These three convergence results and Slutsky's theorem yield the convergence result.

Proof of Theorem 1.5

With the notation $\tilde{f} \triangleq f - \pi_A f$ and (4.2), let us consider the decomposition

$$\int_0^T \tilde{f}(X_t)\,dt = \int_0^{\tau_0} \tilde{f}(X_t)\,dt + \sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} I_k \tilde{f} + \int_{\tau_{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor}}^{T} \tilde{f}(X_t)\,dt \,. \tag{4.5}$$

The $I_k \tilde{f}$ are i.i.d. and are distributed as $\int_0^\tau \tilde{f}(X_t)\,dt$ under $\mathbb{P}_\emptyset$, with expectation 0 and variance $\mathbb{E}_\emptyset(\tau)\sigma^2(f)$, see Theorem 3.6. Since f takes its values in $[a,b]$,

$$\left| \int_0^{\tau_0} \tilde{f}(X_t)\,dt \right| \leq |b-a|\tau_0 \,, \qquad \left| \int_{\tau_{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor}}^{T} \tilde{f}(X_t)\,dt \right| \leq |b-a|\,|T - \tau_{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor}| \,.$$

But

$$T - \tau_{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} = -\tau_0 - \sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} (\tau_k - \tau_{k-1}) + T$$
$$= -\tau_0 - \sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} (\tau_k - \tau_{k-1} - \mathbb{E}_\emptyset(\tau)) + T - \lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor \mathbb{E}_\emptyset(\tau)$$

in which $0 \le T - \lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor \mathbb{E}_\emptyset(\tau) < \mathbb{E}_\emptyset(\tau)$ and the $\tau_k - \tau_{k-1} - \mathbb{E}_\emptyset(\tau)$ are i.i.d., have same law as $\tau - \mathbb{E}_\emptyset(\tau)$ under $\mathbb{P}_\emptyset$, and have expectation 0 and variance $\mathrm{Var}_\emptyset(\tau)$. Thus,

$$\mathbb{P}_\mathfrak{m}\left(\left|\frac{1}{T}\int_0^T f(X_t)\,dt - \pi_A f\right| \ge \varepsilon\right)$$
$$\le \mathbb{P}_\mathfrak{m}\left(\left|\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} I_k\tilde{f}\right| + |b-a|\left(2\tau_0 + \left|\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} (\tau_k - \tau_{k-1} - \mathbb{E}_\emptyset(\tau))\right| + \mathbb{E}_\emptyset(\tau)\right) \ge T\varepsilon\right) .$$

Now, using that

$$T\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau) - 2|b-a|\mathbb{E}_\mathfrak{m}(\tau_0) = 2\frac{(T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau)}{2} + \sqrt{T}\varepsilon - 2|b-a|\mathbb{E}_\mathfrak{m}(\tau_0) ,$$

we obtain that

$$\begin{aligned}
&\mathbb{P}_\mathfrak{m}\left(\left|\frac{1}{T}\int_0^T f(X_t)\,dt - \pi_A f\right| \ge \varepsilon\right) \\
&\quad\le \mathbb{P}_\mathfrak{m}\left(\left|\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} I_k\tilde{f}\right| \ge \frac{(T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau)}{2}\right) \\
&\qquad + \mathbb{P}_\mathfrak{m}\left(\left|\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} (\tau_k - \tau_{k-1} - \mathbb{E}_\emptyset(\tau))\right| \ge \frac{(T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau)}{2|b-a|}\right) \\
&\qquad + \mathbb{P}_\mathfrak{m}\left(\tau_0 - \mathbb{E}_\mathfrak{m}(\tau_0) \ge \frac{\sqrt{T}\varepsilon - 2|b-a|\mathbb{E}_\mathfrak{m}(\tau_0)}{2|b-a|}\right) .
\end{aligned} \tag{4.6}$$

We aim to apply Bernstein's inequality [20, Cor. 2.10 p.25, (2.17), (2.18) p.24] to bound the three terms of the right hand side. We recall that to apply Bernstein inequality to random variables $X_1, \dots X_N$, there should exist constants c and v such that

$$\sum_{k=1}^N \mathbb{E}_\mathfrak{m}\left[X_k^2\right] \le v, \quad \text{and} \quad \sum_{k=1}^N \mathbb{E}_\mathfrak{m}\left[(X_k)_+^n\right] \le \frac{n!}{2} v c^{n-2}, \quad \forall n \ge 3.$$

