Check for updates

J. Appl. Probab. **63,** 607–622 (2026)
*doi:*10.1017/jpr.2025.10047

SCALING LIMITS FOR SUPERCRITICAL NEARLY UNSTABLE HAWKES PROCESSES

CHENGUANG LIU,* *TU Delft*
LIPING XU,** *** AND
AN ZHANG,** **** *Beihang University*

Abstract

We investigate the asymptotic behavior of nearly unstable Hawkes processes whose regression kernel has L^1 norm strictly greater than 1 and close to 1 as time goes to infinity. We find that the scaling size determines the scaling behavior of the processes as in Jaisson and Rosenbaum (2015). Specifically, after a suitable rescale of $(a_T - 1)/T e^{b_T T x}$, the limit of the sequence of Hawkes processes is deterministic. Also, with another appropriate rescaling of $1/T^2$, the sequence converges in law to an integrated Cox–Ingersoll–Ross-like process. This theoretical result may apply to model the recent COVID-19 outbreak in epidemiology and phenomena in social networks.

Keywords: Hawkes processes; limit theorems; point processes

2020 Mathematics Subject Classification: Primary 60G55
Secondary 60F05

1. Introduction and main results

1.1. Introduction

The Hawkes process $(Z_t)_{t\geq 0}$ is a random point process that describes temporal stochastic self-exciting phenomena that was first introduced in 1971 [19]. The evolution of the process is influenced by the timing of the past events. It is determined by its intensity process $(\lambda_t)_{t\geq 0}$, which is of the form $\lambda_t = \mu + \int_0^t \varphi(t-s)\,\mathrm{d}Z_s$, where $\mu > 0$ is interpreted as the 'background rate' or 'exogenous rate' of the process, and φ is the regression kernel, which is the intensity function of a nonhomogeneous Poisson process, and quantifies the influence of past events of the process on the arrival of future events. For more introductory material on Hawkes processes, see [9, 10].

The Hawkes process is not only an interesting subject in mathematics but has also been widely applied to various fields to model both natural and social phenomena, like modeling earthquakes and their aftershocks in seismology [19, 21], brain activity in neuroscience [18], risk estimation, transaction times and midquote changes and so on in mathematical finance

Received 11 August 2025; accepted 11 August 2025.
* Postal address: Delft Institute of Applied Mathematics, EEMCS, TU Delft, 2628 Delft, The Netherlands. Email: C.Liu-13@tudelft.nl
** Postal address: School of Mathematical Sciences, Beihang University, PR China.
*** Email address: xuliping.p6@gmail.com
**** Email address: anzhang@buaa.edu.cn

© The Author(s), 2025. Published by Cambridge University Press on behalf of Applied Probability Trust. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

[2, 3, 20], social network interactions [8, 29], etc. There are substantial numbers of works on Hawkes processes and their applications in various disciplines; here we have given a nonexhaustive reference and we refer the reader to the bibliography and references therein for more details.

From a mathematical point of view, there has been much investigation into the large-time behavior of the Hawkes process, especially the law of large numbers and the central limit theorem, which have been established for different cases of Hawkes processes, like linear, nonlinear, subcritical, critical, and supercritical, under the fixed-rate assumption $\int_0^\infty \varphi \, \mathrm{d}t < \infty$ [1, 6, 27, 31, 33]. The large-deviation behavior for both linear and nonlinear Hawkes processes was established in [4, 15, 16, 32, 34, 35] (which is not an exhaustive list of references).

Recently, the limit theorem for multivariate marked Hawkes processes was established [30]. Meanwhile, functional central limit theorems for subcritical and critical Hawkes processes and convergence for the heavy-tailed Hawkes processes were investigated in [22, 23], respectively. In particular, for the linear regime, because of its tractability (especially the immigration–birth representation), the limit behaviors have been well understood and widely applied in practice. We also restrict our consideration to the linear case. The rate $\int_0^\infty \varphi \, \mathrm{d}t < 1$ refers to the subcritical case, and $\int_0^\infty \varphi \, \mathrm{d}t = 1$ to the critical case; otherwise, we are dealing with the supercritical case. In finance, this fixed rate is the so-called *branching ratio*, which is interpreted as the average proportion of endogenously triggered events. However, in practice, this ratio can't always be fixed: it may vary from time to time caused by many real factors. This is consistent with the statistical estimation results, which often seem to show that only nearly unstable Hawkes processes are able to fit the data properly, see the introduction of [26]. Regarding the nearly unstable processes, first introduced in [26], this means that the L^1 norm of their kernel is close to 1. In [26] a time-dependent kernel is assumed with its L^1 norm less than 1 and close to 1 as time goes to infinity, which is a near instability condition, i.e. $\int_0^\infty \varphi^T \, \mathrm{d}t = a_T < 1$ and $\lim_{T\to\infty} a_T = 1$, where $(a_T)_{T\geq 0}$ is a sequence of positive numbers such that, for all T, $a_T < 1$. Together with the assumption $\int_0^\infty t\varphi^T(t) \, \mathrm{d}t = m < \infty$, [26] then shows that these nearly unstable Hawkes processes, under suitable scaling, converge deterministically if $T(1 - a_T) \to +\infty$ as $T \to \infty$, i.e. (law of large numbers)

$$\sup_{t\in[0,1]} \frac{1 - a_T}{T} \left| Z_{Tt}^T - \mathbb{E}\left[Z_{Tt}^T\right] \right| \to 0 \text{ in probability,}$$

and tend to an integrated Cox–Ingersoll–Ross (CIR) process if $T(1 - a_T) \to \lambda > 0$ as $T \to \infty$, i.e. (central limit theorem)

$$\left(\frac{1 - a_T}{T} Z_{Tt}^T \right)_{t\in[0,1]} \to \left(\int_0^t X_s \, \mathrm{d}s \right)_{t\in[0,1]} \text{ in law}$$

for the Skorohod topology, where $(X_t)_{t\in[0,1]}$ is a CIR process, originally introduced in [7] and used classically to model stochastic volatilities in finance.

1.2. Motivation and comparison

As we noticed, the great bulk of the literature related to the Hawkes process and its applications to geophysics and finance focuses on the subcritical regime, whereas the supercritical regime is little studied but is more germane to epidemics, like modeling the spread of COVID-19 in [17], and to social networks [8, 29].

In this paper, we consider supercritical nearly unstable Hawkes processes, i.e. the kernels of the Hawkes processes depending on T, with L^1 norm strictly larger than 1 and close to 1 as

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

$T \to \infty$. When the norm of the kernel is close to 1 at a lower speed than T goes to infinity (see H(∞)), we prove that the renormalized Hawkes process is asymptotically deterministic in some sense, see Theorem 1. And when the norm of the kernel is close to 1 at a suitable speed (see the assumption H(Λ)), we prove that the limit behavior of our sequence of Hawkes processes, with a suitable scale of $1/T^2$, is an integrated diffusion process, like a CIR process, which is a strong solution of a Brownian driven stochastic differential equation(SDE), see Theorem 2. To do so, we adopt a different approach from [26] in order to get the limiting behavior of a sequence of Hawkes processes with a scale of $(1 - a_T)/T$ with $0 < a_T < 1$. In [26], the limiting behavior of the intensity of the processes was studied. Roughly speaking, for $t \in [0, 1]$ the authors rewrote the intensity λ^T_{Tt} to $C^T_t = \lambda^T_{Tt}(1 - a_T)$, which can be written as a stochastic integral equation, then proved the tightness of C^T_t and passed the coefficients of the equation to the limit to obtain the limiting law. However, in our context, instead of considering the intensity, we directly deal with Hawkes processes whose limit is the same as their cumulated intensity, following the spirit of [33, Section 5.3], which dealt with the limit behavior of Z^T_{Tt}/T^2 for the critical case, i.e. $a_T \equiv 1$ for any T. There, it was proved that the sequence of Z^T_{Tt}/T^2 converges in law, for the Skorohod topology, to an integrated squared Bessel process, i.e. $\int_0^t X_s \, \mathrm{d}s$, with X_t satisfying

$$\mathrm{d}X_t = \frac{\mu}{m} \, \mathrm{d}t + \frac{1}{m}\sqrt{X_t} \, \mathrm{d}B_t.$$

