Limit theorems for Markovian Hawkes processes with a large initial intensity

Xuefeng Gao [1], Lingjiong Zhu [2]

December 8, 2015

Abstract

Hawkes process is a class of simple point processes that is self–exciting and has clustering effect. The intensity of this point process depends on its entire past history. It has wide applications in finance, neuroscience, social networks, criminology, seismology, and many other fields. In this paper, we study the linear Hawkes process with an exponential kernel in the asymptotic regime where the initial intensity of the Hawkes process is large. We derive limit theorems for this asymptotic regime as well as the regime when both the initial intensity and the time are large.

1 Introduction

Let N be a simple point process on $\mathbb{R}$ and let $\mathcal{F}_t^{-\infty} := \sigma(N(C), C \in \mathcal{B}(\mathbb{R}), C \subset (-\infty, t])$ be an increasing family of σ-algebras. Any nonnegative $\mathcal{F}_t^{-\infty}$-progressively measurable process λ_t with

$$\mathbb{E}\left[N(a,b]|\mathcal{F}_a^{-\infty}\right] = \mathbb{E}\left[\int_a^b \lambda_s ds \middle| \mathcal{F}_a^{-\infty}\right],$$

a.s. for all intervals $(a,b]$ is called an $\mathcal{F}_t^{-\infty}$-intensity of N. We use the notation $N_t := N(0,t]$ to denote the number of points in the interval $(0,t]$.

A Hawkes process is a simple point process N admitting an $\mathcal{F}_t^{-\infty}$-intensity

$$\lambda_t := \lambda\left(\int_{-\infty}^{t-} h(t-s)N(ds)\right), \tag{1.1}$$

where $\lambda(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ is locally integrable, left continuous, $h(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ and we always assume that $\|h\|_{L^1} = \int_0^\infty h(t)dt < \infty$. In (1.1), $\int_{-\infty}^{t-} h(t-s)N(ds)$ stands for $\sum_{\tau<t} h(t-\tau)$, where τ are the occurrences of the points before time t. In the literature, $h(\cdot)$ and $\lambda(\cdot)$ are usually referred to as exciting function (or sometimes kernel function) and rate function respectively. A Hawkes process is linear if $\lambda(\cdot)$ is linear and it is nonlinear otherwise.

The Hawkes process was first proposed by Alan Hawkes in 1971 to model earthquakes and their aftershocks [28]. It naturally generalizes the Poisson process and it captures both

[1]Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T. Hong Kong; xfgao@se.cuhk.edu.hk

[2]Corresponding Author. Department of Mathematics, Florida State University, 1017 Academic Way, Tallahassee, FL-32306, United States of America; zhu@math.fsu.edu.

the self–exciting property and the clustering effect. In addition, Hawkes process is a very versatile model which is amenable to statistical analysis. These explain why it has wide applications in neuroscience, genome analysis, criminology, social networks, healthcare, seismology, insurance, finance and many other fields. For a list of references, we refer to [47].

Throughout this paper, we assume an exponential exciting function $h(t) := \alpha e^{-\beta t}$ where $\alpha, \beta > 0$, and a linear rate function $\lambda(z) := \mu + z$ where the base intensity $\mu \geq 0$. That is, we restrict ourselves to the linear Markovian Hawkes process[3]. To see the Markov property, we define

$$Z_t := \int_{-\infty}^{t} \alpha e^{-\beta(t-s)} N(ds) = Z_0 \cdot e^{-\beta t} + \int_{0}^{t} \alpha e^{-\beta(t-s)} N(ds).$$

Then, the process Z is Markovian and satisfies the dynamics:

$$dZ_t = -\beta Z_t dt + \alpha dN_t, \tag{1.2}$$

where N is a Hawkes process with intensity $\lambda_t = \mu + Z_{t-}$ at time t. In addition, the pair (Z, N) is also Markovian. We also assume $Z_0 = Z_{0-}$, i.e., there is no jump at time zero.

In the literature of financial applications of Hawkes processes, the exponential exciting function and thus the Markovian case, together with the linear rate function, is the most widely used due to the tractability of the theoretical analysis as well as the simulations and calibrations. See, e.g., the survey paper Bacry et al. [3] and the references therein.

In this paper we consider an asymptotic regime where $Z_0 = n$, and $n \in \mathbb{R}^+$ is sent to infinity. We derive limit theorems (functional law of large numbers and functional central limit theorems) for linear Markovian Hawkes processes in this asymptotic regime as well as the regime when both the initial intensity and the time are large. We do not study here the non–Markovian case which may require a different approach since knowing the value of Z_0 is not sufficient to describe the future evolution of a non–Markovian Hawkes process.

Our main results (Theorem 1–4) are mainly oriented to develop approximations for the transient behavior of Hawkes processes with large initial intensity λ_0. Note that $\lambda_0 = \mu + Z_0$, so when Z_0 is large, we have λ_0 large. In practice, this means the intensity of arriving events observed at time zero is considerably larger than the base intensity μ. Our results could be potentially useful for applications where Hawkes process is a relevant model. As an example, consider stock trading. Arrivals of trades or orders are usually clustered in time and can be modeled by Markovian Hawkes processes, see, e.g., Bowsher [7], Da Fonseca and Zaatour [11], Hewlett [30]. If the current trading intensity of a stock is high (a given large value), then our asymptotic analysis can help approximate the volume duration, i.e., the waiting time until a predetermined buy (or sell) volume is traded[4]. Such a volume

[3]Note that, the Hawkes process is in general non–Markovian. When the exciting function is given by sum of exponential functions, the Hawkes process is Markovian. For simplicity of presentation, we focus on a single exponential function in this paper.

[4]Assuming trade size is a constant given by the average trade size over a certain time window.

duration is useful in measuring the (time) costs of liquidity in the trading process, see, e.g., Gourieroux et al. [23], Hautsch [27]. As an another example, consider portfolio credit risk. Hawkes processes have been proposed to capture the clustering effect in the arrival of company defaults in a portfolio, see, e.g., Errias et al. [19]. Our asymptotic analysis can help approximate the number of corporate defaults occurring over a given time interval when the initial intensity of defaults is high such as during a financial crisis.

We now explain the difference between our work and the existing literature. Note that, almost all the existing literature on limit theorems for Hawkes processes are for large time asymptotics, e.g., the functional law of large numbers and functional central limit theorems in Bacry et al. [2], the large deviations principle in Bordenave and Torrisi [6], the moderate deviation principle in Zhu [48] for linear Hawkes processes and see, e.g., Karabash and Zhu [35], Zhu [49] for the limit theorems for extensions of linear Hawkes processes; and the functional central limit theorems in Zhu [50], quantitative Gaussian and Poisson approximations in Torrisi [45, 46], the large deviations in Zhu [51, 52] for nonlinear Hawkes processes. When $\lambda(z) = \mu + z$, note that the large-time limit theorems mentioned above hold under the so–called subcritical regime, that is when $\|h\|_{L^1} < 1$. [5] The critical and supercritical regimes are when $\|h\|_{L^1} = 1$ and $\|h\|_{L^1} > 1$ respectively. Those different regimes first appeared in Zhu [47] and the terminology comes from the Galton–Watson trees because the linear Hawkes process has the well-known immigration-birth representation, see, e.g., [29]. The limit theorems for critical, supercritical and other regimes have been studied in Zhu [47]. The limit theorems for the nearly critical, or so–called nearly unstable case was studied in Jaisson and Rosenbaum [33, 34].

Other than the large time asymptotics, the large dimensional asymptotics of Hawkes processes have been studied recently, see e.g. Delattre et al. [16], that is, the asymptotics for the multivariate nonlinear Hawkes process where the number of dimension goes to infinity. The mean-field limit was obtained in [16] for different regimes. See also Chevallier [10], Hodara and Löcherbach [32], Delattre and Fournier [15].

To the best of our knowledge, this is the first paper to study the large initial value asymptotics ($Z_0 = n \to +\infty$) in the context of the Hawkes process. Large initial value asymptotics have been studied in the many different facets of applied probability. For example, in queueing networks, if the initial total queue length is large, one may derive a deterministic model so called fluid limit for the system. Such a fluid limit model is useful in understanding the stability and thus the performance of multiclass queueing networks, see, e.g., Bramson [8], Dai and Meyn [13]. As an another example, consider mathematical biology where the large initial value corresponds to the large initial population. In the context of birth–death processes, the large initial value asymptotics have been obtained as the applications to study the cancer dynamics, see e.g. Foo and Leder [20] and Foo et al. [21] and the references therein.

The rest of the paper is organized as follows. In Section 2, we state our main results.

[5] Note that one can also define different regimes for nonlinear Hawkes processes, see Zhu [47].

In Section 3, we present the proofs of the main results. The proofs of auxiliary results are collected in the appendix.

2 Main results

In this section we state our main results. Note that processes Z and N both depend on the initial condition $Z_0 = n$ and one can use $Z^{(n)}, N^{(n)}$ to emphasize the dependence on $Z_0 = n$. For simplicity of notations, we use Z_t, N_t to denote $Z_t^{(n)}, N_t^{(n)}$ throughout this paper. Write $D[a, b]$ as the space of càdlàg processes on $[a, b] \subset [0, \infty)$ that are equipped with Skorohod J_1 topology (see e.g., Billingsley [4]).

2.1 Limit theorems with large Z_0

In this section, we present limit theorems including functional law of large numbers (FLLN) and functional central limit theorem (FCLT) for processes Z and N in the asymptotic regime where $Z_0 = n \to \infty$.

We first state the FLLN for the processes Z and N.

Theorem 1. *Fix any $T > 0$. As $n \to \infty$, we have*

$$\sup_{0 \le t \le T} \left| \frac{Z_t}{n} - e^{(\alpha-\beta)t} \right| \to 0, \qquad \textit{almost surely}, \tag{2.1}$$

$$\sup_{0 \le t \le T} \left| \frac{N_t}{n} - \psi(t) \right| \to 0, \qquad \textit{almost surely}, \tag{2.2}$$

where

$$\psi(t) := \begin{cases} \frac{e^{(\alpha-\beta)t}-1}{\alpha-\beta}, & \alpha \neq \beta, \\ t, & \alpha = \beta. \end{cases} \tag{2.3}$$

We next state the FCLT for the processes Z and N. Recall that a Gaussian process is called centered if its mean function is identically zero.

Theorem 2. *For any $T > 0$, as $n \to \infty$, the sequence of processes*

$$\left\{ \frac{Z_t - ne^{(\alpha-\beta)t}}{\sqrt{n}} : t \in [0, T] \right\} \tag{2.4}$$

converges in distribution to a centered Gaussian process G on $D[0, T]$, where the covariance function of G is given as follows: for $0 \le s \le t$,

$$Cov(G_s, G_t) = \begin{cases} \frac{\alpha^2}{\alpha-\beta}(e^{(\alpha-\beta)(t+s)} - e^{(\alpha-\beta)t}), & \alpha \neq \beta, \\ \alpha^2 s, & \alpha = \beta. \end{cases} \tag{2.5}$$

Furthermore, the sequence of re–normalized Hawkes processes

$$\left\{ \frac{N_t - n \cdot \psi(t)}{\sqrt{n}} : t \in [0,T] \right\} \tag{2.6}$$

converges in distribution to a centered Gaussian process H on $D[0,T]$, where H is given by

$$H_t := \frac{G_t}{\alpha} + \frac{\beta}{\alpha} \int_0^t G_s ds, \quad t \in [0,T]. \tag{2.7}$$

The covariance function of H is given in (3.27) *and* (3.28).

While we focus on the setting $\alpha, \beta > 0$, it is helpful to consider the case $\alpha = \beta = 0$ to develop some intuition about the above results. In such a special case we have $Z_t \equiv Z_0 = n$ for all t, and the Hawkes process N becomes a homogeneous Poisson process with rate $\mu + n$. The large–intensity limits of Poisson processes are well–known: when μ is fixed and $n \to \infty$, the Poisson process N satisfies the FLLN in (2.2). In addition, the sequence of re–normalized Poisson processes in (2.6) converges in distribution to a standard Brownian motion. This is consistent with our FCLT for the case $\alpha = \beta$ where the limiting Gaussian process H has a covariance function

$$\mathrm{Cov}(H_t, H_s) = s + \alpha s t + \alpha^2 \left(\frac{ts^2}{2} - \frac{s^3}{6} \right), \quad \text{for } 0 \le s \le t,$$

and H becomes a standard Brownian motion when $\alpha = 0$.

We next discuss several properties of the limiting Gaussian processes G and H. Note that the theory of Gaussian processes is rich. We refer the reader to textbooks by, e.g., Hida and Hitsuda [31], Lifshits [38] and Piterbarg [42] for introduction and comprehensive treatment of Gaussian processes.

- It is easy to see from the covariance functions that the limiting Gaussian processes G and H are both non–stationary.

- Recall (see, e.g., Revuz and Yor [43, p.86]) that a centered Gaussian process Υ with covariance function $\Gamma(s,t) := \mathbb{E}[\Upsilon_s \Upsilon_t]$ is Markovian if and only if

  $$\Gamma(s,u)\Gamma(t,t) = \Gamma(s,t)\Gamma(t,u).$$

  for every $0 \le s < t < u$. Thus we can directly check from the covariance functions of G and H that the Gaussian process G is a Markov process and H is not. In addition, we deduce from the Markov property of G and (2.7) that the joint process (G,H) is also Markovian. This is not surprising since we know for each fixed $Z_0 = n$, the process Z is Markovian, the point process N is not Markovian, but the pair (Z,N) is Markovian.

- In the critical case when $\alpha = \beta$, one readily finds from (2.5) that the limiting Gaussian process G is actually a Brownian motion with drift zero and variance α^2.

- In both the sub–critical and super–critical cases where $\alpha \neq \beta$, the limiting Gaussian process G with covariance (2.5) also has a version with continuous sample path on the bounded time interval $[0, T]$. Hence the process H in (2.7) inherits such continuity property. One can readily verify this as follows: note from (2.5) that $G_{t+h} - G_t$ is a Gaussian random variable with mean zero and variance given by

$$\mathbb{E}[(G_{t+h} - G_t)^2] = \frac{\alpha^2}{\alpha - \beta} \cdot [e^{(\alpha-\beta)h} - 1] \cdot e^{(\alpha-\beta)t} \cdot (e^{(\alpha-\beta)(t+h)} + 1 - e^{2(\alpha-\beta)t}).$$

  Using the facts that $0 \leq t \leq t+h \leq T$ and $\frac{1}{\alpha-\beta} \cdot [e^{(\alpha-\beta)h} - 1] \leq Ch$ for some constant C depending on α, β and T only, we deduce that

$$\mathbb{E}[(G_{t+h} - G_t)^4] = 3 \left(\mathbb{E}[(G_{t+h} - G_t)^2]\right)^2 \leq C' \cdot h^2, \tag{2.8}$$

  where C' is a also constant only depending on α, β and T. Hence the continuity result follows from Kolmogrov's continuity criterion (see, e.g., Revuz and Yor [43, p.19]).

