Precise deviations for Hawkes processes

FUQING GAO[1] and LINGJIONG ZHU[2]

[1]*School of Mathematics and Statistics, Wuhan University, Wuhan 430072, People's Republic of China.
E-mail: fqgao@whu.edu.cn*
[2]*Department of Mathematics, Florida State University, 1017 Academic Way, Tallahassee, FL-32306, United States
of America. E-mail: zhu@math.fsu.edu*

Hawkes process is a class of simple point processes with self-exciting and clustering properties. Hawkes process has been widely applied in finance, neuroscience, social networks, criminology, seismology, and many other fields. In this paper, we study precise deviations for Hawkes processes for large time asymptotics, that strictly extends and improves the existing results in the literature. Numerical illustrations will also be provided.

Keywords: Hawkes processes; large deviations; precise deviations

1. Introduction

The Hawkes process is a simple point process N_t with the stochastic intensity given by:

$$\lambda_t = \nu + \int_0^{t-} h(t-s)\, dN_s, \tag{1.1}$$

where $\nu > 0$ and $h : \mathbb{R}_+ \to \mathbb{R}_+$ being locally bounded. We assume that $N_0 = 0$, that is, the Hawkes process starts at time zero with empty history. The Hawkes process is named after Alan Hawkes [23]. In the literature, ν is called the baseline intensity, and h is called the exciting function, or kernel function, encoding the influence of past events on the intensity. Brémaud and Massoulié [9] generalized the dynamics (1.1) and the formula for the intensity (1.1) by a nonlinear function of $\int_0^{t-} h(t-s)\, dN_s$, and hence came the name nonlinear Hawkes processes. The original model (1.1) proposed by Hawkes [23] is thus sometimes referred to as the linear Hawkes process.

For the Hawkes process (1.1), the occurrence of a jump increases the intensity of the point process, and thus increases the likelihood of more future jumps. On the other hand, the intensity declines when there is no occurrence of new jumps. The self-exciting and clustering property makes the Hawkes process very appealing in applications in finance and many other fields. The Hawkes process is widely used in the modeling of the limit order books in high frequency trading, see, for example, Alfonsi and Blanc [4] for optimal execution, and Abergel and Jedidi [1] for ergodicity in Hawkes based limit order books models, and also the modeling of the duration between trades, see, for example, Bauwens and Hautsch [7] or the arrival process of buy and sell orders, see, for example, Bacry et al. [5]. The Hawkes process also finds applications in dark pool trading [16]. In the context of credit risk modeling, Errais et al. [14] used a top down approach using the affine point process, which includes Hawkes process as a special case. Giot [20], Chavez-Demoulin et al. [10] tested Hawkes processes in the risk management context. Aït-Sahalia et al. [2,3] used Hawkes processes to model two key aspects of asset prices: clustering in time and cross sectional contamination between regions. The Hawkes process has also been used to explain the supply and demand microstructure in an interest rate model in Hainaut [22]. The applications other than finance include: neuroscience, see, for example, [31,33], genome analysis, see, for example, [21,33], networks and sociology, see, for example, [12,40], queueing theory, see, for example, [17], insurance, see, for example, [35,41], criminology, see, for example, [28,32], seismology, see, for example, [29,30] and many other fields.

In this paper, we consider the linear Hawkes process N_t with $N_0 = 0$ with the intensity (1.1). We assume throughout this paper that

$$\|h\|_{L^1} = \int_0^\infty h(t)\,dt < 1, \quad \text{and} \quad \int_0^\infty th(t)\,dt < \infty. \tag{1.2}$$

Let us first review the limit theorems for linear Hawkes processes in the literature. It is well known that under the assumption $\|h\|_{L^1} < 1$, there exists a unique stationary Hawkes process defined on $\mathbb{R}$ with the intensity $\lambda_t = \nu + \int_{-\infty}^{t-} h(t-s)\,dN_s$, and we have the law of large numbers $\frac{N_t}{t} \to \frac{\nu}{1-\|h\|_{L^1}}$ as $t \to \infty$. Bacry et al. [6] obtained a functional central limit theorem for multivariate Hawkes process and as a special case of their result,

$$\frac{N_t - \frac{\nu}{1-\|h\|_{L^1}}t}{\sqrt{t}} \to N\left(0, \frac{\nu}{(1-\|h\|_{L^1})^3}\right), \tag{1.3}$$

in distribution as $t \to \infty$ under the assumption that $\int_0^\infty t^{1/2}h(t)\,dt < \infty$. For the central limit theorem (1.3), also see Hawkes and Oakes [24]. Moreover, Bordenave and Torrisi [8] proved that $\mathbb{P}(\frac{N_t}{t} \in \cdot)$ satisfies a large deviation principle, with the rate function:

$$I(x) = x \log\left(\frac{x}{\nu + x\|h\|_{L^1}}\right) - x + x\|h\|_{L^1} + \nu, \tag{1.4}$$

if $x \geq 0$ and $I(x) = +\infty$ otherwise. We refer to [13] for the theory of large deviations. Note that the rate function in the paper [8] is written as the Legendre transform expression, and the formula (1.4) is first mentioned in [42]. Also notice that in [8], the assumption $\int_0^\infty th(t)\,dt < \infty$ is needed, and indeed this assumption is not necessary, see, for example, [27]. A moderate deviation principle is obtained in [42] that fills in the gap between the central limit theorem and the large deviation principle. Other works on the asymptotics of linear Hawkes processes, including the nearly unstable Hawkes processes, that is, when the $\|h\|_{L^1}$ is close to 1, see, for example, [25,26], and the large initial intensity asymptotics for the Markovian case [18,19], and the large baseline intensity asymptotics [17].

For nonlinear Hawkes processes, [43] studies the central limit theorem, and [44] obtains a process-level, that is, level-3 large deviation principle, and hence has the scalar large deviations as a by-product. An alternative expression for the rate function when the system is Markovian is obtained in [45]. Recently, Torrisi [36,37] studies the rate of convergence in the Gaussian and Poisson approximations of the simple point processes with stochastic intensity, which includes as a special case, the nonlinear Hawkes process.

The large deviations [8] and moderate deviations [42] for linear Hawkes processes are of the Donsker–Varadhan type, which only gives the leading order term, but not the higher order ones. More precise deviations are desirable in many applications, which motivates us to study the precise deviations for linear Hawkes processes. In this paper, we will derive the precise deviations for linear Hawkes processes, using the recent mod-ϕ convergence theory developed in [15]. To apply the mod-ϕ convergence theory, two key steps are to show the convergence of the moment generating function at a certain speed, and verify the limit corresponds to an infinitely divisible distribution. The moment generating function for linear Hawkes processes has semi-explicit formula, and then the precise deviations results follow from the mod-ϕ convergence theory after careful analysis and a series of propositions and lemmas. The paper is organized as follows. We will state the main results of our paper in Section 2. In particular, we will give precise large deviations results in Section 2.1 and precise moderate deviations and fluctuation results in Section 2.2. We provide numerical illustrations in Section 3. All the proofs will be given in the Appendix.