However, in our Theorem 2, we extend the above case to the nearly unstable supercritical case, i.e. $a_T > 1$ and $\lim_{T\to\infty} a_T = 1$, and we prove that the limit of Z^T_{Tt}/T^2 is an integrated CIR-like process for the Skorohod topology, i.e. $\int_0^t X_s \, \mathrm{d}s$, with X_t satisfying

$$X_t = \frac{1}{m}\int_0^t (\mu + \lambda X_s) \, \mathrm{d}s + \frac{1}{m}\int_0^t \sqrt{X_s} \, \mathrm{d}B_s.$$

Moreover, the supercritical case was also investigated in [33, Section 5.4], transferring the supercritical case to the critical case using the Malthusian parameter θ, which has often been done when dealing with supercritical branching processes and supercritical Hawkes processes. This proved that the intensity λ_t grows exponentially in t, and $Z_t/\mathrm{e}^{\theta t}$ converges almost surely to a process given by $\mu/(\theta^2\bar{m}) + W/(\theta\bar{m})$, with $W = \int_0^\infty \mathrm{e}^{-\theta t} \, \mathrm{d}M_t$. Finally, we also mention the mean field limit for multivariate Hawkes processes studied in [12, 28] when the dimension goes to infinity.

1.3. Setting

We consider a measurable function $\varphi\colon [0, \infty) \to [0, \infty)$ and a Poisson measure $(\Pi(\mathrm{d}t, \mathrm{d}z))$ on $[0, \infty) \times [0, \infty)$ with intensity $\mathrm{d}t \, \mathrm{d}z$. The sequence $\{a_T\}_{T\geq 1}$ indexed by T is positive and $\lim_{T\to\infty} a_T = 1$. We set $\varphi^T = a_T\varphi$. We consider the following system indexed by T: for all $t \geq 0$,

$$Z^T_t := \int_0^t \int_0^\infty \mathbf{1}_{\{z\leq \lambda^T_s\}} \, \Pi(\mathrm{d}s, \mathrm{d}z), \qquad \text{where } \lambda^T_t := \mu + \int_0^{t-} \varphi^T(t - s) \, \mathrm{d}Z^T_s. \tag{1}$$

In this paper, $\int_0^t$ means $\int_{[0,t]}$, and $\int_0^{t-}$ means $\int_{[0,t)}$. The solution $(Z^T_t)_{t\geq 0}$ is a counting processes. By [12, Proposition 1], see also [5, 13], the system (1) has a unique $(\mathcal{F}_t)_{t\geq 0}$-measurable càdlàg solution, where $\mathcal{F}^T_t = \sigma(\Pi(A)\colon A \in \mathcal{B}([0, t] \times [0, \infty))$, as long as φ^T is locally integrable.

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

1.4. An illustrative example in epidemiology

We first recall the notable branching representation of Hawkes processes in terms of Galton–Watson trees given in [21]. That is, in a population process, the migrants arrive according to a Poisson process with 'exogenous rate' μ. Then each migrant gives birth to children according to an 'endogenous rate' given by φ, and each child gives birth to grandchildren according to the same 'endogenous rate'.

Now let us consider the epidemiological example. Let process $(Z_t^T)_{t\geq 0}$ represent the number of infected individuals over the observation period $[0, T]$. This process accounts for two main mechanisms of infection: *immigration* and *transmission* from already infected individuals:

- Immigration of infected individuals: New infections enter the population independently of current local transmissions, modeled by a homogeneous Poisson process with rate μ. This represents external sources of infection, such as travelers or new arrivals carrying the disease.
- Local transmission dynamics: Once an individual is infected, they can further spread the infection to others. This local transmission is captured by an inhomogeneous Poisson process, where the intensity is given by the kernel function $a_T\varphi(t-s)$. Here,
  - t denotes the current time;
  - s is the time at which an individual was infected;
  - φ is a kernel function, typically assumed to be decreasing. This assumption reflects the reality that the infectivity of an individual often decreases over time due to factors such as recovery or isolation.

The parameter a_T represents the level of social interference, which encompasses various control measures such as medical interventions, government policies, and vaccination efforts. Notably, a_T is a function of time T, and it typically decreases as T increases, reflecting the effectiveness of these interventions in reducing the spread of infection over time. So, from this point of view, $a_T > 1$ is quite consistent with reality.

1.5. Assumptions

We always assume

$$\|\varphi\|_1 = \int_0^\infty \varphi(s)\,\mathrm{d}s = 1 \quad \text{and} \quad \int_0^\infty s\varphi(s)\,\mathrm{d}s = m > 0. \tag{Λ}$$

Also, for $\lambda > 0$, we assume either

$$a_T > 1 \quad \text{and} \quad \lim_{T\to\infty} T(a_T - 1) = \lambda, \tag{H(Λ)}$$

$$a_T > 1 \quad \text{and} \quad \lim_{T\to\infty} T(a_T - 1) = \infty. \tag{H(∞)}$$

The assumption $\int_0^\infty s\varphi(s)\,\mathrm{d}s = m > 0$ is crucial in our context and enables us to approximate Ψ^T by an exponential function in Lemma 1.

1.6. Notation

We first recall the convolution of two functions $f, g\colon [0,\infty) \mapsto \mathbb{R}$, which (if it exists) is defined by $(f * g)(t) = \int_0^t f(t-s)g(s)\,\mathrm{d}s$. $(\varphi)^{*n}$ represents the nth convolution product

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

of the function φ, and $(\varphi)^{*1} = \varphi$. Since $\int_0^\infty \varphi(s)\,\mathrm{d}s = 1$, it's not hard to observe that $\int_0^\infty (\varphi^T)^{*n}(s)\,\mathrm{d}s = (a_T)^n$. We adopt the conventions $\varphi^{*0}(s)\,\mathrm{d}s = \delta_0(\mathrm{d}s)$ and $\varphi^{*0}(t-s)\,\mathrm{d}s = \delta_t(\mathrm{d}s)$, and the convention $\varphi^{*n}(s) = 0$ for $s < 0$. We also recall the L^1 norm of a function φ, denoted by $\|\varphi\|_1 = \int_0^\infty |\varphi(s)|\,\mathrm{d}s$.

Since $1 < a_T = \int_0^\infty \varphi^T(s)\,\mathrm{d}s < +\infty$, we can choose a unique positive sequence $\{b_T\}_{T\geq 1}$ such that $\int_0^\infty \mathrm{e}^{-b_T s}\varphi^T(s)\,\mathrm{d}s = 1/a_T$; b_T is sometimes referred to as the Malthusian parameter in the literature for fixed T [33]. We define the following functions on $\mathbb{R}^+$:

$$\tilde{\varphi}^T(t) = \mathrm{e}^{-b_T t}\varphi^T(t), \qquad \tilde{\Psi}^T(t) = \sum_{n\geq 1} (\tilde{\varphi}^T)^{*n}(t). \tag{2}$$

Note that $\tilde{\Psi}^T(t)$ is well defined on $\mathbb{R}^+$ since $\|\tilde{\varphi}^T\|_1 = 1/a_T < 1$. A direct computation implies that

$$\|\tilde{\Psi}^T\|_1 = \frac{\|\tilde{\varphi}^T\|_1}{1 - \|\tilde{\varphi}^T\|_1} = \frac{1}{a_T - 1}. \tag{3}$$