To illustrate the usefulness of our limit theorems, let us consider the first passage time problem mentioned in the introduction. For a given volume K, we are interested the waiting time τ_K defined by $\tau_K := \inf\{t > 0 : N_t \geq K\}$ where by convention the infimum of an empty set is $+\infty$. When $Z_0 = n$ is large, we can approximate the distribution of τ_K using Theorem 2 as follows: for each $0 < t < \infty$,

$$\mathbb{P}(\tau_K \leq t) = \mathbb{P}(N_t \geq K) \approx \mathbb{P}(n\psi(t) + \sqrt{n}H_t \geq K) = 1 - \Phi\left(\frac{K - n\psi(t)}{\sqrt{n \cdot Var(H_t)}}\right), \tag{2.9}$$

where $\psi(\cdot)$ is given in (2.3), $Var(H_t)$ is the variance of H_t and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable. Thus the probability in (2.9) can be readily computed after one fits a Markovian Hawkes processes to the data and estimates the parameters Z_0, α and β. See, e.g., Ozaki [41], Daley and Vere–Jones [12, Chapter 7] for estimation methods and related issues.

2.2 Limit theorems with large Z_0 and large time

In this section, we present limit theorems for processes Z and N in the asymptotic regime where both $Z_0 = n$ and the time go to infinity. Such limit theorems could provide insights on the 'macroscopic' behavior of the Hawkes processes with large initial intensity.

When the time is sent to infinity, Hawkes processes behave differently depending on the value of $\|h\|_{L^1}$ (see, e.g., Zhu [47]). In our case, the exciting function is exponential: $h(t) = \alpha e^{-\beta t}$. So we have the following different cases/regimes:

- Super–critical: $\alpha > \beta$;
- Sub–critical: $\alpha < \beta$;
- Critical: $\alpha = \beta$;
- Nearly–critical: $\alpha \approx \beta$.

We study each case separately.

We first state the FLLN for the process Z and the point process N when Z_0 and the time are both large.

Theorem 3. *(i) (Super–Critical Case) Assume that $\alpha > \beta > 0$ and let $\tau_n = \frac{\log n}{\alpha - \beta}$. For any $T > 0$, as $n \to \infty$ we have*

$$\sup_{0 \le s \le T} \left| \frac{Z_{s\tau_n}}{n^{1+s}} - 1 \right| \to 0, \qquad \textit{almost surely,} \tag{2.10}$$

$$\sup_{0 \le s \le T} \left| \frac{N_{s\tau_n}}{n^{1+s}} - \frac{1}{\alpha - \beta} + \frac{1}{(\alpha - \beta) n^s} \right| \to 0, \qquad \textit{almost surely.} \tag{2.11}$$

(ii) (Sub–Critical Case) Assume that $\beta > \alpha > 0$ and let $t_n = \frac{\log n}{\beta - \alpha}$. Then, for any $0 < T < 1$, as $n \to \infty$ we have

$$\sup_{0 \le s \le T} \left| \frac{Z_{st_n}}{n^{1-s}} - 1 \right| \to 0, \qquad \textit{in probability,} \tag{2.12}$$

$$\sup_{0 \le s \le T} \left| \frac{N_{st_n}}{n} - \frac{1}{\beta - \alpha} + \frac{1}{(\beta - \alpha) n^s} \right| \to 0, \qquad \textit{in probability.} \tag{2.13}$$

We have almost surely convergence for (2.12) *and* (2.13) *if* $0 < T < \frac{1}{2}$.

A few remarks are in order.

- We choose to speed up the time by a factor of $\log n$ and scale the space accordingly to ensure the convergence to a non–trivial limit. Such a scaling simplifies our presentation and it is natural in view of (2.1) and (2.2) in Theorem 1.
- In the sub–critical case, we restrict $T < 1$. One can easily show that $\mathbb{E}[Z_{st_n}] = n^{1-s}$ (see Proposition 5), which implies that $\mathbb{E}[Z_{st_n}] = 1$ if $s = 1$. Therefore, the process Z starts with a large number $Z_0 = n$, and after t_n unit of time, it drops to one on average (Z is always positive). So it makes sense to consider the law of large numbers for Z_{st_n}/n^{1-s} for $0 \le s \le T$ where $0 < T < 1$. Due to our proof technique, the almost surely convergence here is restricted to $T < \frac{1}{2}$.

- We do not state the FLLN in the critical or nearly–critical case since the limit process is trivial under appropriate time and space scalings, as readily seen from the FCLT result below.

We next state the FCLT for the processes Z and N when $Z_0 = n$ is large and the time is speeded up in an appropriate way.

Theorem 4. *(i)(Critical and Nearly Critical Cases) Fix $\beta > 0$ and $\gamma \in \mathbb{R}$. For $Z_0 = n$, define $\alpha_n = \beta + \frac{\gamma}{n}$, which is positive for all large n. Assume that*

$$dZ_t = -\beta Z_t dt + \alpha_n dN_t,$$

where the point process N has intensity $\mu + Z_{t-}$ at time t. Then as $n \to \infty$, we have the sequence of processes $\left\{\frac{Z_{tn}}{n} : t \in [0,T]\right\}$ converges in distribution to the process X on $D[0,T]$, where X satisfies

$$dX_t = (\beta\mu + \gamma X_t)dt + \beta\sqrt{X_t}dB_t, \qquad X_0 = 1, \tag{2.14}$$

where B is a standard Brownian motion. In addition, we have the sequence of re–normalized Hawkes processes $\left\{\frac{N_{tn}}{n^2} : t \in [0,T]\right\}$ converges in distribution toward the process

$$\left\{\int_0^t X_s ds : t \in [0,T]\right\}$$

on $D[0,T]$.

(ii)(Super–Critical Case) Assume that $\alpha > \beta > 0$ and let $\tau_n = \frac{\log n}{\alpha-\beta}$. For any $0 < t < T < \frac{1}{2}$, as $n \to \infty$, we have the sequence of processes

$$\left\{\frac{Z_{s\tau_n} - n^{1+s}}{\sqrt{n^{1+2s}}}, \quad s \in [t,T]\right\}, \tag{2.15}$$

converges in distribution to the process Y on $D[t,T]$, where $Y_s \equiv \xi$ for $s \in [t,T]$ and ξ is a normal random variable with mean 0 and variance $\frac{\alpha^2}{\alpha-\beta}$. Moreover, we have

$$\left\{\frac{N_{s\tau_n} - \frac{n^{1+s}-n}{\alpha-\beta}}{n^{\frac{1}{2}+s}}, \quad s \in [t,T]\right\}, \tag{2.16}$$

converges in distribution to the process $\frac{Y}{\alpha-\beta}$ on $D[t,T]$.

(iii) (Sub–Critical Case) Assume that $\beta > \alpha > 0$ and let $t_n = \frac{\log n}{\beta-\alpha}$. For any $0 < T < 1$, we have all the finite dimensional distributions of the processes

$$\left\{\frac{Z_{st_n} - n^{1-s}}{\sqrt{n^{1-s}}}, \quad s \in [0,T]\right\}, \tag{2.17}$$

converges in distribution to the corresponding finite dimensional distributions of a centered Gaussian process R with covariance function $Cov(R_u, R_v) = \frac{\alpha^2}{\beta-\alpha}$ if $u = v > 0$ and zero otherwise.

Discussions on Theorem 4

- Since $\beta\mu \geq 0$ and $X_0 = 1$, there is a unique strong solution $X \geq 0$ to the stochastic differential equation (2.14), see, e.g., Chapter IX in Revuz and Yor [43]. The process X is known as a square–root diffusion or a Cox–Ingersoll–Ross (CIR) process, which is widely used in modeling interest rate and stochastic volatility in financial mathematics. In addition, the properties of X and the time integral of X are well studied. See, e.g., Dufresne [17] and Göing–Jaeschke and Yor [22].

- In the critical case where $\gamma = 0$, the process $\frac{4}{\beta^2}X$ with X defined by (2.14) is also known as the square of a $\frac{4\mu}{\beta}$–dimensional Bessel process. See, e.g., Revuz and Yor [43]. See also Hamana and Matsumoto [24] and [25, 26] for further properties of Bessel processes such as first passage times.

- In the nearly critical case when $\gamma < 0$, we have $0 < \alpha_n < \beta$ for each large n and α_n/β approaches one as $n \to \infty$. Such Hawkes processes are called nearly unstable in Jaisson and Rosenbaum [33]. In line with the results in [33], we obtain, in this nearly unstable case, mean–reverting CIR process in the limit for the process Z where Z_0 and the time are both sent to infinity. We also remark that since we consider a different asymptotic regime compared with [33], we obtain a different initial condition for the limiting process X in (2.14). In addition, our result holds for general $\gamma \in \mathbb{R}$.

- In the super–critical case, the weak limit of the sequence of processes in (2.15) is a constant normal random variable ξ. Note that this process–level convergence is restricted to $D[t,T]$ for $T > t > 0$, and such weak convergence can not be extended to the space $D[0,T]$, which is readily seen after noting that $Z_0 = n$.

- In the sub–critical case, the centered Gaussian process R, with covariance function $\mathrm{Cov}(R_u, R_v) = \frac{\alpha^2}{\beta-\alpha}$ if $u = v$ and zero otherwise, is known to exist. But such a Gaussian process does not have a measurable version and hence a càdlàg version, see, e.g. Revuz and Yor [43, p.37]. Therefore, we only have the convergence of finite dimensional distributions for the sequence of processes in (2.17), but we do not have process–level convergence. In other words, the sequence of processes in (2.17) is not tight. Hence we also do not have or state weak convergence result for the Hawkes process N (when both Z_0 and the time are large) in this sub–critical case.

3 Proofs of Main Results

In this section we gather the proofs of our main results Theorem 1–4. For notational simplicity, in all the proofs we use $C > 0$ as a generic constant which may vary line from line. The constant C may depend on α, β, γ and T, but it is independent of n.

Our proof strategy is to first show the functional law of large numbers and functional central limit theorems for processes Z and N when the base intensity μ is zero, and then extend the proofs to the case when $\mu > 0$. Such a strategy relies critically on the observation described in the next section.

3.1 Decomposition of Hawkes processes

This section presents a decomposition result for linear Hawkes processes.

The linear Hawkes process has the well–known immigration birth representation, see, e.g., [29]. That is, the immigrant arrives according to a homogeneous Poisson process with constant rate μ. Each immigrant would produce children and the number of children has Poisson distribution with parameter $\|h\|_{L^1} = \frac{\alpha}{\beta}$. Conditional on the number of the children of an immigrant, the time that a child was born has probability density function $\frac{h(t)}{\|h\|_{L^1}} = \beta e^{-\beta t}$. Each child would produce children according to the same laws independent of other children. All the immigrants produce children independently. The number points of a linear Hawkes process on a time interval $[0, t]$ equals the total number of immigrants and the descendants on the interval $[0, t]$.

Recall that we are interested in a Hawkes process N with intensity $\mu + Z_{t-}$ at time t and $Z_0 = n$, and

$$Z_t = ne^{-\beta t} + \int_0^t \alpha e^{-\beta(t-s)} dN_s.$$

By the immigration-birth representation, we can decompose linear Hawkes process as

$$N_t = N_t^{(0)} + N_t^{(1)}, \quad t \geq 0, \tag{3.1}$$

where $N_t^{(0)}$ is the number of points of the immigrants that arrive according to an inhomogenous Poisson process with rate $ne^{-\beta t}$ and all the descendants on the interval $[0, t]$ and $N_t^{(1)}$ is the number of points of the immigrants that arrive according to a homogeneous Poisson process with rate μ and all the descendants on the interval $[0, t]$. Therefore, $N^{(0)}$ is a simple point process with intensity $Z^{(0)}$, where

$$Z_t^{(0)} = ne^{-\beta t} + \int_0^t \alpha e^{-\beta(t-s)} dN_s^{(0)},$$

and in the differential form,

$$dZ_t^{(0)} = -\beta Z_t^{(0)} dt + \alpha dN_t^{(0)}, \quad Z_0^{(0)} = n.$$

$N^{(1)}$ is a simple point process with intensity $\lambda^{(1)}$ where

$$\lambda_t^{(1)} := \mu + Z_t^{(1)} = \mu + \int_0^t \alpha e^{-\beta(t-s)} dN_s^{(1)},$$

where

$$Z_t^{(1)} = \int_0^t \alpha e^{-\beta(t-s)} dN_s^{(1)},$$

and in the differential form,

$$dZ_t^{(1)} = -\beta Z_t^{(1)} dt + \alpha dN_t^{(1)}, \quad Z_0^{(1)} = 0.$$

In addition, the two point processes $N^{(0)}$ and $N^{(1)}$ are independent of each other. As a result, we also have

$$Z_t = Z_t^{(0)} + Z_t^{(1)}, \quad t \geq 0, \tag{3.2}$$

and the processes $Z^{(0)}$ and $Z^{(1)}$ are also independent of each other. An alternative way to see the validity of decompositions (3.1) and (3.2) is via the Poisson embedding technique often used in the theory of point processes [9].

Now we illustrate the very high level idea of our proofs of limit theorems. The first step is to study the Hawkes process N with $\mu = 0$. In the decomposition $N = N^{(0)} + N^{(1)}$, we note that $N^{(1)}$ is a linear Hawkes process which is empty on $(-\infty, 0]$. Thus when $\mu = 0$, $N_t^{(1)} \equiv 0$ for any $t \geq 0$. So we have $N = N^{(0)}$ and $Z = Z^{(0)}$ in this case. Once we have established limit theorems for $Z^{(0)}$ and $N^{(0)}$, we move to the second step: consider Hawkes process with $\mu > 0$. In view of the decompositions in (3.1) and (3.2), it suffices to have limit theorems for $N^{(1)}$ and $Z^{(1)}$ when $n \to \infty$. This is relatively straightforward: when we consider a finite time horizon, then $N^{(1)}$ and $Z^{(1)}$ do not contribute in the scaling limits as they are independent of $Z_0 = Z_0^{(0)} = n$ which goes to infinity; When we consider the time also goes to infinity, the functional law of large numbers and central limit theorems of $N^{(1)}$ have already been well studied in the literature. So combining the limit theorems we obtained for $N^{(0)}$ and $Z^{(0)}$ in the first step, we have the limit theorems for processes N and Z.

3.2 Preliminaries

This section presents preliminaries for proving the main results. Throughout this section, N is a simple point process and its intensity at time t is given by Z_{t-} where $dZ_t = -\beta Z_t dt + \alpha dN_t$. That is, we consider $\mu = 0$.

We have the following immediate observation. We can express the jump process $N_t = N(0, t]$ in terms of the intensity process Z_t in the following way,

$$N_t = \frac{Z_t - Z_0}{\alpha} + \frac{\beta}{\alpha} \int_0^t Z_s ds. \tag{3.3}$$

This relation is very useful: once we prove limit theorems for process Z, the identity (3.3) allows us to prove limit theorems for process N in a direct way.

Define for $t \geq 0$,

$$M_t := N_t - \int_0^t Z_s ds.$$

Then M is a martingale with respect to the natural filtration, and the predictable quadratic variation $\langle M \rangle_t$ of this martingale is given by $\int_0^t Z_s ds$, see, e.g. [39]. It is readily seen that

$$\begin{aligned} d(Z_t - ne^{(\alpha-\beta)t}) &= -\beta Z_t dt + \alpha dN_t - (\alpha - \beta) ne^{(\alpha-\beta)t} dt \\ &= (\alpha - \beta)(Z_t - ne^{(\alpha-\beta)t}) dt + \alpha dM_t. \end{aligned}$$

Therefore, we have

$$Z_t - ne^{(\alpha-\beta)t} = e^{(\alpha-\beta)t} \alpha \int_0^t e^{-(\alpha-\beta)s} dM_s. \tag{3.4}$$

We next summarize two results on the moments of Z_t and related inequalities. The proofs are given in the appendix.