First,

$$\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} \mathbb{E}_\mathfrak{m}((I_k\tilde{f})^2) = \left\lfloor \frac{T}{\mathbb{E}_\emptyset(\tau)} \right\rfloor \mathbb{E}_\emptyset(\tau)\sigma^2(f) \le T\sigma^2(f)$$

and, for $n \geq 3$,

$$\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} \mathbb{E}_\mathfrak{m}\big((I_k\tilde f)_\pm^n\big) = \Big\lfloor \frac{T}{\mathbb{E}_\emptyset(\tau)}\Big\rfloor \mathbb{E}_\mathfrak{m}\big((I\tilde f)_\pm^n\big)$$

$$\leq \frac{n!}{2} T\sigma^2(f)\bigg(\sup_{k\geq 3}\Big(\frac{2}{k!}\frac{\mathbb{E}_\mathfrak{m}\big((I\tilde f)_\pm^k\big)}{\mathbb{E}_\emptyset(\tau)\sigma^2(f)}\Big)^{\frac{1}{k-2}}\bigg)^{n-2} \triangleq \frac{n!}{2} T\sigma^2(f)(c^\pm(f))^{n-2}\,.$$

Then,

$$\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} \mathbb{E}_\mathfrak{m}\big((\tau_k-\tau_{k-1}-\mathbb{E}_\emptyset(\tau))^2\big) = \Big\lfloor \frac{T}{\mathbb{E}_\emptyset(\tau)}\Big\rfloor \operatorname{Var}_\emptyset(\tau) \leq T\frac{\operatorname{Var}_\emptyset(\tau)}{\mathbb{E}_\emptyset(\tau)}$$

and, for $n \geq 3$,

$$\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} \mathbb{E}_\mathfrak{m}\big((\tau_k-\tau_{k-1}-\mathbb{E}_\emptyset(\tau))_\pm^n\big) = \Big\lfloor T/\mathbb{E}_\emptyset(\tau)\Big\rfloor \mathbb{E}_\emptyset\big((\tau-\mathbb{E}_\emptyset(\tau))_\pm^n\big)$$

$$\leq \frac{n!}{2} T\frac{\operatorname{Var}_\emptyset(\tau)}{\mathbb{E}_\emptyset(\tau)}\bigg(\sup_{k\geq 3}\Big(\frac{2}{k!}\frac{\mathbb{E}_\emptyset\big((\tau-\mathbb{E}_\emptyset(\tau))_\pm^k\big)}{\operatorname{Var}_\emptyset(\tau)}\Big)^{\frac{1}{k-2}}\bigg)^{n-2} \triangleq \frac{n!}{2} T\frac{\operatorname{Var}_\emptyset(\tau)}{\mathbb{E}_\emptyset(\tau)}(c^\pm(\tau))^{n-2}\,.$$

Lastly, $\mathbb{E}_\mathfrak{m}\big((\tau_0-\mathbb{E}_\mathfrak{m}(\tau_0))^2\big) = \operatorname{Var}_\mathfrak{m}(\tau_0)$ and, for $n \geq 3$,

$$\mathbb{E}_\mathfrak{m}\big((\tau_0-\mathbb{E}_\mathfrak{m}(\tau_0))_+^n\big)$$

$$\leq \frac{n!}{2}\operatorname{Var}_\mathfrak{m}(\tau_0)\bigg(\sup_{k\geq 3}\Big(\frac{2}{k!}\frac{\mathbb{E}_\mathfrak{m}\big((\tau_0-\mathbb{E}_\mathfrak{m}(\tau_0))_+^k\big)}{\operatorname{Var}_\mathfrak{m}(\tau_0)}\Big)^{\frac{1}{k-2}}\bigg)^{n-2} \triangleq \frac{n!}{2}\operatorname{Var}_\mathfrak{m}(\tau_0)(c^+(\tau_0))^{n-2}\,.$$

Applying [20, Cor. 2.10 p.25, (2.17), (2.18) p.24] to the r.h.s. of (4.6) yields that

$$\mathbb{P}_\mathfrak{m}\bigg(\Big|\frac{1}{T}\int_0^T f(X_t)\,dt - \pi_A f\Big| \geq \varepsilon\bigg)$$

$$\leq \exp\left(-\frac{((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))^2}{8T\sigma^2(f)+4c^+(f)((T-\sqrt{T})\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau))}\right)$$

$$+\exp\left(-\frac{((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))^2}{8T\sigma^2(f)+4c^-(f)((T-\sqrt{T})\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau))}\right)$$

$$+\exp\left(-\frac{((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))^2}{8T|b-a|^2\frac{\operatorname{Var}_\emptyset(\tau)}{\mathbb{E}_\emptyset(\tau)}+4|b-a|c^+(\tau)((T-\sqrt{T})\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau))}\right)$$

$$+\exp\left(-\frac{((T-\sqrt{T})\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))^2}{8T|b-a|^2\frac{\operatorname{Var}_\emptyset(\tau)}{\mathbb{E}_\emptyset(\tau)}+4|b-a|c^-(\tau)((T-\sqrt{T})\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau))}\right)$$

$$+\exp\left(-\frac{(\sqrt{T}\varepsilon - 2|b-a|\mathbb{E}_\mathfrak{m}(\tau_0))^2}{8|b-a|^2\operatorname{Var}_\mathfrak{m}(\tau_0)+4|b-a|c^+(\tau_0)(\sqrt{T}\varepsilon-2|b-a|\mathbb{E}_\mathfrak{m}(\tau_0))}\right)$$

which is (1.11).