We also define, for $t \geq 0$,

$$\Psi^T(t) := \mathrm{e}^{b_T t}\tilde{\Psi}^T(t), \tag{4}$$

which is also well defined. For $n \geq 1$, it's not hard to find that $\mathrm{e}^{-b_T t}(\varphi^T)^{*n}(t) = (\mathrm{e}^{-b_T t}\varphi^T)^{*n}(t) = (\tilde{\varphi}^T)^{*n}(t)$. Thus, we have

$$\Psi^T(t) = \sum_{n\geq 1} (\varphi^T)^{*n}(t). \tag{5}$$

We recall $(\lambda_t^T)_{t\geq 0}$ defined in (1), and introduce the compensated Poisson measure $\tilde{\Pi}(\mathrm{d}s, \mathrm{d}z) = \Pi(\mathrm{d}s, \mathrm{d}z) - \mathrm{d}s\,\mathrm{d}z$ associated with the Poisson measure Π. For $t \geq 0$, set

$$M_t^T := \int_0^t \int_0^\infty \mathbf{1}_{\{z\leq \lambda_s^T\}}\,\tilde{\Pi}(\mathrm{d}s, \mathrm{d}z), \tag{6}$$

which is a martingale process related to Z_t^T. We then readily find that $M_t^T = Z_t^T - \int_0^t \lambda_s^T\,\mathrm{d}s$. According to [12, Remark 10] (see also [25, Chapter 1, Section 4e] for definitions and properties of pure jump martingales and of their quadratic variations) we then have $\mathbb{E}[M_s^T M_t^T] = \mathbb{E}[Z_{s\wedge t}^T]$ and $[M^T, M^T]_t = Z_t^T$.

1.7. Main results

We first give the law of large numbers in the following sense.

Theorem 1. *Under the assumptions* H(Λ) *and* H(∞), *the sequence of Hawkes processes is asymptotically deterministic, in the sense that the following convergence in* L^2 *holds as* $T \to \infty$*:*

$$\sup_{t\in[0,1]} \frac{a_T - 1}{T\mathrm{e}^{b_T T t}}\left|Z_{Tt}^T - \mathbb{E}\left[Z_{Tt}^T\right]\right| \to 0.$$

Next, we introduce our second main result related to the central limit theorem.

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

Theorem 2. *Under the assumptions* (Λ) *and* H(Λ)*, for the Skorohod topology,*

$$\left(\frac{Z^T_{Tt}}{T^2}\right)_{t\in[0,1]} \xrightarrow{\text{(d)}} \left(\int_0^t X_s\,\mathrm{d}s\right)_{t\in[0,1]} \qquad \text{as } T\to\infty,$$

where $(X_t)_{t\in[0,1]}$ *is the unique strong solution to the SDE*

$$X_t = \frac{1}{m}\int_0^t (\mu+\lambda X_s)\,\mathrm{d}s + \frac{1}{m}\int_0^t \sqrt{X_s}\,\mathrm{d}B_s, \tag{7}$$

where $(B_t)_{t\ge0}$ *is a Brownian motion.*

1.8. Quantity analysis in epidemiology

As stated in Section 1.4, where an informal application of our result is given, the process $(Z^T_t)_{t\ge0}$ represents the number of infected individuals over the observation period $[0, T]$. In this subsection, we will show how we apply Theorem 2 to quantify the number of infected individuals $(Z^T_t)_{t\ge0}$ on $[0, T]$. Throughout the subsection, we fix the positive constants μ, λ, and m, and assume $n := 4\mu m^2 \in \mathbb{N}$. The process $(X_t)_{t\ge0}$ solving (7) introduced in Theorem 2 admits the following distributional representation. There exist n independent and identically distributed (i.i.d.) Gaussian processes $\{(Y^i_t)_{t\in[0,1]}\}_{i=1,\dots n}$ such that

$$(X_t)_{t\in[0,1]} \overset{\text{(d)}}{=} ((Y^1_t)^2+\cdots+(Y^n_t)^2)_{t\in[0,1]},$$

and each component Y^i satisfies the linear SDE

$$\mathrm{d}Y^i_t = \frac{\lambda}{2m}Y^i_t\,\mathrm{d}t + \frac{1}{2m}\,\mathrm{d}W^i_t, \qquad Y^i_0 = 0, \tag{8}$$

where $\{W^i\}_{i=1}^n$ are independent standard Brownian motions. For $t\in[0,1]$, the distribution function of the random variable $\int_0^t (Y^1_s)^2\,\mathrm{d}s$ is denoted by $F_t(x) = \mathbb{P}\big(\int_0^t (Y^1_s)^2\,\mathrm{d}s \le x\big)$ for all $x\in\mathbb{R}$, and the corresponding n-fold convolution is denoted by F^{*n}_t, i.e. the distribution function of $\sum_{i=1}^n \int_0^t (Y^i_s)^2\,\mathrm{d}s$. Then, by Theorem 2, we have the following quantity estimation.

Proposition 1. *Assume that there exists some* $n\in\mathbb{N}$ *such that* $4\mu m^2 = n$*. Then, for any positive constants* $0 < c \le C$*,*

$$\lim_{T\to\infty}\mathbb{P}\left(c \le \frac{Z^T_{Tt}}{T^2} \le C\right) = F^{*n}_t(C) - F^{*n}_t(c),$$

where $(Z^T_{Tt})_{t\ge0}$ *is defined in Theorem 2, and* F_t *is the distribution function of the random variable* $\int_0^t (Y^1_s)^2\,\mathrm{d}s$*, where* Y^1 *is defined in (8).*

Proof. Recall that $n = 4\mu m^2 \in \mathbb{N}$, and that $\{(Y^i_t)_{t\in[0,1]}\}_{i=1,\dots n}$ are n i.i.d. Gaussian processes solving (8). Set $\tilde{X}_t := \sum_{i=1}^n (Y^i_t)^2$. A direct application of Itô's formula together with the relation $n = 4\mu m^2$ yields

$$\begin{aligned}\mathrm{d}\tilde{X}_t = \frac{2}{m}\sum_{i=1}^n Y^i_t\,\mathrm{d}Y^i_t + \frac{n}{4m^2}\,\mathrm{d}t &= \frac{1}{m}\left(\lambda\sum_{i=1}^n (Y^i_t)^2 + \mu\right)\mathrm{d}t + \sum_{i=1}^n \frac{Y^i_t}{m}\,\mathrm{d}W^i_t \\ &= \frac{1}{m}(\lambda\tilde{X}_t+\mu)\,\mathrm{d}t + \frac{\sqrt{\tilde{X}_t}}{m}\,\mathrm{d}\tilde{W}_t,\end{aligned}$$

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

where $\tilde{W}_t = \sum_{i=1}^n \int_0^t (Y_s^i/\sqrt{\tilde{X}_s})\,\mathrm{d}W_s^i$ is a martingale. Moreover, it is easy to check that $[\tilde{W}, \tilde{W}]_t = t$. Following from Lévy's characterization theorem, we conclude that $(\tilde{W}_t)_{t\in[0,1]}$ is a standard Brownian motion. Hence, $(\tilde{X}_t)_{t\in[0,1]} \overset{(\mathrm{d})}{=} (X_t)_{t\in[0,1]}$ with common initial condition 0. In particular, $\int_0^t X_s\,\mathrm{d}s \overset{(\mathrm{d})}{=} \int_0^t \tilde{X}_s\,\mathrm{d}s$. Applying Theorem 2, we therefore obtain the stated limit. □

Remark 1. The processes $\{(Y_t^i)_{t\in[0,1]}\}_{i=1,\dots,n}$ here are in fact n i.i.d. Ornstein–Uhlenbeck processes. The distribution function F_t of $\int_0^t (Y_s^1)^2\,\mathrm{d}s$, or, more generally, the distribution function of the integration of the squared general Ornstein–Uhlenbeck process, has been very well studied, e.g. see [11], where the Laplace transform of the probability density of the integral of the squared Ornstein–Uhlenbeck process was established explicitly. The distribution of X and of its time integral have also been studied in depth, although the general case (without $4\mu m^2 \in \mathbb{N}$) quickly becomes complicated (see, e.g., [14, Chapter 10]).

