Submitted to the Annals of Statistics

NONPARAMETRIC INFERENCE FOR HAWKES PROCESSES WITH A RESCALED STOCHASTIC TIME DEPENDENT BASELINE

BY YOANN POTIRON[1,a],

[1]*Faculty of Business and Commerce, Keio University,* [a]*potiron@fbc.keio.ac.jp*

We develop nonparametric inference for Hawkes processes with a rescaled stochastic and time dependent baseline. The inference procedure is based on the average of the point process. We consider estimation for the average over time of the intensity process. We first show the existence of these point processes. We also show the central limit theorem of the inference procedure. This requires the assumption that the kernel does not have a too fat tail. We also need that the baseline process is Lipschitz continuous with bounded starting value. The main novelty in the proofs is to establish the renewal equation for stochastic processes.

1. Introduction. This paper concerns nonparametric inference for point processes. The main stylized fact in this strand of literature is the presence of event clustering in time. This motivates to rely on the so-called Hawkes mutually exciting processes (see Hawkes (1971a) and Hawkes (1971b)). We define the point process N_t of dimension d as the number of events from the starting time 0 to the final time t and λ_t its intensity. A standard definition of Hawkes mutually exciting processes is given by

$$\lambda_t = \nu + \int_0^t h(t-s)\, dN_s. \tag{1}$$

Here, ν is a Poisson baseline of dimension d and h is a kernel matrix of dimension $d \times d$. The point processes are mutually exciting in the sense that the diagonal components $h^{(i,i)}$ are self-exciting terms for the ith process and non diagonal components $h^{(i,j)}$ are cross exciting terms for the ith process made by events from the jth process. The particular case $h = 0$ corresponds to a classical Poisson process, so that we can view Hawkes processes as a natural extension of Poisson processes.

The main application of Hawkes processes lies in seismology (see Rubin (1972), Ozaki (1979), Vere-Jones and Ozaki (1982) and Ogata (1978), Ogata (1988)). There are also applications in quantitative finance (see Chavez-Demoulin, Davison and McNeil (2005), Embrechts, Liniger and Lin (2011), Bacry et al. (2013), Jaisson and Rosenbaum (2015), Jaisson and Rosenbaum (2016), Clinet and Yoshida (2017)). Some applications are also in financial econometrics (see Chen and Hall (2013), Clinet and Potiron (2018), Kwan, Chen and Dunsmuir (2023), Potiron and Volkov (2025+)). We can also find some applications in biology (see Reynaud-Bouret and Schbath (2010) and Donnet, Rivoirard and Rousseau (2020)). See also Liniger (2009) and Hawkes (2018) with the references therein.

There are many theoretical results for Hawkes processes in statistics. Hawkes and Oakes (1974) provide a Poisson cluster process representation for the Hawkes process. Brémaud and Massoulié (1996) study stability of nonlinear Hawkes processes. Zhu (2013) gives central limit theorem for nonlinear Hawkes processes. Zhu (2015) considers large deviations for Markovian nonlinear Hawkes processes. The microstructure of stochastic volatility models with self-excitation is investigated in Horst and Xu (2022). Horst and Xu (2021) and

MSC2020 subject classifications: Primary 62G05, 62M09; secondary 60G55.
Keywords and phrases: point processes, Hawkes mutually exciting processes, nonparametric inference, central limit theory, time dependent baseline, renewal equation.

Horst and Xu (2024+) give functional limit theorems for Hawkes processes. Xu (2024) studies diffusion approximations for self-excited systems. Karim, Laeven and Mandjes (2025+) introduce compound multivariate Hawkes processes. Potiron (2025+) consider parametric inference.

Empirical evidence with financial applications suggests that the baseline is time dependent during intraday trades. Chen and Hall (2013) report in their empirical study (see Section 5.2, pp. 7–10) that goodness-of-fit results are in favor of their time-varying baseline model compared to a group of alternatives. In Figure 2 (p. 20), they document the time dependent nonrandom function for both polynomial and exponential kernel. This nonrandom path is also visible in Figure 2 (p. 3488) from Clinet and Potiron (2018).

There are also theoretical results in statistics of Hawkes processes with a baseline which is time dependent and possibly random. Chen and Hall (2013), Roueff, von Sachs and Sansonnet (2016), Clinet and Potiron (2018), Roueff and Von Sachs (2019), Cheysson and Lang (2022), Kwan, Chen and Dunsmuir (2023), Mammen and Müller (2023) and Erdemlioglu et al. (2025) study locally stationary Hawkes processes. Potiron et al. (2025a) and Potiron et al. (2025b) introduce Hawkes processes with Itô semimartingale baseline.

In this paper, we consider Hawkes processes, where the kernel has a general form and is nonparametric. We introduce a baseline which is rescaled stochastic and time dependent. The inference procedure is based on the average of the point process. We consider estimation for the rescaled integral of the intensity process λ_t between the starting time 0 and the final time tT for any time $t \in [0,1]$, namely

$$\Lambda_{t,T} = \frac{1}{T}\int_0^{tT} \lambda_s ds. \tag{2}$$

We have applications in management science where the target quantity (2) can be interpreted as the arrival rate in a queuing system (see Kao and Chang (1988) and Leemis (1991)). We also have applications in computer networks for the expected internet traffic (see Kuhl and Wilson (2001), and seismology for the expected number of earthquakes. Finally, we have applications in finance where the intensity of a quote plays an inverse role to the volatility of an asset price (see Potiron et al. (2025a)).

There are some research work on estimation of the target quantity (2). They are restricted to the case when the intensity λ_t is nonrandom. A pioneer work for nonparametric estimation is Leemis (1991). A different nonparametric approach based on kernel estimation is suggested by Lewis and Shedler (1976). A wavelet based nonparametric method can be found in Kuhl and Bhairgond (2000). Parametric methods are also available in Kao and Chang (1988), Lee, Wilson and Crawford (1991), Kuhl, Wilson and Johnson (1997), Kuhl and Wilson (2000). Finally, a semiparametric framework is considered in Kuhl and Wilson (2001).

There are two contributions in this paper. First, we give an existence result in Proposition 1. This complements Theorem 5.1 (p. 3476) from Clinet and Potiron (2018) and Proposition 1 in Erdemlioglu et al. (2025). In particular, our nonparametric kernel framework allows for more general kernels, which is useful for applications. Indeed, there is empirical evidence in finance that the kernel decays as the power distribution (see Bacry, Dayri and Muzy (2012) and Hardiman, Bercot and Bouchaud (2013)). The arguments used in the proofs slightly extend the arguments from Brémaud and Massoulié (1996) and Clinet and Potiron (2018).

Second, our main contribution is the central limit theorem of the inference procedure in Theorem 1. This extends Corollary 1 (p. 2481) from Bacry et al. (2013) which is restricted to nonrandom constant baseline. This requires the assumption that the kernel does not have a too fat tail. We also need that the baseline process is Lipschitz continuous with bounded starting value.

The main novelty in the proofs of the central limit theorem is to establish the renewal equation for stochastic processes. This extends Bacry et al. (2013) in which the function from the renewal equation is nonrandom. Jaisson and Rosenbaum (2015) also uses the renewal equation with a different asymptotics. More generally, renewal techniques for Hawkes processes are studied in Costa et al. (2020) and Graham (2021).

The remainder of this paper is organized as follows. We introduce Hawkes processes with a rescaled stochastic and time dependent baseline and show their existence in Section 2. We consider nonparametric inference for Hawkes processes and prove the central limit theorem in Section 3. The proofs are gathered in Section 4. Finally, we provide concluding remarks in Section 5.

2. Hawkes processes with a rescaled stochastic time dependent baseline. In this section, we introduce Hawkes processes with a rescaled stochastic and time dependent baseline. We give an existence result in Proposition 1. This complements the framework from Clinet and Potiron (2018) and Erdemlioglu et al. (2025). In particular, our nonparametric kernel framework allows for more general kernels, which is useful for applications. The arguments used in the proofs slightly extend the arguments from Brémaud and Massoulié (1996) and Clinet and Potiron (2018).

We start with an introduction to the point process N_t of dimension d. For any index $i = 1,\dots,d$, each component of the point process $N_t^{(i)}$ counts the number of events between 0 and t for the ith process. Here, we denote the ith component of a vector V by $V^{(i)}$. We define $N_t^{(i)}$ as a simple point process on the space $\mathbb{R}^+$, i.e. a family

$$(N^{(i)}(C))_{C\in\mathcal{B}(\mathbb{R}^+)}$$

of random variables with values in the space $\overline{\mathbb{N}} = \mathbb{N}\cup\{+\infty\}$. Here, $\mathcal{B}(S)$ denotes the Borel σ-algebra on the space S for any space S. Moreover,

$$N^{(i)}(C) = \sum_{k\in\mathbb{N}} \mathbf{1}_C(T_k^{(i)})$$

and $\{T_k^{(i)}\}_{k\in\mathbb{N}}$ is a sequence of event times, which are $\mathbb{R}^+$ valued and random.

The definition of simple point process requires some specific assumptions on the point process. We assume that the time of the first event $T_0^{(i)}$ is equal to 0 a.s. and the following times are increasing for each process a.s. Namely, we assume that

$$\mathbb{P}(T_0^{(i)} = 0 \text{ and } T_k^{(i)} < T_{k+1}^{(i)} \text{ for } k\in\mathbb{N}_* \text{ and } i = 1,\cdots,d) = 1. \tag{3}$$

Here, we define for any space S such that $0\in S$ the space without zero as S_*. We also assume that no events happen at the same time for different processes a.s., i.e.

$$\mathbb{P}(T_k^{(i)} \neq T_l^{(j)} \text{ for } k,l\in\mathbb{N}_* \text{ and } i,j = 1,\cdots,d \text{ s.t. } i\neq j) = 1.$$

We introduce the stochastic basis $\mathcal{B} = \big(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in\mathbb{R}^+},\mathbb{P}\big)$, namely a probability space equipped with a filtration. The filtration $\mathcal{F}_t$ represents the information available at the time $t\in\mathbb{R}^+$. We assume that the stochastic basis $\mathcal{B}$ satisfies the usual conditions. We denote the natural filtration generated by some stochastic process X for any time $t\in\mathbb{R}$ as $\mathcal{F}_t^X = \sigma\{X_s : 0\le s\le t\}$. We first introduce the definition of the $\mathcal{F}_t$-intensity for the point process N_t.

DEFINITION 1. Any stochastic process λ_t defined on the real positive numbers $\mathbb{R}^+$ and satisfying the following properties is called an $\mathcal{F}_t$ intensity of the point process N_t. First, we have that

$$\mathbb{E}[N((a,b]) \mid \mathcal{F}_a] = \mathbb{E}\Big[\int_a^b \lambda_s ds \mid \mathcal{F}_a\Big] \text{ a.s.} \tag{4}$$

for all intervals $(a, b] \subset \mathbb{R}^+$. Second, the stochastic process λ_t is $\mathcal{F}_t$ progressively measurable, of dimension d where each component $\lambda_t^{(i)}$ takes its values in the space of nonnegative real numbers $\mathbb{R}^+$.

