CENTRE FOR **STOCHASTIC GEOMETRY**
AND ADVANCED **BIOIMAGING**

Raúl Fierro, Víctor Leiva and Jesper Møller

The Hawkes Process with Different Excitation Functions and its Asymptotoc Behavior

No. 10, December 2013

(12 December 2013)

THE HAWKES PROCESS WITH DIFFERENT EXCITATION FUNCTIONS AND ITS ASYMPTOTIC BEHAVIOR

RAÚL FIERRO,* *Pontificia Universidad Católica de Valparaíso.*

VÍCTOR LEIVA,** *Universidad de Valparaíso.*

JESPER MØLLER,*** *Aalborg University.*

Abstract

The standard Hawkes process is constructed from a homogeneous Poisson process and using the same exciting function for different generations of offspring. We propose an extension of this process by considering different exciting functions. This consideration could be important to be taken into account in a number of fields; e.g. in seismology, where main shocks produce aftershocks with possibly different intensities. The main results are devoted to the asymptotic behavior of this extension of the Hawkes process. Indeed, a law of large numbers and a central limit theorem are stated. These results allow us to analyze the asymptotic behavior of the process when unpredictable marks are considered.

Keywords: central limit theorem; law of large numbers; clustering effect; unpredictable marks.

2010 Mathematics Subject Classification: Primary 60G55
Secondary 60F05

1. Introduction

The standard Hawkes process (HP) is a temporal point process having long memory, clustering effect and the self-exciting property. The standard HP and its extension to a marked point process are of wide interest, partly because of their many important

* Postal address: Brasil 2950. Casilla 4059, Valparaíso, Chile. Email: rfierro@ucv.cl
** Postal address: Gran Bretaña 1111. Casilla 5030, Valparaíso, Chile. Email: victor.leiva@uv.cl
*** Postal address: Fredrik Bajers Vej 7G. DK-9220 Aalborg, Denmark Ø. Email: jm@math.aau.dk

applications and illustrative examples in the theory of non-Markovian point processes constructed by a conditional intensity. The seminal ideas are due to Hawkes [9, 10] and Hawkes and Oakes [11], whereas useful reviews on the topic are provided in Daley and Vere-Jones [4] and Zhu [21]. Its applications include fields such as finance, genetics, neuroscience and seismology; see e.g. Carstensen et al. [3], Embrechts et al. [5], Gusto and Schbath [8], Ogata [16, 17] and Pernice et al. [18].

As mentioned, the standard HP is a cluster process, where the starting points of the clusters are called immigrants and appear according to a homogeneous Poisson process on the non-negative time-axis. Each immigrant is the ancestor of a first generation of offspring, each point of first generation offspring is the ancestor of a second generation point offspring, and so on. Thereby the cluster for an immigrant is a set of generations of offspring. More precisely, for a given ancestor appearing at time s, the associated offspring point process is Poisson with intensity function $\gamma(t-s)$, which is defined for $t > s$ and is not depending on immigrant and offspring points generated before time s. Thus the clusters, conditional to the immigrants, are independent. Note that the same exciting function γ is used for all offspring processes. This is the crucial difference with the extension proposed in our work, where we allow different exciting functions for the different generations of offspring. This extension could be relevant for instance in seismology, where main shocks generate aftershocks with possible different intensities.

The main objective of this work is to investigate the asymptotic behavior of our extension of the HP process. Indeed, a law of large numbers and a central limit theorem are established. Furthermore, by making use of these results, a central limit theorem is proved when unpredictable marks are added to the process. In particular our asymptotic results do not require the complete identification of offspring processes, but only of the integrals of their exciting functions. We also extend a result obtained by Fierro *et al.* in [6]. Recently, functional central limit theorems for linear and non-linear HP have been obtained in [1] and [20], respectively. However, their results are based on the standard HP, while ours, coming from a more general definition of HP, cannot be obtained from these works. Simulation algorithms and statistical methodology for the extension proposed in this paper remain as open problems to be developed in future studies. For details on exact and approximate simulation algorithms for the standard HP with unpredictable marks, see [13, 14].

The paper is organized as follows. The results of this work are introduced in the second section, which is divided into four subsections. In Subsection 2.1, we define the HP with different excitation functions and establish some preliminary facts. In Subsection 2.2, we present two of the main results namely, a law of large numbers and a central limit theorem for the process. In Subsection 2.3, we consider two special cases, one of them is the standard HP and the other concerns the case consisting of a finite number of generations. In Subsection 2.4, we state a central limit theorem for the process with unpredictable marks. The proofs of our results are provided in the third section.

2. The Hawkes process with different excitation functions

2.1. Definition and preliminary results

In the sequel, $\{\gamma_n\}_{n\in\mathbb{N}}$ denotes a sequence of locally integrable functions from $\mathbb{R}_+$ to $\mathbb{R}_+$. Here $\mathbb{R}_+ = [0,\infty)$ is the non-negative time-axis, and $\mathbb{N} = \{0,1,\ldots\}$ the set of non-negative integers.

The following proposition is the basis of what we name the HP with different excitation functions. For concepts related to counting processes and their stochastic intensities, we refer to [2].

Proposition 2.1. *There exist a probability space* $(\Omega, \mathcal{F}, \mathrm{P})$ *and a sequence* $\{N^n\}_{n\in\mathbb{N}}$ *of non-explosive counting processes without common jumps satisfying the following three conditions:*

(A1) N^0 *is a Poisson process with intensity* γ_0*.*

(A2) *For each* $n \geq 1$*,* N^n *has predictable stochastic intensity* λ^n *given by* $\lambda^n_t = \int_0^t \gamma_n(t-s)\,\mathrm{d}N^{n-1}_s$*.*

(A3) *For each* $n \in \mathbb{N}$*, conditional to* $N^0,\ldots,N^n$*,* N^{n+1} *is a non-homogeneous Poisson process with intensity* λ^{n+1}*.*

Definition 2.1. Let $\{N^n\}_{n\in\mathbb{N}}$ be as in Proposition 2.1 and $N = \sum_{n=0}^{\infty} N^n$. We call N^0 the immigrant process, N^n $(n \geq 1)$ the nth generation offspring process and N the HP with excitation functions $\{\gamma_n\}_{n\in\mathbb{N}}$.

Remark 2.1. In the standard HP, $\gamma_0 = \mu$ is constant and all $\gamma_n = \gamma$ for all $n \geq 1$. In this case there is no need of identifying the offspring processes, since N has stochastic intensity λ given by $\lambda_t = \mu + \int_0^t \gamma(t-s)\,\mathrm{d}N_s$.

Remark 2.2. In Proposition 2.1, (A3) allows us to obtain, recursively, the joint distribution of $N^0, \ldots, N^n$, for $n \in \mathbb{N}$. It is easy to see that (A2) and (A3) are equivalent.

Remark 2.3. Notice that N is univocally defined in distribution. Indeed, according to Theorem 3.6 in [12], there exists, on the Skorohod space, a unique counting process having predictable stochastic intensity $\lambda = \gamma_0 + \sum_{n=1}^{\infty} \lambda^n$.

Let Λ^n be the compensator of N^n, that is, for each $n \in \mathbb{N}$ and $t \geq 0$, $\Lambda_t^n = \int_0^t \lambda_s^n\,\mathrm{d}s$, where $\lambda_s^0 = \gamma_0(s)$ is a deterministic function. Thus, for each $n \in \mathbb{N}$, $M^n = N^n - \Lambda^n$ is a $(\mathbb{F}, \mathrm{P})$-martingale, where $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ with $\mathcal{F}_t = \sigma(N_s^0; s \leq t)$ the σ-algebra generated by $\{N_s^0; 0 \leq s \leq t\}$.