Proposition 5. *Suppose $\alpha, \beta > 0$. We have*

$$\mathbb{E}[Z_t | Z_0] = Z_0 e^{(\alpha-\beta)t}.$$

If $\alpha = \beta$, we have

$$\begin{aligned} \mathbb{E}[Z_t^2 | Z_0] &= Z_0^2 + \alpha^2 Z_0 t, \\ \mathbb{E}[Z_t^3 | Z_0] &= Z_0^3 + 3\alpha^2 Z_0^2 t + \frac{3}{2} \alpha^4 Z_0 t^2 + \alpha^3 Z_0 t. \end{aligned}$$

If $\alpha \neq \beta$, we have

$$\begin{aligned} \mathbb{E}[Z_t^2 | Z_0] &= Z_0^2 e^{2(\alpha-\beta)t} + \frac{\alpha^2 Z_0}{\alpha - \beta} (e^{2(\alpha-\beta)t} - e^{(\alpha-\beta)t}), \\ \mathbb{E}[Z_t^3 | Z_0] &= \left(Z_0^3 + \frac{3\alpha^2 Z_0^2}{\alpha - \beta} + \frac{3\alpha^4 Z_0}{(\alpha-\beta)^2} + \frac{\alpha^3 Z_0}{2(\alpha-\beta)} - \frac{3\alpha^4 Z_0}{2(\alpha-\beta)^2} \right) e^{3(\alpha-\beta)t} \\ &\quad - \left(\frac{3\alpha^2 Z_0^2}{\alpha-\beta} + \frac{3\alpha^4 Z_0}{(\alpha-\beta)^2} \right) e^{2(\alpha-\beta)t} - \left(\frac{\alpha^3 Z_0}{2(\alpha-\beta)} - \frac{3\alpha^4 Z_0}{2(\alpha-\beta)^2} \right) e^{(\alpha-\beta)t}. \end{aligned}$$

Proposition 6. *Given $Z_0 = n$ sufficiently large, we have*

$$\sup_{0 \leq t \leq T} \mathbb{E}\left[\left(Z_t - ne^{(\alpha-\beta)t} \right)^4 \right] \leq Cn^2, \tag{3.5}$$

$$\sup_{\delta \leq t \leq T} \mathbb{E}\left[\left(\int_{t-\delta}^t e^{-(\alpha-\beta)s} dM_s \right)^4 \right] \leq Cn^2 \delta^2, \quad \textit{for } \delta \in [0, t]. \tag{3.6}$$

In addition, when $\alpha \neq \beta$ we have for any $t \geq 0$

$$\mathbb{E}\left[\left(\int_0^t e^{-(\alpha-\beta)s} dM_s\right)^4\right] \tag{3.7}$$
$$\leq Ct\left[\frac{n^2}{2(\alpha-\beta)}(1-e^{-2(\alpha-\beta)t}) + \frac{\alpha^2 n}{2(\alpha-\beta)^2}(1-e^{-2(\alpha-\beta)t}) - \frac{\alpha^2 n}{3(\alpha-\beta)^2}(1-e^{-3(\alpha-\beta)t})\right].$$

Here C is a constant that may depend on α, β, γ and T, but it is independent of n and t.

3.2.1 Moment generating function of Z_t

In this section we present a result on moment generating function of Z_t, which can be found in Errais et al. [19] for example. We include it here since it is a critical tool in establishing Gaussian limit in the functional central limit theorems. Indeed, in [19], they only discussed the Laplace transforms. But for the purpose of showing the convergence of distribution, we need convergence of moment generating functions in a neighborhood of zero.

We first note that for a Markovian Hawkes process N with intensity being Z_{t-} at time t, the Markov process Z has the infinitesimal generator

$$\mathcal{A}f(z) = -\beta z \frac{\partial f}{\partial z} + z[f(z+\alpha) - f(z)],$$

where $f \in C_b^1$. Indeed, the operator $\mathcal{A}$ can be applied to unbounded functions f and the boundedness assumption can be relaxed to $\mathbb{E}[f(Z_t)]$ finite for every $t > 0$ in many applications, e.g. Kolmogorov backward equation, see e.g. Remark 6.2.1. and Proposition 6.3.1 in [5].

It is proved in Zhu [50] that there exists some $\theta_c > 0$ so that for any $\theta < \theta_c$, $\mathbb{E}[e^{\theta N_t}] < \infty$. In our case, $Z_t \leq Z_0 + \alpha N_t$ and thus for any $\theta < \theta_c$, $\mathbb{E}[e^{\theta Z_t}] < \infty$. Therefore, we can apply Proposition 6.3.1. in [5] and for $-\theta < \theta_c$, the function $u(z,t) := \mathbb{E}[e^{-\theta Z_t}|Z_0 = z]$ satisfies the Kolmogorov backward equation:

$$\begin{aligned}
&\frac{\partial u}{\partial t} = -\beta z \frac{\partial u}{\partial z} + z[u(z+\alpha, t) - u(z,t)],\\
&u(z,0) = e^{-\theta z}.
\end{aligned}$$

The above equation is affine and we can try the Ansatz $u(z,t) = e^{A(t)z + B(t)}$. Substituting the Ansatz into the equations above we obtain

$$A'(t) = -\beta A(t) + e^{A(t)\alpha} - 1, \tag{3.8}$$
$$A(0) = -\theta, \tag{3.9}$$
$$B'(t) = 0, \qquad B(0) = 0.$$

This implies that $B(t) = 0$ for any $t \geq 0$. For convenience, in the following we will write $A(t, -\theta)$ instead of $A(t)$ to emphasize that A takes value $-\theta$ at time 0. Then we have

$$u(z,t) := \mathbb{E}[e^{-\theta Z_t} | Z_0 = z] = e^{A(t,-\theta)z}. \tag{3.10}$$

We remark that in general the function A is not explicit in the sense that there is no closed-form solution to the differential equation system given by (3.8) and (3.9). One key idea in establishing our central limit theorems is that we exploit perturbation theory (see, e.g., [40]) of differential equations to show the convergence of certain moment generating functions.

3.3 Proof of Theorem 1

We prove Theorem 1 in this section. We first consider the case $\mu = 0$, and then extend the proof to the case $\mu > 0$ using the observation in Section 3.1.

3.3.1 Proof of Theorem 1 when $\mu = 0$

Proof of (2.1). When $\mu = 0$, let us recall from (3.4) that

$$Z_t - ne^{(\alpha-\beta)t} = e^{(\alpha-\beta)t}\alpha \int_0^t e^{-(\alpha-\beta)s} dM_s.$$

Together with Doob's martingale inequality, we have for any $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\leq t\leq T}\left|\frac{Z_t}{n} - e^{(\alpha-\beta)t}\right| \geq \epsilon\right) \\
&\leq \mathbb{P}\left(\alpha(1+e^{|\alpha-\beta|T}) \sup_{0\leq t\leq T}\left|\int_0^t e^{-(\alpha-\beta)s} dM_s\right| \geq n\epsilon\right) \\
&\leq \frac{\alpha^4(1+e^{|\alpha-\beta|T})^4}{n^4\epsilon^4}\mathbb{E}\left[\left(\int_0^T e^{-(\alpha-\beta)s} dM_s\right)^4\right].
\end{aligned}$$

On combining with inequality (3.7) we obtain

$$\mathbb{P}\left(\sup_{0\leq t\leq T}\left|\frac{Z_t}{n} - e^{(\alpha-\beta)t}\right| \geq \epsilon\right) \leq \frac{Cn^2 + Cn}{n^4\epsilon^4}.$$

Since $\sum_{n=1}^{\infty} \frac{Cn^2+Cn}{n^4\epsilon^4}$ is finite, the result (2.1) follows from the Borel-Cantelli lemma. □

Proof of (2.2). By (2.1), we have

$$\sup_{0\leq t\leq T}\left|\frac{\int_0^t Z_s ds}{n} - \psi(t)\right| \to 0, \qquad \text{almost surely as } n \to \infty,$$

where $\psi(t) = \int_0^t e^{(\alpha-\beta)s} ds$. Since $M_t = N_t - \int_0^t Z_s ds$, thus in order to show (2.2), it suffices to show that

$$\sup_{0 \le t \le T} \frac{|M_t|}{n} \to 0, \qquad \text{almost surely as } n \to \infty. \tag{3.11}$$

Similar to the proof of (2.1), we can apply Doob's martingale inequality and (3.7) to show that for any $\epsilon > 0$

$$\mathbb{P}\left(\sup_{0 \le t \le T} \frac{|M_t|}{n} \ge \epsilon \right) \le \frac{C}{n^4 \epsilon^4} \mathbb{E} M_T^4 \le \frac{1}{n^4 \epsilon^4} \cdot C(n^2 + n).$$

Thus (3.11) follows from Borel-Cantelli lemma.

3.3.2 Proof of Theorem 1 when $\mu > 0$

As described in Section 3.1, when $\mu > 0$, we can decompose $Z_t = Z_t^{(0)} + Z_t^{(1)}$ where $Z^{(0)}$ and $Z^{(1)}$ are independent. We have established in the previous section that

$$\sup_{0 \le t \le T} \left| \frac{Z_t^{(0)}}{n} - e^{(\alpha-\beta)t} \right| \to 0, \quad \text{almost surely.}$$

In addition, note that $Z_t^{(1)}$ is independent of $Z_0^{(0)} = n$ and hence $\sup_{0 \le t \le T} Z_t^{(1)}/n \to 0$ almost surely as $n \to \infty$. Now (2.1) immediately follows.

Similarly, we have $N_t = N_t^{(0)} + N_t^{(1)}$ when $\mu > 0$. Since $N_t^{(1)}$ is independent of the parameter n, we obtain $N_T^{(1)}/n \to 0$ almost surely as $n \to \infty$. Thus (2.2) follows.

3.4 Proof of Theorem 2

We prove Theorem 2 in this section.

3.4.1 FCLT for Z when $\mu = 0$

In this section we prove the weak convergence of re–normalized processes of Z in (2.4) on $D[0,T]$ for the case $\mu = 0$. For notational simplicity, we define for $Z_0 = n$ and each $t \ge 0$

$$\tilde{Z}_t := \frac{Z_t - n e^{(\alpha-\beta)t}}{\sqrt{n}}. \tag{3.12}$$

Our approach is to apply Theorem 13.5 in Billingsley [4]. In particular, we verify the following three conditions: (recall that G is a centered Gaussian process with covariance (2.5))

(a) $G_t - G_{t-\Delta}$ converge to zero in distribution as $\Delta \to 0$.

(b) Finite–dimensional distributions of $\tilde{Z}$ converge to those of G.

(c) For $0 \le t - \delta \le t + \delta \le T$, there exists a constant C independent of n such that

$$\mathbb{E}\left[(\tilde{Z}_{t+\delta} - \tilde{Z}_t)^2(\tilde{Z}_t - \tilde{Z}_{t-\delta})^2\right] \le C\delta^2.$$

We first prove (a). When $\alpha = \beta$, we know G is a Brownian motion with mean zero and variance α^2. Hence $\mathbb{E}[(G_t - G_{t-\Delta})^2] = \alpha^2\Delta$ which goes to zero as $\Delta \to 0$. Hence, by Chebychev's inequality, we conclude that $G_t - G_{t-\Delta} \to 0$ in distribution as $\Delta \to 0$. When $\alpha \ne \beta$, we find from (2.8) that $\mathbb{E}[(G_t - G_{t-\Delta})^4] \le C\Delta^2$. Hence $G_t - G_{t-\Delta} \to 0$ in distribution as $\Delta \to 0$ by Chebychev's inequality.

We next prove (b). We first show that for each fixed $t \ge 0$, the sequence of random variables $\{\tilde{Z}_t : n \ge 1\}$ defined in (3.12) converges in distribution as $n \to \infty$. To this end, we study the moment generating function of $\tilde{Z}_t$. Fix $\theta \in \mathbb{R}$. It is immediate from (3.10) that, for any sufficiently large n,

$$\mathbb{E}\left[e^{-\theta\frac{Z_t - ne^{(\alpha-\beta)t}}{\sqrt{n}}}\big|Z_0 = n\right] = \exp\left(A\left(t, -\frac{\theta}{\sqrt{n}}\right)\cdot n + \sqrt{n}\theta e^{(\alpha-\beta)t}\right), \tag{3.13}$$

To show this sequence of moment generating functions converges when $n \to \infty$, we rely on the asymptotic approximations for A which satisfies the ordinary differential equation (3.8) with initial condition $-\frac{\theta}{\sqrt{n}}$. For n large, the quantity $-\frac{\theta}{\sqrt{n}}$ is small so we introduce the following expansion:

$$A\left(t, -\frac{\theta}{\sqrt{n}}\right) = f_0(t) + f_1(t)\cdot\left(-\frac{\theta}{\sqrt{n}}\right) + f_2(t)\cdot\frac{\theta^2}{n} + O(n^{-\frac{3}{2}}), \tag{3.14}$$

where $O(\epsilon)$ is a term bounded by $C_t \cdot \epsilon$ for a positive constant C_t and ϵ small enough. Due to the smooth dependence of the solution A to differential equations (3.8) and (3.9) on the initial value, we infer that f_0, f_1, f_2 and the constant in the big O notation are all uniformly bounded for $t \in [0, T]$. (see, e.g., [40] for background on perturbation theory and asymptotic expansions for differential equations.) Next we use the differential equation (3.8) to determine the unknown functions f_0, f_1 and f_2. First, it is obvious that $f_0(t) = A(t, 0) \equiv 0$, i.e., the solution to (3.8) is zero when we have zero initial condition. This implies that for $t \in [0, T]$, we have $A(t, -\frac{\theta}{\sqrt{n}}) = O(\frac{1}{\sqrt{n}})$ as $n \to \infty$, which further implies that

$$e^{\alpha A\left(t, -\frac{\theta}{\sqrt{n}}\right)} - \beta A\left(t, -\frac{\theta}{\sqrt{n}}\right) - 1 = (\alpha-\beta)A\left(t, -\frac{\theta}{\sqrt{n}}\right) + \frac{1}{2}\alpha^2 A\left(t, -\frac{\theta}{\sqrt{n}}\right)^2 + O(n^{-\frac{3}{2}}).$$

Together with (3.14), the differential equation (3.8) becomes

$$\begin{aligned}
&f_1'(t) \cdot \left(-\frac{\theta}{\sqrt{n}}\right) + f_2'(t) \cdot \frac{\theta^2}{n} + O(n^{-\frac{3}{2}}) \\
&\quad = (\alpha-\beta) f_1(t) \cdot \left(-\tfrac{\theta}{\sqrt{n}}\right) + [\tfrac{1}{2}\alpha^2 f_1(t)^2 + (\alpha-\beta) f_1(t)] \cdot \tfrac{\theta^2}{n} + O(n^{-\frac{3}{2}}).
\end{aligned}$$

Hence equating the coefficients depending on the power of $-\frac{\theta}{\sqrt{n}}$, we have

$$\begin{aligned}
f_1'(t) &= (\alpha-\beta) f_1(t), \\
f_2'(t) &= \frac{1}{2}\alpha^2 f_1(t)^2 + (\alpha-\beta) f_2(t).
\end{aligned}$$

For the initial conditions, we can deduce from the equality $A(0, -\frac{\theta}{\sqrt{n}}) = -\frac{\theta}{\sqrt{n}}$ that

$$f_1(0) = 1, \quad f_2(0) = 0.$$

Solving the above two differential equations we obtain

$$\begin{aligned}
f_1(t) &= e^{(\alpha-\beta)t}, \\
f_2(t) &= \begin{cases} \frac{1}{2}\frac{\alpha^2}{\alpha-\beta}(e^{2(\alpha-\beta)t} - e^{(\alpha-\beta)t}), & \text{if} \quad \alpha \neq \beta \\ \frac{1}{2}\alpha^2 t, & \text{if} \quad \alpha = \beta. \end{cases}
\end{aligned} \tag{3.15}$$

Thus we obtain

$$A\left(t, -\frac{\theta}{\sqrt{n}}\right) = e^{(\alpha-\beta)t} \cdot \left(-\frac{\theta}{\sqrt{n}}\right) + f_2(t) \cdot \frac{\theta^2}{n} + O\left(n^{-\frac{3}{2}}\right), \tag{3.16}$$

which implies that as $n \to \infty$

$$\left(A\left(t, -\frac{\theta}{\sqrt{n}}\right) \cdot n + \sqrt{n}\theta e^{(\alpha-\beta)t}\right) \to \frac{\theta^2}{2} \cdot 2f_2(t).$$

Therefore, the sequences of the moment generating functions in (3.13) converges when $n \to \infty$, and we obtain for fixed $t > 0$,

$$\frac{Z_t - ne^{(\alpha-\beta)t}}{\sqrt{n}} \to G_t, \quad \text{in distribution,}$$

where G_t is a random variable following normal distribution with mean zero and variance $2f_2(t)$ where f_2 is given in (3.15).