Intuitively, the intensity corresponds to the expected number of events given the past information, i.e.

$$\lambda_t = \lim_{u \to 0} \mathbb{E}\Big[\frac{N_{t+u} - N_t}{u}\Big|\mathcal{F}_t\Big] \text{ a.s.}$$

Moreover, we have that the compensated point process defined as

$$M_t = N_t - \int_0^t \lambda_s ds. \tag{5}$$

is an $\mathcal{F}_t$ martingale a.s. Finally, we note that $N((a, b])$ is a.s. finite if and only if $\int_a^b \lambda_s ds$ is a.s. finite. For background on point processes, the reader can consult Jacod (1975), Jacod and Shiryaev (2003), Daley and Vere-Jones (2003), and Daley and Vere-Jones (2008).

The present work is concerned with Hawkes processes featuring a stochastic and time dependent baseline. More specifically, the intensity λ_t of the point process N_t for any time $t \in [0, T]$ follows

$$\lambda_t = b_t + \int_0^t h(t - s)\, dN_s. \tag{6}$$

Here, the kernel h is a matrix of dimension $d \times d$. The point processes are mutually exciting in the sense that the diagonal components $h^{(i,i)}$ are self-exciting terms for the ith process and non-diagonal components $h^{(i,j)}$ are cross-exciting terms for the ith process made by events from the jth process. Moreover, the baseline b_t is a stochastic process of dimension d.

We introduce a rescaled baseline which satisfies $b_t = \nu_{t/T}$ for any time $t \in [0, T]$. Here, ν_t is a stochastic process of dimension d defined on the time interval $[0, 1]$. Then, the intensity λ_{tT} of the point process with rescaled baseline N_{tT} for any time $t \in [0, 1]$ and any final time $T > 0$ follows

$$\lambda_{tT} = \nu_t + \int_0^{tT} h(tT - s)\, dN_s. \tag{7}$$

Here, the point process N_{tT} and its intensity λ_{tT} implicitly depend on the time T. Moreover, the baseline process ν_t is rescaled from the time interval $[0, 1]$ to the time interval $[0, T]$.

We denote the spectral radius of any matrix ϕ by $\rho(\phi)$. Then, we define the L^1 norm matrix for the kernel h of dimension $d \times d$ as

$$\|h\|_1 = \int_0^\infty h(s) ds.$$

We first introduce assumptions required for the existence of Hawkes processes with a rescaled stochastic and time dependent baseline.

ASSUMPTION 1. (a) For any index $i = 1, \ldots, d$, the ith component of the baseline process is a.s. positive on the time interval $[0, 1]$ a.e. , i.e.

$$\mathbb{P}\big(\nu_t^{(i)} > 0 \; \forall t \in [0, 1] \text{ a.e.}\big) = 1.$$

(b) For any index $i = 1, \ldots, d$, the ith component of the baseline process is a.s. integrable on the time interval $[0, 1]$, i.e.

$$\mathbb{P}\Big(\int_0^1 \nu_s^{(i)} ds < \infty\Big) = 1.$$

(c) We have that the point process N_t is generated by a stochastic process $\underline{N}$, which is an $\mathcal{F}_t$ adapted Poisson process of intensity 1 and dimension $2d$. Namely, we have for any index $i = 1, \dots, d$, any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$N_{tT}^{(i)} = \int_{[0,tT]\times\mathbb{R}} \mathbf{1}_{[0,\lambda_s^{(i)}]}(x) \underline{N}^{(2i-1)} * \underline{N}^{(2i)}(ds \times dx).$$

Moreover, we have that the baseline process ν_t is independent from the Poisson process $\underline{N}$. Finally, we have that the filtration is equal to $\mathcal{F}_{tT} = \mathcal{F}_t^{\nu} \vee \mathcal{F}_{tT}^{\underline{N}}$ for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$.

(d) For any index $i = 1, \dots, d$, any index $j = 1, \dots, d$ and any time $t \in \mathbb{R}^+$, the component with ith raw and jth column of the kernel is nonnegative at time t, i.e. $h^{(i,j)}(t) \geq 0$.

(e) The spectral radius of the L^1 norm matrix for the kernel is strictly less than one, i.e., $\rho(\|h\|) < 1$.

Assumption 1 (a) implies that the point process is well-defined. Assumption 1 (b) takes its roots in the simpler case of heterogeneous Poisson processes without a kernel (see Daley and Vere-Jones (2003)). Assumption 1 (c) introduces Poisson imbedding and is already required with traditional Hawkes processes (see Brémaud and Massoulié (1996), Section 3, pp. 1571-1572). In particular, the point process N_t is generated by a Poisson process $\underline{N}$. More specifically, the stochastic process N_t is defined as the point process counting the points of the Poisson process $\underline{N}$ below the intensity curve $t \to \lambda_t$. Assumption 1 (c) also considers independence between the the baseline process ν_t and the Poisson process $\underline{N}$. Moreover, we can deduce from Assumption 1 (c) that for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ the natural filtration generated by the point process N_t is included in the main filtration, i.e. $\mathcal{F}_{tT}^{N} \subset \mathcal{F}_{tT}$. Assumption 1 (d) is restrictive for kernels with inhibitory effects. Finally, Assumption 1 (e) is necessary to obtain a stationary intensity with finite first moment (see Lemma 1 (p. 495) in Hawkes and Oakes (1974) and Theorem 1 (p. 1567) in Brémaud and Massoulié (1996)).

Overall, the constraints on the kernel shape are weaker than the constraints on the kernel shape from Theorem 5.1 in Clinet and Potiron (2018). More specifically, our framework only requires the nonnegativity of the kernel whereas Clinet and Potiron (2018) considers exponential kernels, which are very restrictive for applications. However, Clinet and Potiron (2018) consider locally parametric Hawkes processes, where the baseline and the parameters of the kernels are stochastic and time dependent. See also Condition 1 in Erdemlioglu et al. (2025) for the extension to generalized gamma kernels.

In the proposition that follows, we state the existence of Hawkes processes with a rescaled stochastic and time dependent baseline. The kernel has a general form and is nonparametric. This complements Theorem 5.1 in Clinet and Potiron (2018) and Proposition 1 in Erdemlioglu et al. (2025). In particular, our nonparametric kernel framework allows for more general kernels, which is useful for applications. The arguments used in the proof slightly extends the arguments from Brémaud and Massoulié (1996) and Clinet and Potiron (2018).

PROPOSITION 1. *We assume that Assumption 1 holds. Then, there exists an $\mathcal{F}_{tT}$-adapted point process N_{tT} with an $\mathcal{F}_{tT}$-intensity of the form (7) for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$. Moreover, the intensity process λ_{tT} is a.s. integrable on the space $t \in [0,1]$.*

3. Nonparametric inference for Hawkes processes.

In this section, we develop nonparametric inference for Hawkes processes with a rescaled stochastic and time dependent baseline. The inference procedure is based on the average of the point process. As this is useful for applications, we extend the framework from Bacry et al. (2013) which is restricted to nonrandom constant baseline. We show the central limit theorem of the inference procedure in Theorem 1. This requires the assumption that the kernel does not have a too fat tail. We also need that the baseline process is Lipschitz continuous with bounded starting value. The main novelty in the proofs is to establish the renewal equation for stochastic processes. This extends Bacry et al. (2013) in which the function from the renewal equation is nonrandom.

We consider estimation for the rescaled integral of the intensity process λ_t between the starting time 0 and the final time tT for any time $t \in [0,1]$, namely

$$\Lambda_{t,T} = \frac{1}{T} \int_0^{tT} \lambda_s ds. \tag{8}$$

We have applications in management science where the target quantity (8) can be interpreted as the arrival rate in a queuing system. We also have applications in computer networks for the expected internet traffic, and seismology for the expected number of earthquakes. Finally, we have applications in finance where the intensity of a quote plays an inverse role to the volatility of an asset price.

This estimation procedure is in the sense of a stochastic process starting from the time interval $[0,1]$. In the particular case when $t = 1$, the target quantity (8) corresponds to estimation for the average of the intensity process λ_t between the starting time 0 and the final time T. We denote the limit process of $\Lambda_{t,T}$ for any time $t \in [0,1]$ as the time T increases, i.e. $T \to \infty$, by

$$\Lambda_t = (I - \|h\|_1)^{-1} \int_0^t \nu_s ds. \tag{9}$$

We propose estimation of the limit process Λ_t for any time $t \in [0,1]$ by

$$\widehat{\Lambda}_t = \frac{N_{tT}}{T}. \tag{10}$$

Then, we introduce some quantities required to establish the form of the asymptotic covariance matrix. We define w_t as the stochastic process which is a diagonal matrix of dimension $d \times d$ for any time $t \in [0,1]$. More specifically, we have that the ith diagonal component of the stochastic process w_t at the time $t \in [0,1]$ is equal to

$$w_t^{(i,i)} = \big((I - \|h\|_1)^{-1}\nu_t\big)^{(i)}.$$

Then, we define c_t as the stochastic process of dimension $d \times d$ for any $t \in [0,1]$ which satisfies

$$c_t = (I - \|h\|_1)^{-1} w_t^{1/2}.$$

We have now all the ingredients to derive the form of the asymptotic covariance matrix. We define the asymptotic covariance matrix for any time $t \in [0,1]$ as

$$\Sigma_t^2 = \int_0^t c_s c_s^T ds.$$

We deliver in what follows the assumptions used for the central limit theorem of the nonparametric inference procedure. This is based on Hawkes processes with a rescaled stochastic and time dependent baseline.

ASSUMPTION 2. (a) The kernel satisfies $\int_0^\infty t h(t) dt < \infty$.
(b) The starting point of the baseline is a.s. bounded, i.e. there is a nonrandom constant satisfying $C_0 \geq 0$ and

$$\mathbb{P}\big(\nu_0 \leq C_0\big) = 1.$$

(c) The baseline is a.s. Lipschitz-continuous with nonrandom constant satisfying $C > 0$ on the time interval $[0,1]$, i.e.

$$\mathbb{P}\big(|\nu_t - \nu_s| \leq C|t-s| \forall (t,s) \in [0,1]^2\big) = 1.$$

Assumption 2 (a) puts some restrictions on the kernel shape $h(t)$. This corresponds exactly to Assumption (A2) in Bacry et al. (2013) (p. 2480). This is required to obtain a martingale form of the intensity process. Assumption 2 (a) is also used in Jaisson and Rosenbaum (2015). Assumptions 2 (b) and (c) put restrictions on the baseline process ν_t. More specifically, Assumption 2 (b) relies on an starting point which is bounded a.s. and Assumption 2 (c) considers a baseline process ν_t which is Lipschitz-continuous a.s. Assumptions 2 (b) and (c) are mainly used to obtain a locally bounded baseline process ν_t. This is needed to establish the renewal equation. Finally, Assumption 2 (c) is also used in the proofs based on local estimation.

Overall, the assumptions on the kernel shape from this paper are exactly the same as the assumptions used for Corollary 1 in Bacry et al. (2013). Moreover, the assumptions on the baseline process are novel to the literature. This allows the baseline to be random and time dependent, which is useful for applications.