Proposition 2.2. *For each* $n \in \mathbb{N} \setminus \{0\}$ *and* $t \geq 0$, $\Lambda_t^n = \int_0^t \gamma_n(t-s) N_s^{n-1}\,\mathrm{d}s$.

For two locally integrable functions f and g from $\mathbb{R}_+$ to $\mathbb{R}$, $f * g$ denotes the convolution between f and g, i.e., $(f * g)(t) = \int_0^t f(t-s)g(s)\,\mathrm{d}s$, for $t \geq 0$.

Proposition 2.3. *For each* $t \geq 0$,

$$\mathrm{E}(N_t) = \int_0^t \sum_{n=0}^{\infty} (\gamma_0 * \cdots * \gamma_n)(u)\,\mathrm{d}u.$$

Proposition 2.3 motivates to consider the following condition:

(B) For each $t \geq 0$, the sequence $\{\gamma_n\}_{n\in\mathbb{N}}$ satisfies

$$\int_0^t \sum_{n=0}^{\infty} (\gamma_0 * \cdots * \gamma_n)(u)\,\mathrm{d}u < \infty.$$

Let $M = \sum_{n=0}^{\infty} M^n$. Then the HP N is a counting process with compensator $\Lambda = \sum_{n=0}^{\infty} \Lambda^n$ and, under condition (B), $M = N - \Lambda$ is a $(\mathbb{F}, \mathrm{P})$-martingale.

For any measurable function $h : [0, \infty) \to [0, \infty]$, we denote its Laplace transform by $\mathcal{L}[h]$, i.e., for $s \in \mathbb{R}$, $\mathcal{L}[h](s) = \int_0^\infty \mathrm{e}^{-su}\, h(u)\,\mathrm{d}u$.

Remark 2.4. Under condition (B), N is a non-explosive counting process with predictable compensator Λ.

Proposition 2.4. *Condition (B) is satisfied when one of the following five conditions holds:*

(C1) *There exists* $s_0 > 0$ *such that* $\sup_{n\in\mathbb{N}} \mathcal{L}[\gamma_n](s_0) < 1$.

(C2) $\lim_{s\to\infty} \sup_{k\in\mathbb{N}} \mathcal{L}[\gamma_k](s) = 0$.

(C3) *There exist* $C > 0$ *and* $a > 0$ *such that* $\sup_{k\in\mathbb{N}} \gamma_k(t) \le C\,\mathrm{e}^{at}$.

(C4) $\int_0^\infty \sup_{k\in\mathbb{N}} \gamma_k(s)\,\mathrm{d}s < \infty$.

(C5) $\sup_{k\in\mathbb{N}} \int_0^\infty \gamma_k(s)\,\mathrm{d}s < 1$.

2.2. Asymptotic results

Let $\rho = \sup_{k\in\mathbb{N}} \int_0^\infty \gamma_k(s)\,\mathrm{d}s$. In this subsection, we assume the following condition holds:

(D) There exists $\overline{\gamma_0} = \lim_{t\to\infty} \frac{1}{t}\int_0^t \gamma_0(s)\,\mathrm{d}s$ and $\rho < 1$.

In particular, from Proposition 2.4, condition (B) holds when condition (D) is satisfied.

In the sequel, $m_0 = \overline{\gamma_0}$, for each $p \in \mathbb{N}\setminus\{0\}$, $m_p = \overline{\gamma_0}\prod_{i=1}^p \int_0^\infty \gamma_i(u)\,\mathrm{d}u$ and $m = \sum_{p=0}^\infty m_p$. Notice that, under condition (D), $m < \infty$.

For the standard HP, the condition $\rho < 1$ is usually assumed in order to obtain a non-explosive process (see e.g. [4]).

We have the following law of large numbers.

Theorem 2.1. *As* $t \to \infty$, $\{N_t/t\}_{t>0}$ *and* $\{\Lambda_t/t\}_{t>0}$ *converge* P*-a.s. to* m*, and* $\{M_t/t\}_{t>0}$ *converges in quadratic mean to zero.*

The following central limit theorem is the main result of this work.

Theorem 2.2. *For each* $t > 0$*, let* $X_t = (N_t - m)/\sqrt{t}$ *and*

$$\sigma_N^2 = \sum_{j=0}^{\infty} \left(1 + \sum_{p=1}^{\infty} \prod_{i=j+1}^{p+j} \int_0^{\infty} \gamma_i(u)\,\mathrm{d}u \right)^2 m_j.$$

Then, $\sigma_N^2 < \infty$ *and, as* $t \to \infty$*,* $\{X_t\}_{t>0}$ *converges in distribution to a normal random variable with mean zero and variance* σ_N^2*.*

The proofs of Theorems 2.1 and 2.2, provided in Section 3, involve the following three lemmas.

Lemma 2.1. *Let* h *be a non-negative measurable function defined on* $\mathbb{R}_+$*. Then, for each* $s, t \geq 0$ *with* $s \leq t$*,*

$$\int_s^t (h * \gamma_0)(v)\,\mathrm{d}v \leq \left(\int_0^{\infty} h(r)\,\mathrm{d}r \right) \left(\int_s^t \gamma_0(u)\,\mathrm{d}u \right).$$

Lemma 2.2. *For each* $q \in (0, 2]$ *exists* $C > 0$ *such that*

$$\sum_{j=0}^{\infty} \sup_{t>0} \mathrm{E} \left(\sup_{0 \leq u \leq t} |M_u^j/\sqrt{t}|^q \right) \leq C.$$

Lemma 2.3. *For each integer* $p \geq 1$*,*

$$\Lambda^p = \sum_{j=0}^{p-1} \gamma_p * \cdots * \gamma_{j+1} * M^j + \gamma_p * \cdots * \gamma_1 * \gamma_0 * 1 \tag{1}$$

and

$$\Lambda = \sum_{p=1}^{\infty} \sum_{j=1}^{p-1} \gamma_p * \cdots * \gamma_{j+1} * M^j + \sum_{p=0}^{\infty} \gamma_p * \cdots * \gamma_1 * \gamma_0 * 1. \tag{2}$$

Moreover,

$$\lim_{t \to \infty} \frac{1}{t} \sum_{p=1}^{\infty} \sum_{j=0}^{p-1} \mathrm{E}[(\gamma_p * \cdots * \gamma_{j+1} * |M^j|)_t] = 0 \tag{3}$$

and

$$\lim_{t \to \infty} \sup_{p \in \mathbb{N}} \left| \frac{\Lambda_t^p}{t} - m_p \right| = 0 \quad \mathrm{P} - a.s. \tag{4}$$

2.3. Two particular cases

Below we consider two special cases where condition (D) is satisfied and consequently the process $\{X_t\}_{t>0}$, defined in Theorem 2.2, has asymptotic normality. Thereon two corollaries of Theorem 2.2 are derived.

In the first case, the functions γ_n ($n \in \mathbb{N} \setminus \{0\}$) are assumed to be equal and hence it covers the case of the standard HP.

Corollary 2.1. *Suppose the excitation functions $\gamma_n = \gamma$ do not depend on n, for $n \geq 1$, and the following two conditions hold:*

(E1) *The limit $\overline{\gamma_0} = \lim_{t\to\infty} \frac{1}{t}\int_0^t \gamma_0(s)\,\mathrm{d}s$ exists.*

(E2) $\int_0^\infty \gamma(u)\,\mathrm{d}u < 1.$

Then, as $t \to \infty$, $\{X_t\}_{t>0}$ converges in distribution to a normal random variable with mean zero and variance

$$\sigma_N^2 = \frac{\overline{\gamma_0}}{\left(1 - \int_0^\infty \gamma(u)\,\mathrm{d}u\right)^3}.$$

The second particular case is when there exists $n^* \in \mathbb{N}$ such that $\gamma_{n^*+1} = 0$, a.e., with respect to the Lebesgue measure. Then, there is at most n^* generations of offspring processes. The particular case $n^* = 1$ corresponds to a Neyman-Scott cluster point process where the 'mother point process' (i.e., the immigrant process) is included (see e.g. [15]).