Now to prove (b), it remains to show the finite dimensional distributions of $\tilde{Z}$ in (3.12) converge for dimensions of two and higher. We use mathematical induction and rely on the Markov property of Z. Given $Z_0 = n$, suppose for arbitrary $0 < t_1 < \ldots < t_k$,

$$\left(\tilde{Z}_{t_1}, \ldots, \tilde{Z}_{t_k}\right) \to \left(G_{t_1}, \ldots, G_{t_k}\right), \quad \text{in distribution.}$$

Here we suppose $(G_{t_1},\dots,G_{t_k})$ follows $k-$variate normal distribution with mean zero and $k\times k$ covariance matrix $[\sigma_{ij}]_{k\times k}$ where for $1\le i\le j\le k$

$$\sigma_{ij} := \mathrm{Cov}(G_{t_i},G_{t_j}) = \begin{cases} \frac{\alpha^2}{\alpha-\beta}(e^{(\alpha-\beta)(t_i+t_j)}-e^{(\alpha-\beta)t_j}), & \alpha\neq\beta,\\ \alpha^2 t_i, & \alpha=\beta, \end{cases} \tag{3.17}$$

and $\sigma_{ij}=\sigma_{ji}$ if $i>j$. Hence for $\theta_1,\dots,\theta_k\in\mathbb{R}$, we have as $n\to\infty$

$$\mathbb{E}\left[e^{-\sum_{i=1}^k \theta_i \tilde{Z}_{t_i}}\big|Z_0=n\right]\to \exp\left(\frac{1}{2}\sum_{i=1}^k\sum_{j=1}^k \theta_i\theta_j\sigma_{ij}\right). \tag{3.18}$$

To show for arbitrary $0<t_1<\ldots<t_k<t_{k+1}$,

$$\left(\tilde{Z}_{t_1},\dots,\tilde{Z}_{t_{k+1}}\right)\to\left(G_{t_1},\dots,G_{t_{k+1}}\right),\quad \text{in distribution,}$$

where $\left(G_{t_1},\dots,G_{t_{k+1}}\right)$ follows $(k+1)-$variate normal distribution with mean zero and some covariance matrix consistent with (3.17), we show the sequence of moment generating functions converges as $n\to\infty$. By tower property of conditional expectation and Markov property of Z, for any sufficiently large n, we have

$$\mathbb{E}\left[e^{-\sum_{i=1}^{k+1}\theta_i\tilde{Z}_{t_i}}\middle|Z_0=n\right]=\mathbb{E}\left[e^{-\sum_{i=1}^k\theta_i\tilde{Z}_{t_i}}\cdot\mathbb{E}\left[e^{-\theta_{k+1}\tilde{Z}_{t_{k+1}}}|Z_{t_k}\right]\middle|Z_0=n\right]. \tag{3.19}$$

Similar for (3.13), we first deduce from (3.10) that

$$\begin{aligned} &\mathbb{E}\left[e^{-\theta_{k+1}\tilde{Z}_{t_{k+1}}}|Z_{t_k}\right]\\ =\;& \exp\left(A\left(t_{k+1}-t_k,-\frac{\theta_{k+1}}{\sqrt{n}}\right)Z_{t_k}+\theta_{k+1}\sqrt{n}e^{(\alpha-\beta)t_{k+1}}\right)\\ =\;& \exp\left(A\left(t_{k+1}-t_k,-\frac{\theta_{k+1}}{\sqrt{n}}\right)\cdot(\sqrt{n}\tilde{Z}_{t_k}+ne^{(\alpha-\beta)t_k})+\theta_{k+1}\sqrt{n}e^{(\alpha-\beta)t_{k+1}}\right), \end{aligned}$$

where the last equality follows from (3.12). With this, we deduce from (3.19) that

$$\mathbb{E}\left[e^{-\sum_{i=1}^{k+1}\theta_i\tilde{Z}_{t_i}}\big|Z_0=n\right]=\mathbb{E}\left[e^{-\sum_{i=1}^k\hat{\theta}_i^{(n)}\tilde{Z}_{t_i}}\middle|Z_0=n\right]\cdot\exp(\Gamma^n), \tag{3.20}$$

where $\hat{\theta}_i^{(n)}=\theta_i$ for $i=1,\dots,k-1$, $\hat{\theta}_k^{(n)}=\theta_k-\sqrt{n}A(t_{k+1}-t_k,-\frac{\theta_{k+1}}{\sqrt{n}})$ and

$$\Gamma^n := A\left(t_{k+1}-t_k,-\frac{\theta_{k+1}}{\sqrt{n}}\right)\cdot ne^{(\alpha-\beta)t_k}+\theta_{k+1}\sqrt{n}e^{(\alpha-\beta)t_{k+1}}.$$

From the expansion of A at (3.16) we infer that

$$\begin{aligned}\lim_{n\to\infty}\hat{\theta}_k^{(n)} &= \theta_k+\theta_{k+1}e^{(\alpha-\beta)(t_{k+1}-t_k)},\\ \lim_{n\to\infty}\Gamma^n &= \theta_{k+1}^2 e^{(\alpha-\beta)t_k}\cdot f_2(t_{k+1}-t_k),\end{aligned}$$

where f_2 is the function given in (3.15). In conjunction with (3.18) and (3.20), and simplifying the resulting expression we obtain as $n\to\infty$,

$$\mathbb{E}\left[e^{-\sum_{i=1}^{k+1}\theta_i\tilde{Z}_{t_i}}\big|Z_0=n\right]\to\exp\left(\frac{1}{2}\sum_{i=1}^{k+1}\sum_{j=1}^{k+1}\theta_i\theta_j\sigma_{ij}\right),$$

where

$$\begin{aligned}\sigma_{(k+1)(k+1)} &= 2f_2(t_{k+1}),\\ \sigma_{i(k+1)} &= \begin{cases}\frac{\alpha^2}{\alpha-\beta}(e^{(\alpha-\beta)(t_i+t_{k+1})}-e^{(\alpha-\beta)t_{k+1}}), & \alpha\neq\beta,\\ \alpha^2 t_i, & \alpha=\beta,\end{cases}\end{aligned}$$

and $\sigma_{(k+1)i}=\sigma_{i(k+1)}$ for $i=1,\ldots,k$. Hence we have established that the finite dimension distribution of $\tilde{Z}^n$ converges to that of a centered Gaussian process G with covariance function is given by: for $0\le s\le t$,

$$\mathrm{Cov}(G_t,G_s)=\begin{cases}\frac{\alpha^2}{\alpha-\beta}(e^{(\alpha-\beta)(t+s)}-e^{(\alpha-\beta)t}) & \alpha\neq\beta,\\ \alpha^2 s, & \alpha=\beta.\end{cases}$$

Finally we prove (c). We study the cases $\alpha\neq\beta$ and $\alpha=\beta$ separately. We start with the case $\alpha\neq\beta$. Given Proposition 5, it is straightforward to compute that

$$\begin{aligned}&\mathbb{E}\left[\left(Z_{t+\delta}-ne^{(\alpha-\beta)(t+\delta)}-Z_t+ne^{(\alpha-\beta)t}\right)^2|Z_t\right] \qquad (3.21)\\ &=\mathbb{E}[Z_{t+\delta}^2|Z_t]-2(ne^{(\alpha-\beta)(t+\delta)}-ne^{(\alpha-\beta)t}+Z_t)\mathbb{E}[Z_{t+\delta}|Z_t]\\ &\quad+(ne^{(\alpha-\beta)(t+\delta)}-ne^{(\alpha-\beta)t}+Z_t)^2\\ &=Z_t^2e^{2(\alpha-\beta)\delta}+\frac{\alpha^2Z_t}{\alpha-\beta}(e^{2(\alpha-\beta)\delta}-e^{(\alpha-\beta)\delta})\\ &\quad-2\left[ne^{(\alpha-\beta)t}(e^{(\alpha-\beta)\delta}-1)+Z_t\right]Z_te^{(\alpha-\beta)\delta}\\ &\quad+n^2e^{2(\alpha-\beta)t}(e^{(\alpha-\beta)\delta}-1)^2+Z_t^2+2ne^{(\alpha-\beta)t}(e^{(\alpha-\beta)\delta}-1)Z_t\\ &=(e^{(\alpha-\beta)\delta}-1)^2(Z_t-ne^{(\alpha-\beta)t})^2+(e^{(\alpha-\beta)\delta}-1)\frac{\alpha^2e^{(\alpha-\beta)\delta}}{\alpha-\beta}Z_t.\end{aligned}$$

In addition, by the Markovian property of Z we obtain

$$\begin{aligned}
&\mathbb{E}\left[(\tilde{Z}_{t+\delta}-\tilde{Z}_t)^2(\tilde{Z}_t-\tilde{Z}_{t-\delta})^2\right] \\
&=\mathbb{E}\left[\mathbb{E}\left[(\tilde{Z}_{t+\delta}-\tilde{Z}_t)^2(\tilde{Z}_t-\tilde{Z}_{t-\delta})^2\middle|\sigma(Z_s:0\le s\le t)\right]\right] \\
&=\mathbb{E}\left[(\tilde{Z}_t-\tilde{Z}_{t-\delta})^2\cdot\mathbb{E}\left[(\tilde{Z}_{t+\delta}-\tilde{Z}_t)^2\middle|Z_t\right]\right].
\end{aligned}$$

On combining the above equation with (3.21) we obtain

$$\begin{aligned}
&\mathbb{E}\left[(\tilde{Z}_{t+\delta}-\tilde{Z}_t)^2(\tilde{Z}_t-\tilde{Z}_{t-\delta})^2\right] \\
&=\frac{1}{n^2}(e^{(\alpha-\beta)\delta}-1)^2\mathbb{E}\left[(Z_t-ne^{(\alpha-\beta)t})^2\left(Z_t-ne^{(\alpha-\beta)t}-Z_{t-\delta}+ne^{(\alpha-\beta)(t-\delta)}\right)^2\right] \\
&\quad+\frac{1}{n^2}(e^{(\alpha-\beta)\delta}-1)\frac{\alpha^2e^{(\alpha-\beta)\delta}}{\alpha-\beta}\mathbb{E}\left[Z_t\left(Z_t-ne^{(\alpha-\beta)t}-Z_{t-\delta}+ne^{(\alpha-\beta)(t-\delta)}\right)^2\right].
\end{aligned}\tag{3.22}$$

Next we derive upper bounds for the terms in (3.22). First, let us bound the first term on the right hand side of (3.22). Direct computation yields

$$\begin{aligned}
&\mathbb{E}\left[(Z_t-ne^{(\alpha-\beta)t})^2\left(Z_t-ne^{(\alpha-\beta)t}-Z_{t-\delta}+ne^{(\alpha-\beta)(t-\delta)}\right)^2\right] \\
&\le 2\mathbb{E}\left[(Z_t-ne^{(\alpha-\beta)t})^2\left[\left(Z_t-ne^{(\alpha-\beta)t}\right)^2+\left(Z_{t-\delta}-ne^{(\alpha-\beta)(t-\delta)}\right)^2\right]\right] \\
&=2\mathbb{E}\left[\left(Z_t-ne^{(\alpha-\beta)t}\right)^4\right]+2\mathbb{E}\left[\left(Z_t-ne^{(\alpha-\beta)t}\right)^2\left(Z_{t-\delta}-ne^{(\alpha-\beta)(t-\delta)}\right)^2\right] \\
&\le 2\mathbb{E}\left[\left(Z_t-ne^{(\alpha-\beta)t}\right)^4\right]+2\left[\mathbb{E}\left(Z_t-ne^{(\alpha-\beta)t}\right)^4\right]^{1/2}\left[\mathbb{E}\left(Z_{t-\delta}-ne^{(\alpha-\beta)(t-\delta)}\right)^4\right]^{1/2} \\
&\le 4Cn^2,
\end{aligned}$$

where the second last inequality follows from Cauchy-Schwarz inequality and the last inequality is due to Proposition 6. On combining with the fact that

$$|e^{(\alpha-\beta)\delta}-1|\le C\delta \quad \text{for} \quad \delta\in[0,T], \tag{3.23}$$

we infer that the first term on the right hand side of (3.22) is upper bounded by $C\delta^2$ for some constant C.

We next proceed to bound the second term on the right hand side of (3.22). By (3.4), we have

$$
\begin{aligned}
&Z_t - ne^{(\alpha-\beta)t} - (Z_{t-\delta} - ne^{(\alpha-\beta)(t-\delta)}) \\
&= e^{(\alpha-\beta)t}\alpha \int_0^t e^{-(\alpha-\beta)s} dM_s - e^{(\alpha-\beta)(t-\delta)}\alpha \int_0^{t-\delta} e^{-(\alpha-\beta)s} dM_s \\
&= e^{(\alpha-\beta)t}\alpha \int_{t-\delta}^t e^{-(\alpha-\beta)s} dM_s + (e^{(\alpha-\beta)\delta} - 1)e^{(\alpha-\beta)(t-\delta)}\alpha \int_0^{t-\delta} e^{-(\alpha-\beta)s} dM_s.
\end{aligned}
$$

Therefore, by Cauchy-Schwarz inequality we find

$$
\begin{aligned}
&\mathbb{E}\left[Z_t \left(Z_t - ne^{(\alpha-\beta)t} - Z_{t-\delta} + ne^{(\alpha-\beta)(t-\delta)} \right)^2 \right] \\
&\le [\mathbb{E}Z_t^2]^{1/2} \left[\mathbb{E}\left(Z_t - ne^{(\alpha-\beta)t} - Z_{t-\delta} + ne^{(\alpha-\beta)(t-\delta)} \right)^4 \right]^{1/2} \\
&\le [\mathbb{E}Z_t^2]^{1/2} \Bigg[4\mathbb{E}\left(e^{(\alpha-\beta)t}\alpha \int_{t-\delta}^t e^{-(\alpha-\beta)s} dM_s \right)^4 \\
&\qquad\qquad + 4\mathbb{E}\left((e^{(\alpha-\beta)\delta} - 1)e^{(\alpha-\beta)(t-\delta)}\alpha \int_0^{t-\delta} e^{-(\alpha-\beta)s} dM_s \right)^4 \Bigg]^{1/2} \\
&= 2[\mathbb{E}Z_t^2]^{1/2} \Bigg(e^{4(\alpha-\beta)t}\alpha^4 \mathbb{E}\left(\int_{t-\delta}^t e^{-(\alpha-\beta)s} dM_s \right)^4 \\
&\qquad\qquad + (e^{(\alpha-\beta)\delta} - 1)^4 e^{4(\alpha-\beta)(t-\delta)}\alpha^4 \mathbb{E}\left(\int_0^{t-\delta} e^{-(\alpha-\beta)s} dM_s \right)^4 \Bigg)^{1/2}.
\end{aligned}
$$

Now Proposition 5 implies that $[\mathbb{E}Z_t^2]^{1/2} \le Cn$. In conjunction with Proposition 6 and (3.23), we deduce from the above inequality that

$$
\begin{aligned}
&\mathbb{E}\left[Z_t \left(Z_t - ne^{(\alpha-\beta)t} - Z_{t-\delta} + ne^{(\alpha-\beta)(t-\delta)} \right)^2 \right] \\
&\le Cn \cdot \left(Cn^2\delta^2 + C\delta^4 n^2 \right)^{\frac{1}{2}} \\
&\le Cn^2\delta. \qquad (3.24)
\end{aligned}
$$

Hence it follows that the second term on the right hand side of (3.22) is also upper bounded by $C\delta^2$ for some constant C. Therefore, we have proved (c) for the case $\alpha \neq \beta$.