Proof of Corollary 1.6

Under $\mathbb{P}_\emptyset$, $\tau_0 = 0$ and thus Equation (4.6) reads:

$$\mathbb{P}_\emptyset\left(\left|\frac{1}{T}\int_0^T f(X_t)\,dt - \pi_A f\right| \geq \varepsilon\right) \leq \mathbb{P}_\emptyset\left(\left|\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} I_k\tilde{f}\right| \geq \frac{T\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau)}{2}\right) \tag{4.7}$$
$$+\,\mathbb{P}_\emptyset\left(\left|\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor}(\tau_k - \tau_{k-1} - \mathbb{E}_\emptyset(\tau))\right| \geq \frac{T\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau)}{2|b-a|}\right).$$

Similarly as for the proof of Theorem 1.5 we apply Bernstein inequality for each of the terms in the right hand side. However, in order to simplify the obtained bound, we change the upper bounds of the moments of $I_k\tilde{f}$ and $\tau_k - \tau_{k1} - \mathbb{E}_\emptyset(\tau)$. Namely we use the fact that for all $n \geq 1$,

$$\mathbb{E}_\emptyset(\tau^n) \leq \frac{n!}{\alpha^n}\mathbb{E}_\emptyset(e^{\alpha\tau}) \quad \text{and} \quad \mathbb{E}_\emptyset(|\tau - \mathbb{E}_\emptyset(\tau)|^n) \leq \frac{n!}{\alpha^n}\mathbb{E}_\emptyset(e^{\alpha\tau})e^{\alpha\mathbb{E}_\emptyset(\tau)}.$$

Since τ is a nonnegative random variable, $e^{\alpha\mathbb{E}_\emptyset(\tau)} \geq 1$ and in the sequel it will be more convenient to use the following upper bound: for all $n \geq 1$,

$$\mathbb{E}_\emptyset(\tau^n) \leq \frac{n!}{\alpha^n}\mathbb{E}_\emptyset(e^{\alpha\tau})e^{\alpha\mathbb{E}_\emptyset(\tau)}.$$

Then

$$\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor}\mathbb{E}_\emptyset((I_k\tilde{f})^2) \leq \left\lfloor\frac{T}{\mathbb{E}_\emptyset(\tau)}\right\rfloor\mathbb{E}_\emptyset(\tau^2)(b-a)^2 \leq \frac{2(b-a)^2}{\alpha^2}\left\lfloor\frac{T}{\mathbb{E}_\emptyset(\tau)}\right\rfloor\mathbb{E}_\emptyset(e^{\alpha\tau})e^{\alpha\mathbb{E}_\emptyset(\tau)},$$

and, for $n \geq 3$,

$$\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor}\mathbb{E}_\emptyset(|I_k\tilde{f})|^n) \leq \frac{n!}{2}\left(\left\lfloor\frac{T}{\mathbb{E}_\emptyset(\tau)}\right\rfloor|b-a|^2\frac{2}{\alpha^2}\mathbb{E}_\emptyset(e^{\alpha\tau})e^{\alpha\mathbb{E}_\emptyset(\tau)}\right)\left(\frac{|b-a|}{\alpha}\right)^{n-2}.$$

Setting

$$v = \frac{2(b-a)^2}{\alpha^2}\left\lfloor\frac{T}{\mathbb{E}_\emptyset(\tau)}\right\rfloor\mathbb{E}_\emptyset(e^{\alpha\tau})e^{\alpha\mathbb{E}_\emptyset(\tau)}, \quad \text{and} \quad c = \frac{|b-a|}{\alpha},$$

and applying Bernstein inequality, we obtain that

$$\mathbb{P}_\emptyset\left(\left|\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor} I_k\tilde{f}\right| \geq \frac{T\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau)}{2}\right) \leq 2\exp\left(-\frac{\left(T\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau)\right)^2}{4\left(2v + (T\varepsilon - |b-a|\mathbb{E}_\emptyset(\tau))c\right)}\right).$$

Also

$$\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor}\mathbb{E}_\emptyset\left((\tau_k - \tau_{k-1} - \mathbb{E}_\emptyset(\tau))^2\right) \leq \frac{2}{\alpha^2}\left\lfloor\frac{T}{\mathbb{E}_\emptyset(\tau)}\right\rfloor\mathbb{E}_\emptyset(e^{\alpha\tau})e^{\alpha\mathbb{E}_\emptyset(\tau)},$$

and, for $n \geq 3$,

$$\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor}\mathbb{E}_\emptyset\left(|\tau_k - \tau_{k-1} - \mathbb{E}_\emptyset(\tau)|^n\right) \leq \frac{n!}{2}\left(\left\lfloor\frac{T}{\mathbb{E}_\emptyset(\tau)}\right\rfloor\frac{2}{\alpha^2}\mathbb{E}_\emptyset(e^{\alpha\tau})e^{\alpha\mathbb{E}_\emptyset(\tau)}\right)\frac{1}{\alpha^{n-2}}.$$