Remark 2. (*Implications for epidemic size*.) The assumption $4\mu m^2 \in \mathbb{N}$ does not lose any generality. When estimating the number of infected individuals, we can always bracket the true value by choosing μ slightly larger or smaller so that the assumption holds. In addition, we consider this assumption not only because the general case is more complicated, but also because it enables us to give a concise and clear estimation of the quantity. Indeed, we can solve Y_t^i as

$$\frac{\mathrm{e}^{\lambda t/2m}}{2m}\int_0^t \mathrm{e}^{-\lambda s/2m}\,\mathrm{d}W_s^i$$

and, by a direct computation or the law of large numbers, we deduce that

$$\sum_{i=1}^n \int_0^1 (Y_s^i)^2\,\mathrm{d}s = n\left(\mathbb{E}\left[\int_0^1 (Y_s^1)^2\,\mathrm{d}s\right] + \chi_n\right),$$

where $\chi_n = n^{-1}\sum_{i=1}^n \left(\int_0^1 (Y_s^i)^2\,\mathrm{d}s - \mathbb{E}\left[\int_0^1 (Y_s^i)^2\,\mathrm{d}s\right]\right)$ with $\mathbb{E}[(\chi_n)^2]^{1/2} \le C/\sqrt{n}$ because the $(Y^i)_{i=1,\dots,n}$ are i.i.d. In typical applications we take large $n = 4\mu m^2 \gg 1$, where μ is the immigration rate. Consequently,

$$\lim_{T\to\infty}\frac{Z_T^T}{T^2} \overset{(\mathrm{d})}{=} \int_0^1 X_s\,\mathrm{d}s \overset{(\mathrm{d})}{=} \int_0^1 \tilde{X}_s\,\mathrm{d}s = \sum_{i=1}^n \int_0^1 (Y_s^i)^2\,\mathrm{d}s \approx n\mathbb{E}\left[\int_0^1 (Y_s^1)^2\,\mathrm{d}s\right] = \frac{\mu(\mathrm{e}^{\lambda/m} - 1 - \lambda/m)}{\lambda^2},$$

where '$a_n \approx b_n$' means $\mathbb{E}[|a_n - b_n|]/n \to 0$ as $n \to \infty$. Since the function $(\mathrm{e}^x - x - 1)/x^2$ is increasing exponentially, it's clear that reducing the local transmission rate λ is markedly more effective than lowering the immigration rate μ. The kernel parameter m, being a property of the pathogen, is typically fixed in the absence of mutation.

In the following, C stands for a positive constant whose value may change from line to line. When necessary, we will indicate as subscripts the parameters it depends on.

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

2. Analysis of the convolution function

The aim of this section is to provide some auxiliary results on the kernel function φ. We are going to illustrate the long-term behavior of the function Ψ^T. To do this, we first recall the following proposition.

Proposition 2. ([26, Proposition 2.2].) *Under assumption* (Λ), *consider a sequence* $\{c_T\}_{T\geq 1}$ *with* $0 < c_T < 1$ *and* $\lim_{T\to\infty} c_T = 1$, *and a random variable* $X^T := (1/T)\sum_{i=1}^{I^T} X_i$, *where* (X_i) *are i.i.d. random variables with density* φ *and* I^T *is a geometric random variable independent of* (X_i) *with parameter* $1 - c_T$. *If* $\lim_{T\to\infty} T(1 - c_T) = \lambda$, *then* $X^T \overset{(\mathrm{d})}{\to} X$ *as* $T \to \infty$, *where* X *admits an exponential distribution with parameter* λ/m.

Writing $\hat{\Psi}^T(t) := \sum_{k=1}^{\infty} (c_T)^k (\varphi)^{*k}(t)$, let ρ^T be the density of X^T; then the proof of Proposition 2 in [26] gives

$$\hat{\Psi}^T(Tt) = \rho^T(t)\frac{c_T}{(1 - c_T)T},$$

which implies directly that

$$\lim_{T\to\infty} \hat{\Psi}^T(Tt) = \frac{1}{m}\mathrm{e}^{-\lambda t/m} \tag{9}$$

for $x \geq 0$ in the weak sense.

In the next lemma, we see that Ψ^T in terms of $a_T > 1$ has the same convergence as $\hat{\Psi}^T$ in (9).

Lemma 1. *Under assumptions* (Λ) *and* H(Λ), *and recalling* Ψ^T *defined in (4), we have the following convergence in the weak sense for* $t \geq 0$: $\lim_{T\to\infty} \Psi^T(Tt) = (1/m)\mathrm{e}^{\lambda t/m}$.

Proof. First, recall that the positive sequence $\{b_T\}_{T\geq 0}$ satisfy

$$\int_0^{\infty} \mathrm{e}^{-b_T s}\varphi^T(s)\,\mathrm{d}s = \frac{1}{a_T}$$

and define $m_T := \int_0^{\infty} \mathrm{e}^{-b_T s} s\varphi^T(s)\,\mathrm{d}s$. Since $\lim_{T\to\infty} a_T = 1$, we must have $\lim_{T\to\infty} b_T = 0$. Hence, by the dominated convergence theorem, we have $\lim_{T\to\infty} m_T = m$, and then $\lim_{T\to\infty} (a_T - (1/a_T))/b_T = m$. Indeed,

$$\begin{aligned}\lim_{T\to\infty} \frac{(a_T - (1/a_T))}{b_T} &= \lim_{T\to\infty} a_T \int_0^{\infty} \frac{(1 - \mathrm{e}^{-b_T s})\varphi(s)}{b_T}\,\mathrm{d}s \\ &= \int_0^{\infty} \lim_{T\to\infty} \frac{(1 - \mathrm{e}^{-b_T s})\varphi(s)}{b_T}\,\mathrm{d}s = \int_0^{\infty} s\varphi(s)\,\mathrm{d}s = m.\end{aligned}$$

Due to $\lim_{T\to\infty} T(a_T - 1) = \lambda$, we thus conclude that $\lim_{T\to\infty} Tb_T = 2\lambda/m$. Note that $\int_0^{\infty} s\tilde{\varphi}^T(s)\,\mathrm{d}s = m_T$, and recalling (2) and (9), we have $\lim_{T\to\infty} \tilde{\Psi}^T(Tt) = (1/m)\mathrm{e}^{-\lambda t/m}$. Consequently, according to (4),

$$\lim_{T\to\infty} \Psi^T(Tt) = \lim_{T\to\infty} \mathrm{e}^{Tb_T t}\tilde{\Psi}^T(Tt) = \mathrm{e}^{2\lambda tm}\frac{1}{m}\mathrm{e}^{-\lambda t/m} = \frac{1}{m}\mathrm{e}^{\lambda t/m},$$

which completes the proof. □

We now recall a classical lemma that we will need.

Lemma 2. ([1, Lemma 3], [12, Lemma 8].) *Let* A *be a constant. Consider* $f, g\colon [0,\infty) \mapsto \mathbb{R}$ *locally bounded, and assume that* $\varphi\colon [0,\infty) \mapsto [0,\infty)$ *is locally integrable. If* $f_t = g_t + \int_0^t \varphi(t - s)Af_s\,\mathrm{d}s$ *for all* $t \geq 0$, *then* $f_t = \sum_{n\geq 0} \int_0^t \varphi^{*n}(t - s)A^n g_s\,\mathrm{d}s$.

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

We are now ready to estimate the expectation of the Hawkes process, which is crucial to proving the tightness.