Corollary 2.2. *Suppose condition (E1) and that there exists $n^* \in \mathbb{N}$ such that $\gamma_{n^*+1} = 0$, a.e., with respect to the Lebesgue measure. Then, as $t \to \infty$, $\{X_t\}_{t>0}$ converges in distribution to a normal random variable with mean zero and variance*

$$\sigma_N^2 = \sum_{j=0}^{n^*} \left(1 + \sum_{p=1}^{n^*-j} \prod_{i=j+1}^{p+j} \int_0^\infty \gamma_i(u)\,\mathrm{d}u\right)^2 m_j.$$

2.4. Unpredictable marks

Consider the extension of the standard HP with unpredictable marks defined in [4] and [13] to the case of our HP with different excitation functions, i.e., for each $k \in \mathbb{N}$, we associate a random mark ξ_k to the kth jump time T_k, where these marks are independent, identically distributed and independent of N. Moreover, assume the marks are real-valued random variables with mean ν and variance σ^2. Under these assumptions, we study the asymptotic distribution of the process $\{R_t\}_{t>0}$ defined by

$$R_t = \frac{1}{\sqrt{t}}\left(\sum_{k=0}^{N_t} \xi_k - \nu \mathrm{E}(N_t)\right).$$

Using the notation of Theorem 2.2, we have the following central limit theorem, which extends a result obtained by Fierro *et al.* in [6].

Theorem 2.3. *If condition (D) is satisfied, then $\{R_t\}_{t>0}$ converges in distribution to a normal random variable with mean zero and variance $m\sigma^2 + \nu\sigma_N^2$.*

The proof of Theorem 2.3 uses the following result.

Lemma 2.4. *Let $\{U_t\}_{t>0}$ and $\{V_t\}_{t>0}$ be two real stochastic processes defined on $(\Omega, \mathcal{F}, \mathrm{P})$ and (U, V) be a bivariate random vector defined on the same probability space. Moreover, suppose the following two conditions hold:*

(F1) *For any $\epsilon > 0$, there exists $C_\epsilon > 0$ such that $\sup_{t>0} \mathrm{P}(\max\{|U_t|, |V_t|\} > C_\epsilon) < \epsilon$.*

(F2) *For any bounded functions u and v from $\mathbb{R}$ to $\mathbb{R}$, $\lim_{t\to\infty} \mathrm{E}(u(U_t)v(V_t)) = \mathrm{E}(u(U)v(V))$.*

Then, as $t \to \infty$, $\{(U_t, V_t)\}_{t>0}$ converges in distribution to (U, V).

3. Proofs

Below I_A stands for the indicator function of a set A.

Proof of Proposition 2.1 Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a complete probability space where a Poisson process N^0, with intensity γ_0, is defined. Let $\{\Lambda_t^1\}_{t\geq 0}$ be the increasing and $(\mathbb{F}, \mathrm{P})$-adapted process defined as

$$\Lambda_t^1 = \int_0^t \left(\int_0^u \gamma_1(u-s)\,\mathrm{d}N_s^0\right) \mathrm{d}u.$$

Since Λ^1 is predictable and continuous, it follows from Theorem 3.6 in [12] that there exists a counting process N^1 adapted to the filtration $\mathbb{F}$ with compensator Λ^1. Consequently, for any predictable process $\{C_s\}_{s\geq 0}$, we have

$$\mathrm{E}\left(\int_0^\infty C_s\,\mathrm{d}N_s^1\right) = \mathrm{E}\left(\int_0^\infty C_s\,\mathrm{d}\Lambda_s^1\right) = \mathrm{E}\left(\int_0^\infty C_s \lambda_s^1\,\mathrm{d}s\right),$$

where $\lambda^1_u = \int_0^u \gamma_1(u-s)\,\mathrm{d}N^0_s$. This proves λ^1 is a stochastic intensity for N^1. Because N^0 is non-explosive, for each $t \geq 0$, $\Lambda^1_t < \infty$, P-a.s., which implies N^1 is non-explosive.

Next, suppose $N^1, \ldots, N^n$ are non-explosive counting processes having stochastic intensities $\lambda^1, \ldots, \lambda^n$, respectively, given by

$$\lambda^m_t = \int_0^t \gamma_m(t-s)\,\mathrm{d}N^{m-1}, \qquad 1 \leq m \leq n,$$

and let $\{\Lambda^{n+1}_t\}_{t\geq 0}$ be the $(\mathbb{F}, \mathrm{P})$-adapted and increasing process defined as

$$\Lambda^{n+1}_t = \int_0^t \left(\int_0^u \gamma_1(u-s)\,\mathrm{d}N^n_s \right) \mathrm{d}u.$$

We have Λ^{n+1} is predictable and continuous, and as before, Theorem 3.6 in [12] implies there exists an $(\mathbb{F}, \mathrm{P})$-adapted counting process N^{n+1} with compensator Λ^{n+1}. Accordingly, for any predictable process $\{C_s\}_{s\geq 0}$, we have

$$\mathrm{E}\left(\int_0^\infty C_s\,\mathrm{d}N^{n+1}_s \right) = \mathrm{E}\left(\int_0^\infty C_s\,\mathrm{d}\Lambda^{n+1}_s \right) = \mathrm{E}\left(\int_0^\infty C_s \lambda^{n+1}_s\,\mathrm{d}s \right),$$

where $\lambda^{n+1}_u = \int_0^u \gamma_{n+1}(u-s)\,\mathrm{d}N^n_s$. This proves λ^{n+1} is a stochastic intensity for N^{n+1}. Since N^n is non-explosive, for each $t \geq 0$, $\Lambda^{n+1}_t < \infty$, P-a.s., which implies N^{n+1} is non-explosive. Hence by induction, $\{N^n\}_{n\in\mathbb{N}}$ is a sequence of non-explosive counting processes satisfying (A1) and (A2).

Let $n, p \in \mathbb{N}$ with $p > 0$. Since λ^{n+p} depends on $\omega \in \Omega$ only through $N^{n+p-1}(\omega)$, conditional to $N^0, \ldots, N^{n+p-1}$, N^{n+p} is distributed as a Poisson process with intensity λ^{n+p}. In particular, (A3) holds. Let us prove that N^n and N^{n+p} have no common jumps. Suppose T is a stopping time such that $\Delta N^n_T = 1$, P-a.s. Hence T is measurable with respect to the σ-algebra generated by N^n and thus

$$\begin{aligned} \mathrm{E}(\Delta N^{n+p}_T | N^{n+p-1}) &= \mathrm{E}(\textstyle\int_0^\infty \mathrm{I}_{\{T\}}(u)\,\mathrm{d}N^{n+p}_u | N^{n+p-1}) \\ &= \mathrm{E}(\textstyle\int_0^\infty \mathrm{I}_{\{T\}}(u)\lambda^{n+p}_u\,\mathrm{d}u | N^{n+p-1}) \\ &= \textstyle\int_0^\infty \mathrm{I}_{\{T\}}(u)\mathrm{E}(\lambda^{n+p}_u | N^{n+p-1})\,\mathrm{d}u \\ &= 0 \end{aligned}$$

because for each $\omega \in \Omega$, the Lebesgue measure of $\{T(\omega)\}$ equals 0. Consequently, $\mathrm{E}(\Delta N^n_T \Delta N^{n+p}_T) = \mathrm{E}(\Delta N^n_T \mathrm{E}(\Delta N^{n+p}_T | N^{n+p-1}_T)) = 0$, and therefore, $\Delta N^n_T \Delta N^{n+p}_T = 0$, P-a.s., which completes the proof.