In the theorem that follows, we state the central limit theorem of the nonparametric inference procedure. This is based on Hawkes processes with a rescaled stochastic and time dependent baseline. The kernel has a general form and is nonparametric. The inference procedure is based on the average of the point process. We consider asymptotics when the final time diverges to infinity, i.e. $T \to +\infty$. This is the main result of this paper. This extends Corollary 1 in Bacry et al. (2013) which is restricted to nonrandom constant baseline. The main novelty in the proofs is to establish the renewal equation for stochastic processes. This extends Bacry et al. (2013) in which the function from the renewal equation is nonrandom. Moreover, the convergence rate is $\sqrt{T}$. Finally, we denote by $\overset{\mathcal{D}-s}{\longrightarrow}$ the $\mathcal{F}_t$-stable weak convergence for the Skorokhod space $\mathbb{D}([0,1],\mathbb{R}^d)$ equipped with its topology.

THEOREM 1. *We assume that Assumptions 1 and 2 hold. Then, there is a canonical d-dimensional standard Brownian extension of the stochastic basis $\mathcal{B}$. This extension includes the canonical standard Brownian motion W_t which satisfies as $T \to \infty$ that*

$$\sqrt{T}\big(\widehat{\Lambda}_t - \Lambda_t\big) \overset{\mathcal{D}-s}{\longrightarrow} \int_0^t c_s dW_s. \tag{11}$$

4. Proofs. We begin this section with some general guidelines that we use extensively during the proofs. First, we use C for any generic constant, and the value of the constant can change from one line to the next. In addition, any operation with two vectors of the same dimension means the operation component by component.

4.1. *Proof of existence.* In this part, we focus on the proof of the existence of Hawkes processes with a rescaled stochastic and time dependent baseline. This corresponds to the proof of Proposition 1.

Prior to the first lemma, we introduce some notation. First, we define the point process N_{tT} conditioned by the information from the baseline process for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ as

$$N_{tT,\nu} = \mathbb{E}\big[N_{tT}\big|\mathcal{F}_1^\nu\big]. \tag{12}$$

We also define the intensity process λ_{tT} conditioned by the information from the baseline process for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ as

$$\lambda_{tT,\nu} = \mathbb{E}\big[\lambda_{tT}\big|\mathcal{F}_1^\nu\big]. \tag{13}$$

Finally, we define the filtration $\mathcal{F}_{tT}$ conditioned by the information from the baseline process for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ as

$$\mathcal{F}_{tT,\nu} = \mathbb{E}\big[\mathcal{F}_{tT}\big|\mathcal{F}_1^{\nu}\big]. \tag{14}$$

This first lemma shows that the stochastic process $\lambda_{tT,\nu}$ is the $\mathcal{F}_{tT,\nu}$ intensity of the point process $N_{tT,\nu}$ in the sense of Definition 1. This extends Lemma 10.1 (p. 2) from the supplementary materials of Clinet and Potiron (2018).

LEMMA 1. *We assume that Assumption 1 (c) hold. Then, the stochastic process $\lambda_{tT,\nu}$ is the $\mathcal{F}_{tT,\nu}$ intensity of the point process $N_{tT,\nu}$ in the sense of Definition 1.*

PROOF OF LEMMA 1. To prove the lemma, we verify that the properties from Definition 1 are satisfied. First, we have for any interval $(a,b] \subset [0,T]$, any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ by Definitions (12) and (14) that

$$\mathbb{E}\big[N_{tT,\nu}((a,b])\big|\mathcal{F}_{a,\nu}\big] = \mathbb{E}\Big[\mathbb{E}\big[N((a,b])\big|\mathcal{F}_1^{\nu}\big]\Big|\mathbb{E}\big[\mathcal{F}_a\big|\mathcal{F}_1^{\nu}\big]\Big]. \tag{15}$$

Then, we can rewrite the right side of Equation (15) by conditional expectation properties for any interval $(a,b] \subset [0,T]$, any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ as

$$\mathbb{E}\Big[\mathbb{E}\big[N((a,b])\big|\mathcal{F}_1^{\nu}\big]\Big|\mathbb{E}\big[\mathcal{F}_a\big|\mathcal{F}_1^{\nu}\big]\Big] = \mathbb{E}\Big[\mathbb{E}\big[N((a,b])\big|\mathcal{F}_a\big]\Big|\mathcal{F}_1^{\nu}\Big]. \tag{16}$$

In addition, we obtain by Equation (4) from Definition 1 for any interval $(a,b] \subset [0,T]$, any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\mathbb{E}\Big[\mathbb{E}\big[N((a,b])\big|\mathcal{F}_a\big]\Big|\mathcal{F}_1^{\nu}\Big] = \mathbb{E}\Big[\mathbb{E}\Big[\int_a^b \lambda_s ds\Big|\mathcal{F}_a\Big]\Big|\mathcal{F}_1^{\nu}\Big] \text{ a.s.} \tag{17}$$

Moreover, we get by conditional expectation properties for any interval $(a,b] \subset [0,T]$, any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\mathbb{E}\Big[\mathbb{E}\Big[\int_a^b \lambda_s ds\Big|\mathcal{F}_a\Big]\Big|\mathcal{F}_1^{\nu}\Big] = \mathbb{E}\Big[\mathbb{E}\Big[\int_a^b \lambda_s ds\Big|\mathcal{F}_1^{\nu}\Big]\Big|\mathbb{E}\big[\mathcal{F}_a\big|\mathcal{F}_1^{\nu}\big]\Big]. \tag{18}$$

Finally, we deduce by Tonelli's theorem, Definitions (13) and (14) for any interval $(a,b] \subset [0,T]$, any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\mathbb{E}\Big[\mathbb{E}\Big[\int_a^b \lambda_s ds\Big|\mathcal{F}_1^{\nu}\Big]\Big|\mathbb{E}\big[\mathcal{F}_a\big|\mathcal{F}_1^{\nu}\big]\Big] = \mathbb{E}\Big[\int_a^b \lambda_{s,\nu} ds\Big|\mathcal{F}_{a,\nu}\Big]. \tag{19}$$

Thus, Equations (15), (16), (17), (18) and (19) for any interval $(a,b] \subset [0,T]$, any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ yield

$$\mathbb{E}\big[N_{tT,\nu}((a,b])\big|\mathcal{F}_{a,\nu}\big] = \mathbb{E}\Big[\int_a^b \lambda_{s,\nu} ds\Big|\mathcal{F}_{a,\nu}\Big] \text{ a.s.} \tag{20}$$

This means that we have shown Equation (4). Second, the process $\lambda_{tT,\nu}$ is $\mathcal{F}_{tT,\nu}$-progressively measurable, of dimension d where each component $\lambda_{t,\nu}^{(i)}$ takes its values in the space of non-negative real numbers $\mathbb{R}^+$. Thus, we have shown Definition 1. □

We now give the proof of the existence of Hawkes processes with a rescaled stochastic and time dependent baseline. It extends the proof of Theorem 7 (pp. 1585-1587) in Brémaud and Massoulié (1996). It complements the proof of Theorem 5.1 (pp. 3-4) in the supplement of Clinet and Potiron (2018) and the proof of Proposition 1 in Erdemlioglu et al. (2025).

PROOF OF PROPOSITION 1. The strategy of the proof consists in defining a suitable sequence of point processes and intensity $(N_{tT,k}, \lambda_{tT,k})_{k\in\mathbb{N}}$ for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$. Then, we show that their limit defined as $(N_{tT}, \lambda_{tT}) = \lim_{k\to\infty}(N_{tT,k}, \lambda_{tT,k})$ exists and that the point process N_{tT} admits λ_{tT} as an $\mathcal{F}_{tT}$-intensity given by Equation (7).

We first introduce for any time $t \in [0,1]$, any final time $T \in \mathbb{R}^+$ and any index $i = 1,\dots,d$ the ith component of the initial intensity process

$$\lambda^{(i)}_{tT,0} = \nu^{(i)}_t.$$

We also introduce the ith component of the initial point process $N^{(i)}_{tT,0}$ for any time $t \in [0,1]$, any final time $T \in \mathbb{R}^+$ and any index $i = 1,\dots,d$. It is defined as the point process counting the points of the Poisson process $\underline{N}^{(2i-1)} * \underline{N}^{(2i)}$ below the curve $t \to \lambda^{(i)}_{tT,0}$, namely

$$N^{(i)}_{tT,0} = \int_{[0,tT]\times\mathbb{R}} \mathbf{1}_{[0,\lambda^{(i)}_{s,0}]}(x)\underline{N}^{(2i-1)} * \underline{N}^{(2i)}(ds \times dx).$$

We then define recursively the sequence of point process and its intensity $(N^{(i)}_{tT,k}, \lambda^{(i)}_{tT,k})$ for any time $t \in [0,1]$, any final time $T \in \mathbb{R}^+$ and any index as

$$\lambda_{tT,k+1} = \nu_t + \int_0^{tT} h(tT-s)dN_{s,k}, \tag{21}$$

$$N^{(i)}_{tT,k+1} = \int_{[0,tT]\times\mathbb{R}} \mathbf{1}_{[0,\lambda^{(i)}_{s,k+1}]}(x)\underline{N}^{(2i-1)} * \underline{N}^{(2i)}(ds \times dx) \text{ for any } i = \dots, d.$$

First, we have that the stochastic process $\lambda^{(i)}_{tT,k}$ is positive for any time $t \in [0,1]$, any final time $T \in \mathbb{R}^+$ and any index $i = 1,\dots,d$ a.s. by Assumptions 1 (a) and (d). Thus, the stochastic process $\lambda_{tT,k}$ is a well-defined intensity. Then, an extension to the stochastic time dependent baseline case of the arguments from Lemma 3 and Example 4 (pp. 1571-1572) in Brémaud and Massoulié (1996) yields that the point process $N_{tT,k}$ is $\mathcal{F}_{tT}$-adapted. It also gives that the stochastic process $\lambda_{tT,k}$ is $\mathcal{F}_{tT}$-predictable and an $\mathcal{F}_{tT}$-intensity of $N_{tT,k}$ in the sense of Definition 1. Moreover, Assumption 1 (d) implies that $(N^{(i)}_{tT,k}, \lambda^{(i)}_{tT,k})$ is componentwise increasing with k and thus converges to some limit $(N^{(i)}_{tT}, \lambda^{(i)}_{tT})$ a.s. for any time $t \in [0,1]$, any final time $T \in \mathbb{R}^+$ and any index $i = 1,\dots,d$.