Applying Bernstein inequality again, we obtain that

$$\mathbb{P}_\emptyset\left(\left|\sum_{k=1}^{\lfloor T/\mathbb{E}_\emptyset(\tau)\rfloor}(\tau_k-\tau_{k-1}-\mathbb{E}_\emptyset(\tau))\right|\geq\frac{T\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau)}{2|b-a|}\right)\leq 2\exp\left(-\frac{\left(T\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau)\right)^2}{4\left(2v+(T\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau))c\right)}\right).$$

Equation (4.7) gives that

$$\mathbb{P}_\emptyset\left(\left|\frac{1}{T}\int_0^T f(X_t)\,dt-\pi_A f\right|\geq\varepsilon\right)\leq 4\exp\left(-\frac{\left(T\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau)\right)^2}{4\left(2v+(T\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau))c\right)}\right).$$

To prove the second part of Corollary 1.6 we have to solve

$$\eta=4\exp\left(-\frac{\left(T\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau)\right)^2}{4\left(2v+(T\varepsilon-|b-a|\mathbb{E}_\emptyset(\tau))c\right)}\right)\tag{4.8}$$

by expressing ε as function of η, for any $\eta\in(0,1)$.

Let us define the following decreasing bijection from $\mathbb{R}_+$ into $\mathbb{R}_-$:

$$\varphi(x)=-\frac{x^2}{4(2v+cx)}.$$

The solution of (4.8) is then $\varepsilon_\eta=(|b-a|\mathbb{E}_\emptyset(\tau)+x_0)/T$ where x_0 is the unique positive solution of

$$\varphi(x)=\log\left(\frac{\eta}{4}\right)\quad\Leftrightarrow\quad x^2+4c\log\left(\frac{\eta}{4}\right)x+8v\log\left(\frac{\eta}{4}\right)=0.$$

Computing the roots of this second order polynomial, we can show that there always exist one negative and one positive root as soon as $\eta<4$. More precisely,

$$x_0=-2c\log\left(\frac{\eta}{4}\right)+\sqrt{4c^2\log^2\left(\frac{\eta}{4}\right)-8v\log\left(\frac{\eta}{4}\right)},$$

which concludes the proof.

A Appendix

A.1 Return time for $M/G/\infty$ queues

We now state a general result for the tail behavior of the time of return to zero $\mathcal{T}_1$ of a $M/G/\infty$ queue with a service time admitting exponential moments. The queues considered in this section all start with zero customers.

The result is based on the computation of the Laplace transform $\mathbb{E}(\mathrm{e}^{-s\mathcal{T}_1})$ on the half-plane $\{s\in\mathbb{C}:\Re(s)>0\}$ by Takács [28, 29]. We extend analytically this Laplace transform to $\{s\in\mathbb{C}:\Re(s)>s_c\}$ for an appropriate $s_c<0$, which yields exponential moments.

This result has an interest in itself, independently of the Hawkes processes considered in the paper.

Theorem A.1. *Consider a $M/G/\infty$ queue with arrival rate $\lambda > 0$ and generic service duration H satisfying for some $\gamma > 0$ that, for $t \geq 0$,*

$$\mathbb{P}(H > t) \triangleq 1 - G(t) = O(\mathrm{e}^{-\gamma t})\,.$$

Let V_1 denote the arrival time of the first customer, $\mathcal{T}_1$ the subsequent time of return of the queue to zero, and $B = \mathcal{T}_1 - V_1$ the corresponding busy period.

a) If $\beta < \gamma$ then $\mathbb{E}(\mathrm{e}^{\beta B}) < \infty$. In particular $\mathbb{P}(B \geq t) = O(\mathrm{e}^{-\beta t})$.

b) If $\lambda < \gamma$, then $\mathbb{P}(\mathcal{T}_1 \geq t) = O(\mathrm{e}^{-\lambda t})$. If $\gamma \leq \lambda$, for $\alpha < \gamma$ then $\mathbb{P}(\mathcal{T}_1 \geq t) = O(\mathrm{e}^{-\alpha t})$.