Proof of Proposition 2.2 By the Fubini theorem and a change of variable, we have

$$\begin{aligned} \Lambda_t^n &= \int_0^t \left(\int_0^u \gamma_n(u-s)\,\mathrm{d}N_s^{n-1} \right) \mathrm{d}u \\ &= \int_0^t \left(\int_0^{t-s} \gamma_n(u)\,\mathrm{d}u \right) \mathrm{d}N_s^{n-1} \\ &= \int_0^t F_n(t-s)\,\mathrm{d}N_s^{n-1}, \end{aligned}$$

where $F_n(t) = \int_0^t \gamma_n(u)\,\mathrm{d}u$. Integrating by parts, we obtain

$$\int_0^t F_n(t-s)\,\mathrm{d}N_s^{n-1} = F_n(0)N_t^{n-1} - F_n(t)N_0^{n-1} + \int_0^t \gamma_n(t-s)N_s^{n-1}\,\mathrm{d}s$$

and hence $\Lambda_t^n = \int_0^t \gamma_n(t-s)N_s^{n-1}\,\mathrm{d}s$, which concludes the proof.

Proof of Proposition 2.3 Let $\mu_0 = \gamma_0$ and, for each $n \geq 1$ and $t \geq 0$, $\mu_n(t) = \mathrm{E}(\lambda_t^n)$. From Proposition 2.2, we have

$$\mu_n(t) = \mathrm{E}\left(\int_0^t \gamma_n(t-s)\,\mathrm{d}N_s^{n-1} \right) = \int_0^t \gamma_n(t-s)\mathrm{E}(\lambda_s^{n-1})\,\mathrm{d}s = (\gamma_n * \mu_{n-1})(t).$$

It follows by induction that $\mu_n = \gamma_0 * \gamma_1 * \cdots * \gamma_n$ and hence

$$\sum_{n=0}^{\infty} \mathrm{E}(N_t^n) = \int_0^t \sum_{n=0}^{\infty} (\gamma_0 * \cdots * \gamma_n)(u)\,\mathrm{d}u,$$

which concludes the proof.

Proof of Proposition 2.4 Let $H(t) = \mathrm{E}(N_t)$, $r = \sup_{n\in\mathbb{N}} \mathcal{L}[\gamma_n](s_0)$ and suppose (C1) holds. By Proposition 2.3,

$$\mathcal{L}[H](s_0) \leq \frac{1}{s_0} \sum_{n=0}^{\infty} r^{n+1} = \frac{r}{s_0(1-r)} < \infty.$$

Consequently, $H < \infty$ a.e. with respect to the Lebesgue measure, and since H is continuous, for each $t \geq 0$, $H(t) < \infty$, which implies (B).

Note that (C2) implies there exists $s_0 > 0$ such that $\sup_{k\in\mathbb{N}} \mathcal{L}[\gamma_k](s_0) < 1$. Hence (C2) implies (C1) and consequently (B) is satisfied. Under (C3), we have

$$0 \leq \sup_{k\in\mathbb{N}} \mathcal{L}[\gamma_k](s) \leq C \int_0^{\infty} \mathrm{e}^{-(s-a)u}\,\mathrm{d}u = \frac{C}{s-a},$$

whenever $s > a$, and thus (C3) implies (C2) and consequently also (B).

By the Dominated Convergence Theorem (DCT), (C4) implies (C2) and hence (B) holds.

Finally,

$$\begin{aligned}\int_0^\infty (\gamma_0 * \cdots * \gamma_n)(u)\,\mathrm{d}u &= \left(\int_0^\infty \gamma_0(u)\,\mathrm{d}u\right)\cdots\left(\int_0^\infty \gamma_n(u)\,\mathrm{d}u\right)\\ &\le \left(\sup_{k\in\mathbb{N}}\int_0^\infty \gamma_k(s)\,\mathrm{d}s\right)^{n+1}\end{aligned}$$

and therefore (C5) implies (B), concluding the proof.

Proof of Lemma 2.1 We have

$$\begin{aligned}\int_s^t (h*\gamma_0)(v)\,\mathrm{d}v &= \int_s^t\left(\int_0^v h(v-u)\gamma_0(u)\,\mathrm{d}u\right)\mathrm{d}v\\ &= \int_s^t \gamma_0(u)\left(\int_0^{t-u} h(r)\,\mathrm{d}r\right)\mathrm{d}u\\ &\le \left(\int_0^\infty h(r)\,\mathrm{d}r\right)\left(\int_s^t \gamma_0(u)\,\mathrm{d}u\right),\end{aligned}$$

which concludes the proof.

Proof of Lemma 2.2 Since $\lambda^j = \gamma_j * \cdots * \gamma_1 * \gamma_0$, from Lemma 2.1, we have

$$\mathrm{E}(\Lambda_t^j) = \int_0^t (\gamma_j * \cdots * \gamma_1 * \gamma_0)(u)\,\mathrm{d}u \le \rho^j \int_0^t \gamma_0(u)\,\mathrm{d}u.$$

Hence the Jensen and Doob inequalities imply

$$\mathrm{E}\left(\sup_{0\le u\le t}|M_u^j|^q\right) \le \mathrm{E}\left(\sup_{0\le u\le t}|M_u^j|^2\right)^{q/2} \le 2^q\mathrm{E}(\Lambda_t^j)^{q/2} \le 2^q\rho^{jq/2}\left(\int_0^t \gamma_0(u)\,\mathrm{d}u\right)^{q/2}.$$

Thus,

$$\sup_{t>0}\mathrm{E}\left(\sup_{0\le u\le t}|M_u^j/\sqrt{t}|^q\right) \le 2^q\rho^{jq/2}\sup_{t>0}\left(\frac{1}{t}\int_0^t \gamma_0(u)\,\mathrm{d}u\right)^{q/2}$$

and consequently

$$\sum_{j=0}^\infty \sup_{t>0}\mathrm{E}\left(\sup_{0\le u\le t}|M_u^j/\sqrt{t}|^q\right) \le C,$$

where $C = 2^q \sup_{t>0}\left(\frac{1}{t}\int_0^t \gamma_0(u)\,\mathrm{d}u\right)^{q/2} / \left(1-\rho^{q/2}\right)$. This completes the proof.

Proof of Lemma 2.3 For each $p \in \mathbb{N}$, $N^p = M^p + \Lambda^p$, and for each $p \geq 1$, $\Lambda^p = \gamma * N^{p-1}$. Hence (1) follows by induction and (2) is obtained from (1).

Let $F(t) = \frac{1}{t}\sum_{p=1}^{\infty}\sum_{j=0}^{p-1}(\gamma_p * \cdots * \gamma_{j+1} * |M^j|)_t$ for $t > 0$. Then

$$\begin{aligned} |F(t)| &= \left| \frac{1}{t}\sum_{j=0}^{\infty}(|M^j| * \sum_{p=j+1}^{\infty} \gamma_p * \cdots * \gamma_{j+1})_t \right| \\ &\leq \frac{1}{t}\sum_{j=0}^{\infty} \sup_{0\leq u\leq t} |M_u^j| \int_0^{\infty} \sum_{p=j+1}^{\infty} (\gamma_p * \cdots * \gamma_{j+1})(u)\,\mathrm{d}u \\ &= \frac{1}{t}\sum_{j=0}^{\infty} \sup_{0\leq u\leq t} |M_u^j| \sum_{p=j+1}^{\infty} \prod_{i=j+1}^{p} \int_0^{\infty} \gamma_i(u)\,\mathrm{d}u \\ &= \frac{\rho}{1-\rho}\sum_{j=0}^{\infty} \sup_{0\leq u\leq t} \frac{|M_u^j|}{t}, \end{aligned}$$

and from Lemma 2.2, we have $\lim_{t\to\infty} \mathrm{E}(|F(t)|) = 0$, which proves (3).