2. Main results

In this paper, we apply the recently developed mod-ϕ convergence method [15] to obtain precise large deviations for linear Hawkes processes for the large time asymptotic regime. We will also obtain the precise moderate deviations and some fluctuations results.

Let us first recall the definition of mod-ϕ convergence, see, for example, Definition 1.1.1. in [15]. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real-valued random variables and $\mathbb{E}[e^{zX_n}]$ exist in a strip $\mathcal{S}_{(c,d)} := \{z \in \mathbb{C} : c < \mathcal{R}(z) < d\}$, with $c < 0 < d$ extended real numbers, that is, we allow $c = -\infty$ and $d = +\infty$ and $\mathcal{R}(z)$ denotes the real part of $z \in \mathbb{C}$ throughout this paper. We assume that there exists a non-constant infinitely divisible distribution ϕ with $\int_{\mathbb{R}} e^{zx}\phi(dx) = e^{\eta(z)}$, which is well defined on $\mathcal{S}_{(c,d)}$, and an analytic function $\psi(z)$ that does not vanish on the real part of $\mathcal{S}_{(c,d)}$ such that locally uniformly in $z \in \mathcal{S}_{(c,d)}$,

$$e^{-t_n\eta(z)}\mathbb{E}\big[e^{zX_n}\big] \to \psi(z), \tag{2.1}$$

where $t_n \to +\infty$ as $n \to \infty$. Then we say that X_n converges mod-ϕ on $\mathcal{S}_{(c,d)}$ with parameters $(t_n)_{n\in\mathbb{N}}$ and limiting function ψ.

Assume that ϕ is a lattice distribution, that is, a distribution with support included in $\gamma + \lambda\mathbb{Z}$ for some constants $\gamma, \lambda > 0$. Also assume that the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ converges mod-ϕ at speed $O((t_n)^{-v})$ (Definition 2.1.1. in [15]), that is,

$$\sup_{z\in K}\big|e^{-t_n\eta(z)}\mathbb{E}\big[e^{zX_n}\big] - \psi(z)\big| \le C_K(t_n)^{-v}, \tag{2.2}$$

where $C_K > 0$ is some constant, for any compact set $K \subset \mathcal{S}_{(c,d)}$. Then Theorem 3.2.2. in [15] states that for any $x \in \mathbb{R}$ in the interval $(\eta'(c), \eta'(d))$ such that $t_nx \in \mathbb{N}$, we have

$$\mathbb{P}(X_n = t_nx) = \frac{e^{-t_nF(x)}}{\sqrt{2\pi t_n\eta''(\theta^*)}}\left(\psi(\theta^*) + \frac{a_1}{t_n} + \frac{a_2}{t_n^2} + \cdots + \frac{a_{v-1}}{t_n^{v-1}} + O\left(\frac{1}{t_n^v}\right)\right), \tag{2.3}$$

as $n \to \infty$, where θ^* is defined via $\eta'(\theta^*) = x$, and $F(x) := \sup_{\theta\in\mathbb{R}}\{\theta x - \eta(\theta)\}$ is the Legendre transform of $\eta(\cdot)$, and if $x \in \mathbb{R}$ and $x \in (\eta'(0), \eta'(d))$, then, as $n \to \infty$,

$$\mathbb{P}(X_n \ge t_nx) = \frac{e^{-t_nF(x)}}{\sqrt{2\pi t_n\eta''(\theta^*)}}\frac{1}{1 - e^{-\theta^*}}\left(\psi(\theta^*) + \frac{b_1}{t_n} + \frac{b_2}{t_n^2} + \cdots + \frac{b_{v-1}}{t_n^{v-1}} + O\left(\frac{1}{t_n^v}\right)\right), \tag{2.4}$$

where $(a_k)_{k=1}^\infty$, $(b_k)_{k=1}^\infty$ are rational fractions in the derivatives of η and ψ at θ^*.

The mod-ϕ convergence developed in [15] also implies the precise moderate deviations and extended central limit theorem, see Theorem 3.3.1. and Corollary 3.3.5. in [15]. More precisely, the second part of Corollary 3.3.5. in [15] states that assuming the mod-ϕ convergence holds, then for any $y = o((t_n)^{1/2-1/m})$, where $m \ge 3$, $m \in \mathbb{N}$, one has

$$\mathbb{P}\big(X_n \ge t_n\eta'(0) + \sqrt{t_n\eta''(0)}\,y\big) = \frac{(1+o(1))}{y\sqrt{2\pi}} \cdot e^{-\sum_{i=2}^{m-1}\frac{F^{(i)}(\eta'(0))}{i!}\frac{(\eta''(0))^{i/2}y^i}{(t_n)^{(i-2)/2}}}. \tag{2.5}$$

The precise moderate deviations and extended central limit theorem then follow from (2.5). See [15] for further discussions. Note that the precise deviations in (2.3) and (2.4) require mod-ϕ convergence at speed $O((t_n)^{-v})$, while the fluctuations result in (2.5) only requires mod-ϕ convergence.

Finally, [15] suggested in their Remark 3.2.5. how $(a_k)_{k=1}^\infty$, $(b_k)_{k=1}^\infty$ in (2.3) and (2.4) can be computed without providing detailed formulas. For the sake of applications and implementations, we will spell out the explicit formulas for $(a_k)_{k=1}^\infty$ and $(b_k)_{k=1}^\infty$ in the next proposition and the proof will be provided in the Appendix.