Proposition 3. *Consider the solution* $(Z_t^T)_{t\geq 0}$ *to (1).*

(i) *Under the assumptions* (Λ) *and* H(Λ), *for* $0 \leq t \leq 1$,

$$\lim_{T\to\infty} \frac{\mathbb{E}\big[Z_{Tt}^T\big]}{T^2} = \frac{\mu e^{\lambda t/m}}{m} \int_0^t v e^{-\lambda v/m}\, \mathrm{d}v = -\frac{\mu}{\lambda} + \frac{m\mu}{\lambda^2}(e^{\lambda t/m} - 1).$$

In particular, $\mathbb{E}\big[Z_{Tt}^T\big] \leq CT^2$ *for some constant* $C > 0$.

(ii) *Under assumptions* (Λ) *and* H(∞),

$$\sup_{t\in[0,1]} \frac{a_T - 1}{T e^{b_T T t}} \mathbb{E}\big[Z_{Tt}^T\big] \leq \mu a_T.$$

In particular,

$$\sup_{t\in[0,T]} \frac{a_T - 1}{T e^{b_T t}} \mathbb{E}\big[Z_t^T\big] \leq \mu a_T.$$

Proof. The proof is inspired by [12, Lemma 11], which deals with multivariate Hawkes processes. We first rewrite the Hawkes process $(Z_t^T)_{t\geq 0}$. Starting from (1) and using (6), it's not difficult to find that

$$Z_t^T = M_t^T + \int_0^t \lambda_s^T\, \mathrm{d}s = M_t^T + \mu t + \int_0^t \int_0^s \varphi^T(s-v)\, \mathrm{d}Z_v^T\, \mathrm{d}s.$$

Using [13, Lemma 22], we have $\int_0^t \int_0^s \varphi^T(s-v)\, \mathrm{d}Z_v^T\, \mathrm{d}s = \int_0^t \varphi^T(t-s) Z_s^T\, \mathrm{d}s$. Accordingly,

$$Z_t^T = M_t^T + \mu t + \int_0^t \varphi^T(t-s) Z_s^T\, \mathrm{d}s. \tag{10}$$

Since $(M_t^T)_{t\geq 0}$ is martingale, we see that $\mathbb{E}[M_t^T] = 0$ for all $t \geq 0$. Hence, for $t \geq 0$,

$$\mathbb{E}[Z_t^T] = \mu t + \int_0^t \varphi^T(t-s)\mathbb{E}[Z_s^T]\, \mathrm{d}s = \mu t + \int_0^t a_T \varphi(t-s)\mathbb{E}[Z_s^T]\, \mathrm{d}s. \tag{11}$$

Applying Lemma 2 and noting (5) and that $\varphi^{*0}(t-s)\, \mathrm{d}s = \delta_t(\mathrm{d}s)$, we thus conclude that

$$\mathbb{E}[Z_t^T] = \mu \sum_{n=0}^{\infty} \int_0^t (a_T)^n \varphi^{*n}(t-s) s\, \mathrm{d}s = \mu t + \mu \int_0^t s\Psi^T(t-s)\, \mathrm{d}s. \tag{12}$$

Now, we obtain

$$\mathbb{E}\big[Z_{Tt}^T\big] = \mu T t + \mu \int_0^{Tt} s\Psi^T(Tt-s)\, \mathrm{d}s \overset{s=Tv}{=} \mu T t + \mu T^2 \int_0^t v\Psi^T(T(t-v))\, \mathrm{d}v.$$

Following from Lemma 1, we have

$$\lim_{T\to\infty} \frac{\mathbb{E}\big[Z_{Tt}^T\big]}{T^2} = \lim_{T\to\infty} \left[\frac{\mu t}{T} + \mu \int_0^t v\Psi^T(T(t-v))\, \mathrm{d}v\right] = \frac{\mu e^{\lambda t/m}}{m} \int_0^t v e^{-\lambda v/m}\, \mathrm{d}v.$$

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

It's thus clear that $\lim_{T\to\infty} \mathbb{E}\big[Z_{Tt}^T\big]/T^2 \le \mu e^{\lambda/m}/m$. Whence, there exists some constant $C > 0$ such that $\mathbb{E}\big[Z_{Tt}^T\big]/T^2 \le C$, which completes the proof of point (i).

For point (ii), the second inequality is nothing but the first one since $t \in [0, 1]$. We thus only need to prove the first inequality. Let's first recall from (2) that $\tilde{\Psi}^T(t) = \sum_{n\ge1} (\tilde{\varphi}^T)^{*n}(t)$ and from (3) that $\int_0^\infty \tilde{\Psi}^T(t)\,\mathrm{d}t = 1/(a_T - 1)$. Using (12) again, replacing t by Tt for $0 \le t \le 1$,

$$\begin{aligned}\frac{a_T - 1}{T}\mathbb{E}\big[Z_{Tt}^T\big] &= \frac{\mu(a_T-1)}{T}\left[Tt + \int_0^{Tt} s e^{b_T(Tt-s)} e^{-b_T(Tt-s)} \Psi^T(Tt-s)\,\mathrm{d}s\right] \\ &\le \frac{\mu(a_T-1)}{T}\left[T + Te^{b_T Tt}\int_0^\infty \tilde{\Psi}^T(s)\,\mathrm{d}s\right] \qquad \text{(using } 0 \le t \le 1 \text{ and (12))} \\ &\le T\frac{\mu(a_T-1)e^{b_T Tt}}{T}\left(1 + \frac{1}{a_T-1}\right) = \mu a_T e^{b_T Tt},\end{aligned}$$

which concludes the proof. □

3. Proof of the asymptotic behavior of the process

In this section, we prove the law of large numbers and the central limit theorem for the process.

Proof of Theorem 1. Noting from (10), (11), and (5) that $\Psi^T(t) = \sum_{n\ge1} (\varphi^T)^{*n}(t)$ and applying Lemma 2, we have

$$Z_{Tt}^T - \mathbb{E}\big[Z_{Tt}^T\big] = \sum_{n\ge0} (a_T)^n \int_0^{Tt} \varphi^{*n}(Tt-s) M_s^T\,\mathrm{d}s = M_{Tt}^T + \int_0^{Tt} \Psi^T(Tt-s) M_s^T\,\mathrm{d}s,$$

since $\mathbb{E}[M_s^T M_t^T] = \mathbb{E}[Z_{s\wedge t}^T]$. Then, it is easy to conclude from Proposition 3(ii) that, for any $0 \le s, t \le T$,

$$\mathbb{E}\big[\big(M_{s\wedge t}^T\big)^2\big] = \mathbb{E}[Z_{s\wedge t}^T] \le \frac{\mu T a_T}{a_T - 1} e^{b_T s\wedge t}.$$

Also, noting from (4) that $\Psi^T(t) = e^{b_T t}\tilde{\Psi}^T(t)$, we observe that, for any $t \in [0, 1]$,

$$\begin{aligned}&\mathbb{E}\left[\left(\int_0^{Tt} \Psi^T(Tt-s) M_s^T\,\mathrm{d}s\right)^2\right] \\ &\quad= \mathbb{E}\left[\int_0^{Tt}\int_0^{Tt} \Psi^T(Tt-s)\Psi^T(Tt-u) M_s^T M_u^T\,\mathrm{d}s\,\mathrm{d}u\right] \\ &\quad= \int_0^{Tt}\int_0^{Tt} \Psi^T(Tt-s)\Psi^T(Tt-u)\mathbb{E}[M_s^T M_u^T]\,\mathrm{d}s\,\mathrm{d}u \\ &\quad= \int_0^{Tt}\int_0^{Tt} \Psi^T(Tt-s)\Psi^T(Tt-u)\mathbb{E}[Z_{s\wedge u}^T]\,\mathrm{d}s\,\mathrm{d}u \\ &\quad= e^{2b_T Tt}\int_0^{Tt}\int_0^{Tt} \tilde{\Psi}^T(Tt-s)\tilde{\Psi}^T(Tt-u)e^{-b_T(s+u)}\mathbb{E}[Z_{s\wedge u}^T]\,\mathrm{d}s\,\mathrm{d}u \\ &\quad\le e^{2b_T Tt}\int_0^{Tt}\int_0^{Tt} \tilde{\Psi}^T(Tt-s)\tilde{\Psi}^T(Tt-u)e^{-b_T(s\wedge u)}\mathbb{E}[Z_{s\wedge u}^T]\,\mathrm{d}s\,\mathrm{d}u \\ &\quad\le \frac{\mu T a_T}{a_T-1} e^{2b_T Tt}\left(\int_0^\infty \tilde{\Psi}^T(s)\,\mathrm{d}s\right)^2 = \mu T a_T e^{2b_T Tt}(a_T-1)^{-3}.\end{aligned}$$