Let $h_p = \gamma_p * \cdots * \gamma_1$ and $h = \sum_{p=1}^{\infty} h_p$. We have

$$\begin{aligned} \frac{1}{t}(\gamma_p * \cdots * \gamma_1 * \gamma_0 * 1)(t) &= \frac{1}{t}\int_0^t (h_p * \gamma_0)(u)\,\mathrm{d}u \\ &= -\int_0^t h_p(s)\left(\frac{1}{t}\int_{t-s}^t \gamma_0(u)\,\mathrm{d}u\right)\mathrm{d}s \\ &+ \frac{1}{t}\int_0^t \gamma_0(u)\,\mathrm{d}u \int_0^t h_p(s)\,\mathrm{d}s. \end{aligned}$$

Hence

$$\begin{aligned} \left|\frac{1}{t}(\gamma_p * \cdots * \gamma_1 * \gamma_0 * 1)(t) - m_p\right| &\leq \int_0^{\infty} h(s)\left(\frac{1}{t}\int_{t-s}^t \gamma_0(u)\,\mathrm{d}u\right)\mathrm{d}s \\ &+ \left|\frac{1}{t}\int_0^t \gamma_0(u)\,\mathrm{d}u \int_0^t h_p(s)\,\mathrm{d}s - m_p\right|. \end{aligned}$$

By the DCT,

$$\lim_{t\to\infty}\int_0^{\infty} h(s)\left(\frac{1}{t}\int_{t-s}^t \gamma_0(u)\,\mathrm{d}u\right)\mathrm{d}s = 0$$

and

$$\begin{aligned}\left|\frac{1}{t}\int_0^t \gamma_0(u)\,\mathrm{d}u\int_0^t h_p(s)\,\mathrm{d}s - m_p\right| &= \left|\left(\frac{1}{t}\int_0^t \gamma_0(u)\,\mathrm{d}u - \overline{\gamma_0}\right)\int_0^t h_p(s)\,\mathrm{d}s\right.\\ &\quad \left. - \ \overline{\gamma_0}\int_t^\infty h_p(s)\,\mathrm{d}s\right|\\ &\le \left|\frac{1}{t}\int_0^t \gamma_0(u)\,\mathrm{d}u - \overline{\gamma_0}\right|\int_0^\infty h(s)\,\mathrm{d}s\\ &\quad + \ \overline{\gamma_0}\int_t^\infty h(s)\,\mathrm{d}s.\end{aligned}$$

Since $\int_0^\infty h(s)\,\mathrm{d}s \le \rho/(1-\rho) < \infty$, we have

$$\lim_{t\to\infty}\sup_{p\in\mathbb{N}}\left|\frac{1}{t}(\gamma_p * \cdots * \gamma_1 * \gamma_0 * 1)(t) - m_p\right| = 0. \tag{5}$$

From (1), (3) and (5), we obtain (4).

Proof of Theorem 2.1 We have

$$\mathrm{E}(M_t^2) = \sum_{j=0}^\infty \mathrm{E}(\Lambda_t^j) = \sum_{j=0}^\infty \mathrm{E}(|M_t^j|^2).$$

Hence from Lemma 2.2 and the DCT, we obtain

$$\lim_{t\to\infty}\mathrm{E}(|M_t/t|^2) = \sum_{j=0}^\infty \lim_{t\to\infty}\mathrm{E}(|M_t^j/t|^2) = 0,$$

which proves $\{M_t/t\}_{t>0}$ converges in quadratic mean to zero.

From (2), for each $t > 0$, we have

$$\frac{\Lambda_t}{t} = \frac{1}{t}\sum_{p=1}^\infty\sum_{j=0}^{p-1}\gamma_p * \cdots * \gamma_{j+1} * M^j + \frac{1}{t}\sum_{p=1}^\infty \gamma_p * \cdots * \gamma_1 * \gamma_0 * 1.$$

Hence from (3) and the Fatôu lemma, in order to prove $\{\Lambda_t/t\}_{t>0}$ converges P-a.s. to zero, it suffices to prove that

$$\lim_{t\to\infty}\frac{1}{t}\sum_{p=1}^\infty \gamma_p * \cdots * \gamma_1 * \gamma_0 * 1 = m. \tag{6}$$

Lemma 2.1 implies

$$\begin{aligned}\sum_{p=1}^\infty \sup_{t>0}\frac{1}{t}(\gamma_p * \cdots * \gamma_1 * \gamma_0 * 1)(t) &\le \sup_{t>0}\left(\frac{1}{t}\int_0^t \gamma_0(u)\,\mathrm{d}u\right)\sum_{p=1}^\infty\int_0^\infty (\gamma_p * \cdots * \gamma_1)(r)\,\mathrm{d}r\\ &\le \sup_{t>0}\left(\frac{1}{t}\int_0^t \gamma_0(u)\,\mathrm{d}u\right)\frac{\rho}{1-\rho}\\ &< \infty.\end{aligned}$$

Hence (6) follows from the DCT along with (5). Since $\{M_t/t\}_{t>0}$ is uniformly integrable, $\{M_t/t\}_{t>0}$ converges P-a.s. to zero. Thus, $\{N_t/t\}_{t>0}$ converges P-a.s. to m and the proof is complete.

Proof of Theorem 2.2 From (2), for each $t > 0$,

$$X_t = \frac{1}{\sqrt{t}} M_t + \frac{1}{\sqrt{t}} \sum_{p=1}^{\infty} \sum_{j=0}^{p-1} \int_0^t (\gamma_p * \cdots * \gamma_{j+1})(u) M^j_{t-u} \, \mathrm{d}u.$$

Let

$$Y_t = \frac{1}{\sqrt{t}} M_t + \sum_{p=1}^{\infty} \sum_{j=0}^{p-1} \frac{M^j_t}{\sqrt{t}} \int_0^{\infty} (\gamma_p * \cdots * \gamma_{j+1})(u) \, \mathrm{d}u$$

and $D_t = X_t - Y_t$ for $t > 0$. Notice that $D_t = D_{1,t} + D_{2,t}$, where

$$D_{1,t} = \frac{1}{\sqrt{t}} \sum_{p=1}^{\infty} \sum_{j=0}^{p-1} \int_0^t (\gamma_p * \cdots * \gamma_{j+1})(u) (M^j_{t-u} - M^j_t) \, \mathrm{d}u$$

and

$$D_{2,t} = \sum_{p=1}^{\infty} \sum_{j=0}^{p-1} \frac{M^j_t}{\sqrt{t}} \int_t^{\infty} (\gamma_p * \cdots * \gamma_{j+1})(u) \, \mathrm{d}u.$$

We need to prove $\{D_{1,t}\}_{t>0}$ and $\{D_{2,t}\}_{t>0}$ converge in probability to zero.

We have

$$\mathrm{E}(|D_{1,t}|) \leq \sum_{p=1}^{\infty} \sum_{j=0}^{p-1} \int_0^{\infty} (\gamma_p * \cdots * \gamma_{j+1})(u) \mathrm{E}(|M^j_{t-u} - M^j_t|/\sqrt{t}) \, \mathrm{d}u$$

and, since

$$|M^j_{t-u} - M^j_t|/\sqrt{t} \leq 2 \sup_{0 \leq u \leq t} |M^j_u|/\sqrt{t},$$

we have $(\gamma_p * \cdots * \gamma_{j+1})(u)\mathrm{E}(|M^j_{t-u} - M^j_t|/\sqrt{t})$ is bounded by

$$C_{p,j}(u) = 2(\gamma_p * \cdots * \gamma_{j+1})(u) \sup_{t>0} \mathrm{E}\left(\sup_{0 \leq u \leq t} |M^j_u|/\sqrt{t} \right).$$