Proposition 1. *For any function f, denote its k-th order derivative by $f^{(k)}$. Let $\mathcal{S}_n$ be the set consisting of all the n-tuples of non-negative integers $(m_1, \ldots, m_n)$ satisfying: $1 \cdot m_1 + 2 \cdot m_2 + 3 \cdot m_3 + \cdots + n \cdot m_n = n$.*

(i) *For every $k \geq 1$,*

$$a_k = \sum_{\ell=0}^{2k} \frac{\psi^{(2k-\ell)}(\theta^*)}{(2k-\ell)!} \sum_{\mathcal{S}_\ell} \frac{(-1)^{m_1+\cdots+m_\ell}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_\ell! \ell!^{m_\ell}}$$
$$\cdot \prod_{j=1}^{\ell} \left(\frac{1}{\eta''(\theta^*)} \frac{\eta^{(j+2)}(\theta^*)}{(j+2)(j+1)} \right)^{m_j} \frac{(-1)^k (2(k+m_1+\cdots+m_\ell)-1)!!}{(\eta''(\theta^*))^k}. \tag{2.6}$$

(ii) *For every $k \geq 1$,*

$$b_k = \sum_{n=0}^{2k} \sum_{\mathcal{S}_n} \frac{e^{-(m_1+\cdots+m_n)\theta^*}(m_1+\cdots+m_n)!(1-e^{-\theta^*})^{-(m_1+\cdots+m_n)-1}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_n! n!^{m_n}} \cdot \prod_{j=1}^{n} (-1)^{j \cdot m_j}$$
$$\cdot \sum_{\ell=0}^{2k-n} \frac{\psi^{(2k-n-\ell)}(\theta^*)}{(2k-n-\ell)!} \sum_{\mathcal{S}_\ell} \frac{(-1)^{m_1+\cdots+m_\ell}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_\ell! \ell!^{m_\ell}}$$
$$\cdot \prod_{j=1}^{\ell} \left(\frac{1}{\eta''(\theta^*)} \frac{\eta^{(j+2)}(\theta^*)}{(j+2)(j+1)} \right)^{m_j} \frac{(-1)^k (2(k+m_1+\cdots+m_\ell)-1)!!}{(\eta''(\theta^*))^k}. \tag{2.7}$$

2.1. Precise large deviations

Our main results for the precise large deviations for the Hawkes process is stated as follows. It provides the full expansion to arbitrary order in the large time asymptotic regime, which generalizes the large deviations result in [8].

Theorem 2. *Given $v \in \mathbb{N}$. Assume (1.2) and the following condition hold:*

$$\int_0^\infty s^{v+1} h(s)\,ds < \infty. \tag{2.8}$$

(i) *For any $x > 0$, and $tx \in \mathbb{N}$, as $t \to \infty$,*

$$\mathbb{P}(N_t = tx) = e^{-tI(x)} \sqrt{\frac{I''(x)}{2\pi t}} \left(\psi(\theta^*) + \frac{a_1}{t} + \frac{a_2}{t^2} + \cdots + \frac{a_{v-1}}{t^{v-1}} + O\left(\frac{1}{t^v}\right) \right), \tag{2.9}$$

where for any $\mathcal{R}(z) \leq \|h\|_{L^1} - 1 - \log \|h\|_{L^1}$,

$$\psi(z) := e^{v\varphi(z)}, \quad and \quad \varphi(z) := \int_0^\infty \left[F(s; z) - x(z) \right] ds, \tag{2.10}$$

which is analytic in $\mathcal{S}_{(-\infty, \|h\|_{L^1}-1-\log\|h\|_{L^1})}$, and F is the unique solution that satisfies

$$F(t; z) = e^{z + \int_0^t (F(t-s; z)-1)h(s)\,ds}, \tag{2.11}$$

with the constraint $|F(t;z)| \leq \frac{1}{\|h\|_{L^1}}$, and it is analytic in $\mathcal{S}_{(-\infty, \|h\|_{L^1}-1-\log\|h\|_{L^1})}$, and $x(z) := F(\infty;z)$ exists and it satisfies the equation

$$x(z) = e^{z+\|h\|_{L^1}(x(z)-1)}, \tag{2.12}$$

and it is analytic in $\mathcal{S}_{(-\infty, \|h\|_{L^1}-1-\log\|h\|_{L^1})}$. And $I(x)$ is defined in (1.4), $I''(x) = \frac{v^2}{x(v+\|h\|_{L^1}x)^2}$, and

$$\theta^* = \log\left(\frac{x}{v+\|h\|_{L^1}x}\right) - \frac{\|h\|_{L^1}x}{v+\|h\|_{L^1}x} + \|h\|_{L^1}, \tag{2.13}$$

where $(a_k)_{k=1}^\infty$ are rational fractions in the derivatives of η and ψ at θ^* whose formulas are given in (2.6), where $\eta(z) := v(x(z)-1)$.

(ii) *For any* $x > \frac{v}{1-\|h\|_{L^1}}$ *and* $tx \in \mathbb{N}$, *as* $t \to \infty$,

$$\mathbb{P}(N_t \geq tx) = e^{-tI(x)}\sqrt{\frac{I''(x)}{2\pi t}}\frac{1}{1-e^{-\theta^*}}\left(\psi(\theta^*) + \frac{b_1}{t} + \frac{b_2}{t^2} + \cdots + \frac{b_{v-1}}{t^{v-1}} + O\left(\frac{1}{t^v}\right)\right), \tag{2.14}$$

where $(b_k)_{k=1}^\infty$ are rational fractions in the derivatives of η and ψ at θ^* whose formulas are given in (2.7).

Note that the mod-ϕ convergence theory developed in [15] works for the a sequence of random variables X_n associated with discrete times t_n, see (2.1). In Theorem 2, although N_t is a continuous-time process, the restriction $tx \in \mathbb{N}$ guarantees that t takes values along $t_n := \frac{n}{x}$ so that t takes discrete values and we can define $X_n := N_{n/x}$ to apply the mod-ϕ convergence theory developed for discrete times in [15].

In order to apply and implement Theorem 2, we need to compute the sequences $(a_k)_{k=1}^\infty$ and $(b_k)_{k=1}^\infty$ whose formulas are provided in Proposition 1 which rely on the derivatives of η and ψ at θ^*. Next, we provide recursive formulas to compute the derivatives of η and ψ at θ^* of any order.

Proposition 3.

(i) $\eta(\theta^*) = v(x(\theta^*)-1)$, and for $k \geq 1$, $\eta^{(k)}(\theta^*) = vx^{(k)}(\theta^*)$. For $k \geq 1$, $x^{(k)}(\theta^*)$ can be computed recursively as:

$$x^{(k)}(\theta^*) = \frac{x(\theta^*)}{1-\|h\|_{L^1}x(\theta^*)}\sum_{\mathcal{T}_k}\frac{k!\cdot\|h\|_{L^1}^{m_1+\cdots+m_{k-1}}}{m_1!1!^{m_1}m_2!2!^{m_2}\cdots m_{k-1}!(k-1)!^{m_{k-1}}}\cdot\prod_{j=1}^{k-1}\left(x^{(j)}(\theta^*)\right)^{m_j}$$

$$+ \frac{x(\theta^*)}{1-\|h\|_{L^1}x(\theta^*)}\sum_{\ell=0}^{k-1}\binom{k}{\ell}\sum_{\mathcal{S}_\ell}\frac{\ell!\cdot\|h\|_{L^1}^{m_1+\cdots+m_\ell}}{m_1!1!^{m_1}m_2!2!^{m_2}\cdots m_\ell!\ell!^{m_\ell}}\cdot\prod_{j=1}^{\ell}\left(x^{(j)}(\theta^*)\right)^{m_j},$$

where $\mathcal{T}_k$ denotes the set of $(k-1)$-tuples of non-negative integers $(m_1,\ldots,m_{k-1})$ satisfying the constraint $1\cdot m_1 + 2\cdot m_2 + 3\cdot m_3 + \cdots + (k-1)\cdot m_{k-1} = k$.