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

Whence, letting $T \to \infty$,

$$\frac{(a_T-1)^2}{T^2 \mathrm{e}^{2b_T Tt}} \mathbb{E}\big[\big(Z_{Tt}^T - \mathbb{E}[Z_{Tt}^T]\big)^2\big] \le \frac{Ca_T}{T(a_T-1)} \to 0,$$

where C is a positive constant. □

Next, we prove our second main result related to the central limit theorem. The idea of the proof follows from [33, Section 5.3].

Proof of Theorem 2. We divide the proof into several steps.

Step 1. We first rewrite the cumulated intensity of the Hawkes process. Recalling $M_t^T = Z_t^T - \int_0^t \lambda_s^T \, \mathrm{d}s$ and λ_s^T introduced in (1), we have

$$\begin{aligned}\int_0^{Tt} \lambda_s^T \, \mathrm{d}s &= \mu Tt + \int_0^{Tt} \int_0^s \varphi^T(s-m) \, \mathrm{d}Z_m^T \, \mathrm{d}s \\ &= \mu Tt + \int_0^{Tt} \int_0^s \varphi^T(s-m) \, \mathrm{d}M_m^T \, \mathrm{d}s + \int_0^{Tt} \int_0^s \varphi^T(s-m) \lambda_m^T \, \mathrm{d}m \, \mathrm{d}s.\end{aligned}$$

Rearranging, we then get

$$\int_0^{Tt} \lambda_s^T \, \mathrm{d}s - \int_0^{Tt} \int_0^s \varphi^T(s-m) \lambda_m^T \, \mathrm{d}m \, \mathrm{d}s = \mu Tt + \int_0^{Tt} \int_0^s \varphi^T(s-m) \, \mathrm{d}M_m^T \, \mathrm{d}s.$$

By Fubini's theorem and then changing the variable to $u = s - m$, we have

$$\int_0^{Tt} \lambda_s^T \, \mathrm{d}s - \int_0^{Tt} \left(\int_0^{Tt-u} \varphi^T(s) \, \mathrm{d}s \right) \lambda_u^T \, \mathrm{d}u = \mu Tt + \int_0^{Tt} \left(\int_0^{Tt-u} \varphi^T(s) \, \mathrm{d}s \right) \mathrm{d}M_u^T. \tag{13}$$

Define $\Phi^T(t) := \int_0^t \varphi^T(s) \, \mathrm{d}s$ for $t \ge 0$. Then, we have $\Phi^T(t) = 0$ for $t < 0$. Thus, the equality in (13) can be rewritten equivalently as

$$\int_0^{Tt} (1 - \Phi^T(Tt-u)) \lambda_u^T \, \mathrm{d}u = \mu Tt + \int_0^{Tt} \Phi^T(Tt-u) \, \mathrm{d}M_u^T.$$

By a change of variables to $u = Ts$, we get

$$\int_0^t T(1 - \Phi^T(T(t-s))) \lambda_{Ts}^T \, \mathrm{d}s = \mu Tt + \int_0^t \Phi^T(T(t-s)) \, \mathrm{d}M_{Ts}^T.$$

Setting

$$B_t^T = \frac{1}{\sqrt{T}} \int_0^{tT} \frac{\mathrm{d}M_s^T}{\sqrt{\lambda_s^T}}$$

and dividing by T on both sides, we can write

$$\int_0^t T(1 - \Phi^T(T(t-s))) \frac{\lambda_{Ts}^T}{T} \, \mathrm{d}s = \mu t + \int_0^t \Phi^T(T(t-s)) \sqrt{\frac{\lambda_{Ts}^T}{T}} \, \mathrm{d}B_s^T. \tag{14}$$

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

Step 2. In this step, we verify that, for the Skorohod topology,

$$\lim_{T\to\infty}(B_t^T)_{t\in[0,1]} \xrightarrow{\text{(d)}} (B_t)_{t\in[0,1]},$$

where $(B_t)_{t\in[0,1]}$ is a Brownian motion. To prove this, by [25, Theorem VIII-3.11], it suffices to verify, as $T\to\infty$,

(i) the quadratic variation of $(B_t^T)_{t\in[0,1]}$, i.e. $[B^T, B^T]_t \to t$ in probability for all $t\in[0,1]$ fixed;

(ii) $\sup_{t\in[0,1]} |B_t^T - B_{t-}^T| \to 0$ in probability.

It is not difficult to check (ii): for $t\in[0,1]$, we use that the jumps of M^T are always equal to 1 (because the jumps of M^T are counted by Z^T, which are all of size 1), and that λ_t^T are always bigger than $\mu>0$. Hence, we have $\sup_{t\in[0,1]} |B_t^T - B_{t-}^T| \le 1/\sqrt{\mu T}$, which tends to 0 as $T\to\infty$.

Concerning (i), recall that the quadratic variation $[M^T, M^T]_t = Z_t^T$ for $t\in[0,1]$. Then for any $t\in[0,1]$, we have

$$[B^T, B^T]_t = \frac{1}{T}\int_0^{tT} \frac{\mathrm{d}Z_s^T}{\lambda_s^T} = \frac{1}{T}\int_0^{tT} \frac{\mathrm{d}M_s^T}{\lambda_s^T} + \frac{1}{T}\int_0^{tT} \frac{\lambda_s^T}{\lambda_s^T}\,\mathrm{d}s = \frac{1}{T}\int_0^{tT} \frac{\mathrm{d}M_s^T}{\lambda_s^T} + t.$$

Hence, to get (i), it suffices to prove that

$$\mathbb{E}\left[\left(\frac{1}{T}\int_0^{tT} \frac{\mathrm{d}M_s^T}{\lambda_s^T}\right)^2\right] \to 0 \qquad \text{as } T\to\infty.$$

In fact, the Burkholder–Davis–Gundy inequality implies that, for any $t\in[0,1]$,

$$\begin{aligned}\mathbb{E}\left[\left(\frac{1}{T}\int_0^{tT} \frac{\mathrm{d}M_s^T}{\lambda_s^T}\right)^2\right] &\le \frac{C}{T^2}\mathbb{E}\left[\int_0^{tT} \frac{\mathrm{d}Z_s^T}{(\lambda_s^T)^2}\right] \\ &= \frac{C}{T^2}\mathbb{E}\left[\int_0^{tT} \frac{\mathrm{d}M_s^T}{(\lambda_s^T)^2} + \int_0^{tT} \frac{\lambda_s^T\,\mathrm{d}s}{(\lambda_s^T)^2}\right] \\ &\le \frac{C}{T^2}\mathbb{E}\left[\int_0^{tT} \frac{\mathrm{d}M_s^T}{(\lambda_s^T)^2}\right] + \frac{CtT}{\mu T^2} \le \frac{C}{\mu T}.\end{aligned}$$

The last inequality follows from

$$\mathbb{E}\left[\int_0^{tT} \frac{\mathrm{d}M_s^T}{(\lambda_s^T)^2}\right] = 0.$$

Now letting $T\to\infty$, we complete this step.