Proof. We have $\mathcal{T}_1 = V_1 + B$, and the strong Markov property of the Poisson process yields that V_1 and B are independent. Since V_1 is exponential of parameter λ, we need mainly to study B. Takács has proved in [28, Eq. (37)], see also [29, Theorem 1 p. 210], that the Laplace transform of $\mathcal{T}_1$ satisfies

$$\mathbb{E}(\mathrm{e}^{-s\mathcal{T}_1}) = 1 - \frac{1}{\lambda + s}\frac{1}{\int_0^\infty \mathrm{e}^{-st-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\,\mathrm{d}t}\,, \qquad s \in \mathbb{C}\,,\ \Re(s) > 0\,. \tag{A.1}$$

Since the Laplace transform of V_1 is $\frac{\lambda}{\lambda+s}$, the Laplace transform of B satisfies

$$\mathbb{E}(\mathrm{e}^{-sB}) = \frac{\lambda + s}{\lambda} - \frac{1}{\lambda}\frac{1}{\int_0^\infty \mathrm{e}^{-st-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\,\mathrm{d}t}\,, \qquad s \in \mathbb{C}\,,\ \Re(s) > 0\,. \tag{A.2}$$

There is an apparent singularity in the r.h.s. of (A.1) and of (A.2), since the integral term increases to infinity as s decreases to 0. This is normal, since these formulæ remain valid for heavy tailed service. Moreover, (A.1) is proved in [28, 29] using the Laplace transform of a measure with infinite mass. We shall remove this apparent singularity and compute the abscissa of convergence of the Laplace transform in the l.h.s. of (A.2) .

The main point to prove is that the abscissa of convergence σ_c of the Laplace transform in the l.h.s. of (A.2) satisfies $\sigma_c \leq -\gamma$. In order to remove the apparent singularity in the r.h.s. of (A.2), we use integration by parts: on the half-line $\{s \in \mathbb{R} : s > 0\}$,

$$\begin{aligned}\int_0^\infty \mathrm{e}^{-st-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\,\mathrm{d}t &= \left[\frac{\mathrm{e}^{-st}}{-s}\mathrm{e}^{-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\right]_{t=0}^\infty \\ &\qquad - \int_0^\infty \frac{\mathrm{e}^{-st}}{-s}(-\lambda[1-G(t)])\,\mathrm{e}^{-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\,\mathrm{d}t \\ &= \frac{1}{s} - \frac{\lambda}{s}\int_0^\infty [1-G(t)]\,\mathrm{e}^{-st-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\,\mathrm{d}t\,.\end{aligned} \tag{A.3}$$

After inspection of the integral on the r.h.s., since $1 - G(t) = O(\mathrm{e}^{-\gamma t})$ and

$$\lambda\int_0^\infty [1-G(t)]\,\mathrm{e}^{-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\,\mathrm{d}t = \left[-\mathrm{e}^{-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\right]_{t=0}^\infty = 1 - \mathrm{e}^{-\lambda\mathbb{E}(H)} < 1,$$

we are able to define a constant $\theta < 0$ and an analytic function f by setting

$$\theta = \inf\left\{s \leq 0 : \lambda\int_0^\infty [1-G(t)]\,\mathrm{e}^{-st-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\,\mathrm{d}t < 1\right\} \vee (-\gamma)\,,$$

$$f(s) = \frac{\lambda + s}{\lambda} - \frac{s}{\lambda}\frac{1}{1 - \lambda\int_0^\infty [1-G(t)]\,\mathrm{e}^{-st-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\,\mathrm{d}t}\,, \qquad s \in \mathbb{C}\,,\ \Re(s) > \theta\,. \tag{A.4}$$

The Laplace transform in the l.h.s. of (A.2) has an abcissa of convergence $\sigma_c \le 0$ and is analytic in the half-plane $\{s \in \mathbb{C} : \Re(s) > \sigma_c\}$, see Widder [31, Theorem 5a p. 57]. Both this Laplace transform and f are analytic in the domain $\{s \in \mathbb{C} : \Re(s) > \max(\theta, \sigma_c)\}$, and since these two analytic functions coincide there on the half-line $\{s \in \mathbb{R} : s > 0\}$ they must coincide in the whole domain, see Rudin [27, Theorem 10.18 p. 208], so that

$$\mathbb{E}(\mathrm{e}^{-sB}) = f(s)\,, \qquad s \in \mathbb{C}\,,\ \Re(s) > \max(\theta, \sigma_c)\,.$$

This Laplace transform must have an analytic singularity at $s = \sigma_c$, see Widder [31, Theorem 5b p. 58], and since f is analytic in $\{s \in \mathbb{C} : \Re(s) > \theta\}$ necessarily $\sigma_c \le \theta$.

Since $\theta < 0$, by monotone convergence

$$\lim_{s\to\theta^+} f(s) = \frac{\lambda+\theta}{\lambda} - \frac{\theta}{\lambda}\frac{1}{1-\lambda\int_0^\infty [1-G(t)]\,\mathrm{e}^{-\theta t-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\,\mathrm{d}t} = \mathbb{E}(\mathrm{e}^{-\theta B}) \in [1,\infty]\,,$$

which implies that $\lambda\int_0^\infty [1-G(t)]\,\mathrm{e}^{-\theta t-\lambda\int_0^t [1-G(u)]\,\mathrm{d}u}\,\mathrm{d}t < 1$, and thus that $\theta = -\gamma$.