(ii) *For every $k \geq 1$,*

$$\psi^{(k)}(\theta^*) = \sum_{\mathcal{S}_k} \frac{k! \cdot \nu^{m_1 + \cdots + m_k} \cdot \psi(\theta^*)}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_k! k!^{m_k}}$$

$$\cdot \prod_{j=1}^{k} \left(\int_0^\infty \left[F^{(j)}(s; \theta^*) - x^{(j)}(\theta^*) \right] ds \right)^{m_j}, \tag{2.15}$$

where $F^{(k)}(\cdot; \theta^*)$, $k \geq 1$, can be computed recursively as

$$F^{(k)}(t; \theta^*)$$

$$= F(t; \theta^*) \cdot \int_0^t F^{(k)}(t - s; \theta^*) h(s) \, ds$$

$$+ F(t; \theta^*) \cdot \sum_{\mathcal{T}_k} \frac{k!}{m_1! 1!^{m_1} \cdots m_{k-1}! (k-1)!^{m_{k-1}}} \prod_{j=1}^{k-1} \left(\int_0^t F^{(j)}(t - s; \theta^*) h(s) \, ds \right)^{m_j}$$

$$+ F(t; \theta^*) \cdot \sum_{\ell=0}^{k-1} \binom{k}{\ell} \sum_{\mathcal{S}_\ell} \frac{\ell!}{m_1! 1!^{m_1} \cdots m_\ell! \ell!^{m_\ell}}$$

$$\times \prod_{j=1}^{\ell} \left(\int_0^t F^{(j)}(t - s; \theta^*) h(s) \, ds \right)^{m_j}. \tag{2.16}$$

The main strategy of the proof of Theorem 2 is to show the mod-ϕ convergence and apply Theorem 3.2.2. in [15]. The proof of Theorem 2 relies on a series of lemmas and propositions that we will state later. We will first recall and discuss some well-known properties for the linear Hawkes process that will be used extensively in our proofs later.

Hawkes and Oakes [24] first discovered that a linear Hawkes process has an immigration-birth representation. The immigrants (roots) arrive according to a standard Poisson process $\bar{N}$ with intensity $\nu > 0$. Each immigrant generates children according to a Galton–Watson tree. Indeed, for an immigrant that arrives at time s, children are generated according to a Poisson process with intensity $h(t - s)$ at time $t \geq s$, and each child will generate children independently following the same mechanism. It turns out that the total number of children generated by the immigrant follows a Poisson distribution with parameter $\|h\|_{L^1}$ and conditional on the total number of children generated by the immigrant, say K, then the birth times of the children subtracting the arrival time of the immigrant are distributed as the order statistics of k i.i.d. random variables with density $h(\cdot)/\|h\|_{L^1}$. Finally, the Hawkes process N_t is the number of all the immigrants and their descendants that arrive on the time interval $[0, t]$.

By using the immigration-birth representation for linear Hawkes processes, it is known that (see, e.g., [16,27,42])

$$\mathbb{E}\left[e^{zN_t} \right] = e^{\nu \int_0^t (F(s; z) - 1) \, ds}, \tag{2.17}$$

where F satisfies the equation:

$$F(t; z) = e^{z + \int_0^t (F(t-s; z) - 1) h(s) \, ds}, \quad t \geq 0, \tag{2.18}$$

which holds for any $z \in \mathbb{C}$, such that

$$\mathcal{R}(z) \leq \theta_c := \|h\|_{L^1} - 1 - \log \|h\|_{L^1}. \tag{2.19}$$

Note that by the immigration-birth representation, we can interpret $F(t; z)$ as $F(t; z) = \mathbb{E}[e^{zS_t}]$, where S_t is the number of all the descendants of an immigrant that arrives at time 0, on the time interval $[0, t]$ including the immigrant. Moreover, let us define $x(z) = \mathbb{E}[e^{zS_\infty}]$. It is well known that $x(z)$ satisfies the algebraic equation, see, for example, [27]:

$$x(z) = e^{z + \|h\|_{L^1}(x(z)-1)}. \tag{2.20}$$

This algebraic equation may have more than one solution. It is known that for $z \in \mathbb{R}$, there are at most two solutions of this algebraic equation and $\mathbb{E}[e^{zS_\infty}]$ is the smaller solution, see, for example, [27]. By dominated convergence theorem, for $\mathcal{R}(z) \leq \theta_c$, where θ_c is defined in (2.19), we have $F(t, z) \to x(z)$, as $t \to \infty$. The limit $x(\cdot)$ has the following properties (Part (ii) of the following Proposition 4 follows from Theorem 3.2.1. in [8]).

Proposition 4. *For any $\theta \in \mathbb{R}$, and $\theta \leq \theta_c$, where θ_c is defined in (2.19), we have*

(i) $x(\theta)\|h\|_{L^1} \leq 1$.
(ii) $x'(\theta) \to \infty$ as $\theta \uparrow \theta_c$.

We know that $\mathbb{E}[e^{zS_t}]$ satisfies the equation (2.11). Thus, as a by-product of the immigration-birth representation for linear Hawkes processes, we get the existence of the solution of (2.11). Let us notice that for any $\mathcal{R}(z) \leq \theta_c$,

$$\left|F(t; z)\right| = \left|\mathbb{E}[e^{zS_t}]\right| \leq \mathbb{E}[\left|e^{zS_t}\right|] = \mathbb{E}[e^{\mathcal{R}(z)S_t}] \leq \frac{1}{\|h\|_{L^1}}. \tag{2.21}$$

Therefore, it suffices to consider the solution of the equation (2.11) that satisfies the constraint $|F(t; z)| \leq \frac{1}{\|h\|_{L^1}}$. With this additional constraint, the equation (2.11) has a unique solution.