We next proceed to prove (c) for the case $\alpha = \beta$. The proof is similar as for the case

$\alpha \neq \beta$, so we only outline the key steps. Given $Z_0 = n$, we have $\tilde{Z}_t = \frac{Z_t - n}{\sqrt{n}}$. Then we have

$$
\begin{aligned}
&\mathbb{E}\left[(\tilde{Z}_{t+\delta} - \tilde{Z}_t)^2(\tilde{Z}_t - \tilde{Z}_{t-\delta})^2\right] \\
&= \frac{1}{n^2}\mathbb{E}\left[(Z_{t+\delta} - Z_t)^2(Z_t - Z_{t-\delta})^2\right] \\
&= \frac{1}{n^2}\mathbb{E}\left[(Z_t - Z_{t-\delta})^2 \cdot \mathbb{E}[(Z_{t+\delta} - Z_t)^2 | Z_t]\right] \\
&= \frac{\alpha^2\delta}{n^2}\mathbb{E}\left[Z_t(Z_t - Z_{t-\delta})^2\right],
\end{aligned}
\tag{3.25}
$$

where the second equality follows from the fact that Z is Markovian, and the third equality follows from a similar argument as for (3.21) in the case $\alpha \neq \beta$. To bound the last term in (3.25) we note that

$$
\begin{aligned}
\mathbb{E}\left[Z_t(Z_t - Z_{t-\delta})^2\right] &= \mathbb{E}[Z_t^3] - 2\mathbb{E}[Z_t^2 Z_{t-\delta}] + \mathbb{E}[Z_t Z_{t-\delta}^2] \\
&= \mathbb{E}[Z_{t-\delta}^3] + 3\alpha^2\delta\mathbb{E}[Z_{t-\delta}^2] + \frac{3}{2}\alpha^4\delta^2\mathbb{E}[Z_{t-\delta}] + \alpha^3\delta\mathbb{E}[Z_{t-\delta}] \\
&\quad - 2\mathbb{E}[Z_{t-\delta}^3] - 2\alpha^2\delta\mathbb{E}[Z_{t-\delta}^2] + \mathbb{E}[Z_{t-\delta}^3] \\
&= \alpha^2\mathbb{E}[Z_{t-\delta}^2] + \left(\frac{3}{2}\alpha^4\delta^2 + \alpha^3\delta\right)\mathbb{E}[Z_{t-\delta}] \\
&= \alpha^2\delta(n^2 + \alpha^2 n(t-\delta)) + \left(\frac{3}{2}\alpha^4\delta^2 + \alpha^3\delta\right)n,
\end{aligned}
$$

where the first equality follows from Proposition 5 and tower property of conditional expectation, and the last equality follows from Proposition 5 and the fact that $Z_0 = n$. Hence,

$$
\mathbb{E}\left[(\tilde{Z}_{t+\delta} - \tilde{Z}_t)^2(\tilde{Z}_t - \tilde{Z}_{t-\delta})^2\right] = \alpha^4\delta^2\left[1 + \frac{\alpha^2}{n}(t-\delta) + \frac{3}{2n}\alpha^2\delta + \frac{\alpha}{n}\right] \leq C\delta^2,
$$

where C is a positive constant that is independent of n. Hence we have also established (c) when $\alpha = \beta$. The proof is therefore complete.

3.4.2 FCLT for N when $\mu = 0$

In this section we prove the weak convergence of the sequence of re–normalized Hakwes processes (2.6) when $\mu = 0$. Let us recall that the intensity process Z_t satisfies the dynamics

$$
dZ_t = -\beta Z_t dt + \alpha dN_t, \qquad Z_0 = n,
$$

where N_t is a simple point process with intensity Z_{t-}. We can express the jump process $N_t = N(0,t]$ in terms of the intensity process Z_t in the following way,

$$
N_t = \frac{Z_t - Z_0}{\alpha} + \frac{\beta}{\alpha}\int_0^t Z_s ds.
$$

This immediately yields that

$$
\begin{aligned}
\frac{N_t - n\int_0^t e^{(\alpha-\beta)s}ds}{\sqrt{n}} &= \frac{1}{\sqrt{n}}\left[\frac{Z_t - n}{\alpha} + \frac{\beta}{\alpha}\int_0^t Z_s ds - n\int_0^t e^{(\alpha-\beta)s}ds\right] \\
&= \frac{1}{\sqrt{n}}\left[\frac{Z_t - ne^{(\alpha-\beta)t}}{\alpha} + \frac{\beta}{\alpha}\int_0^t (Z_s - ne^{(\alpha-\beta)s})ds\right] \qquad (3.26)
\end{aligned}
$$

Note that we have shown that for any $T > 0$, as $n \to \infty$,

$$
\left\{\frac{Z_t - ne^{(\alpha-\beta)t}}{\sqrt{n}} : t \in [0, T]\right\} \to G,
$$

weakly on $D[0, T]$ equipped with Skorohod J_1 topology, where G is a centered Gaussian process. Thus, using (3.26), we conclude from [36, Theorem 2.2] that as $n \to \infty$

$$
\left\{\frac{N_t - n\int_0^t e^{(\alpha-\beta)s}ds}{\sqrt{n}} : t \in [0, T]\right\} \to H,
$$

weakly on $D[0, T]$, where

$$
H_t = \frac{G_t}{\alpha} + \frac{\beta}{\alpha}\int_0^t G_s ds \quad \text{for } t \geq 0.
$$

Since G is a centered Gaussian process, it is then readily seen that H is a also centered Gaussian process. The covariance function of H is given as follows: for $0 \leq s \leq t$, when $\alpha \neq \beta$,

$$
\begin{aligned}
&\mathrm{Cov}(H_t, H_s) \qquad (3.27) \\
&= \frac{1}{\alpha^2}\mathbb{E}[G_t G_s] + \frac{\beta}{\alpha^2}\int_0^t \mathbb{E}[G_u G_s]du \\
&\quad + \frac{\beta}{\alpha^2}\int_0^s \mathbb{E}[G_v G_t]dv + \frac{\beta^2}{\alpha^2}\int_0^t\int_0^s \mathbb{E}[G_u G_v]dudv \\
&= \frac{1}{\alpha-\beta}(e^{(\alpha-\beta)(t+s)} - e^{(\alpha-\beta)t}) + \frac{\beta}{\alpha-\beta}\left[\frac{e^{2(\alpha-\beta)s} - e^{(\alpha-\beta)s}}{(\alpha-\beta)} - se^{(\alpha-\beta)s}\right] \\
&\quad + \frac{\beta}{(\alpha-\beta)^2}(e^{(\alpha-\beta)s} - 1)(e^{(\alpha-\beta)t} - e^{(\alpha-\beta)s}) \\
&\quad + \frac{\beta}{\alpha-\beta}\left[\frac{e^{(\alpha-\beta)(t+s)} - e^{(\alpha-\beta)t}}{(\alpha-\beta)} - se^{(\alpha-\beta)t}\right] \\
&\quad + \frac{\beta^2}{(\alpha-\beta)^3}\left(e^{(\alpha-\beta)t} - e^{(\alpha-\beta)s}\right)\left(e^{(\alpha-\beta)s} - 1 - s(\alpha-\beta)\right) \\
&\qquad + \frac{\beta^2}{(\alpha-\beta)^2}(e^{(\alpha-\beta)s} - 1)^2 - \frac{2\beta^2}{(\alpha-\beta)^2}se^{(\alpha-\beta)s} + \frac{2\beta^2}{(\alpha-\beta)^3}(e^{(\alpha-\beta)s} - 1).
\end{aligned}
$$

When $\alpha = \beta$, for $0 \le s \le t$, we have

$$\mathrm{Cov}(H_t, H_s) = s + \alpha st + \alpha^2 \left(\frac{ts^2}{2} - \frac{s^3}{6} \right). \tag{3.28}$$

3.4.3 FCLT for Z and N when $\mu > 0$

In this section we prove Theorem 2 for the case $\mu > 0$ using the observation in Section 3.1.

Decompose $N_t = N_t^{(0)} + N_t^{(1)}$. Note for any $T > 0$, we have $\sup_{0 \le t \le T} N_t^{(1)}$ is finite and independent of the parameter n. This implies that as $n \to \infty$

$$\sup_{0 \le t \le T} \frac{N_t^{(1)}}{\sqrt{n}} \to 0, \quad \text{almost surely.} \tag{3.29}$$

Since we have established FCLT for $N_t^{(0)}$, the result for N then readily follows from (3.29).

Similarly, we can decompose $Z_t = Z_t^{(0)} + Z_t^{(1)}$ and note that $Z_t^{(1)}$ is independent of n and hence $\sup_{0 \le t \le T} Z_t^{(1)}/\sqrt{n} \le \alpha \sup_{0 \le t \le T} N_t^{(1)}/\sqrt{n} \to 0$ almost surely as $n \to \infty$. Hence the FCLT for Z follows.

3.5 Proof of Theorem 3

We prove Theorem 3 in this section. We also first prove Theorem 3 when $\mu = 0$, and then prove it when $\mu > 0$ using the observation in Section 3.1.

3.5.1 Proof of Theorem 3 when $\mu = 0$

Proof of part (i): super–critical case. We first show (2.10). Observe first from (3.4) that for any $n \in \mathbb{N}$,

$$\frac{Z_{s\tau_n} - n^{1+s}}{n^{1+s}} = \frac{\alpha}{n} \int_0^{s\tau_n} e^{-(\alpha-\beta)u} dM_u$$

is a martingale. Therefore, Doob's martingale inequality implies that for any $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0 \le s \le T} \left| \frac{Z_{s\tau_n} - n^{1+s}}{n^{1+s}} \right| \ge \epsilon \right) \\
&= \mathbb{P}\left(\sup_{0 \le s \le T} \left| \int_0^{s\tau_n} e^{-(\alpha-\beta)u} dM_u \right| \ge \frac{n\epsilon}{\alpha} \right) \\
&\le \frac{C\alpha^4}{n^4 \epsilon^4} \mathbb{E}\left[\left(\int_0^{T\tau_n} e^{-(\alpha-\beta)u} dM_u \right)^4 \right].
\end{aligned}$$

Now inequality (3.7) implies that for $\alpha > \beta$ and $\tau_n = \frac{\log n}{\alpha - \beta}$,

$$\mathbb{E}\left[\left(\int_0^{T\tau_n} e^{-(\alpha-\beta)u} dM_u\right)^4\right] \leq CT\tau_n \cdot \left(n^2(1-n^{-2T}) + n(1-n^{-2T}) + n(1-n^{-3T})\right).$$

Thus we obtain

$$\mathbb{P}\left(\sup_{0\leq s\leq T}\left|\frac{Z_{s\tau_n} - n^{1+s}}{n^{1+s}}\right| \geq \epsilon\right) \leq \frac{CT\log n}{n^2\epsilon^4}.$$

Hence, by Borel-Cantelli lemma,

$$\sup_{0\leq s\leq T}\left|\frac{Z_{s\tau_n}}{n^{1+s}} - 1\right| \to 0,$$

almost surely as $n \to \infty$.

We next prove (2.11). Recall that the intensity process Z_t satisfies the dynamics

$$dZ_t = -\beta Z_t dt + \alpha dN_t, \qquad Z_0 = n,$$

where N_t is a simple point process with intensity Z_{t-} (since $\mu = 0$). We can therefore express the jump process $N_t = N(0,t]$ in terms of the intensity process Z_t in the following way,

$$N_t = \frac{Z_t - Z_0}{\alpha} + \frac{\beta}{\alpha}\int_0^t Z_s ds.$$

This implies that for any $t > 0$,

$$N_t - n\int_0^t e^{(\alpha-\beta)s} ds = \frac{Z_t - ne^{(\alpha-\beta)t}}{\alpha} + \frac{\beta}{\alpha}\int_0^t (Z_s - ne^{(\alpha-\beta)s}) ds.$$

Hence, for $\alpha > \beta$ and $\tau_n = \frac{\log n}{\alpha-\beta}$, we obtain

$$\begin{aligned}
\left|\frac{N_{s\tau_n} - \frac{n^{1+s}-n}{\alpha-\beta}}{n^{1+s}}\right| &= \left|\frac{Z_{s\tau_n} - n^{1+s}}{\alpha n^{1+s}} + \frac{\beta}{\alpha}\tau_n \int_0^s n^{u-s}\frac{Z_{u\tau_n} - n^{1+u}}{n^{1+u}} du\right| \\
&\leq \left|\frac{Z_{st_n} - n^{1+s}}{\alpha n^{1+s}}\right| + \frac{\beta}{\alpha}\tau_n \sup_{0\leq u\leq s}\left|\frac{Z_{ut_n} - n^{1+u}}{n^{1+u}}\right| \cdot \int_0^s n^{u-s} ds. \\
&\leq \frac{1}{\alpha}\cdot\left|\frac{Z_{st_n} - n^{1+s}}{n^{1+s}}\right| + \frac{\beta}{\alpha(\alpha-\beta)}\sup_{0\leq u\leq s}\left|\frac{Z_{ut_n} - n^{1+u}}{n^{1+u}}\right|.
\end{aligned}$$

Then (2.10) implies the desired result (2.11). $\square$

Proof of part (ii): sub–critical case. The proof is similar as in the super–critical case, so we only outline the key steps.