Step 3. In the third step, we verify that the sequence $(1/T)\int_0^t \lambda_{Ts}^T\,\mathrm{d}s$ is tight in $D[0,1]$ equipped with the Skorohod topology. It is obvious that $(1/T)\int_0^t \lambda_{Ts}^T\,\mathrm{d}s$ is continuous for $t\in[0,1]$. Recalling Z^T and λ^T introduced in (1), we find that

$$\mathbb{E}[Z_{Tt}^T] = \mathbb{E}\left[\int_0^{Tt} \lambda_s^T\,\mathrm{d}s\right] \overset{s=Tv}{=} \mathbb{E}\left[T\int_0^t \lambda_{Tv}^T\,\mathrm{d}v\right].$$

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

Whence, again using Proposition 3(i) for any fixed $x \in [0, 1]$, we conclude that there exists a constant $C > 0$ such that, for any T and any $t \in [0, 1]$,

$$\mathbb{E}\left[\frac{1}{T}\int_0^t \lambda^T_{Ts}\,\mathrm{d}s\right] = \mathbb{E}\left[\frac{Z^T_{Tt}}{T^2}\right] \le \mathbb{E}\left[\frac{Z^T_T}{T^2}\right] \le C.$$

By Prokhorov's theorem, there exists some increasing sequence $\{T_i\}_{i\ge1}$ satisfying $\lim_{i\to\infty} T_i = \infty$ and a continuous process $\phi(t)$ such that $\lim_{i\to\infty}(1/T_i)\int_0^t \lambda^{T_i}_{T_i s}\,\mathrm{d}s = \phi(t)$.

Step 4. In the last step, we prove the result. For any positive smooth function $K(\cdot)$ supported on $\mathbb{R}^+$, taking the convolutions of the both sides of (14) we get

$$\int_0^t\int_0^s K(t-s)T(1-\Phi^T(T(s-s')))\frac{\lambda^T_{Ts'}}{T}\,\mathrm{d}s'\,\mathrm{d}s \tag{15}$$

$$= \mu\int_0^t K(t-s)s\,\mathrm{d}s + \int_0^t\int_0^s K(t-s)\Phi^T(T(s-s'))\sqrt{\frac{\lambda^T_{Ts'}}{T}}\,\mathrm{d}B^T_{s'}\,\mathrm{d}s. \tag{16}$$

For (15), using Fubini's theorem and taking the substitution $s = k + s'$ means it is equal to

$$\begin{aligned}
&\int_0^t\int_{s'}^t K(t-s)T(1-\Phi^T(T(s-s')))\frac{\lambda^T_{Ts'}}{T}\,\mathrm{d}s\,\mathrm{d}s' \\
&\quad= \int_0^t\int_0^{t-s'} K(t-s'-k)T(1-\Phi^T(Tk))\frac{\lambda^T_{Ts'}}{T}\,\mathrm{d}k\,\mathrm{d}s' \\
&\quad\overset{k'=Tk}{=} \int_0^t\int_0^{T(t-s')} K\left(t-s'-\frac{k'}{T}\right)\left(1-\Phi^T(k')\right)\frac{\lambda^T_{Ts'}}{T}\,\mathrm{d}k'\,\mathrm{d}s'.
\end{aligned}$$

Recalling $\Phi^T(t) := \int_0^t \varphi^T(s)\,\mathrm{d}s$ and using Fubini's theorem again, we can easily find that

$$\int_0^\infty (a_T - \Phi^T(k'))\,\mathrm{d}k' = \int_0^\infty\int_{k'}^\infty \varphi^T(s)\,\mathrm{d}s\,\mathrm{d}k' = a_T\int_0^\infty s\varphi(s)\,\mathrm{d}s = ma_T.$$

Hence, by the dominated convergence theorem, we have

$$\begin{aligned}
&\lim_{T\to\infty}\int_0^{T(t-s')} K\left(t-s'-\frac{k'}{T}\right)(1-\Phi^T(k'))\,\mathrm{d}k' \\
&= \lim_{T\to\infty}\left[\int_0^{T(t-s')} K\left(t-s'-\frac{k'}{T}\right)\left(a_T-\Phi^T(k')\right)\mathrm{d}k' + \int_0^{t-s'} K(t-s'-k)T(1-a_T)\,\mathrm{d}k\right] \\
&= K(t-s')\int_0^\infty \varphi(k')\,\mathrm{d}k' - \lambda\int_0^{t-s'} K(t-s'-k)\,\mathrm{d}k \\
&= mK(t-s') - \lambda\int_0^{t-s'} K(u)\,\mathrm{d}u.
\end{aligned}$$

Recall that $(1/T_i)\int_0^t \lambda^{T_i}_{T_i s'}\,\mathrm{d}s' \to \phi_t$ as $i\to\infty$, and using integration by parts and noticing that $\phi_0 = 0$, we have $\int_0^t\left(\int_0^{t-s} K(u)\,\mathrm{d}u\right)\mathrm{d}\phi_s = \int_0^t \phi_s K(t-s)\,\mathrm{d}s$. Hence, taking the limit, (15) turns

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

into

$$\lim_{i\to\infty}\int_0^t\int_0^{t-s'} K(t-s'-k)T_i(1-\Phi_i^T(T_ik))\frac{\lambda_{T_is'}^{T_i}}{T_i}\,\mathrm{d}k\,\mathrm{d}s' = m\int_0^t K(t-s)\,\mathrm{d}\phi_s - \lambda\int_0^t K(t-s)\phi_s\,\mathrm{d}s.$$

Next, we analyze (16). We observe that

$$\lim_{T\to\infty}\int_0^t K(t-s)\Phi^T(T(s-s'))\,\mathrm{d}s = \lim_{T\to\infty}\int_{s'}^t K(t-s)\Phi^T(T(s-s'))\,\mathrm{d}s = \int_{s'}^t K(t-s)\,\mathrm{d}s,$$

since $\lim_{T\to\infty}\Phi^T(T(s-s'))=1$ when $s\geq s'$. Hence,

$$\lim_{i\to\infty}\int_0^t\int_0^s K(t-s)\Phi^{T_i}(T_i(s-s'))\sqrt{\frac{\lambda_{T_is'}^{T_i}}{T_i}}\,\mathrm{d}B_{s'}^{T_i}\,\mathrm{d}s = \int_0^t\left(\int_0^s\sqrt{\frac{\mathrm{d}\phi_{s'}}{\mathrm{d}s'}}\,\mathrm{d}B_{s'}\right)K(t-s)\,\mathrm{d}s.$$

Since (15) = (16), we thus conclude that, for any smooth function K,

$$m\int_0^t K(t-s)\,\mathrm{d}\phi_s - \lambda\int_0^t K(t-s)\phi_s\,\mathrm{d}s = \mu\int_0^t K(t-s)s\,\mathrm{d}s + \int_0^t\left(\int_0^s\sqrt{\frac{\mathrm{d}\phi_{s'}}{\mathrm{d}s'}}\,\mathrm{d}B_{s'}\right)K(t-s)\,\mathrm{d}s,$$

which implies that ϕ_t satisfies

$$m\,\mathrm{d}\phi_t = (\mu t+\lambda\phi_t)\,\mathrm{d}t + \left(\int_0^t\sqrt{\frac{\mathrm{d}\phi_{s'}}{\mathrm{d}s'}}\,\mathrm{d}B_{s'}\right)\mathrm{d}t.$$

Let $X_t = \mathrm{d}\phi_t/\mathrm{d}t$; then

$$X_t = \frac{1}{m}\int_0^t(\mu+\lambda X_s)\,\mathrm{d}s + \frac{1}{m}\int_0^t\sqrt{X_s}\,\mathrm{d}B_s. \tag{17}$$