Proposition 5. *Let $z \in \mathbb{C}$ and $\mathcal{R}(z) \leq \theta_c$. The equation (2.11) with the constraint $|F(t; z)| \leq \frac{1}{\|h\|_{L^1}}$ has a unique solution.*

The key to prove the main result Theorem 2 is to verify the mod-ϕ convergence. More precisely, we need to show that

$$e^{-t\eta(z)}\mathbb{E}[e^{zN_t}] \to \psi(z) := e^{v\varphi(z)}, \tag{2.22}$$

as $t \to \infty$ locally uniformly in z for $\mathcal{R}(z) < \|h\|_{L^1} - 1 - \log \|h\|_{L^1}$, where $\varphi(z) = \int_0^\infty [F(s; z) - x(z)]\,ds$, is analytic in z and $\eta(z) = v(x(z) - 1)$, and $e^{\eta(z)} = e^{v(x(z)-1)} = \mathbb{E}[e^{zY}]$, for some random variable Y, where Y has an infinitely divisible distribution. We will show the mod-ϕ convergence below via a series of lemmas.

First, we show that Y is infinitely divisible. The infinite divisibility is a limitation of the method of mod-ϕ convergence. Fortunately, the limiting distribution in the case of the linear Hawkes process is indeed infinitely divisible.

Lemma 6. *Y has an infinitely divisible distribution.*

To show the mod-ϕ convergence, the main technical lemma is given as follows.

Lemma 7. *For any* $\mathcal{R}(z) < \theta_c$, *where* θ_c *is defined in* (2.19),

$$\varphi(z) = \int_0^\infty \big[F(s; z) - x(z)\big] ds \tag{2.23}$$

is well-defined and analytic, and as $t \to \infty$,

$$e^{-t(\nu(x(z)-1))}\mathbb{E}\big[e^{zN_t}\big] \to e^{\nu\varphi(z)}, \tag{2.24}$$

locally uniformly in z. *In addition, if* $\int_0^\infty s^{\nu+1}h(s)\,ds < \infty$, *then for any compact set* K, *there exists some* $C_K > 0$ *such that* $\sup_{z \in K} |e^{-t(\nu(x(z)-1))}\mathbb{E}[e^{zN_t}] - e^{\nu\varphi(z)}| \le C_K t^{-\nu}$.

To this end, we have established the mod-ϕ convergence for the linear Hawkes process for the large time limit. The proofs of all the propositions, lemmas and Theorem 2.1 will be given in the Appendix.

2.2. Precise moderate deviations and fluctuations

The mod-ϕ convergence also implies the precise moderate deviations and extended central limit theorem, see Theorem 3.3.1. and Corollary 3.3.5. in [15]. By using Corollary 3.3.5. in [15], which is re-stated in our (2.5), we have the following result.

Theorem 8. *Assume* (1.2) *holds. If* $y = o(t^{1/2-1/m})$, *where* $m \ge 3$, *then as* $t \to \infty$,

$$\mathbb{P}\left(N_t \ge \frac{\nu}{1 - \|h\|_{L^1}}t + \sqrt{t}\frac{\sqrt{\nu}}{(1 - \|h\|_{L^1})^{3/2}}y\right) = \frac{(1 + o(1))}{y\sqrt{2\pi}}e^{-\sum_{i=2}^{m-1}\frac{I^{(i)}(\eta'(0))}{i!}\frac{(\eta''(0))^{i/2}y^i}{t^{(i-2)/2}}}, \tag{2.25}$$

where $I(\cdot)$ *is defined in* (1.4) *and for any* $i \ge 2$,

$$I^{(i)}(x) = (i - 2)!(-1)^{i-2}x^{1-i}\left((i - 1)\left(\frac{\|h\|_{L^1}x}{\nu + \|h\|_{L^1}x}\right)^i - i\left(\frac{\|h\|_{L^1}x}{\nu + \|h\|_{L^1}x}\right)^{i-1} + 1\right), \tag{2.26}$$

and $\eta'(0) = \frac{\nu}{1 - \|h\|_{L^1}}$, *and* $\eta''(0) = \frac{\nu}{(1 - \|h\|_{L^1})^3}$.

Remark 9. By letting $m = 3$ in Theorem 8 and $I''(\eta'(0)) = 1/\eta''(0)$, we get an extended central limit theorem, that is, for any $y = o(t^{1/6})$, as $t \to \infty$,

$$\mathbb{P}\left(N_t \ge \frac{\nu}{1 - \|h\|_{L^1}}t + \sqrt{t}\frac{\sqrt{\nu}}{(1 - \|h\|_{L^1})^{3/2}}y\right) = \overline{\Phi}(y)\big(1 + o(1)\big), \tag{2.27}$$

where $\overline{\Phi}(y) := \int_y^\infty \frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx$. Note that (2.27) extends the univariate case of the Hawkes process central limit theorem considered in [6] since we allow $y = o(t^{1/6})$.

Remark 10. By Theorem 3.3.1. in [15], we can also study the moderate deviations result. If $1 \ll y \ll \sqrt{t}$ for $t \to \infty$,[1] that is, the moderate deviations regime, and if we let:

$$s_t := \frac{\nu}{1 - \|h\|_{L^1}} + \frac{\sqrt{\nu}}{(1 - \|h\|_{L^1})^{3/2}} \frac{y}{\sqrt{t}}, \tag{2.28}$$

then, with $\eta'(\theta^*) = s_t$, as $t \to \infty$,

$$\mathbb{P}\left(N_t \geq \frac{\nu}{1 - \|h\|_{L^1}} t + \sqrt{t} \frac{\sqrt{\nu}}{(1 - \|h\|_{L^1})^{3/2}} y\right) = \frac{e^{-tI(s_t)}}{\theta^* \sqrt{2\pi t \eta''(\theta^*)}} (1 + o(1)). \tag{2.29}$$

Remark 11. By letting $m = 4$ in Theorem 8 and $I''(\eta'(0)) = 1/\eta''(0)$ and $I'''(\eta'(0)) = -\eta'''(0)/(\eta''(0))^3$, we get a precise moderate deviation result, that is, for any $y = o(t^{1/4})$, as $t \to \infty$,

$$\mathbb{P}\left(N_t \geq \frac{\nu}{1 - \|h\|_{L^1}} + \sqrt{t} \frac{\sqrt{\nu}}{(1 - \|h\|_{L^1})^{3/2}} y\right) = \frac{(1 + o(1))}{y\sqrt{2\pi}} e^{-\frac{y^2}{2}} e^{\frac{\eta'''(0)}{6(\eta''(0))^{3/2}} \frac{y^3}{\sqrt{t}}}. \tag{2.30}$$