We first show (2.12). For $\alpha < \beta$, we observe that for $0 \le s \le T < 1$,

$$\frac{Z_{st_n} - n^{1-s}}{n^{1-s}} = \frac{\alpha}{n} \int_0^{st_n} e^{-(\alpha-\beta)s} dM_s$$

is a martingale. Thus we have, for $t_n = \frac{\log n}{\beta - \alpha}$ and any $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0 \le s \le T} \left| \frac{Z_{st_n} - n^{1-s}}{n^{1-s}} \right| \ge \epsilon \right) \\
&\le \frac{C}{n^4 \epsilon^4} \mathbb{E}\left[\left(\int_0^{Tt_n} e^{-(\alpha-\beta)u} dM_u \right)^4 \right] \\
&\le \frac{C}{n^4 \epsilon^4} T t_n \cdot \left(n^2 (n^{2T} - 1) + n(n^{2T} - 1) + n(n^{3T} - 1) \right),
\end{aligned}$$

where the last inequality follows from (3.7). Therefore, for $t_n = \frac{\log n}{\beta - \alpha}$, we obtain

$$\mathbb{P}\left(\sup_{0 \le s \le T} \left| \frac{Z_{st_n} - n^{1-s}}{n^{1-s}} \right| \ge \epsilon \right) \le \frac{CT \log n}{n^4 \epsilon^4} \left[n^2 n^{2T} + n \cdot n^{3T} \right],$$

which goes to 0 as $n \to \infty$ for any $T < 1$. Thus we have $\sup_{0 \le s \le T} \left| \frac{Z_{st_n} - n^{1-s}}{n^{1-s}} \right|$ convergences to zero in probability for any $T < 1$. Furthermore, if $T < \frac{1}{2}$, we have

$$\sum_{n=1}^{\infty} \frac{\log n}{n^4 \epsilon^4} \left[n^2 n^{2T} + n \cdot n^{3T} \right] < +\infty.$$

Hence the almost sure convergence follows from Borel-Cantelli lemma.

We next prove (2.13). When $\alpha < \beta$ and $t_n = \frac{\log n}{\beta - \alpha}$, we have

$$\begin{aligned}
\left| \frac{N_{st_n} - \frac{1}{\beta-\alpha}(n - n^{1-s})}{n} \right| &= \left| \frac{Z_{st_n} - n^{1-s}}{\alpha n^{1-s}} + \frac{\beta}{\alpha} t_n \int_0^s n^{-u} \frac{Z_{ut_n} - n^{1-u}}{n^{1-u}} du \right| \\
&\le \left| \frac{Z_{st_n} - n^{1-s}}{\alpha n^{1-s}} \right| + \frac{\beta}{\alpha(\beta - \alpha)} \sup_{0 \le u \le s} \left| \frac{Z_{ut_n} - n^{1-u}}{n^{1-u}} \right|.
\end{aligned}$$

The desired result (2.13) then follows. $\square$

3.5.2 Proof of Theorem 3 when $\mu > 0$

In this section we prove Theorem 3 for the case $\mu > 0$ using the observation in Section 3.1.

Proof of (2.10): Note that $Z_t^{(1)} \leq \alpha N_t^{(1)}$. In Zhu [47], it was showed that $\lim_{t\to\infty} \frac{1}{t} \log N_t^{(1)} = \alpha - \beta$ for $\alpha > \beta > 0$. Thus, for any $0 < T < 1$,

$$\sup_{0\leq s\leq T} \frac{Z^{(1)}_{s\tau_n}}{n^{1+s}} \leq \alpha \sup_{0\leq s\leq T} \frac{N^{(1)}_{s\tau_n}}{n^{1+s}} \leq \frac{N^{(1)}_{T\tau_n}}{n} \to 0,$$

almost surely as $n \to \infty$ since $\tau_n = \frac{\log n}{\alpha-\beta}$. Then (2.10) follows from the decomposition $Z_t = Z_t^{(0)} + Z_t^{(1)}$ and the result proved in the previous section.

Proof of (2.11): Similarly, it follows from the decomposition $N_t = N_t^{(0)} + N_t^{(1)}$ and that for any $0 < T < 1$,

$$\sup_{0\leq s\leq T} \frac{N^{(1)}_{s\tau_n}}{n^{1+s}} \leq \frac{N^{(1)}_{T\tau_n}}{n} \to 0,$$

almost surely as $n \to \infty$ since $\tau_n = \frac{\log n}{\alpha-\beta}$.

Proof of (2.12): Note that for $\beta > \alpha > 0$, $\frac{N_t^{(1)}}{t} \to \frac{\mu}{1-\frac{\alpha}{\beta}}$ almost surely as $t \to \infty$. Therefore, for $t_n = \frac{\log n}{\beta-\alpha}$ we deduce that for any $0 < T < 1$,

$$\sup_{0\leq s\leq T} \frac{Z^{(1)}_{st_n}}{n^{1-s}} \leq \frac{\alpha N^{(1)}_{Tt_n}}{n^{1-T}} \to 0,$$

almost surely as $n \to \infty$.

Proof of (2.13): Similarly, for $t_n = \frac{\log n}{\beta-\alpha}$, we have

$$\sup_{0\leq s\leq T} \frac{N^{(1)}_{st_n}}{n} = \frac{N^{(1)}_{Tt_n}}{n} \to 0,$$

almost surely as $n \to \infty$ since $\frac{N_t^{(1)}}{t} \to \frac{\mu}{1-\frac{\alpha}{\beta}}$ almost surely as $t \to \infty$. $\square$

3.6 Proof of Theorem 4

This section is devoted to the proof of Theorem 4.

3.6.1 Proof of critical and nearly–critical cases

Proof. The proof is based on diffusion approximations. In particular, we apply Theorem 4.1 in Ethier and Kurtz [18, Chapter 7] and verify their conditions (4.1)–(4.7). Fix $Z_0 = n$.

Define for $t \geq 0$ and $\mu \geq 0$,

$$\mathbb{X}_t^n := \frac{Z_{nt}}{n}, \quad \mathbb{B}_t^n := \int_0^t (\alpha_n \mu + \gamma \mathbb{X}_s^n) ds, \quad \mathbb{A}_t^n := \alpha_n^2 \int_0^t \left(\mathbb{X}_s^n + \frac{\mu}{n} \right) ds. \tag{3.30}$$

One readily checks that the two processes

$$\{\mathbb{X}_t^n - \mathbb{B}_t^n : t \in [0,T]\}, \quad \text{and} \quad \left\{ (\mathbb{X}_t^n - \mathbb{B}_t^n)^2 - \mathbb{A}_t^n : t \in [0,T] \right\}$$

are martingales with respect to the filtration generated by $\mathbb{X}^n$. Since Z can only make jumps of size α_n, we deduce that $\mathbb{X}^n$ can only make jumps of size $\frac{\alpha_n}{n}$. This implies condition (4.3) of Theorem 4.1 in [18, Chapter 7] holds. Condition (4.4) and (4.5) hold trivially since $\mathbb{B}^n$ and $\mathbb{A}^n$ have continuous sample paths. In view of (3.30) and $\lim_{n\to\infty} \alpha_n = \beta$, Condition (4.6) also holds if we set $b(x) = \beta\mu + \gamma x$ and $a(x) = \beta^2 x$. Now if we define

$$dX_t = (\beta\mu + \gamma X_t)dt + \beta\sqrt{X_t} dB_t, \qquad X_0 = 1$$

where B is a standard Brownian motion, then this stochastic differential equation (SDE) has a pathwise unique strong solution. Since $\mathbb{X}_0^n \equiv 1$ for all n, we deduce that $\mathbb{X}^n$ converges weakly to X by Theorem 4.1 in [18, Chapter 7] and the well–known equivalence of SDE and martingale problems (see, e.g., [37]).

We next establish the weak convergence of the sequence of re–normalized Hawkes processes

$$\left\{ \frac{N_{tn}}{n^2} : t \in [0,T] \right\}.$$

Since for $\mu \geq 0$, the point process N has intensity $\mu + Z_{t-}$ at time t, where $dZ_t = -\beta Z_t dt + \alpha_n dN_t$ for $Z_0 = n$. So we still have

$$N_t = \frac{Z_t - Z_0}{\alpha_n} + \frac{\beta}{\alpha_n} \int_0^t Z_s ds,$$

which yields

$$\frac{N_{tn}}{n^2} = \frac{Z_{tn} - Z_0}{n^2 \cdot \alpha_n} + \frac{\beta}{\alpha_n} \int_0^t \frac{Z_{sn}}{n} ds.$$

Since $\lim_{n\to\infty} \alpha_n = \beta$ and we have the weak convergence of the sequence of processes $\left\{ \frac{Z_{tn}}{n} : t \in [0,T] \right\}$, the result then follows from [36, Theorem 2.2]. $\square$

3.6.2 Proof of super–critical case when $\mu = 0$

Proof. The proof is based on Aldous's result on weak convergence of a sequence of martingales to a continuous martingale limit [1].

We first note that in the super-critical case $\alpha > \beta > 0$ where $\tau_n := \frac{\log n}{\alpha - \beta}$, we have

$$Z_{s\tau_n} - n^{1+s} = \alpha n^s \int_0^{s\tau_n} e^{-(\alpha-\beta)u} dM_u.$$

Therefore, $\left\{\frac{Z_{s\tau_n}-n^{1+s}}{n^{\frac{1}{2}+s}} : s \in [0,T]\right\}$ is a martingale for any fixed n. Moreover, it is straightforward to compute that

$$\begin{aligned}
\sup_{n\in\mathbb{N}} \mathbb{E}\left[\left(\frac{Z_{s\tau_n}-n^{1+s}}{n^{\frac{1}{2}+s}}\right)^2\right] &= \sup_{n\in\mathbb{N}} \frac{\alpha^2}{n}\mathbb{E}\left[\int_0^{s\tau_n} e^{-2(\alpha-\beta)u} Z_u du\right] \\
&= \sup_{n\in\mathbb{N}} \frac{\alpha^2}{n}\int_0^{s\tau_n} e^{-(\alpha-\beta)u} Z_0 du \\
&= \sup_{n\in\mathbb{N}} \frac{\alpha^2}{\alpha-\beta}\left(1-\frac{1}{n^s}\right) \\
&= \frac{\alpha^2}{\alpha-\beta}.
\end{aligned}$$

Therefore, for each s, $\left\{\frac{Z_{s\tau_n}-n^{1+s}}{n^{\frac{1}{2}+s}} : n \geq 1\right\}$ is uniformly integrable.

We next establish the convergence of finite-dimensional distributions. To this end, we first show for fixed $s > 0$, the sequence of moment generating functions

$$\mathbb{E}\left[e^{-\theta\frac{Z_{s\tau_n}-n^{1+s}}{\sqrt{n^{1+2s}}}} \big| Z_0 = n\right]$$

converges when $n \to \infty$. It is straightforward to compute that

$$\mathbb{E}\left[e^{-\theta\frac{Z_{s\tau_n}-n^{1+s}}{\sqrt{n^{1+2s}}}} \big| Z_0 = n\right] = \exp\left(A\left(s\tau_n, -\frac{\theta}{\sqrt{n^{1+2s}}}\right)\cdot n + \sqrt{n}\theta\right), \tag{3.31}$$

where A solves the differential equation in (3.8). Noting that $e^{(\alpha-\beta)s\tau_n} = n^s$, which implies

$$e^{(\alpha-\beta)s\tau_n}\cdot\frac{-\theta}{\sqrt{n^{1+2s}}} = \frac{-\theta}{\sqrt{n}} \ll 1, \quad \text{for } n \text{ sufficiently large.}$$

Hence we obtain from (3.16) the following asymptotic expansion:

$$A\left(s\tau_n, -\frac{\theta}{\sqrt{n^{1+2s}}}\right) = n^s\cdot\left(-\frac{\theta}{\sqrt{n^{1+2s}}}\right) + \frac{1}{2}\frac{\alpha^2}{\alpha-\beta}(n^{2s}-n^s)\frac{\theta^2}{n^{1+2s}} + o\left(\frac{1}{n}\right), \tag{3.32}$$

where $o\left(\frac{1}{n}\right)$ is a term that is smaller in magnitude than any multiple of $1/n$ as $n \to \infty$. We then immediately obtain from (3.31) that

$$\lim_{n\to\infty}\mathbb{E}\left[e^{-\theta\frac{Z_{s\tau_n}-n^{1+s}}{\sqrt{n^{1+2s}}}} \big| Z_0 = n\right] = \exp\left(\frac{1}{2}\frac{\alpha^2}{\alpha-\beta}\theta^2\right),$$

which implies that for fixed $s > 0$,

$$\frac{Z_{s\tau_n} - n^{1+s}}{\sqrt{n^{1+2s}}} \to \xi, \quad \text{weakly as } n \to \infty,$$

where ξ is a normal random variable with mean zero and variance $\frac{\alpha^2}{\alpha-\beta}$.

After establishing the convergence of one-dimensional marginal distributions, we proceed to consider the dimension of two. The general case of finite dimensions follows similarly and the proof is omitted. Fix $\theta_1, \theta_2 \in \mathbb{R}$. For any $0 < u < v$, and for sufficiently large n, the moment generating function is given by

$$\mathbb{E}\left[e^{-\frac{\theta_1}{\sqrt{n}n^u} Z_{u\tau_n} - \frac{\theta_2}{\sqrt{n}n^v} Z_{v\tau_n}} \big| Z_0 = n \right] \cdot e^{\theta_1\sqrt{n} + \theta_2\sqrt{n}}. \tag{3.33}$$

It can be directly computed that

$$\begin{aligned} &\mathbb{E}\left[e^{-\frac{\theta_1}{\sqrt{n}n^u} Z_{u\tau_n} - \frac{\theta_2}{\sqrt{n}n^v} Z_{v\tau_n}} \big| Z_0 = n \right] \\ = \ & \mathbb{E}\left[e^{-\frac{\theta_1}{\sqrt{n}n^u} Z_{u\tau_n}} \cdot \mathbb{E}\left[e^{-\frac{\theta_2}{\sqrt{n}n^v} Z_{v\tau_n}} \big| Z_{u\tau_n} \right] \Big| Z_0 = n \right] \\ = \ & \mathbb{E}\left[e^{-\frac{\theta_1}{\sqrt{n}n^u} Z_{u\tau_n}} \cdot \left[e^{A((v-u)\tau_n, -\frac{\theta_2}{\sqrt{n}n^v}) \cdot Z_{u\tau_n}} \right] \Big| Z_0 = n \right], \end{aligned}$$

hence the moment generating function in (3.33) becomes,

$$\mathbb{E}\left[e^{-\hat{\theta}_1^{(n)} \frac{Z_{u\tau_n} - n^{1+u}}{\sqrt{n^{1+2u}}}} \big| Z_0 = n \right] \cdot \exp(\Gamma^n),$$

where

$$\begin{aligned} \hat{\theta}_1^{(n)} \ &:= \ \theta_1 - \sqrt{n^{1+2u}} \cdot A\left((v-u)\tau_n, -\frac{\theta_2}{\sqrt{n}n^v} \right), \\ \Gamma^n \ &:= \ \left(\theta_2 + \sqrt{n^{1+2u}} \cdot A\left((v-u)\tau_n, -\frac{\theta_2}{\sqrt{n}n^v} \right) \right) \cdot \sqrt{n}. \end{aligned}$$

Similar as in (3.32), we have the following asymptotic expansion for A:

$$\begin{aligned} &A\left((v-u)\tau_n, -\frac{\theta_2}{\sqrt{n}n^v} \right) \\ = \ & n^{v-u} \cdot \left(-\frac{\theta_2}{\sqrt{n^{1+2v}}} \right) + \frac{1}{2}\frac{\alpha^2}{\alpha-\beta}(n^{2(v-u)} - n^{v-u})\frac{\theta_2^2}{n^{1+2v}} + o\left(n^{-1-2u}\right) \\ = \ & -\frac{\theta_2}{\sqrt{n^{1+2u}}} + \frac{1}{2}\frac{\alpha^2}{\alpha-\beta}\frac{1}{n^{1+2u}} + o\left(n^{-1-2u}\right), \end{aligned}$$

where $o\left(n^{-1-2u}\right)$ is a term that is smaller in magnitude than any multiple of n^{-1-2u} as $n \to \infty$. This suggests that

$$\begin{aligned} \lim_{n\to\infty} \hat{\theta}_1^{(n)} &= \theta_1 + \theta_2, \\ \lim_{n\to\infty} \Gamma^n &= 0. \end{aligned}$$

It immediately follows that

$$\lim_{n\to\infty} \mathbb{E}\left[e^{-\frac{\theta_1}{\sqrt{n}n^u} Z_{u\tau_n} - \frac{\theta_2}{\sqrt{n}n^v} Z_{v\tau_n}} | Z_0 = n \right] \cdot e^{\theta_1\sqrt{n} + \theta_2\sqrt{n}} = \exp\left(\frac{1}{2} \frac{\alpha^2}{\alpha - \beta} (\theta_1 + \theta_2)^2 \right),$$

which implies that

$$\left(\frac{Z_{u\tau_n} - n^{1+u}}{\sqrt{n^{1+2u}}}, \frac{Z_{v\tau_n} - n^{1+v}}{\sqrt{n^{1+2v}}} \right) \to (Y(u), Y(v)), \quad \text{weakly as } n \to \infty,$$

where $Y(u) = Y(v) := \xi$ for $v > u > 0$, and ξ is a normal random variable with mean zero and variance $\frac{\alpha^2}{\alpha-\beta}$. It is clear that Y is a continuous process on $D[t,T]$ for any $0 < t < T$. Thus by Proposition 1.2 in Aldous [1], we deduce that

$$\left\{ \frac{Z_{s\tau_n} - n^{1+s}}{\sqrt{n^{1+2s}}} : s \in [t,T] \right\} \to Y,$$

weakly on $D[t,T]$.