We now introduce the sequence of vector processes $\rho_{tT,k}$ for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ defined as

$$\rho_{tT,k} = \mathbb{E}[\lambda_{tT,k} - \lambda_{tT,k-1}|\mathcal{F}^\nu_1]. \tag{22}$$

First, we get by its definition (22) for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\rho_{tT,k+1} = \mathbb{E}[\lambda_{tT,k+1} - \lambda_{tT,k}|\mathcal{F}^\nu_1]. \tag{23}$$

Then, we obtain by inserting the intensity definition (7) into Equation (23) for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\rho_{tT,k+1} = \mathbb{E}\Big[\int_0^{tT} h(tT-s)(dN_{s,k+1} - dN_{s,k})\Big|\mathcal{F}^\nu_1\Big]. \tag{24}$$

Moreover, we get by an application of Lemma 1 with point process properties for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\rho_{tT,k+1} = \mathbb{E}\Big[\int_0^{tT} h(tT-s)(\lambda_{s,k+1} - \lambda_{s,k})ds\Big|\mathcal{F}^\nu_1\Big]. \tag{25}$$

In addition, we can deduce by Tonelli's theorem and Definition (22) that

$$\rho_{tT,k+1} = \int_0^{tT} h(tT-s)\rho_{s,k}ds. \tag{26}$$

We define $\Phi_{tT,k}$ as the integral of the stochastic process $\rho_{s,k}$ from the starting time 0 to the final time tT for any time $t \in [0,1]$ and any time $T \in \mathbb{R}^+$, namely $\Phi_{tT,k} = \int_0^{tT} \rho_s^k ds$. Then, we have by another application of Tonelli's theorem that for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ a.s. that

$$\Phi_{tT,k+1} = \int_0^{tT} \Big(\int_0^{tT-s} h(u)du \Big) \rho_{s,k} ds. \tag{27}$$

Then, Assumption 1 (e) implies that $|\Phi_{tT,k+1}| \leq r|\Phi_{tT,k}|$ a.s. in which $r = \rho(\|h\|)$. Thus, we can deduce that $G : \Phi_{tT,k} \to \Phi_{tT,k+1}$ is a.s. a contraction function. It turns out that the limit of the telescopic series $(\sum_{l=0}^k \Phi_{tT,l})_{k\in\mathbb{N}}$ exists by arguments used in Banach fixed-point theorem. Working with the telescopic series and applying the monotone convergence theorem to the series yield for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\mathbb{E}\Big[\int_0^{tT} \lambda_s ds \Big| \mathcal{F}_1^\nu \Big] \leq \int_0^t \nu_s ds + r\mathbb{E}\Big[\int_0^{tT} \lambda_s ds \Big| \mathcal{F}_1^\nu \Big]. \tag{28}$$

By rearranging the terms in Expression (28), we get for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\mathbb{E}\Big[\int_0^{tT} \lambda_s ds \Big| \mathcal{F}_1^\nu \Big] \leq (1-r)^{-1} \int_0^t \nu_s ds. \tag{29}$$

Given Condition 1 (b), the expression in the right side of Expression (29) is finite a.s. As its conditional expectation is finite, we can deduce that $\int_0^{tT} \lambda_s ds$ is finite a.s. for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$. Moreover, the stochastic process λ_{tT} is $\mathcal{F}_{tT}$-predictable as a limit of such processes. The point process $N_{tT}^{(i)}$ counts the points of the Poisson process $\underline{N}^{(2i-1)} * \underline{N}^{(2i)}$ under the curve $t \mapsto \lambda_{tT}^{(i)}$ by an application of the monotone convergence theorem. Thus, the point process N_{tT} admits the stochastic process λ_{tT} as an $\mathcal{F}_{tT}$-intensity in the sense of Definition 1 by an extension to the stochastic time dependent baseline case of the arguments from Lemma 3 (p. 1571) in Brémaud and Massoulié (1996). It implies that the point process N_t is finite a.s. for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$.

Finally, it remains to show that the intensity process λ_{tT} is of the form (7) for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$. The monotonicity properties of the point process $N_{tT,k}^{(i)}$ and the intensity process $\lambda_{tT,k}^{(i)}$ ensure for any index $k \in \mathbb{N}$, any time $t \in [0,1]$, any final time $T \in \mathbb{R}^+$ and any index $i = 1, \ldots, d$ that

$$\lambda_{tT,k}^{(i)} \leq \nu_t^{(i)} + \Big(\int_0^{tT} h(tT-s)dN_s \Big)^{(i)},$$

$$\lambda_{tT}^{(i)} \geq \nu_t^{(i)} + \Big(\int_0^{tT} h(tT-s)dN_{s,k} \Big)^{(i)}. \tag{30}$$

This gives Equation (7) by taking the limit $k \to +\infty$ in both inequalities. □

4.2. *Proof of the central limit theorem.* In this part, we focus on the proof of the central limit theorem to estimate the rescaled integral of the intensity process based on Hawkes processes. This corresponds to the proof of Theorem 1.

We start with the discretization in time of the statistical problem. For any final time $T \in \mathbb{R}^+$, we consider $M = \lfloor 1/\Delta \rfloor$ intervals with equal length $\Delta = 1/T$ such that $\bigcup_{i=1}^{M}[(i-1)\Delta, i\Delta) \subset [0,1)$. Here, $\lfloor \cdot \rfloor$ denotes the floor function. For any index $i = 1, \dots, M$, we define an estimator for local Poisson estimates on the ith interval $[(i-1)\Delta, i\Delta)$ as

$$\widehat{\lambda}_i = \frac{1}{\Delta}(N_{i\Delta-} - N_{(i-1)\Delta}). \tag{31}$$

Before giving the first lemma in the proof of Theorem 1, we introduce some definition. First, we define the Laplace transform of the kernel h of dimension $d \times d$ at the frequency $s \in \mathbb{R}^+$ as

$$\widehat{h}(s) = \int_0^\infty e^{-st} h(t) dt. \tag{32}$$

Then, we define the convolution of f and g for f and g two integrable functions defined on the space of positive real numbers $\mathbb{R}^+$ at time $t \in \mathbb{R}^+$ as

$$f * g_t = \int_0^t f(t-s) g(s)\, ds. \tag{33}$$

Moreover, we define recursively f^{*k} for $k \in \mathbb{N}$ as $f^{*1} = f$ and f^{*k} is the convolution product of $f^{*(k-1)}$ and the function f for $k \geq 2$. Similarly, we define the convolution of f and X for an integrable function f and a stochastic process X defined on the space of positive real numbers $\mathbb{R}^+$ at time $t \in \mathbb{R}^+$ as

$$f * dX_t = \int_0^t f(t-s)\, dX_s. \tag{34}$$

Furthermore, we denote by $\psi \colon \mathbb{R}^+ \to \mathbb{R}_d^+$ the resolvent kernel at the time $t \in \mathbb{R}^+$ of the kernel h which satisfies

$$\psi(t) = h(t) + h * \psi_t. \tag{35}$$

Then, we define the integral of the resolvent kernel ψ from the starting time 0 to the final time t as

$$\Psi(t) = \int_0^t \psi(s) ds. \tag{36}$$

We also define the integral of $\psi(s - tT)$ between the starting time $(i-1)\Delta$ and the final time $i\Delta$ for any time $t \in [0,1]$, any final time $T \in \mathbb{R}^+$ and any index $i = 1, \dots, M$ as

$$\Delta_i \Psi(-tT) = \int_{(i-1)\Delta}^{i\Delta} \psi(s - tT)\, ds. \tag{37}$$

Moreover, we denote the uniform big O by $\underline{O}$. It is defined through

$$f(tT) = \underline{O}\big(g(tT)\big) \iff |f(tT)| \leq C g(tT)$$

for any time $t \in [0,1]$, any final time $T \in \mathbb{R}^+$ and some constant $C \in \mathbb{R}_+$ which does not depend on the final time T and the time t. Finally, we introduce $a \wedge b$ which is the minimum between two real numbers $a \in \mathbb{R}$ and $b \in \mathbb{R}$.

The first lemma gives the asymptotic properties of the resolvent kernel, which can be expressed as a Laplace transform of the kernel.

LEMMA 2. *We assume that Assumption 1 holds. Then, we have for any time* $t \in [0,1]$ *and any final time* $T \in \mathbb{R}^+$ *that*

$$\psi(tT) \geq 0. \tag{38}$$

We also assume that Assumption 2 (a) holds. Then, we have that

$$\Psi(tT) = \widehat{h}(0) + \underline{O}\left(1 \wedge \frac{1}{tT}\right). \tag{39}$$

Moreover, we have for any index $i = 1, \ldots, M$ *that*

$$\Delta_i \Psi(-tT) = \underline{O}\left(1 \wedge \frac{1}{((i-1)\Delta - tT)}\right). \tag{40}$$

PROOF OF LEMMA 2. Since $\rho(\|h\|_1) < 1$ by Assumption 1 (e), the function

$$\theta \colon f(tT) \mapsto (\nu_t + h * f_{tT})$$

is a contraction function for any final time $T \in \mathbb{R}^+$. Thus, we can apply Banach fixed-point theorem to get a fixed-point $\psi = f_\infty$ with recursion $f_k = \theta(f_{k-1})$. Then, we obtain recursively by the definitions of the function θ and the function f_k that

$$\begin{aligned} f_k(tT) = \theta(f_{k-1}(tT)) &= \nu_t + h * (f_{k-1})_{tT} = \nu_t + h * \theta(f_{k-2})_{tT} \\ &= \nu_t + h * (\nu_t + h * (f_{k-2})_{tT})_{tT} \\ &= \nu_t + \sum_{l=1}^{k-2} h^{*l} * \nu_t + h^{*(k-1)} * (f_1)_{tT}. \end{aligned}$$

For the initial value, we can choose $f_1 = 0$. Then, $f_k(tT)$ is nonnegative for any $k \in \mathbb{N}$ such that $k > 1$, any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$. This is due to the fact that the baseline is a.s. positive on the time interval $[0,1]$ a.e. by Assumption 1 (a) and the kernel h is nonnegative by Assumption 1 (d). Thus, we have shown Expression (38).

To show Equation (39), we first have for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\begin{aligned} \widehat{h}(0) - \widehat{h}\left(\frac{1}{tT}\right) &= \int_0^\infty \left(1 - e^{-\frac{s}{tT}}\right) h(s) ds \\ &\geq \int_{tT}^\infty \left(1 - e^{-\frac{s}{tT}}\right) h(s) ds \\ &\geq (1 - e^{-1}) \int_{tT}^\infty \psi(s) ds. \end{aligned} \tag{41}$$

Here, we use the definition of the Laplace transform in the equality, Assumption 1 (d) in the first inequality, Definition (35) and Expression (38) in the last inequality.

From Assumption 2 (a), we obtain that the Laplace transform of the kernel $\widehat{h}$ is continuously differentiable at the time 0 and that its derivative is equal to

$$\widehat{h}'(0) = \int_0^\infty t h(t) dt < \infty.$$

Then, we can apply the mean value theorem for the function $\widehat{\psi}(s) = \widehat{h}(s)/(1 - \widehat{h}(s))$ and we obtain that

$$\widehat{\psi}(s) = \widehat{\psi}(0) + \widehat{\psi}'(0)s + r(s)s. \tag{42}$$

Here, $r(s)$ is the remainder which satisfies $\lim_{s \to 0} r(s) = 0$.