Note that (8) gives a precise moderate deviation result, and it provides a more precise tail estimate than [42]. To see this, let $y = \frac{(1 - \|h\|_{L^1})^{3/2}}{\sqrt{\nu}} \frac{a(t)}{\sqrt{t}} x$, where x is a constant independent of t. Then for $t^{1/2} \ll a(t) \ll t^{3/4}$, as $t \to \infty$, we have

$$\mathbb{P}\left(N_t \geq \frac{\nu}{1 - \|h\|_{L^1}} t + a(t)x\right)$$

$$= \frac{\sqrt{\nu}(1 + o(1))}{(1 - \|h\|_{L^1})^{3/2} \sqrt{2\pi}} \frac{\sqrt{t}}{a(t)x} e^{-\frac{(1 - \|h\|_{L^1})^3}{2\nu} \frac{a(t)^2}{t} x^2} e^{\frac{\eta'''(0)}{6(\eta''(0))^3} \frac{a(t)^3}{t^2} x^3}$$

$$= \frac{\sqrt{\nu}(1 + o(1))}{(1 - \|h\|_{L^1})^{3/2} \sqrt{2\pi}} \frac{\sqrt{t}}{a(t)x} e^{-\frac{(1 - \|h\|_{L^1})^3}{2\nu} \frac{a(t)^2}{t} x^2} e^{\frac{(1 + 2\|h\|_{L^1})(1 - \|h\|_{L^1})^4}{6\nu^2} \frac{a(t)^3}{t^2} x^3},$$

where $\eta'(0) = \frac{\nu}{1 - \|h\|_{L^1}}$, $\eta''(0) = \frac{\nu}{(1 - \|h\|_{L^1})^3}$, and $\eta'''(0) = \nu \frac{1 + 2\|h\|_{L^1}}{(1 - \|h\|_{L^1})^5}$.[2]

3. Numerical illustrations

In this section, we illustrate our precise deviations results by comparing the approximation of the tail probability $\mathbb{P}(N_t \geq xt)$ by using our formulas and by using Monte Carlo simulations. Since the event $\{N_t \geq xt\}$ we are interested in is a rare event, we will first develop the importance sampling. Rare event simulations using importance sampling have been studied for affine point processes, a generalization of the linear Hawkes process when the exciting function is exponential, see [38,39]. We are interested in the importance sampling for linear Hawkes processes with general exciting function, which are non-Markovian in general, and is not covered in [38,39]. The importance sampling has also been used to estimate the ruin probability in a risk model where the arrival process of the claims follows a Hawkes process [35].

---

[1]For any two functions f and g, $f(t) \ll g(t)$ as $t \to \infty$ means $\frac{f(t)}{g(t)} \to 0$ as $t \to \infty$.
[2]The derivation of $\eta'''(0)$ will be given in the proof of Theorem 8.

We are interested to estimate right tail probability $\mathbb{P}(N_t \geq xt)$, where $x > \frac{v}{1-\|h\|_{L^1}}$. Note that $\{N_t \geq xt\}$ is a rare event for $x > \frac{v}{1-\|h\|_{L^1}}$. The idea of the importance sampling is to change the measure from $\mathbb{P}$ to a new measure $\hat{\mathbb{P}}$ under which the event $\{N_t \geq xt\}$ becomes a typical event.

Let us define a new probability measure $\hat{\mathbb{P}}$ under which the N_t process is again a linear Hawkes process, but with baseline intensity γv and the exciting function $\gamma h(\cdot)$, where γ is a positive constant to be chosen later. Under the new measure, the N_t process has the intensity $\hat{\lambda}_t = \gamma \lambda_t$. We have the following result which will be proved in the Appendix.

Proposition 12.

$$\mathbb{P}(N_t \geq xt) = e^{-tI(x)} \hat{\mathbb{E}}\left[1_{\{N_t \geq xt\}} \cdot e^{((\gamma-1)\|h\|_{L^1} - \log \gamma)(N_t - xt) - (\gamma - 1)\int_0^t H(t-u)\,dN_u}\right], \tag{3.1}$$

where $H(t) := \int_t^\infty h(s)\,ds$, and $I(x)$ is given in (1.4).

Our numerical illustrations include three different methods: (1) importance sampling; (2) precise deviations up to the first-order approximation; (3) precise deviations up to the second-order approximation.

(1) *Importance sampling.* We simulate under $\hat{\mathbb{E}}$ a Hawkes process N_t with intensity:

$$\hat{\lambda}_t = \gamma v + \gamma \int_0^{t-} h(t-s)\,dN_s, \tag{3.2}$$

where $\gamma = \frac{x}{v + \|h\|_{L^1}x}$. Note that under the new measure $\hat{\mathbb{P}}$, we have $\hat{\mathbb{P}}(\frac{N_t}{t} \to \frac{\gamma v}{1 - \gamma \|h\|_{L^1}}) = 1$. By choosing γ as $\gamma = \frac{x}{v + \|h\|_{L^1}x}$, we have $\frac{\gamma v}{1 - \gamma \|h\|_{L^1}} = x$, and it follows that $\{N_t \geq xt\}$ is a typical event under the new measure $\hat{\mathbb{P}}$.

Using importance sampling Monte Carlo method (Proposition 12), we estimate

$$e^{-tI(x)} \hat{\mathbb{E}}\left[1_{\{N_t \geq xt\}} \cdot e^{((\gamma-1)\|h\|_{L^1} - \log \gamma)(N_t - xt) - (\gamma - 1)\int_0^t H(t-u)\,dN_u}\right], \tag{3.3}$$

where $H(t) := \int_t^\infty h(s)\,ds$ denotes the right tail of the exciting function, and $I(x) = x\log(\frac{x}{v + x\|h\|_{L^1}}) - x + x\|h\|_{L^1} + v$ as is given in (1.4).

(2) *Precise deviations up to the first-order approximation.* We approximate $\mathbb{P}(N_t \geq xt)$ by the formula:

$$e^{-tI(x)}\sqrt{\frac{I''(x)}{2\pi t}}\,c_0, \tag{3.4}$$

where $c_0 = \frac{1}{1 - e^{-\theta^*}}\psi(\theta^*)$, and $I''(x) = \frac{v^2}{x(v + \|h\|_{L^1}x)^2}$ is the second derivative of $I(x)$ defined in (1.4), and $\theta^* = \log(\frac{x}{v + \|h\|_{L^1}x}) - \frac{\|h\|_{L^1}x}{v + \|h\|_{L^1}x} + \|h\|_{L^1}$, where

$$\psi(\theta) = e^{v\int_0^\infty [F(s;\theta) - x(\theta)]\,ds}, \tag{3.5}$$

where F is the unique solution that satisfies

$$F(t;\theta) = e^{\theta + \int_0^t (F(t-s;\theta) - 1)h(s)\,ds}, \tag{3.6}$$

Table 1. Numerical illustration of $\mathbb{P}(N_t \geq xt)$ for $h(t) = e^{-2t}$, $\nu = 1$ and $x = 4, 5$