We next prove (2.16), i.e., the functional central limit theorem for rescaled jump process N in the super–critical case. Recall from (3.3) that

$$\frac{N_{s\tau_n} - \frac{n^{1+s}-n}{\alpha-\beta}}{n^{\frac{1}{2}+s}} = \frac{Z_{s\tau_n} - n^{1+s}}{\alpha n^{\frac{1}{2}+s}} + \frac{\beta}{\alpha}\tau_n \int_0^s n^{u-s} \frac{Z_{u\tau_n} - n^{1+u}}{n^{\frac{1}{2}+u}} du,$$

and that $Y_s \equiv \xi$ for $s > 0$, where ξ is a normal random variable with mean 0 and variance $\frac{\alpha^2}{\alpha-\beta}$. In addition, note that

$$\tau_n \int_0^s n^{u-s} du = \frac{1}{\alpha - \beta}(1 - n^{-s}).$$

Therefore we deduce that

$$\left\{ \frac{N_{s\tau_n} - \frac{n^{1+s}-n}{\alpha-\beta}}{n^{\frac{1}{2}+s}} : s \in [t,T] \right\} \to \frac{\xi}{\alpha} + \frac{\beta}{\alpha} \cdot \frac{1}{\alpha-\beta}\xi,$$

weakly on $D[t,T]$. The weak limit is simply $\frac{1}{\alpha-\beta}\xi$. □

3.6.3 Proof of super–critical case when $\mu > 0$

Proof. Zhu [47] showed that when $\alpha > \beta$, $\lim_{t\to\infty} \frac{1}{t} \log N_t^{(1)} = \alpha - \beta$. Note that $Z_t^{(1)} \leq \alpha N_t^{(1)}$ and for any $0 < T < \frac{1}{2}$,

$$\sup_{0\leq s\leq T} \frac{Z_{s\tau_n}^{(1)}}{\sqrt{n^{1+2s}}} \leq \frac{\alpha N_{T\tau_n}^{(1)}}{\sqrt{n}} \to 0,$$

almost surely as $n \to \infty$. Similarly, we have

$$\sup_{0\leq s\leq T} \frac{N_{s\tau_n}^{(1)}}{\sqrt{n^{1+2s}}} \leq \frac{N_{T\tau_n}^{(1)}}{\sqrt{n}} \to 0,$$

almost surely as $n \to \infty$. Hence the results follow from the decompositions $Z = Z^{(0)} + Z^{(1)}$ and $N = N^{(0)} + N^{(1)}$. $\square$

3.6.4 Proof of sub–critical case when $\mu = 0$

Proof. We prove the convergence of finite dimensional distributions of the rescaled Z processes. To this end, we first show for fixed $s < 1$, the sequence of moment generating functions

$$\mathbb{E}\left[e^{-\theta \frac{Z_{st_n} - n^{1-s}}{\sqrt{n^{1-s}}}} \big| Z_0 = n\right]$$

converges when $n \to \infty$. It is straightforward to compute that for any fixed $\theta \in \mathbb{R}$, and for any sufficiently large n,

$$\mathbb{E}\left[e^{-\theta \frac{Z_{st_n} - n^{1-s}}{\sqrt{n^{1-s}}}} \big| Z_0 = n\right] = \exp\left(A\left(st_n, -\frac{\theta}{\sqrt{n^{1-s}}}\right) \cdot n + \theta\sqrt{n^{1-s}}\right).$$

Noting that $e^{(\alpha-\beta)st_n} = n^{-s}$. This term, together with the quantity $-\frac{\theta}{\sqrt{n^{1-s}}}$, goes to zero as $n \to \infty$. Hence we obtain from (3.16) the following asymptotic expansion:

$$A\left(st_n, -\frac{\theta}{\sqrt{n^{1-s}}}\right) = n^{-s} \cdot \left(-\frac{\theta}{\sqrt{n^{1-s}}}\right) + \frac{1}{2}\frac{\alpha^2}{\alpha - \beta}(n^{-2s} - n^{-s})\frac{\theta^2}{n^{1-s}} + o\left(n^{-1}\right).$$

We then immediately obtain that

$$\mathbb{E}\left[e^{-\theta \frac{Z_{st_n} - n^{1-s}}{\sqrt{n^{1-s}}}} \big| Z_0 = n\right] = \exp\left(\frac{1}{2}\frac{\alpha^2}{\beta - \alpha}\theta^2\right),$$

which implies that for fixed $s \in (0, 1)$,

$$\frac{Z_{st_n} - n^{1-s}}{\sqrt{n^{1-s}}} \to R_s, \quad \text{weakly as } n \to \infty$$

where R_s is a normal random variable with mean zero and variance $\frac{\alpha^2}{\beta-\alpha}$.

We proceed to consider the dimension of two. The proof for the general case of finite dimensions follows from a similar argument and hence is omitted. Fix $\theta_1, \theta_2 \in \mathbb{R}$. For any $0 < u < v < 1$, and any sufficiently large n, the moment generating function is given by

$$\mathbb{E}\left[e^{-\frac{\theta_1}{\sqrt{n^{1-u}}}Z_{ut_n}-\frac{\theta_2}{\sqrt{n^{1-v}}}Z_{vt_n}}|Z_0=n\right]\cdot e^{\theta_1\sqrt{n^{1-u}}+\theta_2\sqrt{n^{1-v}}}. \tag{3.34}$$

It can be directly computed that

$$\begin{aligned}
&\mathbb{E}\left[e^{-\frac{\theta_1}{\sqrt{n^{1-u}}}Z_{ut_n}-\frac{\theta_2}{\sqrt{n^{1-v}}}Z_{vt_n}}|Z_0=n\right]\\
&= \mathbb{E}\left[e^{-\frac{\theta_1}{\sqrt{n^{1-u}}}Z_{ut_n}}\cdot\mathbb{E}\left[e^{-\frac{\theta_2}{\sqrt{n^{1-v}}}Z_{vt_n}}|Z_{ut_n}\right]\Big|Z_0=n\right]\\
&= \mathbb{E}\left[e^{-\frac{\theta_1}{\sqrt{n^{1-u}}}Z_{ut_n}}\cdot\left[e^{A((v-u)t_n,-\frac{\theta_2}{\sqrt{n^{1-v}}})\cdot Z_{ut_n}}\right]\Big|Z_0=n\right].
\end{aligned}$$

After rewriting, the moment generating function in (3.34) becomes,

$$\mathbb{E}\left[e^{-\hat{\theta}_1^{(n)}\frac{Z_{ut_n}-n^{1-u}}{\sqrt{n^{1-u}}}}\Big|Z_0=n\right]\cdot\exp(\Gamma^n)$$

where

$$\begin{aligned}
\hat{\theta}_1^{(n)} &:= \theta_1-\sqrt{n^{1-u}}\cdot A\left((v-u)t_n,-\frac{\theta_2}{\sqrt{n^{1-v}}}\right),\\
\Gamma^n &:= n^{1-u}\cdot A\left((v-u)t_n,-\frac{\theta_2}{\sqrt{n^{1-v}}}\right)+\theta_2\cdot\sqrt{n^{1-v}}.
\end{aligned}$$

Note from (3.16) that we have the following asymptotic expansion for A:

$$\begin{aligned}
&A\left((v-u)t_n,-\frac{\theta_2}{\sqrt{n^{1-v}}}\right)\\
&= n^{u-v}\cdot\left(-\frac{\theta_2}{\sqrt{n^{1-v}}}\right)+\frac{1}{2}\frac{\alpha^2}{\alpha-\beta}(n^{2(v-u)}-n^{v-u})\frac{\theta_2^2}{n^{1-v}}+o\left(\frac{1}{n^{1-u}}\right)\\
&= -\frac{\theta_2}{n^{\frac{1+v}{2}-u}}+\frac{1}{2}\frac{\alpha^2\theta_2^2}{\beta-\alpha}\frac{1}{n^{1-u}}+o\left(\frac{1}{n^{1-u}}\right).
\end{aligned}$$

This suggests that

$$\begin{aligned}
\lim_{n\to\infty}\hat{\theta}_1^{(n)} &= \theta_1,\\
\lim_{n\to\infty}\Gamma^n &= \frac{1}{2}\frac{\alpha^2}{\beta-\alpha}\theta_2^2.
\end{aligned}$$

It immediately follows that

$$\lim_{n\to\infty}\mathbb{E}\left[e^{-\frac{\theta_1}{\sqrt{n^{1-u}}}Z_{ut_n}-\frac{\theta_2}{\sqrt{n^{1-v}}}Z_{vt_n}}|Z_0=n\right]\cdot e^{\theta_1\sqrt{n^{1-u}}+\theta_2\sqrt{n^{1-v}}}=\exp\left(\frac{1}{2}\frac{\alpha^2}{\beta-\alpha}(\theta_1^2+\theta_2^2)\right),$$

which implies that

$$\left(\frac{Z_{ut_n}-n^{1-u}}{\sqrt{n^{1-u}}},\frac{Z_{vt_n}-n^{1-v}}{\sqrt{n^{1-v}}}\right)\to(R_u,R_v),\quad\text{weakly as } n\to\infty,$$

where R_u and R_v are independent normal random variables, both with mean zero and variance $\frac{\alpha^2}{\beta-\alpha}$. □

3.6.5 Proof of sub–critical case $\mu > 0$

Proof. Note that for $\beta>\alpha>0$, $\frac{N_t^{(1)}}{t}\to\frac{\mu}{1-\frac{\alpha}{\beta}}$ almost surely as $t\to\infty$. Therefore, for any $0<T<1$,

$$\sup_{0\le s\le T}\frac{Z_{st_n}^{(1)}}{\sqrt{n^{1-s}}}\le\frac{\alpha N_{Tt_n}^{(1)}}{\sqrt{n^{1-T}}}\to 0,$$

almost surely as $n\to\infty$. Hence the result follows from the decomposition $Z=Z^{(0)}+Z^{(1)}$. □

4 Appendix

4.1 Proof of Proposition 5

Proof. The process Z is a piecewise deterministic Markov process as defined in Davis [14]; see also Chapter 11 in Rolski et al. [44]. Suppose f is any continuous function with at most polynomial growth on $(0,\infty)$, i.e., there exists some integer k such that $|f(x)|\le x^k$ for all $x\in(0,\infty)$. Recall that $dZ_t=-\beta Z_t dt+\alpha dN_t$. Define

$$\mathcal{A}f(z)=-\beta z\frac{\partial f}{\partial z}+z[f(z+\alpha)-f(z)]. \tag{4.1}$$

We first verify that f is in the domain of the full generator of the Markov process $\{Z_t:t\ge 0\}$[6]. That is,

$$\left\{f(Z_t)-f(Z_0)-\int_0^t\mathcal{A}f(Z_s)ds:t\ge 0\right\}$$

is a martingale with respect to the natural filtration generated by Z. (See Section 11.1.4 in Rolski et al. (2009) for the definition of full generator and further details). We use Theorem

[6]The domain of the infinitesimal generator of a Markov process is always contained in the domain of its full generator.

11.2.2 in Rolski et al. [44] and check the three conditions there. Since the boundary set of the piecewise deterministic Markov process Z is empty (Z never hits zero), and the sample path of Z is absolutely continuous between jumps, it suffices to check that for each $t \geq 0$,

$$\mathbb{E}\left(\sum_{\tau_i \leq t} |f(Z_{\tau_i}) - f(Z_{\tau_i -})|\right) < \infty,$$

where $\sigma_i's$ are the jump epochs of the process Z. Direct computation yields

$$\begin{aligned}
\mathbb{E}\left(\sum_{\tau_i \leq t} |f(Z_{\tau_i}) - f(Z_{\tau_i -})|\right) &= \mathbb{E}\left[\int_0^t |f(Z_{s-} + \alpha) - f(Z_{s-})| dN_s\right] \\
&\leq 2\mathbb{E}\left[\int_0^t (Z_{s-} + \alpha)^k dN_s\right] \\
&\leq 2\mathbb{E}\left[\int_0^t (Z_0 + \alpha N_{s-} + \alpha)^k dN_s\right] \\
&\leq 2^k\mathbb{E}\left[\int_0^t ((Z_0 + \alpha)^k + \alpha^k N_{s-}^k) dN_s\right] \\
&\leq 2^k \left((Z_0 + \alpha)^k \mathbb{E}[N_t] + \alpha^k \mathbb{E}[N_t^{k+1}]\right) < \infty,
\end{aligned}$$

where we have used the facts that f is of at most polynomial growth, and $Z_t \leq Z_0 + \alpha N_t$ for all t. Write $\mathbb{E}_{Z_0}[\cdot] = \mathbb{E}[\cdot | Z_0]$. The martingale property then implies

$$\mathbb{E}_{Z_0}[f(Z_t)] = f(Z_0) + \mathbb{E}_{Z_0}\left[\int_0^t \mathcal{A}f(Z_s) ds\right], \tag{4.2}$$

where $\mathcal{A}f$ is given in (4.1).