Then, we can deduce for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$
\begin{aligned}
0 &\le \widehat{\psi}(0) - \widehat{\psi}\Big(\frac{1}{tT}\Big) \\
&= -\Big(\widehat{\psi}'(0) + r\Big(\frac{1}{tT}\Big)\Big)\frac{1}{tT} \\
&\le \big(|\widehat{\psi}'(0)| + \big|r\big(\frac{1}{tT}\big)\big|\big)\frac{1}{tT}\mathbf{1}_{\{tT\ge 1\}} + \widehat{\psi}(0)\mathbf{1}_{\{tT<1\}}.
\end{aligned}
$$

Here, we use the fact that the function $\widehat{\psi}$ is decreasing in the first inequality, Equation (42) in the equality and the definition of $|\ |$ in the last inequality. As

$$\sup_{x\in[0,1]} |r(x)| < \infty,$$

we obtain

$$\Big|\widehat{\psi}(0) - \widehat{\psi}\Big(\frac{1}{tT}\Big)\Big| \le C\Big(1 \wedge \frac{1}{tT}\Big). \tag{43}$$

Finally, we get for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$
\begin{aligned}
\Psi(tT) &= \int_0^{tT} \psi(s)\,ds \\
&= \widehat{\psi}(0) - \int_{tT}^{\infty} \psi(s)\,ds \\
&\le \widehat{\psi}(0) + \frac{1}{1-e^{-1}}\Big(\widehat{\psi}(0) - \widehat{\psi}\Big(\frac{1}{tT}\Big)\Big) \\
&\le \widehat{\psi}(0) + C\Big(1 \wedge \frac{1}{tT}\Big).
\end{aligned}
$$

Here, we use Definition (36) in the first equality, the definition of $\widehat{\psi}(0)$ in the second equality, Expression (41) in the first inequality and Expression (43) in the last inequality. Thus, we have proven Equation (39). With the same arguments, we can also show that Equation (40) holds. □

We next state the renewal equation for stochastic processes in the following lemma. This extends Lemma 3 from Bacry et al. (2013) in which the function from the renewal equation is nonrandom. This is required as the baseline ν_t is stochastic and time dependent.

LEMMA 3. *We assume that Assumption 1 holds. We introduce an $\mathcal{F}_t$ adapted stochastic process $g : \mathbb{R}^+ \times \Omega \to \mathbb{R}_d$ which is locally bounded a.s. Then, there exists a stochastic process $f : \mathbb{R}^+ \times \Omega \to \mathbb{R}_d$ which is locally bounded a.s. and solution to the renewal equation for any time $t \in \mathbb{R}^+$ and a.s.*

$$f(t) = g(t) + h * f_t. \tag{44}$$

The solution is given for any time $t \in \mathbb{R}^+$ and a.s. by

$$f_g(t) = g(t) + \psi * g_t. \tag{45}$$

Moreover, the solution $f_g(t)$ is unique in the a.s. and a.e. sense. Namely, we have

$$\mathbb{P}\big(f_g(t) = f(t) \text{ for any } t \in \mathbb{R}^+ \text{ a.e.}\big) = 1 \tag{46}$$

for any stochastic process f satisfying the renewal equation (44).

PROOF OF LEMMA 3. First, we have that the resolvent kernel of the kernel ψ is integrable by Assumption 1 (e) and resolvent kernel properties. We also have that the stochastic process g is locally bounded a.s. by assumption of the lemma. Thus, we can deduce that the stochastic process f_g defined in (45) is locally bounded a.s.

Moreover, we show in what follows that f_g satisfies the renewal equation (44). It is sufficient to prove for any time $t \in \mathbb{R}^+$ and a.s. that

$$h * (f_g)_t = \psi * g_t \tag{47}$$

First, we get by Definition (45) for any time $t \in \mathbb{R}^+$ and a.s. that

$$h * (f_g)_t = h * (g(t) + \psi * g_t)_t. \tag{48}$$

Moreover, we obtain by the definition of the resolvent kernel (35) for any time $t \in \mathbb{R}^+$ that

$$h(t) = \psi(t) - h * \psi_t. \tag{49}$$

Finally, Equations (48) and (49) yield Equation (47).

We show now that the solution $f_g(t)$ is unique in the a.s. and a.e. sense. Namely, we show Equation (46) for any stochastic process $f(t)$ satisfying the renewal equation (44). First, we get as both processes $f_g(t)$ and $f(t)$ satisfy the renewal equation (44) for any time $t \in \mathbb{R}^+$ and a.s. that

$$f(t) - f_g(t) = h * (f - f_g)_t. \tag{50}$$

We introduce the stochastic process v of dimension d such that its ith component for any index $i = 1, \dots, d$ is equal to

$$v^{(i)}(t) = |f^{(i)}(t) - f_g^{(i)}(t)|.$$

Then, we can deduce from Equation (50) for any time $t \in \mathbb{R}^+$ and a.s. that

$$v(t) = h * v_t. \tag{51}$$

This yields for any time $t \in \mathbb{R}^+$ and a.s. that

$$\int_0^\infty v(t)dt \leq ||h||_1 \int_0^\infty v(t)dt. \tag{52}$$

Finally, we get by Assumption 1 (d) for any time $t \in \mathbb{R}^+$ and a.s. that

$$\int_0^\infty v(t)dt < \int_0^\infty v(t)dt. \tag{53}$$

Thus, we have shown Equation (46) for any stochastic process $f(t)$ satisfying the renewal equation (44).

□

Before introducing the next lemma, we give some definition. First, we denote the sum of the baseline and the convolution of the resolvent kernel and the baseline for any time $t \in [0,1]$ by

$$\mu_t = \nu_t + (\psi_{Tt}) * \nu_t. \tag{54}$$

We have that the stochastic process μ_t is $\mathcal{F}^\nu$ adapted by Definition (34). Moreover, we define the limit of the stochastic process μ_t for any time $t \in [0,1]$ as

$$\mu_{t,L} = \big((I - \|h\|_1)^{-1} \mathbf{1}_{\{t \in (0,1]\}} + \mathbf{1}_{\{t=0\}}\big)\nu_t \tag{55}$$

Finally, we define the notation small tau in probability as $Y_T = o_{\mathbb{P}}(Z_T)$, which means that $\frac{Y_T}{Z_T}\mathbf{1}_{\{Z_T \neq 0\}} \xrightarrow{\mathbb{P}} 0$ as $T \to \infty$ for Y_T and Z_T which are random variables.

The following lemma exhibits a martingale representation of the intensity λ_{tT}. It is based on the convolution of the resolvent kernel and the martingale M_{tT} defined in (5). The proof is based on the application of the renewal equation obtained in Lemma 3. It extends Lemma 4 in Bacry et al. (2013) and Proposition 2.1 (p. 606) in Jaisson and Rosenbaum (2015), who consider a constant nonrandom baseline.

LEMMA 4. *We assume that Assumptions 1 and 2 hold. Then, we have for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that the intensity λ_{tT} has the martingale representation*

$$\lambda_{tT} = \mu_t + \psi * dM_{tT}. \tag{56}$$

We also have for any time $t \in [0,1]$ when the final time $T \to \infty$ that

$$\sup_{T^{-\frac{1}{2}} \leq t \leq 1} \sqrt{T}\big(\mu_t - \mu_{t,L}\big) \xrightarrow{\mathbb{P}} 0. \tag{57}$$

PROOF OF LEMMA 4. We first reexpress the intensity process as the renewal equation. We have for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\begin{aligned}\lambda_{tT} &= \nu_t + h * dN_{tT} \\ &= \nu_t + h * (\lambda_{tT} + dM_{tT}) \\ &= \big(\nu_t + h * dM_{tT}\big) + h * \lambda_{tT}.\end{aligned} \tag{58}$$

Here, we use Definition (7) and Definition (34) in the first equality, Definition (5) in the second equality and algebraic manipulation in the third equality.

Thus, the intensity λ_{tT} is solution to the renewal equation (44) a.s. by Equation (58). More specifically, we consider the stochastic processes $g(tT)$ defined as

$$g(tT) = \nu_t + h * dM_{tT}$$

for any time $t \in [0,1]$ and a fixed $T \in \mathbb{R}^+$. To obtain the form of the intensity λ_{tT}, we apply Lemma 3. First, we have that $g(tT)$ are $\mathcal{F}_{tT}$ adapted stochastic processes which are locally bounded a.s. by Definition 1, Assumptions 2 (b) and 2 (c). Then, we have for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\begin{aligned}\lambda_{tT} &= g(tT) + \psi * g_{tT} \\ &= \nu_t + h * dM_{tT} + (\psi_{Tt}) * (\nu_t + h * dM_{tT}) \\ &= (\nu_t + (\psi_{Tt}) * \nu_t) + (h + \psi * h_{tT}) * dM_{tT}.\end{aligned}$$

Here, we use Equation (45) from Lemma 3 in the first equality, the definition of the stochastic process $g(tT)$ in the second equality, algebraic manipulation in the third equality. Then, we have for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\begin{aligned}\lambda_{tT} &= (\nu_t + (\psi_{Tt}) * \nu_t) + (h + \psi * h_{tT}) * dM_{tT} \\ &= (\nu_t + (\psi_{Tt}) * \nu_t) + \psi * dM_{tT}. \\ &= \mu_t + \psi * dM_{tT}.\end{aligned}$$

Here, we use Definition (35) in the second equality and Definition (54) in the third equality. Thus, we can obtain Equation (56).

We show now Equation (57). First, we get for any time $t \in (0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\begin{aligned}\mu_t - \mu_{t,L} &= \nu_t + (\psi_{Tt}) * \nu_t - (I - \|h\|_1)^{-1}\nu_t \\ &= \nu_t + (\psi_{Tt}) * \nu_t - (1 + \widehat{\psi}(0))\nu_t \\ &= (\psi_{Tt}) * \nu_t - \widehat{\psi}(0)\nu_t \\ &= \int_0^{tT} \psi(s)\nu_{t-\frac{s}{T}} ds - \widehat{\psi}(0)\nu_t.\end{aligned}$$

Here, we use Definition (54) and Definition (55) in the first equality, Definition (35) in the second equality, algebraic manipulation in the third equality and Definition (33) in the fourth equality. Then, we have for any time $t \in (0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$\begin{aligned}\mu_t - \mu_{t,L} &= \int_0^{tT} \psi(s)\nu_{t-\frac{s}{T}} ds - \widehat{\psi}(0)\nu_t \\ &= \int_0^{tT} \psi(s)\nu_{t-\frac{s}{T}} ds - \int_0^{\infty} \psi(s) ds \nu_t \\ &= \int_0^{tT} \psi(s)\big(\nu_{t-\frac{s}{T}} - \nu_t\big) ds - \int_{tT}^{\infty} \psi(s) ds. \end{aligned} \tag{59}$$

Here, we use Definition (32) in the second equality and algebraic manipulation in the third equality. Moreover, we obtain that

$$\begin{aligned}\sup_{T^{-\frac{1}{2}} \leq t \leq 1} \sqrt{T}\big|\mu_t - \mu_{t,L}\big| &= \sup_{T^{-\frac{1}{2}} \leq t \leq 1} \sqrt{T}\Big|\int_0^{tT} \psi(s)(\nu_{t-\frac{s}{T}} - \nu_t) ds - \int_{tT}^{\infty} \psi(s) ds\Big| \\ &\leq \sup_{T^{-\frac{1}{2}} \leq t \leq 1} \sqrt{T}\Big(\Big|\int_0^{tT} \psi(s)(\nu_{t-\frac{s}{T}} - \nu_t) ds\Big| + \Big|\int_{tT}^{\infty} \psi(s) ds\Big|\Big) \\ &\leq \sup_{T^{-\frac{1}{2}} \leq t \leq 1} \sqrt{T}\Big|\int_0^{tT} \psi(s)(\nu_{t-\frac{s}{T}} - \nu_t) ds\Big| \\ &\quad + \sup_{T^{-\frac{1}{2}} \leq t \leq 1} \sqrt{T}\Big|\int_{tT}^{\infty} \psi(s) ds\Big|. \end{aligned} \tag{60}$$

Here, we use Equation (59) in the first equality, the triangular inequality in the first inequality and supremum properties in the second inequality.