We conclude that $\sigma_c \le -\gamma$. Thus, if $\beta < \gamma$ then $\mathbb{E}(\mathrm{e}^{\beta B}) < \infty$, and $\mathbb{P}(B \ge t) = O(\mathrm{e}^{-\beta t})$ using the Markov inequality. Moreover, if $\mathbb{P}(B \ge t) = O(\mathrm{e}^{-\alpha t})$ then

$$\begin{aligned}\mathbb{P}(\mathcal{T}_1 \ge t) = \mathbb{P}(B + V_1 \ge t) &= \mathrm{e}^{-\lambda t} + \lambda\int_0^t \mathrm{e}^{-\lambda u}\mathbb{P}(B \ge t-u)\,\mathrm{d}u \\ &\le \mathrm{e}^{-\lambda t} + C\int_0^t \mathrm{e}^{-\lambda u-\alpha(t-u)}\,\mathrm{d}u\,,\end{aligned}$$

hence if $\lambda < \gamma$ then choosing $\lambda < \alpha < \gamma$ yields that

$$\mathbb{P}(\mathcal{T}_1 \ge t) \le \mathrm{e}^{-\lambda t} + C\mathrm{e}^{-\lambda t}\int_0^t \mathrm{e}^{-(\alpha-\lambda)(t-u)}\,\mathrm{d}u \le [1 + C/(\alpha-\lambda)]\mathrm{e}^{-\lambda t},$$

and if $\alpha < \gamma \le \lambda$ then

$$\mathbb{P}(\mathcal{T}_1 \ge t) \le \mathrm{e}^{-\lambda t} + C\mathrm{e}^{-\alpha t}\int_0^t \mathrm{e}^{-(\lambda-\alpha)u}\,\mathrm{d}u \le \Big[1 + \frac{C}{\lambda-\alpha}\Big]\mathrm{e}^{-\alpha t}\,. \qquad \square$$

We now provide a corollary to the previous result.

Proposition A.2. *Consider a $M/G/\infty$ queue with arrival rate $\lambda > 0$ and generic service duration H satisfying for some $\gamma > 0$ that*

$$\mathbb{P}(H > t) = O(\mathrm{e}^{-\gamma t})\,.$$

Let Y_t denote the number of customers at time $t \ge 0$, and for each $E \ge 0$ let

$$\tau_E = \inf\{t \ge E : Y_t = 0\} \tag{A.5}$$

be the first hitting time of zero after E. If $\lambda < \gamma$ then let $\alpha = \lambda$, and if $\gamma \le \lambda$ then let $0 < \alpha < \gamma$. Then there exists a constant $C < \infty$ such that

$$\mathbb{P}(\tau_E \ge t) \le \lambda C E\ e^{-\alpha(t-E)}\,, \quad \forall t \ge E\,.$$

Proof. The successive return times to zero $(\mathcal{T}_k)_{k\geq 0}$ of the process $(Y_t)_{t\geq 0}$ have been defined in (2.13). The events $\{\mathcal{T}_{k-1} \leq E,\ \mathcal{T}_k > E\}$ for $k \geq 1$ define a partition of Ω and, for $t > E$,

$$\begin{aligned}
\mathbb{P}(\tau_E \geq t) &= \sum_{k=1}^{+\infty} \mathbb{P}(\tau_E \geq t,\, \mathcal{T}_{k-1} \leq E,\, \mathcal{T}_k > E) \\
&= \sum_{k=1}^{+\infty} \mathbb{P}(\mathcal{T}_{k-1} \leq E,\, \mathcal{T}_k \geq t) \\
&= \sum_{k=1}^{+\infty} \mathbb{E}\Big(\mathbb{1}_{\{\mathcal{T}_{k-1}\leq E\}} \mathbb{P}(\mathcal{T}_k \geq t \,|\, \mathcal{F}_{\mathcal{T}_{k-1}})\Big) \\
&\leq \sum_{k=1}^{+\infty} \mathbb{E}\Big(\mathbb{1}_{\{\mathcal{T}_{k-1}\leq E\}} \mathbb{P}(\mathcal{T}_k - \mathcal{T}_{k-1} \geq t - E \,|\, \mathcal{F}_{\mathcal{T}_{k-1}})\Big)
\end{aligned}$$

so that, since $\mathcal{T}_k - \mathcal{T}_{k-1}$ is independent of $\mathcal{F}_{\mathcal{T}_{k-1}}$ and distributed as $\mathcal{T}_1$,

$$\mathbb{P}(\tau_E \geq t) \leq \sum_{k=1}^{+\infty} \mathbb{E}\Big(\mathbb{1}_{\{\mathcal{T}_{k-1}\leq E\}}\Big) \mathbb{P}(\mathcal{T}_1 \geq t - E) = \mathbb{P}(\mathcal{T}_1 \geq t - E)\mathbb{E}\left(\sum_{k=1}^{+\infty} \mathbb{1}_{\{\mathcal{T}_{k-1}\leq E\}}\right).$$