To compute the first moment, we apply $f(z) = z$ in (4.2) and obtain

$$\mathbb{E}_{Z_0}[Z_t] = Z_0 + \mathbb{E}_{Z_0}\left[\int_0^t (\alpha - \beta) Z_s ds\right] = Z_0 + \int_0^t (\alpha - \beta) \cdot \mathbb{E}_{Z_0}[Z_s] ds,$$

which implies

$$\frac{d}{dt}\mathbb{E}_{Z_0}[Z_t] = (\alpha - \beta) \cdot \mathbb{E}_{Z_0}[Z_t].$$

Thus we have

$$\mathbb{E}_{Z_0}[Z_t] = Z_0 e^{(\alpha - \beta)t}.$$

Now we compute second moments. Applying $f(z) = z^2$ in (4.2), we find

$$\begin{aligned}
\mathbb{E}_{Z_0}[Z_t^2] &= Z_0^2 + 2(\alpha - \beta) \int_0^t \mathbb{E}_{Z_0}[Z_s^2] ds + \alpha^2 \int_0^t \mathbb{E}_{Z_0}[Z_s] ds \\
&= Z_0^2 + 2(\alpha - \beta) \int_0^t \mathbb{E}_{Z_0}[Z_s^2] ds + \alpha^2 Z_0 \int_0^t e^{(\alpha - \beta)s} ds.
\end{aligned}$$

Solving this equation yields

$$\mathbb{E}_{Z_0}[Z_t^2] = \begin{cases} Z_0^2 + \alpha^2 Z_0 t, & \text{if} \quad \alpha = \beta, \\ Z_0^2 e^{2(\alpha-\beta)t} + \frac{\alpha^2 Z_0}{\alpha-\beta}(e^{2(\alpha-\beta)t} - e^{(\alpha-\beta)t}), & \text{if} \quad \alpha \neq \beta \end{cases}$$

Next we compute the third moments. We apply $f(z) = z^3$ in (4.2). If $\alpha = \beta$, we immediately get

$$\begin{aligned} \mathbb{E}_{Z_0}[Z_t^3] &= Z_0^3 + 3\alpha^2 \int_0^t \mathbb{E}_{Z_0}[Z_s^2]ds + \alpha^3 \int_0^t \mathbb{E}_{Z_0}[Z_s]ds \\ &= Z_0^3 + 3\alpha^2 Z_0^2 t + \frac{3}{2}\alpha^4 Z_0 t^2 + \alpha^3 Z_0 t. \end{aligned}$$

When $\alpha \neq \beta$, we have

$$\begin{aligned} \mathbb{E}_{Z_0}[Z_t^3] &= Z_0^3 + 3(\alpha-\beta)\int_0^t \mathbb{E}_{Z_0}[Z_s^3]ds + 3\alpha^2 \int_0^t \mathbb{E}_{Z_0}[Z_s^2]ds + \alpha^3 \int_0^t \mathbb{E}_{Z_0}[Z_s]ds \\ &= Z_0^3 + 3(\alpha-\beta)\int_0^t \mathbb{E}_{Z_0}[Z_s^3]ds + 3\alpha^2 Z_0^2 \int_0^t e^{2(\alpha-\beta)s}ds \\ &\quad + 3\alpha^2 \frac{\alpha^2 Z_0}{\alpha-\beta}\int_0^t (e^{2(\alpha-\beta)s} - e^{(\alpha-\beta)s})ds + \alpha^3 Z_0 \int_0^t e^{(\alpha-\beta)s}ds \end{aligned}$$

Therefore,

$$\frac{d}{dt}\mathbb{E}_{Z_0}[Z_t^3] = 3(\alpha-\beta)\mathbb{E}_{Z_0}[Z_t^3] + \left(3\alpha^2 Z_0^2 + \frac{3\alpha^4 Z_0}{\alpha-\beta}\right)e^{2(\alpha-\beta)t} + \left(\alpha^3 Z_0 - \frac{3\alpha^4 Z_0}{\alpha-\beta}\right)e^{(\alpha-\beta)t},$$

which yields that

$$\begin{aligned} \mathbb{E}_{Z_0}[Z_t^3]e^{-3(\alpha-\beta)t} - Z_0^3 &= \left(\frac{3\alpha^2 Z_0^2}{\alpha-\beta} + \frac{3\alpha^4 Z_0}{(\alpha-\beta)^2}\right)(1 - e^{-(\alpha-\beta)t}) \\ &\quad + \left(\frac{\alpha^3 Z_0}{2(\alpha-\beta)} - \frac{3\alpha^4 Z_0}{2(\alpha-\beta)^2}\right)(1 - e^{-2(\alpha-\beta)t}), \end{aligned}$$

which implies that

$$\begin{aligned} \mathbb{E}_{Z_0}[Z_t^3] &= \left(Z_0^3 + \frac{3\alpha^2 Z_0^2}{\alpha-\beta} + \frac{3\alpha^4 Z_0}{(\alpha-\beta)^2} + \frac{\alpha^3 Z_0}{2(\alpha-\beta)} - \frac{3\alpha^4 Z_0}{2(\alpha-\beta)^2}\right)e^{3(\alpha-\beta)t} \\ &\quad - \left(\frac{3\alpha^2 Z_0^2}{\alpha-\beta} + \frac{3\alpha^4 Z_0}{(\alpha-\beta)^2}\right)e^{2(\alpha-\beta)t} \\ &\quad - \left(\frac{\alpha^3 Z_0}{2(\alpha-\beta)} - \frac{3\alpha^4 Z_0}{2(\alpha-\beta)^2}\right)e^{(\alpha-\beta)t}. \end{aligned}$$

The proof is therefore complete. $\square$

4.2 Proof of Proposition 6

Proof. First, recall from (3.4) that

$$Z_t - ne^{(\alpha-\beta)t} = e^{(\alpha-\beta)t}\alpha \int_0^t e^{-(\alpha-\beta)s} dM_s. \tag{4.3}$$

Note that $\int_0^t e^{-(\alpha-\beta)s} dM_s$ is a martingale with predictable quadratic variation $\int_0^t e^{-2(\alpha-\beta)s} d\langle M\rangle_s$, where $\langle M\rangle_t = \int_0^t Z_s ds$ is the predictable quadratic variation of the martingale M. By Burkholder-Davis-Gundy inequality, we have

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^t e^{-(\alpha-\beta)s} dM_s\right)^4\right] &\leq \mathbb{E}\left[\left(\sup_{0\leq s\leq t}\left|\int_0^s e^{-(\alpha-\beta)u} dM_u\right|\right)^4\right] \\
&\leq C \cdot \mathbb{E}\left[\left(\int_0^t e^{-2(\alpha-\beta)s} d\langle M\rangle_s\right)^2\right],
\end{aligned}$$

where C is a positive constant. Hence, when $\alpha \neq \beta$, we deduce from Proposition 5 that

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t e^{-(\alpha-\beta)s} dM_s\right)^4\right] \\
&\leq C \cdot \mathbb{E}\left[\left(\int_0^t e^{-2(\alpha-\beta)s} Z_s ds\right)^2\right] \text{(Burkholder-Davis-Gundy inequality)} \\
&\leq Ct\mathbb{E}\left[\int_0^t e^{-4(\alpha-\beta)s} Z_s^2 ds\right] \text{(Cauchy-Schwarz inequality)} \\
&= Ct\int_0^t e^{-4(\alpha-\beta)s}\mathbb{E}[Z_s^2] ds \\
&= Ct\int_0^t e^{-4(\alpha-\beta)s}\left[Z_0^2 e^{2(\alpha-\beta)s} + \frac{\alpha^2 Z_0}{\alpha-\beta}(e^{2(\alpha-\beta)s} - e^{(\alpha-\beta)s})\right] ds \\
&= Ct\left[\frac{n^2}{2(\alpha-\beta)}(1-e^{-2(\alpha-\beta)t}) + \frac{\alpha^2 n}{2(\alpha-\beta)^2}(1-e^{-2(\alpha-\beta)t}) - \frac{\alpha^2 n}{3(\alpha-\beta)^2}(1-e^{-3(\alpha-\beta)t})\right].
\end{aligned}$$

Hence we have established (3.7).

A similar argument yields that

$$\begin{aligned}
&\sup_{\delta\leq t\leq T}\mathbb{E}\left[\left(\int_{t-\delta}^t e^{-(\alpha-\beta)s} dM_s\right)^4\right] \\
&\leq C\delta \sup_{\delta\leq t\leq T}\int_{t-\delta}^t e^{-4(\alpha-\beta)s}\left[Z_0^2 e^{2(\alpha-\beta)s} + \frac{\alpha^2 Z_0}{\alpha-\beta}(e^{2(\alpha-\beta)s} - e^{(\alpha-\beta)s})\right] ds \\
&\leq Cn^2\delta^2.
\end{aligned}$$

Thus we have proved (3.6).

Finally we prove (3.5). When $\alpha = \beta$, a similar argument leads to

$$\mathbb{E}\left[\left(\int_0^t e^{-(\alpha-\beta)s} dM_s\right)^4\right] \le Ct \int_0^t e^{-4(\alpha-\beta)s}\mathbb{E}[Z_s^2]ds$$
$$= Ct\left(Z_0^2 t + \frac{1}{2}\alpha^2 t^2 Z_0\right) = Ct\left(n^2 t + \frac{1}{2}\alpha^2 t^2 n\right).$$

Together with (3.7), we infer that the inequality (3.5) follows in view of (4.3). The proof is therefore complete. □

Acknowledgements

Xuefeng Gao acknowledges support from CUHK Direct Grants for Research with project codes 4055035 and 4055054. We would like to thank Peter Carr, Jasmine Foo and Frederi G. Viens for helpful discussions.

References

[1] Aldous, D. (1989). Stopping times and tightness. II. *The Annals of Probability.* **17(2)**, 586-595.

[2] Bacry, E, S. Delattre, M. Hoffmann and J. F. Muzy. (2013). Scaling limits for Hawkes processes and application to financial statistics. *Stochastic Processes and their Applications* **123**, 2475-2499.

[3] Bacry, E., Mastromatteo, I., and Muzy, J. F. (2015). Hawkes processes in finance. *arXiv:1502.04592*.

[4] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd edition. Wiley-Interscience, New York.

[5] Björk, T. *An Introduction to Point Processes from a Martingale Point of View.* KTH, 2011.

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

[7] Bowsher, C. G. (2007). Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics.* **141(2)**, 876-912.

[8] Bramson, M. (2008) *Stability of queueing networks.* Springer Berlin Heidelberg.

[9] Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Probab.*, **24**, 1563-1588.

[10] Chevallier, J. Mean-field limit of generalized Hawkes processes. *arXiv:1510.05620*.

[11] Da Fonseca, J., and Zaatour, R. (2014). Hawkes process: Fast calibration, application to trade clustering, and diffusive limit. *Journal of Futures Markets*. **34(6)**, 548-579.

[12] Daley, D. J. and Vere-Jones, D. *An Introduction to the Theory of Point Processes*, Volume I and II, 2nd edition. Springer-Verlag, New York, 2003.

[13] Dai, J. G., and Meyn, S. P. (1995). Stability and convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Transactions on Automatic Control*. **40**, 1889-1904.

[14] Davis, M. H. (1984). Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society. Series B (Methodological)*. 353-388.

[15] Delattre, S., Fournier, N. (2015). Statistical inference versus mean field limit for Hawkes processes. Preprint.

[16] Delattre, S., Fournier, N. and M. Hoffmann. (2014). High dimensional Hawkes processes. *arXiv:1403.5764*.

[17] Dufresne, D. (2001). The integrated square-root process. Working Paper, University of Montreal.

[18] Ethier, S. N., and Kurtz, T. G. (2009). *Markov processes: characterization and convergence (Vol. 282)*. John Wiley and Sons.

[19] Errais, E., Giesecke, K. and Goldberg, L. (2010). Affine point processes and portfolio credit risk. *SIAM J. Financial Math.* **1**, 642-665.

[20] Foo, J. and Leder, K. (2013). Dynamics of cancer recurrence. *Annals of Applied Probability* **23**, 1437-1468.

[21] Foo, J., Leder, K., and J. Zhu. (2014). Escape times for branching processes with random mutational fitness effects. *Stochastic Processes and their Applications*. **124**, 3661-3697.

[22] Göing-Jaeschke, A. and M. Yor. (2003) A survey and some generalizations of Bessel processes. *Bernoulli*. **9(2)**, 313–349.

[23] Gourieroux, C., Jasiak, J., and Le Fol, G. (1999). Intra-day market activity. *Journal of financial markets*. **2(3)**, 193-226.

[24] Hamana, Y. and H. Matsumoto. (2013). The probability distributions of the first hitting times of Bessel processes. *Trans. Amer. Math. Soc.* **365**, 5237-5257.

[25] Hamana, Y. and M. Hiroyuki. (2012). The probability densities of the first hitting times of Bessel processes. *J. Math-for-Ind.* **4B**, 91-95.

[26] Hamana, Y. and M. Hiroyuki. (2014). Asymptotics of the probability distributions of the first hitting times of Bessel processes *Electronic Communications in Probability.* **19**, Article 5.

[27] Hautsch, N. (2004). *Modelling irregularly spaced financial data: theory and practice of dynamic duration models.* Springer Science and Business Media.

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

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

[30] Hewlett, P. (2006). Clustering of order arrivals, price impact and trade path optimization. In *Workshop on Financial Modeling with Jump processes.* Ecole Polytechnique (pp. 6-8).

[31] Hida, T., and Hitsuda, M. (1993). *Gaussian processes* (Vol. 120). American Mathematical Society.

[32] Hodara, P. and E. Löcherbach. (2014). Hawkes processes with variable length memory and an infinite number of components. *arXiv:1410.5235.*

[33] Jaisson, T. and M. Rosenbaum. (2015). Limit theorems for nearly unstable Hawkes processes. *Annals of Applied Probability.* **25**, 600-631.

[34] Jaisson, J. and M. Rosenbaum (2015). Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *arXiv: 1504.03100.*

[35] Karabash, D. and L. Zhu. (2015) Limit theorems for marked Hawkes processes with application to a risk model. *Stochastic Models.* **31**, 433-451.

[36] Kurtz, Thomas G and Protter, Philip. (1991) Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability.* 1035–1070.

[37] Kurtz, T. G. (2011). Equivalence of stochastic equations and martingale problems. In *Stochastic Analysis 2010 (pp. 113-130).* Springer Berlin Heidelberg.

[38] Lifshits, M. A. (2013). *Gaussian random functions (Vol. 322).* Springer Science and Business Media.

[39] Lipster, R. S. and A. N. Shiryaev. *Statistics of Random Processes II. Applications*, 2nd Edition, Springer, 2001.

[40] Murdock, James A. (1999) *Perturbations: theory and methods.* **27**, SIAM.

[41] Ozaki, T. (1979). Maximum likelihood estimation of Hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics.* **31(1)**, 145-155.

[42] Piterbarg, V. I. (2012). *Asymptotic methods in the theory of Gaussian processes and fields (Vol. 148).* American Mathematical Society.

[43] Revuz, D. and M. Yor. (1998) *Continuous Martingales and Brownian Motion.* Springer, 3rd Edition.

[44] Rolski, T., Schmidli, H., Schmidt, V., and Teugels, J. (2009). *Stochastic processes for insurance and finance (Vol. 505).* John Wiley and Sons.

[45] Torrisi, G. L. Gaussian approximation of nonlinear Hawkes processes. Preprint.

[46] Torrisi, G. L. Poisson approximation of point processes with stochastic intensity, and application to nonlinear Hawkes processes. Preprint.

[47] Zhu, L. (2013). *Nonlinear Hawkes Processes.* PhD thesis, New York University.

[48] Zhu, L. (2013). Moderate deviations for Hawkes processes. *Statistics & Probability Letters.* **83**, 885-890.

[49] Zhu, L. (2014). Limit theorems for a Cox-Ingersoll-Ross process with Hawkes jumps. *Journal of Applied Probability.* **51**, 699-712.

[50] Zhu, L. (2013). Central limit theorem for nonlinear Hawkes processes. *Journal of Applied Probability.* **50** 760-771.

[51] Zhu, L. (2015). Large deviations for Markovian nonlinear Hawkes Processes. *Annals of Applied Probability.* **25**, 548-581.

[52] Zhu, L. (2014). Process-level large deviations for nonlinear Hawkes point processes. *Annales de l'Institut Henri Poincaré-Probabilités et Statistiques.* **50**, 845-871.