We introduce

$$I = \sup_{T^{-\frac{1}{2}} \leq t \leq 1} \sqrt{T}\Big|\int_0^{tT} \psi(s)(\nu_{t-\frac{s}{T}} - \nu_t) ds\Big|.$$

First, we can deduce from integral and norm properties that

$$I \leq \sup_{T^{-\frac{1}{2}} \leq t \leq 1} \sqrt{T} \int_0^{tT} \psi(s)\big|\nu_{t-\frac{s}{T}} - \nu_t\big| ds.$$

Then, we get from Assumption 2 (c) a.s. that

$$I \leq \sup_{T^{-\frac{1}{2}} \leq t \leq 1} \sqrt{T} \int_0^{tT} \psi(s) \frac{Cs}{T} ds.$$

We also obtain from algebraic manipulation a.s. that

$$I \leq \sup_{T^{-\frac{1}{2}} \leq t \leq 1} \frac{C}{\sqrt{T}} \int_0^{tT} \psi(s) s ds.$$

Moreover, supremum properties yield a.s. that

$$I \leq \frac{C}{\sqrt{T}} \int_0^{\infty} \psi(s) s ds.$$

Finally, we obtain by Assumption 2 (a) and Definition (35) when the final time $T \to \infty$ that

$$I \xrightarrow{\mathbb{P}} 0. \tag{61}$$

We introduce now

$$II = \sup_{T^{-\frac{1}{2}} \leq t \leq 1} \sqrt{T} \Big| \int_{tT}^{\infty} \psi(s) ds \Big|.$$

First, we can deduce from supremum properties and Expression (38) from Lemma 2 that

$$II = \sqrt{T} \Big| \int_{\sqrt{T}}^{\infty} \psi(s) ds \Big|.$$

Then, we obtain by algebraic manipulation that

$$II = \Big| \int_{\sqrt{T}}^{\infty} \sqrt{T} \psi(s) ds \Big|.$$

In addition, we get by supremum properties that

$$II \leq \Big| \int_{\sqrt{T}}^{\infty} s \psi(s) ds \Big|.$$

Finally, we obtain by Assumption 2 (a) and Definition (35) when the final time $T \to \infty$ that

$$II \xrightarrow{\mathbb{P}} 0. \tag{62}$$

Thus, Expressions (60), (61) and (62) yield Equation (57).

□

We introduce some notation prior to the next lemma. First, we denote by X_i the process X evaluated at the end of the ith interval for any interval number $i = 1, \dots, M$, i.e. $X_i = X_{i\Delta}$. We also define $\bar{X}_i$ as the average of X_t on the ith interval, i.e.

$$\bar{X}_i = \Delta^{-1} \int_{(i-1)\Delta}^{i\Delta} X_t \, dt. \tag{63}$$

In addition, we define the increment of the martingale M_t on the ith interval as

$$\varepsilon_i = \frac{1}{\Delta} \int_{(i-1)\Delta}^{i\Delta} dM_t. \tag{64}$$

Moreover, we denote the increment related to the Hawkes component on the ith interval by

$$\epsilon_i = \frac{1}{\Delta} \Big(\int_0^{(i-1)\Delta} \Delta_i \Psi(-t) \, dM_t + \int_{(i-1)\Delta}^{i\Delta} \Psi(i\Delta - t) \, dM_t \Big). \tag{65}$$

Finally, we define the sum of ε_i and ϵ_i as u_i, namely

$$u_i = \varepsilon_i + \epsilon_i. \tag{66}$$

The following lemma is a decomposition of the estimation error u_i as the sum of the error originating from the martingale ε_i and another related to the Hawkes component ϵ_i.

LEMMA 5. *We assume that Assumptions 1 and 2 hold. Then, we have for any final time $T \in \mathbb{R}^+$ and any interval index $i = 1, \dots, M$ the decomposition*

$$\widehat{\lambda}_i = \bar{\mu}_i + u_i. \tag{67}$$

We also have when the final time $T \to \infty$ that

$$\sup_{i \in \mathbb{N} \text{ s.t. } \lfloor \sqrt{T} \rfloor < i \leq M} \sqrt{T}\big(\widehat{\lambda}_i - \bar{\mu}_{i,L} - u_i\big) \xrightarrow{\mathbb{P}} 0. \tag{68}$$

PROOF OF LEMMA 5. First, we have by Definition (31) for any index $i = 1, \dots, M$ that

$$\widehat{\lambda}_i = \frac{1}{\Delta}(N_{i\Delta -} - N_{(i-1)\Delta}).$$

Then, we obtain by Definition (5) for any index $i = 1, \dots, M$ that

$$\widehat{\lambda}_i = \frac{1}{\Delta}(M_{i\Delta -} - M_{(i-1)\Delta}) + \bar{\lambda}_i.$$

In addition, we get by Definition (64) for any index $i = 1, \dots, M$ that

$$\widehat{\lambda}_i = \varepsilon_i + \bar{\lambda}_i.$$

Moreover, we get by Equation (56) from Lemma 4 for any index $i = 1, \dots, M$ that

$$\widehat{\lambda}_i = \varepsilon_i + \bar{\mu}_i + \frac{1}{\Delta}\psi * (M_{i\Delta -} - M_{(i-1)\Delta}).$$

Furthermore, we can deduce by Definitions (34) , (36) and (37) for any index $i = 1, \dots, M$ that

$$\widehat{\lambda}_i = \varepsilon_i + \bar{\mu}_i + \frac{1}{\Delta}\Big(\int_0^{(i-1)\Delta} \Delta_i \Psi(-t)\, dM_t + \int_{(i-1)\Delta}^{i\Delta} \Psi(i\Delta - t)\, dM_t\Big).$$

This yields Equation (67) with the use of Definition (65). Finally, we obtain Equation (68) by Equation (57) from Lemma 4. □

We introduce the rescaled error process X_t which for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ satisfies

$$X_t = \sqrt{T}\big(\widehat{\Lambda}_t - \Lambda_t\big). \tag{69}$$

We also introduce ξ_i which is a random vector of dimension d defined for any interval index $i = 1, \dots, M$ by

$$\xi_i = \sqrt{T} u_i. \tag{70}$$

Moreover, we define the time discretized filtration for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ as

$$\mathbf{F}_t = \mathcal{F}_{\Delta \lfloor tT \rfloor}.$$

Furthermore, we use $\mathbb{E}_{i-1}$, Var_{i-1} and Cov_{i-1} instead of $\mathbb{E}[\cdot|\mathcal{F}_{(i-1)\Delta}]$, $\mathrm{Var}[\cdot|\mathcal{F}_{(i-1)\Delta}]$ and $\mathrm{Cov}[\cdot|\mathcal{F}_{(i-1)\Delta_n}]$ for any interval number $i = 1, \dots, M$. Finally, we define the notation small tau in probability uniformly for any time $t \in [0,1]$ as $Y_T(t) = o^u_{\mathbb{P}}(Z_T(t))$, which means that

$$\sup_{0 \leq t \leq 1} \left| \frac{Y_T(t)}{Z_T(t)} \mathbf{1}_{\{Z_T(t) \neq 0\}} \right| \xrightarrow{\mathbb{P}} 0$$

when the final time $T \to \infty$ for $Y_T(t)$ and $Z_T(t)$ which are stochastic processes.

The following lemma gives a discretization in time of the rescaled error process X_t based on the random vectors ξ_i. It also reexpresses the rescaled error process X_t as the sum of an $\mathbf{F}_t$-martingale and another random variable. To get the $\mathbf{F}_t$-martingale, we compensate the random variables ξ_i by their conditional expectations $\mathbb{E}_{i-1}[\xi_i]$.

LEMMA 6. *We assume that Assumptions 1 and 2 hold. Then, we can discretize the rescaled error process X_t for any time $t \in [0,1]$ when the final time $T \to \infty$ as*

$$X_t = \sum_{i=1}^{\lfloor tT \rfloor} \xi_i + o_{\mathbb{P}}^u(1). \tag{71}$$

We can also reexpress the rescaled error process X_t for any time $t \in [0,1]$ when the final time $T \to \infty$ as

$$X_t = \sum_{i=1}^{\lfloor tT \rfloor} \big(\xi_i - \mathbb{E}_{i-1}[\xi_i]\big) + \sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}[\xi_i] + o_{\mathbb{P}}^u(1). \tag{72}$$

PROOF OF LEMMA 6. First, we get by Definition (69) for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$X_t = \sqrt{T}\big(\widehat{\Lambda}_t - \Lambda_t\big).$$

We also can deduce by Definition (10) for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$X_t = \sqrt{T}\bigg(\frac{N_{tT}}{T} - \Lambda_t\bigg).$$

Then, this can be rewritten by an algebraic manipulation for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ as

$$X_t = \sqrt{T}\Bigg(\sum_{i=1}^{\lfloor tT \rfloor} \frac{(N_{i\Delta-} - N_{(i-1)\Delta})}{T} - \Lambda_t\Bigg).$$

In addition, we obtain by Definition (31) for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$X_t = \sqrt{T}\Bigg(\sum_{i=1}^{\lfloor tT \rfloor} \widehat{\lambda}_i - \Lambda_t\Bigg).$$

Moreover, we get by an algebraic manipulation for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$X_t = \sqrt{T}\Bigg(\sum_{i=1}^{\lfloor \sqrt{T} \rfloor} \widehat{\lambda}_i + \sum_{i=\lfloor \sqrt{T} \rfloor + 1}^{\lfloor tT \rfloor} \widehat{\lambda}_i - \Lambda_t\Bigg).$$

Finally, we get by another algebraic manipulation for any time $t \in [0,1]$ and any final time $T \in \mathbb{R}^+$ that

$$X_t = \sqrt{T}\Bigg(\sum_{i=1}^{\lfloor \sqrt{T} \rfloor} \widehat{\lambda}_i - \Lambda_{\sqrt{T}}\Bigg) + \sqrt{T}\Bigg(\sum_{i=\lfloor \sqrt{T} \rfloor + 1}^{\lfloor tT \rfloor} \widehat{\lambda}_i - \Lambda_t + \Lambda_{\sqrt{T}}\Bigg).$$