By Theorem A.1, under the assumptions there exists a constant C such that

$$\mathbb{P}(\mathcal{T}_1 \geq t - E) \leq C\mathrm{e}^{\alpha(t-E)}\,.$$

Moreover $\sum_{k=1}^{+\infty} \mathbb{1}_{\{\mathcal{T}_{k-1}\leq E\}}$ is the number of returns to zero before time E. It is bounded by the number of arrivals between times 0 and E, which follows a Poisson law of parameter and expectation λE. This leads to the announced inequality. □

A.2 Strong Markov property for homogeneous Poisson point process

In this appendix, we prove a strong Markov property for homogeneous Poisson point processes on the line. This classic result is stated in [26, Proposition 1.18 p.18] when the filtration is the canonical filtration generated by the Poisson point process. Here, the filtration $(\mathcal{F}_t)_{t\geq 0}$ may contain additional information, for example coming from configurations on $\mathbb{R}_-$.

Lemma A.3. *Let Q be a $(\mathcal{F}_t)_{t\geq 0}$-Poisson point process on $(0, +\infty) \times (0, +\infty)$ with unit intensity. Then Q is a strong $(\mathcal{F}_t)_{t\geq 0}$-Markov process in the following sense: for any stopping time T for $(\mathcal{F}_t)_{t\geq 0}$, conditionally on $T < \infty$ the shifted process $S_T Q$ defined by* (3.3) *is a $(\mathcal{F}_{T+t})_{t\geq 0}$-Poisson point process with unit intensity.*

Proof. It is enough to prove that, for any stopping time T and $h, a > 0$, conditionally on $T < \infty$ the random variable $Q((T, T + h] \times (0, a])$ is $\mathcal{F}_{T+h}$-measurable, independent of $\mathcal{F}_T$, and Poisson of parameter ha. Indeed, in order to prove the strong Markov property at a given stopping time T, it is enough to apply the above to the stopping times $T + t$ for $t > 0$ in order to see that $S_T Q$ satisfies that for every $t, h, a > 0$, the random variable $Q((t, t + h] \times (0, a])$ is $\mathcal{F}_{t+h}$-measurable, independent of $\mathcal{F}_t$, and Poisson of parameter ha.

We first prove this for an arbitrary stopping time T with finite values belonging to an increasing deterministic sequence $(t_n)_{n\ge 1}$. For each B in $\mathcal{F}_T$ and $k \ge 0$,

$$\begin{aligned}
&\mathbb{P}(B \cap \{T < \infty\} \cap \{Q((T, T+h] \times (0,a]) = k\}) \\
&\quad = \sum_{n\ge 1} \mathbb{P}(B \cap \{T = t_n\} \cap \{Q((t_n, t_n+h] \times (0,a]) = k\})
\end{aligned}$$

in which, by definition of $\mathcal{F}_T$ and since $\mathcal{F}_{t_{n-1}} \subset \mathcal{F}_{t_n}$,

$$B \cap \{T = t_n\} = (B \cap \{T \le t_n\}) - (B \cap \{T \le t_{n-1}\}) \in \mathcal{F}_{t_n}\,.$$

The $(\mathcal{F}_t)_{t\ge 0}$-Poisson point process property then yields that

$$\mathbb{P}(B \cap \{T = t_n\} \cap \{Q((t_n, t_n+h] \times (0,a]) = k\}) = \mathbb{P}(B \cap \{T = t_n\})\, \mathrm{e}^{-ha} \frac{(ha)^k}{k!}$$

and summation of the series that

$$\mathbb{P}(B \cap \{T < \infty\} \cap \{Q((T, T+h] \times (0,a]) = k\}) = \mathbb{P}(B \cap \{T < \infty\})\, \mathrm{e}^{-ha} \frac{(ha)^k}{k!}\,.$$

Hence $Q((T, T+h] \times (0,a])$ is independent of $\mathcal{F}_T$ and Poisson of parameter ha. Moreover, for $k \ge 0$, similarly

$$\begin{aligned}
&\{T < \infty,\, Q((T, T+h] \times (0,a]) = k\} \cap \{T + h \le t\} \\
&\quad = \bigcup_{n\ge 1} \{T = t_n,\, Q((t_n, t_n+h] \times (0,a]) = k\} \cap \{t_n + h \le t\} \subset \mathcal{F}_t
\end{aligned}$$

and hence $Q((T, T+h] \times (0,a])$ is $\mathcal{F}_{T+h}$-measurable.