$x = 4$	IS	1st Order	2nd Order	$x = 5$	IS	1st Order	2nd Order
$t = 5$	4.71E−02	9.19E−02	−2.08E−02	$t = 5$	1.37E−02	2.68E−02	1.40E−03
$t = 10$	2.06E−02	3.06E−02	1.18E−02	$t = 10$	3.19E−03	4.59E−03	2.41E−03
$t = 25$	1.62E−03	2.02E−03	1.52E−03	$t = 25$	3.49E−05	4.14E−05	3.35E−05
$t = 40$	1.46E−04	1.66E−04	1.41E−04	$t = 40$	4.20E−07	4.66E−07	4.11E−07
$t = 50$	2.93E−05	3.29E−05	2.89E−05	$t = 50$	2.14E−08	2.45E−08	2.22E−08

with the constraint $|F(t; \theta)| \leq \frac{1}{\|h\|_{L^1}}$, and $x(\theta) = F(\infty; \theta)$ is the unique solution to:

$$x(\theta) = e^{\theta + (x(\theta) - 1)\|h\|_{L^1}}, \tag{3.7}$$

with the constraint $|x(\theta)| \leq \frac{1}{\|h\|_{L^1}}$. Note that we need to solve (3.6) and (3.7) numerically.

(3) *Precise deviations up to the second-order approximation.* We approximate $\mathbb{P}(N_t \geq xt)$ by adding a higher-order term to Eqn. (3.4) as follows:

$$e^{-tI(x)} \sqrt{\frac{I''(x)}{2\pi t}} \left(c_0 + \frac{c_1}{t} \right), \tag{3.8}$$

where $c_0 = \frac{\psi(\theta^*)}{1 - e^{-\theta^*}}$, and $c_1 = \frac{b_1}{1 - e^{-\theta^*}}$, where the formula for b_1 can be computed by applying Proposition 1 and Proposition 3. The details for the computations of b_1 can be found in the Appendix.

In our numerical illustrations, we take baseline intensity $\nu = 1$, and consider two different exciting functions: $h(t) = e^{-2t}$; $h(t) = \frac{1}{(1+t)^3}$. In both cases, $\|h\|_{L^1} = \frac{1}{2}$, and in the first case, $H(t) = \int_t^\infty h(s)\,ds = \frac{1}{2}e^{-2t}$, and in the second case, $H(t) = \int_t^\infty h(s)\,ds = \frac{1}{2(1+t)^2}$. We then compare the three different methods (1), (2) and (3) by comparing (3.3), (3.4) and (3.8). We summarize the results in Table 1 when the exciting function has exponential decay $h(t) = e^{-2t}$ and Table 2 when the exciting function has polynomial decay $h(t) = \frac{1}{(1+t)^3}$. In Table 1, we take $x = 4$ and $x = 5$, and consider the times $t = 5, 10, 25, 40, 50$. Numerically, we compute that $c_0 = 4.8$ and $c_1 = -22$. In Table 2, we take $x = 4$ and $x = 5$, and consider the times $t = 5, 10, 25, 40, 50$. Numerically, we compute that $c_0 = 3.51$ and $c_1 = -24$. In both tables, the column IS provides the numerical results using the importance sampling; the column 1st Order provides the numerical results using the precise deviations formula up to the first-order approximation; the column 2nd Order provides the numerical results using the precise deviations formula up to the second-order approximation. In both tables, we observe that as time t gets larger, the approximations get better. First-order approximation tend to overestimate the tail probability

Table 2. Numerical illustration of $\mathbb{P}(N_t \geq xt)$ for $h(t) = 1/(1+t)^3$, $\nu = 1$ and $x = 4, 5$

$x = 4$	IS	1st Order	2nd Order	$x = 5$	IS	1st Order	2nd Order
$t = 5$	3.04E−02	7.39E−02	−4.37E−02	$t = 5$	7.50E−03	1.94E−02	−7.60E−03
$t = 10$	1.33E−02	2.46E−02	5.02E−02	$t = 10$	1.63E−03	3.33E−03	1.01E−03
$t = 25$	1.12E−03	1.62E−03	1.11E−03	$t = 25$	1.95E−05	2.99E−05	2.16E−05
$t = 40$	1.03E−04	1.34E−04	1.07E−04	$t = 40$	2.39E−07	3.38E−07	2.79E−07
$t = 50$	2.28E−05	2.65E−05	2.23E−05	$t = 50$	1.28E−08	1.78E−08	1.53E−08

while second-order approximation tend to underestimate the tail probability in Table 1 while that is not the case in Table 2. This is due to the positivity of c_0 and negativity of c_1. As a result, when t is small, second-order approximation could give negative values which are unrealistic, for example, when $t = 5$ in Table 1 and Table 2. In both tables, as t becomes larger, second-order approximation provides better approximation than the first-order approximation.

Appendix: Proofs

A.1. Proofs of the results in Section 2.1

Proof of Proposition 4. Consider $\theta \in \mathbb{R}$, and $x(\theta) = e^{\theta + \|h\|_{L^1}(x(\theta) - 1)}$. Then $x(\theta)$ is increasing in θ by $x(\theta) = \mathbb{E}[e^{\theta S_\infty}]$ and the definition of S_∞. Moreover, for $\theta = \theta_c = \|h\|_{L^1} - 1 - \log\|h\|_{L^1}$, we have $x(\theta_c) = \frac{1}{\|h\|_{L^1}} e^{\|h\|_{L^1} - 1 + \|h\|_{L^1}(x(\theta_c) - 1)}$, which implies that $x(\theta_c)\|h\|_{L^1} = 1$. Thus, for any $\theta \leq \theta_c$, $x(\theta)\|h\|_{L^1} \leq 1$, which proves (i). We also notice that $x'(\theta) = (1 + \|h\|_{L^1}x'(\theta))x(\theta)$, and therefore as $\theta \uparrow \theta_c$, $x'(\theta) = \frac{x(\theta)}{1 - \|h\|_{L^1}x(\theta)} \to \infty$, that proves (ii), which also follows from Theorem 3.2.1. in [8]. $\square$

Proof of Proposition 5. Suppose F_1 and F_2 are two solutions that satisfy (2.11) with $|F_j| \leq \frac{1}{\|h\|_{L^1}}$, for $j = 1, 2$. Note that for any $z_1, z_2 \in \mathbb{C}$,

$$
\begin{aligned}
\left|e^{z_1} - e^{z_2}\right| &= \left|\sum_{n=1}^{\infty} \frac{z_1^n - z_2^n}{n!}\right| \\
&\leq |z_1 - z_2| \sum_{n=1}^{\infty} \frac{|z_1|^{n-1} + |z_1|^{n-2}|z_2| + \cdots + |z_1||z_2|^{n-2} + |z_2|^{n-1}}{n!} \\
&\leq |z_1 - z_2| \sum_{n=1}^{\infty} \frac{n(|z_1| + |z_2|)^{n-1}}{n!} = |z_1 - z_2|e^{|z_1| + |z_2|}.
\end{aligned}
$$