Now, this yields as $1/\sqrt{T}$ is negligible when the final time $T \to \infty$ for any time $t \in [0,1]$ that

$$X_t = \sqrt{T}\Bigg(\sum_{i=\lfloor \sqrt{T} \rfloor + 1}^{\lfloor tT \rfloor} \widehat{\lambda}_i - \Lambda_t + \Lambda_{\sqrt{T}}\Bigg) + o_{\mathbb{P}}^u(1).$$

This can be rewritten by Definitions (9) and (55) when the final time $T \to \infty$ for any time $t \in [0,1]$ as

$$X_t = \sqrt{T} \sum_{i=\lfloor \sqrt{T} \rfloor + 1}^{\lfloor tT \rfloor} \left(\widehat{\lambda}_i - \bar{\mu}_{i,L} \right) + o_{\mathbb{P}}^u(1).$$

Then, we can deduce by Expression (68) from Lemma 5 when the final time $T \to \infty$ for any time $t \in [0,1]$ that

$$X_t = \sqrt{T} \sum_{i=\lfloor \sqrt{T} \rfloor + 1}^{\lfloor tT \rfloor} u_i + o_{\mathbb{P}}^u(1).$$

Moreover, this yields as $1/\sqrt{T}$ is negligible when the final time $T \to \infty$ for any time $t \in [0,1]$ that

$$X_t = \sqrt{T} \sum_{i=1}^{\lfloor tT \rfloor} u_i + o_{\mathbb{P}}^u(1).$$

Finally, we get Equation (71) from Definition (70). Then, we can deduce Equation (72) by algebraic manipulation and since u_i is integrable from Lemma 2, Assumptions 2 (b) and (c). □

As the rescaled error process X_t takes a martingale form in Equation (72) from Lemma 6, we can use the toolkit from central limit theorems relying on martingales. More specifically, the proof of Theorem 1 is based on an application of Theorem IX.7.28 (pp. 590-591) in Jacod and Shiryaev (2003).

We first show that Condition (7.27) holds with $B_t = 0$ in the following proposition. This proves that the sum of the conditional expectations $\mathbb{E}_{i-1}[\xi_i]$ converges to 0 in probability and uniformly in time.

PROPOSITION 2. *We assume that Assumptions 1 and 2 hold. Then, we have when the final time $T \to \infty$ that*

$$\sup_{0 \leq t \leq 1} \left| \sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}[\xi_i] \right| \xrightarrow{\mathbb{P}} 0. \tag{73}$$

PROOF OF PROPOSITION 2. We first obtain by Definition (70) that

$$\sup_{0 \leq t \leq 1} \left| \sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}[\xi_i] \right| = \sup_{0 \leq t \leq 1} \left| \sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}\left[\sqrt{T} u_i \right] \right|. \tag{74}$$

Then, we can deduce by Definition (66) that

$$\sup_{0 \leq t \leq 1} \left| \sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}\left[\sqrt{T} u_i \right] \right| = \sup_{0 \leq t \leq 1} \left| \sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}\left[\sqrt{T} (\varepsilon_i + \epsilon_i) \right] \right|. \tag{75}$$

We introduce

$$I = \sup_{0 \leq t \leq 1} \left| \sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}\left[\sqrt{T} \varepsilon_i \right] \right|$$

and

$$II = \sup_{0\leq t\leq 1}\Big|\sum_{i=1}^{\lfloor tT\rfloor}\mathbb{E}_{i-1}\Big[\sqrt{T}\epsilon_i\Big]\Big|$$

We get by supremum properties that

$$\sup_{0\leq t\leq 1}\Big|\sum_{i=1}^{\lfloor tT\rfloor}\mathbb{E}_{i-1}\Big[\sqrt{T}u_i\Big]\Big| \leq I + II. \tag{76}$$

For the first term I, we can deduce by Definition (64) that

$$I = \sup_{0\leq t\leq 1}\Big|\sum_{i=1}^{\lfloor tT\rfloor}\mathbb{E}_{i-1}\Big[\frac{\sqrt{T}}{\Delta}\int_{(i-1)\Delta}^{i\Delta} dM_t\Big]\Big|.$$

This leads as M_t is an $\mathcal{F}_t$-martingale to

$$I = 0. \tag{77}$$

For the second term II, we can deduce by Definition (65) that

$$II = \sup_{0\leq t\leq 1}\Big|\sum_{i=1}^{\lfloor tT\rfloor}\mathbb{E}_{i-1}\Big[\frac{\sqrt{T}}{\Delta}\Big(\int_0^{(i-1)\Delta}\Delta_i\Psi(-t)\,dM_t + \int_{(i-1)\Delta}^{i\Delta}\Psi(i\Delta - t)\,dM_t\Big)\Big]\Big|.$$

This leads as M_t is an $\mathcal{F}_t$-martingale and by Lemma 2 when the final time $T\to\infty$ to

$$II \xrightarrow{\mathbb{P}} 0. \tag{78}$$

Moreover, we get from Expressions (76), (77) and (78) when the final time $T\to\infty$ that

$$\sup_{0\leq t\leq 1}\Big|\sum_{i=1}^{\lfloor tT\rfloor}\mathbb{E}_{i-1}\Big[\sqrt{T}u_i\Big]\Big| \xrightarrow{\mathbb{P}} 0. \tag{79}$$

Finally, Expressions (74), (75) and (79) yield the proposition. □

We show that Condition (7.28) holds in the following proposition. This proves that the sum of the covariances converges to the asymptotic covariance in probability.

PROPOSITION 3. *We assume that Assumptions 1 and 2 hold. Then, we have for any time* $t\in[0,1]$ *when the final time* $T\to\infty$ *that*

$$\sum_{i=1}^{\lfloor tT\rfloor}\mathrm{Cov}_{i-1}[\xi_i] \xrightarrow{\mathbb{P}} \int_0^t c_u c_u^T\,du. \tag{80}$$

PROOF OF PROPOSITION 3. First, the random vectors ξ_i^2 are integrable from Lemma 2, Assumptions 2 (b) and (c). Then, covariance properties yield for any time $t\in[0,1]$ and any final time $T\in\mathbb{R}^+$ that

$$\sum_{i=1}^{\lfloor tT\rfloor}\mathrm{Cov}_{i-1}[\xi_i] = \sum_{i=1}^{\lfloor tT\rfloor}\mathbb{E}_{i-1}[\xi_i\xi_i^T] - \sum_{i=1}^{\lfloor tT\rfloor}\mathbb{E}_{i-1}[\xi_i]\mathbb{E}_{i-1}[\xi_i^T]. \tag{81}$$

In addition, we obtain for the first term in the right side of Equation (81) by an extension of the arguments from the proof of Corollary 1 in Bacry et al. (2013) with Assumptions 1 (e) and 2 for any time $t \in [0,1]$ when the final time $T \to \infty$ that

$$\sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}[\xi_i \xi_i^T] \xrightarrow{\mathbb{P}} \int_0^t c_u c_u^T \, du. \tag{82}$$

To deal with the second term in the right side of Equation (81), we use Burkholder-Davis-Gundy inequalites (see Expression (2.1.32) in Jacod and Protter (2012) (p. 39)) with Assumptions 2 (b) and (c). This yields for any time $t \in [0,1]$ when the final time $T \to \infty$ that

$$\sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}[\xi_i] \mathbb{E}_{i-1}[\xi_i^T] \xrightarrow{\mathbb{P}} 0. \tag{83}$$

Finally, Expressions (81), (82) and (83) lead to the proposition. □

We show now that Condition (7.30) holds in the proposition that follows. This proves the Lindeberg condition for the central limit theorem.

PROPOSITION 4. *We assume that Assumptions 1 and 2 hold. Then, we have for any time $t \in [0,1]$ and any $u \in \mathbb{R}^+$ satisfying $u > 0$ when the final time $T \to \infty$ that*

$$\sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}\left[|\xi_i|^2 \mathbf{1}_{\{|\xi_i| > u\}} \right] \xrightarrow{\mathbb{P}} 0. \tag{84}$$

PROOF OF PROPOSITION 4. Since convergence in L^1 implies convergence in probability, it is sufficient to show the convergence in L^1. More specifically, we prove for any time $t \in [0,1]$ and any $u \in \mathbb{R}^+$ satisfying $u > 0$ when the final time $T \to \infty$ that

$$\mathbb{E}\left[\sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}\left[|\xi_i|^2 \mathbf{1}_{\{|\xi_i| > u\}} \right] \right] \to 0. \tag{85}$$

First, we have by linearity of the expectation for any time $t \in [0,1]$ and any $u \in \mathbb{R}^+$ satisfying $u > 0$ that

$$\mathbb{E}\left[\sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}\left[|\xi_i|^2 \mathbf{1}_{\{|\xi_i| > u\}} \right] \right] = \sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}\Big[\mathbb{E}_{i-1}\left[|\xi_i|^2 \mathbf{1}_{\{|\xi_i| > u\}} \right] \Big]. \tag{86}$$

Then, we get by conditional expectation properties for any time $t \in [0,1]$ and any $u \in \mathbb{R}^+$ satisfying $u > 0$ that

$$\sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}\Big[\mathbb{E}_{i-1}\left[|\xi_i|^2 \mathbf{1}_{\{|\xi_i| > u\}} \right] \Big] = \sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}\left[|\xi_i|^2 \mathbf{1}_{\{|\xi_i| > u\}} \right]. \tag{87}$$

Moreover, we can deduce by Hölder's inequality for any time $t \in [0,1]$ and any $u \in \mathbb{R}^+$ satisfying $u > 0$ that

$$\sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}\left[|\xi_i|^2 \mathbf{1}_{\{|\xi_i| > u\}} \right] \leq \sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}\left[|\xi_i|^{2+\eta} \right]^{\frac{1}{2+\eta}} \mathbb{E}\left[\mathbf{1}_{\{|\xi_i| > u\}} \right]^{\frac{1}{q}}. \tag{88}$$

Here, we have that $q \in \mathbb{R}$ which satisfies

$$\frac{1}{2+\eta} + \frac{1}{q} = 1.$$

Also, the random variables $|\xi_i|^{2+\eta}$ are integrable from Lemma 2, Assumptions 2 (b) and (c). Finally, we can conclude by an application of Burkholder-Davis-Gundy inequalites with Assumptions 2 (b) and (c). □

We show that Condition (7.31) holds in the following proposition.