In order to extend this to a general stopping time T, we approximate it by above by the discrete stopping times

$$T_n = \sum_{k=1}^{+\infty} \frac{k}{2^n} \mathbb{1}_{\{\frac{k-1}{2^n} < T \le \frac{k}{2^n}\}}\,, \qquad n \ge 1\,.$$

Letting n go to infinity, the right continuity of $t \mapsto Q((0,t] \times (0,a])$ and of $(\mathcal{F}_t)_{t\ge 0}$ allows to conclude. $\square$

References

[1] E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy. Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Process. Appl.*, 123(7):2475–2499, 2013.

[2] E. Bacry, S. Delattre, M. Hoffmann, and J.F. Muzy. Modeling microstructure noise with mutually exciting point processes. *Quantitative finance*, 13(1):65–77, 2013.

[3] E. Bacry and J-F. Muzy. First- and second-order statistics characterization of Hawkes processes and non-parametric estimation. *IEEE Trans. Inform. Theory*, 62(4):2184–2202, 2016.

[4] H.C.P. Berbee. *Random walks with stationary increments and renewal theory*, volume 112 of *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam, 1979.

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

[6] P. Brémaud, G. Nappo, and G.L. Torrisi. Rate of convergence to equilibrium of marked Hawkes processes. *Journal of Applied Probability*, 39:123–136, 2002.

[7] J. Chevallier. Mean-field limit of generalized hawkes processes. *Stochastic Processes and their Applications*, 2017. To appear.

[8] J. Chevallier, M. J. Caceres, M. Doumic, and P. Reynaud-Bouret. Microscopic approach of a time elapsed neural model. *Mathematical Models and Methods in Applied Sciences*, 25:2669, 2015.

[9] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2003. Elementary theory and methods.

[10] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. II.* Probability and its Applications (New York). Springer, New York, second edition, 2008. General theory and structure.

[11] S. Delattre and N. Fournier. Statistical inference versus mean field limit for Hawkes processes. *Electronic Journal of Statistics*, 10:1223–1295, 2016.

[12] S. Delattre, N. Fournier, and M. Hoffmann. Hawkes processes on large networks. *Annals of Applied Probability*, 26(1):216–261, 2016.

[13] S. Ditlevsen and E. Löcherbach. On oscillating systems of interacting neurons. *Stochastic Processes and their Applications*, 127:1840–1869, 2017.

[14] A. Duarte, E. Löcherbach, and G. Ost. Stability, convergence to equilibrium and simulation of non-linear hawkes processes with memory kernels given by the sum of erlang kernels. 2016. arxiv.org.abs/1610.03300.

[15] W. Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[16] A. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

[17] A.G. Hawkes and L. Adamopoulos. Cluster models for earthquakes: Regional comparisons. *Bull. Int. Statist. Inst.*, 45:454–461, 1973.

[18] A.G. Hawkes and D. Oakes. A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11:493–503, 1974.

[19] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.

[20] P. Massart. *Concentration inequalities and model selection*, volume 1896 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007. Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6–23, 2003, With a foreword by Jean Picard.

[21] L. Massoulié. Stability results for a general class of interacting point processes dynamics, and applications. *Stochastic Processes and their Applications*, 75:1–30, 1998.

[22] Y. Ogata. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of American Statistical Association*, 83(401):9–27, 1988.

[23] P. Reynaud-Bouret, V. Rivoirard, and C. Tuleau-Malot. Inference of functional connectivity in Neurosciences via Hawkes processes. In *1st IEEE Global Conference on Signal and Information Processing*, Austin, Texas, 2013.

[24] P. Reynaud-Bouret and E. Roy. Some non asymptotic tail estimates for hawkes processes. *Bulletin of the Belgian Mathematical Society-Simon Stevin*, 13(5):883–896, 2007.

[25] P. Reynaud-Bouret and S. Schbath. Adaptive estimation for Hawkes processes; application to genome analysis. *Annals of Statistics*, 38(5):2781–2822, 2010.

[26] Ph. Robert. *Stochastic Networks and Queues.* Springer, Berlin-Heidelberg, 2003.

[27] W. Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, third edition, 1987.

[28] L. Takács. On a probability problem arising in the theory of counters. *Proc. Cambridge Philos. Soc.*, 52:488–498, 1956.

[29] L. Takács. *Introduction to the theory of queues.* University Texts in the Mathematical Sciences. Oxford University Press, New York, 1962.

[30] H. Thorisson. *Coupling, stationarity, and regeneration.* Probability and its Applications (New York). Springer-Verlag, New York, 2000.

[31] D. V. Widder. *The Laplace Transform.* Princeton Mathematical Series, v. 6. Princeton University Press, Princeton, N. J., 1941.