Thus, by Lemma 2.2,

$$\begin{aligned} \sum_{p=1}^{\infty} \sum_{j=0}^{p-1} \int_0^{\infty} C_{p,j}(u) \, \mathrm{d}u &= \sum_{j=0}^{\infty} \sum_{p=j+1}^{\infty} \int_0^{\infty} C_{p,j}(u) \, \mathrm{d}u \\ &\leq \frac{2\rho}{1-\rho} \sum_{j=0}^{\infty} \sup_{t>0} \mathrm{E}\left(\sup_{0 \leq u \leq t} |M^j_u|/\sqrt{t} \right) \\ &< \infty. \end{aligned}$$

Consequently,

$$\limsup_{t\to\infty} \mathrm{E}(|D_{1,t}|) \le \sum_{p=1}^{\infty}\sum_{j=0}^{p-1} \int_0^{\infty} (\gamma_p * \cdots * \gamma_{j+1})(u) \limsup_{t\to\infty} \mathrm{E}(|M^j_{t-u} - M^j_t|/\sqrt{t})\,\mathrm{d}u.$$

Let $h_j = \gamma_j * \cdots * \gamma_1$ and $t^* > 0$ such that $\frac{1}{t}\int_0^t \gamma_0(v)\,\mathrm{d}v < \overline{\gamma_0} + 1$ if $t > t^*$. By the Jensen inequality, for each $u \ge 0$,

$$\begin{aligned}
\mathrm{E}(|M^j_{t-u} - M^j_t|/\sqrt{t})^2 &\le \mathrm{E}(|M^j_{t-u} - M^j_t|^2/t)\\
&= \mathrm{E}[(\Lambda^j_t - \Lambda^j_{t-u})/t]\\
&= \frac{1}{t}\int_{t-u}^{t} (h_j * \gamma_0)(v)\,\mathrm{d}v\\
&= \int_0^{t-u} h_j(s)\left(\frac{1}{t}\int_{t-s-u}^{t-s} \gamma_0(r)\,\mathrm{d}r\right)\mathrm{d}s\\
&\quad + \int_{t-u}^{t} h_j(s)\left(\frac{1}{t}\int_{t}^{t-s} \gamma_0(r)\,\mathrm{d}r\right)\mathrm{d}s\\
&\le \int_0^{\infty} h_j(s)\left(\frac{1}{t}\int_{t-s-u}^{t-s} \gamma_0(r)\,\mathrm{d}r\right)\mathrm{d}s\\
&\quad + \int_0^{\infty} h_j(s)\left(\frac{1}{t}\int_{t}^{t-s} \gamma_0(r)\,\mathrm{d}r\right)\mathrm{d}s.
\end{aligned}$$

Since

$$\lim_{t\to\infty} \frac{1}{t}\int_{t-s-u}^{t-s} \gamma_0(r)\,\mathrm{d}r = \lim_{t\to\infty} \frac{1}{t}\int_{t}^{t-s} \gamma_0(r)\,\mathrm{d}r = 0$$

and

$$\int_0^{\infty} h_j(s)\,\mathrm{d}s < \infty,$$

it follows from the DCT that $\lim_{t\to\infty} \mathrm{E}(|M^j_{t-u} - M^j_t|/\sqrt{t}) = 0$, which proves that $\limsup_{t\to\infty} \mathrm{E}(|D_{1,t}|) = 0$.

We have

$$\mathrm{E}(|D_{2,t}|) \le \sum_{p=1}^{\infty}\sum_{j=0}^{p-1} \mathrm{E}(|M^j_t|/\sqrt{t}) \int_t^{\infty} (\gamma_p * \cdots * \gamma_{j+1})(u)\,\mathrm{d}u$$

and

$$\mathrm{E}(|M^j_t|/\sqrt{t}) \sum_{p=j+1}^{\infty} \int_t^{\infty} (\gamma_p * \cdots * \gamma_{j+1})(u)\,\mathrm{d}u \le \sup_{t>0} \mathrm{E}\left(\sup_{0\le u\le t} |M^j_u|/\sqrt{t}\right)\rho^{p-j}.$$

Since, by Lemma 2.2,

$$\sum_{p=1}^{\infty}\sum_{j=0}^{p-1}\sup_{t>0}\mathrm{E}\left(\sup_{0\le u\le t}|M_u^j|/\sqrt{t}\right)\rho^{p-j}=\frac{\rho}{1-\rho}\sum_{j=0}^{\infty}\sup_{t>0}\mathrm{E}\left(\sup_{0\le u\le t}|M_u^j|/\sqrt{t}\right)<\infty,$$

we obtain

$$\lim_{t\to\infty}\mathrm{E}(|D_{2,t}|)\le\sum_{p=1}^{\infty}\sum_{j=0}^{p-1}\lim_{t\to\infty}\mathrm{E}(|M_t^j|/\sqrt{t})\int_t^{\infty}(\gamma_p*\cdots*\gamma_{j+1})(u)\,\mathrm{d}u.$$

But $\sup_{t>0}\mathrm{E}(|M_t^j|/\sqrt{t})<\infty$ and $\int_0^{\infty}(\gamma_p*\cdots*\gamma_{j+1})(u)\,\mathrm{d}u<\infty$. Consequently $\lim_{t\to\infty}\mathrm{E}(|D_{2,t}|)=0$.

Due to $\{D_{1,t}\}_{t>0}$ and $\{D_{2,t}\}_{t>0}$ converge in probability to zero, it only remains to prove $\{Y_t\}_{t>0}$ converges in distribution to a normal random variable with mean zero and variance σ_N^2. To this purpose, we use Theorem 1 in [19] (Chapter 8). For each $j\in\mathbb{N}$, let

$$\alpha_j=1+\sum_{p=1}^{\infty}\prod_{i=j+1}^{p+j}\int_0^{\infty}\gamma_i(u)\,\mathrm{d}u$$

and note that $Y_t=Z_t/\sqrt{t}$, where $Z=\{Z_t\}_{t\ge0}$ is given by $Z_t=\sum_{j=0}^{\infty}\alpha_jM_t^j$. Since $\sup_{j\in\mathbb{N}}\alpha_j<\infty$, we have

$$\mathrm{E}(Z_t^2)\le\sup_{j\in\mathbb{N}}\alpha_j^2\sum_{j=0}^{\infty}\mathrm{E}(|M_t^j|^2)=\sup_{j\in\mathbb{N}}\alpha_j^2\mathrm{E}(N_t)<\infty.$$

Moreover, the martingales M^j ($j\in\mathbb{N}$) have no common jumps. Hence the predictable quadratic variation of the martingale $\{Z_t\}_{t\ge0}$ is given, for each $t\ge0$, by

$$\langle Z\rangle_t=\sum_{j=0}^{\infty}\alpha_j^2\langle M^j\rangle_t=\sum_{j=0}^{\infty}\alpha_j^2\Lambda_t^j.$$

As usual, $[t]$ denotes the integer part of t ($t>0$). By making use of Lemma 2.2, it is easy to see that $\{Y_t-Y_{[t]}\}_{t>0}$ converges in probability to zero. Consequently, in order to prove the convergence of $\{Y_t\}_{t>0}$, it suffices to prove $\{Y_n\}_{n\in\mathbb{N}\setminus\{0\}}$ converges in distribution to a normal random variable with mean zero and variance σ_N^2.

For $n\ge1$, define $\xi_{n,k}=(Z_k-Z_{k-1})/\sqrt{n}$ ($k=1,\dots,n$). Hence $\{\xi_{n,k}\}_{0\le k\le n}$ is a martingale-difference array with respect to $\{\mathcal{E}_{n,k}\}_{0\le k\le n}$, where for each $n\in\mathbb{N}$, $\mathcal{E}_{n,k}=\mathcal{F}_k$, i.e., $\xi_{n,k}$ is $\mathcal{E}_{n,k}$ measurable and $\mathrm{E}(\xi_{n,k}|\mathcal{E}_{n,k-1})=0$.