PROPOSITION 5. *We assume that Assumptions 1 and 2 hold. Then, we have for any time $t \in [0,1]$ and for any bounded $\mathcal{F}_t$-martingale M' of dimension d when the final time $T \to \infty$ that*

$$\sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}[\xi_i^T \Delta_i M'] \xrightarrow{\mathbb{P}} 0. \tag{89}$$

PROOF OF PROPOSITION 5. We first have that the rescaled error process X_t is a purely discontinuous $\mathcal{F}_t$ martingale a.s. in the sense of Definition I.4.11 (b) (p. 40) from Jacod and Shiryaev (2003). This is obtained by Definition (69) and Definition 1. Thus, we can deduce that the product of the martingales $X_t M'_t$ is an $\mathcal{F}_t$ martingale a.s. for any bounded $\mathcal{F}_t$-martingale M' of dimension d by Definition I.4.11 (a) (p. 40) from Jacod and Shiryaev (2003). Thus, we can deduce for any time $t \in [0,1]$ and for any bounded $\mathcal{F}_t$-martingale M' of dimension d a.s. that

$$\sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}[\Delta_i X^T \Delta_i M'] = 0.$$

Finally, this yields by Equation (71) from Lemma 6 for any time $t \in [0,1]$ and for any bounded $\mathcal{F}_t$-martingale M' of dimension d when the final time $T \to \infty$ that

$$\sum_{i=1}^{\lfloor tT \rfloor} \mathbb{E}_{i-1}[\xi_i^T \Delta_i M'] \xrightarrow{\mathbb{P}} 0.$$

Thus, the proposition is shown. □

In what follows, we deliver the proof of Theorem 1, which is based on an application of Theorem IX.7.28 (pp. 590-591) in Jacod and Shiryaev (2003).

PROOF OF THEOREM 1. This is based on an application of Theorem IX.7.28 (pp. 590-591) in Jacod and Shiryaev (2003). We now verify that Conditions (7.27) to (7.31) are satisfied. First, we set the reference martingale $Z_t = 0$ which is a square-integrable $\mathcal{F}_t$-martingale. Thus, Condition (7.29) is directly satisfied. In addition, we have that Condition (7.27) holds by Proposition 2. Moreover, we can deduce that Condition (7.28) is satisfied with the use of Proposition 3. We also get that Condition (7.30) holds by Proposition 4. Finally, we obtain that Condition (7.31) is satisfied by an application of Proposition 5. □

5. Conclusion.

We have developed nonparametric inference for Hawkes processes with a rescaled stochastic and time dependent baseline. The inference procedure was based on the average of the point process. We have considered estimation for the average over time of the intensity process. We have showed the existence of these point processes. We have also showed the central limit theorem of the inference procedure.

Funding. The author was supported in part by Japanese Society for the Promotion of Science Grants-in-Aid for Scientific Research (B) 23H00807.

REFERENCES

BACRY, E., DAYRI, K. and MUZY, J.-F. (2012). Non-parametric kernel estimation for symmetric Hawkes processes. Application to high frequency financial data. *The European Physical Journal B* **85** 1–12.

BACRY, E., DELATTRE, S., HOFFMANN, M. and MUZY, J.-F. (2013). Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Processes and their Applications* **123** 2475–2499.

BRÉMAUD, P. and MASSOULIÉ, L. (1996). Stability of nonlinear Hawkes processes. *Annals of Probability* 1563–1588.

CHAVEZ-DEMOULIN, V., DAVISON, A. and MCNEIL, A. (2005). Estimating value-at-risk: a point process approach. *Quantitative Finance* **5** 227–234.

CHEN, F. and HALL, P. (2013). Inference for a nonstationary self-exciting point process with an application in ultra-high frequency financial data modeling. *Journal of Applied Probability* **50** 1006–1024.

CHEYSSON, F. and LANG, G. (2022). Spectral estimation of Hawkes processes from count data. *Annals of Statistics* **50** 1722–1746.

CLINET, S. and POTIRON, Y. (2018). Statistical inference for the doubly stochastic self-exciting process. *Bernoulli* **24** 3469–3493.

CLINET, S. and YOSHIDA, N. (2017). Statistical inference for ergodic point processes and application to limit order book. *Stochastic Processes and their Applications* **127** 1800–1839.

COSTA, M., GRAHAM, C., MARSALLE, L. and TRAN, V. C. (2020). Renewal in Hawkes processes with self-excitation and inhibition. *Advances in Applied Probability* **52** 879–915.

DALEY, D. J. and VERE-JONES, D. (2003). *An introduction to the theory of point processes: Elementary theory and methods* **1**, 2nd ed. Springer New York, NY.

DALEY, D. J. and VERE-JONES, D. (2008). *An introduction to the theory of point processes: General theory and structure* **2**, 2nd ed. Springer New York, NY.

DONNET, S., RIVOIRARD, V. and ROUSSEAU, J. (2020). Nonparametric Bayesian estimation for multivariate Hawkes processes. *Annals of Statistics* **48** 2698–2727.

EMBRECHTS, P., LINIGER, T. and LIN, L. (2011). Multivariate Hawkes processes: an application to financial data. *Journal of Applied Probability* **48** 367–378.

ERDEMLIOGLU, D., POTIRON, Y., VOLKOV, V. and XU, T. (2025). Estimation of time-dependent latency with locally stationary Hawkes processes. *Working paper available at https://www.fbc.keio.ac.jp/∼ potiron/Erdemlioglu2025workingpaper.pdf*.

GRAHAM, C. (2021). Regenerative properties of the linear Hawkes process with unbounded memory. *Annals of Applied Probability* **31** 2844–2863.

HARDIMAN, S. J., BERCOT, N. and BOUCHAUD, J.-P. (2013). Critical reflexivity in financial markets: a Hawkes process analysis. *The European Physical Journal B* **86** 1–9.

HAWKES, A. G. (1971a). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58** 83–90.

HAWKES, A. G. (1971b). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society. Series B (Methodological)* 438–443.

HAWKES, A. (2018). Hawkes processes and their applications to finance: a review. *Quantitative Finance* **18** 193–198.

HAWKES, A. G. and OAKES, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability* **11** 493–503.

HORST, U. and XU, W. (2021). Functional limit theorems for marked Hawkes point measures. *Stochastic Processes and their Applications* **134** 94–131.

HORST, U. and XU, W. (2022). The microstructure of stochastic volatility models with self-exciting jump dynamics. *Annals of Applied Probability* **32** 4568–4610.

HORST, U. and XU, W. (2024+). Functional limit theorems for Hawkes processes. *To appear in Probability Theory and Related Fields*.

JACOD, J. (1975). Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **31** 235–253.

JACOD, J. and PROTTER, P. E. (2012). *Discretization of processes*. Springer Berlin, Heidelberg.

JACOD, J. and SHIRYAEV, A. (2003). *Limit theorems for stochastic processes*, 2nd ed. Springer Berlin, Heidelberg.

JAISSON, T. and ROSENBAUM, M. (2015). Limit theorems for nearly unstable Hawkes processes. *Annals of Applied Probability* **25** 600–631.

JAISSON, T. and ROSENBAUM, M. (2016). Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *Annals of Applied Probability* **26** 2860–2882.

KAO, E. P. and CHANG, S.-L. (1988). Modeling time-dependent arrivals to service systems: A case in using a piecewise-polynomial rate function in a nonhomogeneous Poisson process. *Management Science* **34** 1367–1379.

KARIM, R. S., LAEVEN, R. J. and MANDJES, M. R. (2025+). Compound Multivariate Hawkes Processes: Large Deviations and Rare Event Simulation. *To appear in Bernoulli*.

KUHL, M. E. and BHAIRGOND, P. S. (2000). Nonparametric estimation of nonhomogeneous Poisson processes using wavelets. In *2000 Winter Simulation Conference Proceedings (Cat. No. 00CH37165)* **1** 562–571. IEEE.

KUHL, M. E., WILSON, J. R. and JOHNSON, M. A. (1997). Estimating and simulating Poisson processes having trends or multiple periodicities. *IIE transactions* **29** 201–211.

KUHL, M. E. and WILSON, J. R. (2000). Least squares estimation of nonhomogeneous Poisson processes. *Journal of Statistical Computation and Simulation* **67** 699–712.

KUHL, M. E. and WILSON, J. R. (2001). Modeling and simulating Poisson processes having trends or non-trigonometric cyclic effects. *European Journal of Operational Research* **133** 566–582.

KWAN, T., CHEN, F. and DUNSMUIR, W. (2023). Alternative asymptotic inference theory for a nonstationary Hawkes process. *Journal of Statistical Planning and Inference* **227** 75–90.

LEE, S., WILSON, J. R. and CRAWFORD, M. M. (1991). Modeling and simulation of a nonhomogeneous Poisson process having cyclic behavior. *Communications in Statistics-Simulation and Computation* **20** 777–809.

LEEMIS, L. M. (1991). Nonparametric estimation of the cumulative intensity function for a nonhomogeneous Poisson process. *Management Science* **37** 886–900.

LEWIS, P. A. and SHEDLER, G. S. (1976). Statistical analysis of non-stationary series of events in a data base system. *IBM Journal of Research and Development* **20** 465–482.

LINIGER, T. (2009). Multivariate Hawkes processes, PhD thesis, ETH Zurich.

MAMMEN, E. and MÜLLER, M. (2023). Nonparametric estimation of locally stationary Hawkes processes. *Bernoulli* **29** 2062–2083.

OGATA, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics* **30** 243–261.

OGATA, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical association* **83** 9–27.

OZAKI, T. (1979). Maximum likelihood estimation of Hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics* **31** 145–155.

POTIRON, Y. (2025+). Parametric inference for Hawkes processes with a general kernel. *Revise and Resubmit for Annals of the Institute of Statistical Mathematics*.

POTIRON, Y. and VOLKOV, V. (2025+). Mutually exciting point processes with latency. *To appear in Journal of the American Statistical Association*.

POTIRON, Y., SCAILLET, O., VOLKOV, V. and YU, S. (2025a). High-frequency estimation of Itô semimartingale baseline for Hawkes processes. *Working paper available at https://www.fbc.keio.ac.jp/∼ $potiron/Potiron2025workingpaper.pdf$*.

POTIRON, Y., SCAILLET, O., VOLKOV, V. and YU, S. (2025b). Estimation of branching ratio for Hawkes processes with Itô semimartingale baseline. *Working paper available at https://www.fbc.keio.ac.jp/∼ $potiron/Potiron2025estimationworkingpaper.pdf$*.

REYNAUD-BOURET, P. and SCHBATH, S. (2010). Adaptive estimation for Hawkes processes; application to genome analysis. *Annals of Statistics* **38** 2781–2822.

ROUEFF, F., VON SACHS, R. and SANSONNET, L. (2016). Locally stationary Hawkes processes. *Stochastic Processes and their Applications* **126** 1710–1743.

ROUEFF, F. and VON SACHS, R. (2019). Time-frequency analysis of locally stationary Hawkes processes. *Bernoulli* **25** 1355–1385.

RUBIN, I. (1972). Regular point processes and their detection. *IEEE Transactions on Information Theory* **18** 547–557.

VERE-JONES, D. and OZAKI, T. (1982). Some examples of statistical estimation applied to earthquake data: I. Cyclic Poisson and self-exciting models. *Annals of the Institute of Statistical Mathematics* **34** 189–207.

XU, W. (2024). Diffusion approximations for self-excited systems with applications to general branching processes. *Annals of Applied Probability* **34** 2650–2713.

ZHU, L. (2013). Central limit theorem for nonlinear Hawkes processes. *Journal of Applied Probability* **50** 760–771.

ZHU, L. (2015). Large deviations for Markovian nonlinear Hawkes processes. *Annals of Applied Probability* **25** 548–581.