Moreover, since

$$\frac{\int_0^t\lambda_{Ts}^T\,\mathrm{d}s}{T} = \frac{\int_0^{Tt}\lambda_s^T\,\mathrm{d}s}{T^2},$$

we find that the difference,

$$\frac{Z_{Tt}^T}{T^2} - \frac{\int_0^t\lambda_{Ts}^T\,\mathrm{d}s}{T} = \frac{Z_{Tt}^T}{T^2} - \frac{\int_0^{Tt}\lambda_s^T\,\mathrm{d}s}{T^2} = \frac{M_{Tt}^T}{T^2},$$

is a martingale. Thus, for any $\varepsilon>0$,

$$\mathbb{P}\left(\sup_{t\in[0,1]}\left|\frac{M_{Tt}^T}{T^2}\right|\geq\epsilon\right)\leq\frac{C}{T^4}\mathbb{E}\left[\int_0^T\lambda_s^T\,\mathrm{d}s\right]\to 0$$

as $T\to\infty$. In fact, using Doob's martingale inequality,

$$\mathbb{E}\left[\sup_{t\in[0,1]}\left|\frac{M_{Tt}^T}{T^2}\right|^2\right]\leq\frac{4}{T^4}\mathbb{E}[(M_T^T)^2] = \frac{4}{T^4}\mathbb{E}[N_T^T] = \frac{4}{T^4}\mathbb{E}\left[\int_0^T\lambda_s^T\,\mathrm{d}s\right].$$

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

Consequently, for the Skorohod topology,

$$\lim_{T\to\infty}\frac{Z_{Tt}^T}{T^2}=\lim_{T\to\infty}\frac{\int_0^t\lambda_{Ts}^T\,\mathrm{d}s}{T}\xrightarrow{\text{(d)}}\int_0^t X_s\,\mathrm{d}s\qquad\text{for } t\in[0,1].$$

Finally, we can readily find that the SDE (17) admits a unique strong solution by using [24, Theorem 3.2]. Indeed, the drift coefficient $b(x):=(\mu/m)+(\lambda x/m)$ is Lipschitz continuous and the diffusion coefficient $\sigma(x):=(1/m)\sqrt{x}$ is $\frac{1}{2}$-Hölder continuous on $[0,\infty)$. Together with the previous steps, this concludes the proof. □

Acknowledgements

We would like to greatly thank Sylvain Delattre for fruitful discussions. The authors also thank the anonymous referee for her/his very careful readings and helpful comments that greatly improved the quality of the manuscript.

Funding information

LX is supported by the National Natural Science Foundation of China (12101028). AZ is partially supported by the Beijing Natural Science Foundation (1242009), the National Key R&D Program of China (2024YFA1015300), the National Natural Science Foundation of China (11801536), the China Scholarship Council (202506020208), and the Fundamental Research Funds for the Central Universities (Beihang QYJC Fund and the S & T Innovation Team Support Program – Young Scientist Team).

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

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

[2] Bacry, E. and Muzy, J.-F. (2014). Hawkes model for price and trades high-frequency dynamics. *Quant. Finance* **14**, 1147–1166.

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

[4] Bordenave, C. and Torrisi, G. L. (2007). Large deviations of Poisson cluster processes. *Stoch. Models* **23**, 593–625.

[5] Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Prob.* **24**, 1563–1588.

[6] Cattiaux, P., Colombani, L. and Costa, M. (2022). Limit theorems for Hawkes processes including inhibition. *Stoch. Process. Appl.* **149**, 404–426.

[7] Cox, J. C., Ingersoll Jr., J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica* **53**, 385–407.

[8] Crane, R. and Sornette, D. (2008). Robust dynamic classes revealed by measuring the response function of a social system. *Proc. Nat. Acad. Sci.* **105**, 15649–15653.

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

[10] Daley, D. J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes*, 2nd edn., Vol. **II**. Springer, New York.

[11] Dankel Jr., T. (1991). On the distribution of the integrated square of the Ornstein–Uhlenbeck process. *SIAM J. Appl. Math.* **51**, 568–574.

[12] Delattre, S. and Fournier, N. (2016). Statistical inference versus mean field limit for Hawkes processes. *Electron. J. Statist.* **10**, 1223–1295.

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press

[13] DELATTRE, S., FOURNIER, N. AND HOFFMANN, M. (2016). Hawkes processes on large networks. *Ann. Appl. Prob.* **26**, 216–261.
[14] FILIPOVIĆ, D. (2009). *Term-Structure Models*. Springer, Berlin.
[15] GAO, F. AND ZHU, L. (2018). Some asymptotic results for nonlinear Hawkes processes. *Stoch. Process. Appl.* **128**, 4051–4077.
[16] GAO, F. AND ZHU, L. (2021), Precise deviations for Hawkes processes. *Bernoulli* **27**, 221–248.
[17] GARETTO, M., LEONARDI, E. AND TORRISI, G. L. (2021). A time-modulated Hawkes process to model the spread of COVID-19 and the impact of countermeasures. *Ann. Rev. Control* **51**, 551–563.
[18] HANSEN, N. R., REYNAUD-BOURET, P. AND RIVOIRARD, V. (2015). Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli* **21**, 83–143.
[19] HAWKES, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58**, 83–90.
[20] HAWKES, A. G. (2018). Hawkes processes and their applications to finance: A review. *Quant. Finance* **18**, 193–198.
[21] HAWKES, A. G. AND OAKES, D. (1974). A cluster process representation of a self-exciting process. *J. Appl. Prob.* **11**, 493–503.
[22] HORST, U. AND XU, W. (2024). Functional limit theorems for Hawkes processes. Preprint, arXiv:2401.11495.
[23] HORST, U., XU, W. AND ZHANG, R. (2023). Convergence of heavy-tailed Hawkes processes and the microstructure of rough volatility. Preprint, arXiv:2312.08784.
[24] IKEDA, N. AND WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd edn. (North-Holland Math. Library **24**). North-Holland, Amsterdam.
[25] JACOD, J. AND SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes* (Grundlehren der mathematischen Wissenschaften **288**). Springer, Berlin.
[26] JAISSON, T. AND ROSENBAUM, M. (2015). Limit theorems for nearly unstable Hawkes processes. *Ann. Appl. Prob.* **25**, 600–631.
[27] KARABASH, D. AND ZHU, L. (2015). Limit theorems for marked Hawkes processes with application to a risk model. *Stoch. Models* **31**, 433–451.
[28] LIU, C. (2020). Statistical inference for a partially observed interacting system of Hawkes processes. *Stoch. Process. Appl.* **130**, 5636–5694.
[29] RIZOI, M.-A., MISHRA, S., KONG, Q., CARMAN, M. AND XIE, L. (2018). SIR-Hawkes: Linking epidemic models and Hawkes processes to model diffusions in finite populations. In *WWW '18: Proc. 2018 World Wide Web Conf.*, pp. 419–428.
[30] XU, W. (2024). Diffusion approximations for self-excited systems with applications to general branching processes. *Ann. Appl. Prob.* **34**, 2650–2713.
[31] ZHU, L. (2013). Central limit theorem for nonlinear Hawkes processes. *J. Appl. Prob.* **50**, 760–771.
[32] ZHU, L. (2013). Moderate deviations for Hawkes processes. *Statist. Prob. Lett.* **83**, 885–890.
[33] ZHU, L. (2013). Nonlinear Hawkes processes. PhD thesis, New York University.
[34] ZHU, L. (2014). Process-level large deviations for nonlinear Hawkes point processes. *Ann. Inst. H. Poincaré Prob. Statist.* **50**, 845–871.
[35] ZHU, L. (2015). Large deviations for Markovian nonlinear Hawkes processes. *Ann. Appl. Prob.* **25**, 548–581.

https://doi.org/10.1017/jpr.2025.10047 Published online by Cambridge University Press