Note that

$$\sum_{k=1}^{n} \mathrm{E}(\xi_{n,k}^2|\mathcal{E}_{n,k-1}) = \sum_{k=1}^{n}\sum_{j=0}^{\infty} \alpha_j^2(\Lambda_k^j - \Lambda_{k-1}^j)/n = \sum_{j=0}^{\infty} \alpha_j^2 \Lambda_n^j/n$$

and

$$\sum_{k=1}^{n} \mathrm{E}(\xi_{n,k}^2|\mathcal{E}_{n,k-1}) - \sigma_N^2 = \sum_{j=0}^{\infty} \alpha_j^2 \left(\frac{\Lambda_n^j}{n} - m_j\right).$$

Thus,

$$\mathrm{E}\left|\sum_{k=1}^{n} \mathrm{E}(\xi_{n,k}^2|\mathcal{E}_{n,k-1}) - \sigma_N^2\right| \leq \sum_{j=0}^{\infty} \alpha_j^2 \mathrm{E}\left|\frac{\Lambda_n^j}{n} - m_j\right|.$$

Notice that if $m_{j^*} = 0$ for some $j^* \in \mathbb{N}$, from (4) we have

$$\lim_{n\to\infty} \sum_{j=0}^{\infty} \alpha_j^2 \mathrm{E}\left|\frac{\Lambda_n^j}{n} - m_j\right| = \lim_{n\to\infty} \sum_{j=0}^{j^*-1} \alpha_j^2 \mathrm{E}\left|\frac{\Lambda_n^j}{n} - m_j\right| = 0.$$

Next, assume $m_j \neq 0$ for all $j \in \mathbb{N}$. This implies that $\overline{\gamma_0} \neq 0$ and from (1) and Lemma 2.1 we obtain

$$\begin{aligned}
\sup_{n\geq 1, j\in\mathbb{N}} \mathrm{E}\left(\frac{\Lambda_n^j}{nm_j}\right) &\leq \sup_{n\geq 1, j\in\mathbb{N}} \left(\frac{1}{m_j}\int_0^n (\gamma_j * \cdots * \gamma_1)(u)\,\mathrm{d}u\right)\left(\frac{1}{n}\int_0^n \gamma_0(u)\,\mathrm{d}u\right) \\
&\leq \sup_{n\geq 1, j\in\mathbb{N}} \left(\frac{1}{m_j}\prod_{i=1}^{j}\int_0^\infty \gamma_i(u)\,\mathrm{d}u\right)\left(\frac{1}{n}\int_0^n \gamma_0(u)\,\mathrm{d}u\right) \\
&= \sup_{n\geq 1, j\in\mathbb{N}} \frac{1}{\overline{\gamma_0} n}\int_0^n \gamma_0(u)\,\mathrm{d}u \\
&< \infty.
\end{aligned}$$

Since

$$\mathrm{E}\left|\frac{\Lambda_n^j}{n} - m_j\right| \leq m_j \mathrm{E}\left|\frac{\Lambda_n^j}{m_j n} - 1\right| \leq (C+1)m_j,$$

where $C = \sup_{n\geq 1, j\in\mathbb{N}} \mathrm{E}(\Lambda_n^j/nm_j)$, and $\sum_{j=0}^{\infty} m_j = m < \infty$, from (4) in Lemma 2.3 we obtain

$$\lim_{n\to\infty} \sum_{j=0}^{\infty} \alpha_j^2 \mathrm{E}\left|\frac{\Lambda_n^j}{n} - m_j\right| = \sum_{j=0}^{\infty} \alpha_j^2 \lim_{n\to\infty} \mathrm{E}\left|\frac{\Lambda_n^j}{n} - m_j\right| = 0.$$

Hence

$$\lim_{n\to\infty} \mathrm{E}\left|\sum_{k=1}^{n} \mathrm{E}(\xi_{n,k}^2|\mathcal{E}_{n,k-1}) - \sigma_N^2\right| = 0.$$

To complete the proof, we need to verify that $\{\xi_{n,k}\}_{0\leq k\leq n}$ satisfies the Lindeberg condition stated in Theorem 1 in [19] (Chapter 8). For this purpose, we prove that the sequence $\{\max_{0\le k\le n}\xi_{n,k}\}_{n\in\mathbb{N}\setminus\{0\}}$ is uniformly integrable and converges in probability to zero (see e.g. pages 314-315 in [7]).

Let $k^*=\min\{k\le n:\xi^2_{n,k}=\max_{0\le k\le n}\xi^2_{n,k}\text{ or }k=n\}$. Hence by the Doob Optional Sampling Theorem along with (1) and Lemma 2.1, we have

$$\begin{aligned}
\mathrm{E}\left(\max_{0\le k\le n}\xi^2_{n,k}\right) &= \mathrm{E}\left(\xi^2_{n,k^*}\right)\\
&= \frac{1}{n}\sum_{j=0}^{\infty}\alpha_j^2\mathrm{E}\left(\Lambda^j_{k^*}-\Lambda^j_{k^*-1}\right)\\
&= \frac{1}{n}\sum_{j=0}^{\infty}\alpha_j^2\mathrm{E}\left(\int_{k^*-1}^{k^*}(\gamma_j*\cdots*\gamma_1*\gamma_0)(u)\,\mathrm{d}u\right)\\
&\le \frac{1}{n}\sum_{j=0}^{\infty}\alpha_j^2\rho^j\mathrm{E}\left(\int_{k^*-1}^{k^*}\gamma_0(u)\,\mathrm{d}u\right)\\
&\le \frac{1}{n}\sup_{t>0}\frac{1}{t}\int_0^t\gamma_0(u)\,\mathrm{d}u\sum_{j=0}^{\infty}\alpha_j^2\rho^j.
\end{aligned}$$

Since $\sup_{t>0}\frac{1}{t}\int_0^t\gamma_0(u)\,\mathrm{d}u\sum_{j=0}^{\infty}\alpha_j^2\rho^j<\infty$, we obtain $\lim_{n\to\infty}\mathrm{E}(\max_{0\le k\le n}\xi^2_{n,k})=0$. Thus, the sequence $\{\max_{0\le k\le n}\xi_{n,k}\}_{n\in\mathbb{N}\setminus\{0\}}$ is uniformly integrable and converges in probability to zero. This concludes the proof.

Proof of Lemma 2.4 For each $C>0$, let φ_C be the function from $\mathbb{R}$ to $\mathbb{R}$ defined as

$$\varphi_C(x)=\begin{cases}-C, & \text{if } x<-C,\\ x, & \text{if } C\le x\le C,\\ C, & \text{if } x>C.\end{cases}$$

Due to (F1), it suffices to prove that, for each $C>0$, $\{(\varphi_C(U_t),\varphi_C(V_t))\}_{t>0}$ converges in distribution to $(\varphi_C(U),\varphi_C(V))$. Fix $C>0$ and let f be a bounded and continuous function from $\mathbb{R}^2$ to $\mathbb{R}$ and $\epsilon>0$. From the Stone-Weierstrass theorem, there exist $u_1,\dots,u_r$ and $v_1,\dots,v_r$, real continuous functions, defined on $K=[-C,C]\times[-C,C]$ such that

$$\sup_{(x,y)\in K}|f(x,y)-\sum_{i=1}^{r}u_i(x)v_i(y)|<\epsilon.$$

Hence

$$|\mathrm{E}[f(\varphi_C(U_t)),\varphi_C(V_t)] - \mathrm{E}[f(\varphi_C(U)),\varphi_C(V)]| \leq \left|\sum_{i=1}^{r} \mathrm{E}[u_i(\varphi_C(U_t))v_i(\varphi_C(V_t))] - \mathrm{E}[u_i(\varphi_C(U))v_i(\varphi_C(V))]\right| + 2\epsilon$$

and from (F2), we obtain

$$\limsup_{t\to\infty} |\mathrm{E}[f(\varphi_C(U_t)),\varphi_C(V_t)] - \mathrm{E}[f(\varphi_C(U)),\varphi_C(V)]| \leq 2\epsilon.$$

Since $\epsilon > 0$ is arbitrary, the proof is complete.

Proof of Theorem 2.3 For each $n \in \mathbb{N}\setminus\{0\}$ and $t > 0$, let

$$X_n = \frac{1}{\sqrt{n}}\sum_{k=0}^{n}(\xi_k - \nu) \quad \text{and} \quad Y_t = \nu\left(\frac{N_t - \mathrm{E}(N_t)}{\sqrt{t}}\right).$$

We have $\{X_n\}_{n\in\mathbb{N}\setminus\{0\}}$ and $\{Y_t\}_{t>0}$ are independent and

$$R_t = \sqrt{\frac{N_t}{t}}X_{N_t} + Y_t. \tag{7}$$

By the standard Central Limit Theorem and Theorem 2.2, $\{X_n\}_{n\in\mathbb{N}\setminus\{0\}}$ and $\{Y_t\}_{t>0}$ converge in distribution to two normal random variables X and Y, respectively. We assume X and Y are defined on $(\Omega,\mathcal{F},\mathrm{P})$ and hence they are independent. By Theorem 2.1, (7) and the Slutzky theorem, it suffices to prove $\{(X_{N_t},Y_t)\}_{t>0}$ converges in distribution to (X,Y). For this purpose, we use Lemma 2.4. Since $\{X_{N_t}\}_{t>0}$ and $\{Y_t\}_{t>0}$ are convergent in distribution, we have $\{(X_{N_t},Y_t)\}_{t>0}$ satisfies (F1). Let u and v be continuous and bounded functions from $\mathbb{R}$ to $\mathbb{R}$, $c_u = \sup_{x\in\mathbb{R}}|u(x)|$ and $c_v = \sup_{x\in\mathbb{R}}|v(x)|$. Since $\{X_t\}_{t>0}$ converges in distribution to X, there exists $t^* \geq 0$ such that $|\mathrm{E}[u(X_t) - u(X)]| < \epsilon$, for all $t > t^*$.

Since X is independent of $\{Y_t\}_{t>0}$ and Y, we have

$$\begin{aligned}
|\mathrm{E}(u(X_{N_t})v(Y_t) - u(X)v(Y))| &\leq |\mathrm{E}([u(X_{N_t}) - u(X)]v(Y_t))| \\
&\quad + |\mathrm{E}(u(X)[v(Y_t) - v(Y)]| \\
&\leq \left|\mathrm{E}[(u(X_{N_t}) - u(X))v(Y_y)\mathrm{I}_{\{N_t>t^*\}}]\right| \\
&\quad + 2c_uc_v\mathrm{P}(N_t \leq t^*) + c_u\left|\mathrm{E}[v(Y_t) - v(Y)]\right|.
\end{aligned}$$

For each $\omega \in \{N_t > t^*\}$, we have

$$
\begin{aligned}
&\left|\mathrm{E}[(u(X_{N_t}) - u(X))v(Y_t)\mathrm{I}_{\{N_t>t^*\}}|N_t](\omega)\right| \\
&\quad = \left|v(Y_t(\omega))\mathrm{E}[(u(X_{N_t}) - u(X))\mathrm{I}_{\{N_t>t^*\}}|N_t](\omega)\right| \\
&\quad \leq c_v \left|\mathrm{E}[(u(X_{N_t(\omega)}) - u(X))\mathrm{I}_{\{N_t>t^*\}}|N_t](\omega)\right| \\
&\quad = c_v \left|\mathrm{E}[(u(X_{N_t(\omega)}) - u(X))]\right| \mathrm{I}_{\{N_t>t^*\}}(\omega) \\
&\quad < c_v\epsilon.
\end{aligned}
$$

Consequently,

$$
|\mathrm{E}(u(X_{N_t})v(Y_t) - u(X)v(Y))| \leq c_v\epsilon + 2c_uc_v\mathrm{P}(N_t \leq t^*) + c_u\left|\mathrm{E}[v(Y_t) - v(Y)]\right|.
$$

But $\epsilon > 0$ is arbitrary and $\lim_{t\to\infty}\{2c_uc_v\mathrm{P}(N_t \leq t^*) + c_u\left|\mathrm{E}[v(Y_t) - v(Y)]\right|\} = 0$. Therefore, $\lim_{t\to\infty}|\mathrm{E}[u(X_{N_t})v(Y_t) - u(X)v(Y)]| = 0$ and, by Lemma 2.4, the proof is complete.

Acknowledgements

The research work of Raúl Fierro and Víctor Leiva was partially supported by Chilean Council for Scientific and Technological Research, grant FONDECYT 1120879. The research work of Jesper Møller was supported by the Danish Council for Independent Research-Natural Sciences, grant 12-124675, "Mathematical and Statistical Analysis of Spatial Data", and by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation.

References

[1] Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J.F. (2011). Some limit theorems for Hawkes processes and application to financial statistics. *Stoch. Process. Appl.* **123,** 2475-2499.

[2] Brémaud, P. (1981). *Point Processes and Queues. Martingale Dynamics.* Springer-Verlag, New York.

[3] Carstensen, L., Sandelin, A., Winther, O. and Hansen N.R. (2010). Multivariate Hawkes process models of the occurrence of regulatory elements. *BMC Bioinformatics* **11,** 456.

[4] Daley, D.J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes*, 2nd edn. Springer-Verlag, New York.

[5] Embrechts, P., Liniger, J.T. and Lu, L. (2011). Multivariate Hawkes processes: an application to financial data. *J. Appl. Probab.* **48A,** 367-378.

[6] Fierro, R., Leiva, V., Ruggeri, F. and Sanhueza, A. (2013). On a Birnbaum-Saunders distribution arising from a non-homogeneous Poisson process. *Stat. Probab. Lett.* **83,** 1233-1239.

[7] Gaenssler, P. and Haeusler, E. (1986). On martingale central limit theory. In: *Dependence in Probability and Statistics. A Survey of Recent Results.* Eds. Eberlein, E. and Taqqu, M.S. Birkhuser Boston Inc., 303-334.

[8] Gusto, G. and Schbath, S. (2005). FADO: A statistical method to detect favored or avoided distances between occurrences of motifs using the Hawkes' model. *Stat. Appl. Genet. Mol. Biol.* **4,** Article 24.

[9] Hawkes, A.G. (1971). Point spectra of some mutually exciting point processes. *J. Roy. Stat. Soc. Ser. B* **33,** 438-443.

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

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

[12] Jacod, J. (1975). Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales. *Z. Wahrsch. Verw. Gebiete* **31,** 235-253.

[13] Møller, J. and Rasmussen, J.G. (2005). Perfect simulation of Hakwes Processes. *J. Appl. Probab.* **37,** 629-646.

[14] Møller, J. and Rasmussen, J.G. (2006). Approximate simulation of Hakwes Processes. *Meth. Comput. Appl. Probab.* **8,** 53-64.

[15] Møller, J. and Waagepetersen, R.W. (2004). *Statistical Inference and Simulation for Spatial Point Processes.* Chapman and Hall/CRC, Boca Raton.

[16] Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.* **83,** 9-27.

[17] Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Ann. I. Stat. Math.* **83,** 9-27.

[18] Pernice, V., Staude, B., Carndanobile, S. and Rotter, S. (2012). Recurrent interactions in spiking networks with arbitrary topology. *Phys. Rev. E* **85,** 031916 [7 pages].

[19] Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer-Verlag, New York.

[20] Zhu, L. (2013). Central limit theorem for nonlinear Hawkes processes. *J. Appl. Probab.* **50,** 760-771.

[21] Zhu, L. (2013). Nonlinear Hawkes processes. Doctoral Thesis, New York University.