Hence, for any $T > 0$ and for any $0 \leq t \leq T$,

$$
\begin{aligned}
\left|F_1(t; z) - F_2(t; z)\right| &= \left|e^z\right|\left|e^{\int_0^t (F_1(t-s;z)-1)h(s)\,ds} - e^{\int_0^t (F_2(t-s;z)-1)h(s)\,ds}\right| \\
&\leq \left|e^z\right|\left|\int_0^t \left(F_1(t-s; z) - F_2(t-s; z)\right)h(s)\,ds\right| \\
&\quad \cdot e^{\left|\int_0^t (F_1(t-s;z)-1)h(s)\,ds\right| + \left|\int_0^t (F_2(t-s;z)-1)h(s)\,ds\right|} \\
&\leq \left|e^z\right|\|h\|_{L^\infty[0,T]} e^{2(\frac{1}{\|h\|_{L^1}}+1)\|h\|_{L^1}} \int_0^t \left|F_1(s; z) - F_2(s; z)\right|\,ds.
\end{aligned}
$$

Note that $F_1(0; z) = F_2(0; z) = e^z$ from (2.11). By Gronwall's inequality, we conclude that $F_1 \equiv F_2$. $\square$

Proof of Lemma 6. Note that $Y = \sum_{i=1}^{K} Z_i$ in distribution, where K is Poisson random variable with parameter ν and Z_i are i.i.d. random variable interpreted as the total number of the nodes in a Galton–Watson tree with the number of children being born in each generation Poisson distributed with parameter $\|h\|_{L^1}$, independent of K. A compound Poisson random variable is infinitely divisible,

which is evident from the expression of its characteristic function. Hence, Y has an infinitely divisible distribution. $\qquad\square$

Proof of Lemma 7. Firstly, it is obvious that for any $s > 0$, and $t > 0$, $F(s, z)$, $x(z)$, and $\int_0^t (F(s, z) - x(z))\,ds$ are analytic in $\mathcal{S}_{(-\infty, \theta_c)}$. Since $\|h\|_{L^1} < 1$ and $x(0) = 1$, for $\mathcal{R}(z) < \|h\|_{L^1} - 1 - \log\|h\|_{L^1}$, we have $|x(z)|\|h\|_{L^1} \leq x(\mathcal{R}(z))\|h\|_{L^1} < 1$. Therefore, for any compact subset $K \subset \{z \in \mathbb{C}; \mathcal{R}(z) < \|h\|_{L^1} - 1 - \log\|h\|_{L^1}\}$, we have

$$\sup_{z \in K}|x(z)|\|h\|_{L^1} \leq x(\mathcal{R}(z))\|h\|_{L^1} < 1.$$

Since $\int_s^\infty h(u)\,du \to 0$ as $s \to \infty$ and $F(s; z) = \mathbb{E}[e^{z S_s}] \to x(z) = \mathbb{E}[e^{z S_\infty}]$, we have that

$$\int_0^s \sup_{z \in K}[F(s - u; z) - x(z)]h(u)\,du \to 0, \quad \text{as } s \to \infty.$$

Note that

$$F(s; z) - x(z) = x(z)\big[e^{\int_0^s [F(s-u;z) - x(z)]h(u)\,du - (x(z)-1)\int_s^\infty h(u)\,du} - 1\big],$$

and for any fixed $\delta > 0$ such that $(1 + \delta)\sup_{z \in K}|x(z)|\|h\|_{L^1} < 1$, there exists $M > 0$, so that for any $s \geq M$ and $z \in K$, we have

$$\big|F(s; z) - x(z)\big| \leq (1 + \delta)\big|x(z)\big|\bigg[\int_0^s \big|F(s - u; z) - x(z)\big|h(u)\,du + \big|x(z) - 1\big|\int_s^\infty h(u)\,du\bigg].$$

Therefore, we get that for any $T > M$,

$$\int_M^T \sup_{z \in K}\big|F(s; z) - x(z)\big|\,ds \leq (1 + \delta)\sup_{z \in K}\big|x(z)\big|\int_M^T \int_0^s \sup_{z \in K}\big|F(s - u; z) - x(z)\big|h(u)\,du\,ds$$

$$+ (1 + \delta)\sup_{z \in K}\big|x(z)\big|\big|x(z) - 1\big|\int_M^T \int_s^\infty h(u)\,du\,ds$$

$$\leq (1 + \delta)\sup_{z \in K}\big|x(z)\big|\int_0^T \int_0^s \sup_{z \in K}\big|F(s - u; z) - x(z)\big|h(u)\,du\,ds$$

$$+ (1 + \delta)\sup_{z \in K}\big|x(z)\big|\big|x(z) - 1\big|\int_0^\infty \int_s^\infty h(u)\,du\,ds,$$

which implies that

$$\int_0^T \sup_{z \in K}\big|F(s; z) - x(z)\big|\,ds$$

$$\leq \int_0^M \sup_{z \in K}\big|F(s; z) - x(z)\big|\,ds + (1 + \delta)\sup_{z \in K}\big|x(z)\big|\int_0^T \int_0^s \sup_{z \in K}\big|F(s - u; z) - x(z)\big|h(u)\,du\,ds$$

$$+ (1 + \delta)\sup_{z \in K}\big|x(z)\big|\big|x(z) - 1\big|\int_0^\infty \int_s^\infty h(u)\,du\,ds$$

$$= \int_0^M \sup_{z \in K}\big|F(s; z) - x(z)\big|\,ds + (1 + \delta)\sup_{z \in K}\big|x(z)\big|\int_0^T \bigg[\int_u^T \sup_{z \in K}\big|F(s - u; z) - x(z)\big|\,ds\bigg]h(u)\,du$$

$$+ (1 + \delta) \sup_{z \in K} |x(z)| |x(z) - 1| \int_0^\infty \int_s^\infty h(u) \, du \, ds$$

$$\leq \int_0^M \sup_{z \in K} |F(s; z) - x(z)| \, ds + (1 + \delta) \sup_{z \in K} |x(z)| \|h\|_{L^1} \int_0^T |F(s; z) - x(z)| \, ds$$

$$+ (1 + \delta) \sup_{z \in K} |x(z)| \left[\sup_{z \in K} |x(z)| + 1 \right] \int_0^\infty s h(s) \, ds,$$

which holds for any $T > M$, and thus we have

$$\int_0^\infty \sup_{z \in K} |F(s; z) - x(z)| \, ds$$

$$\leq \frac{\int_0^M \sup_{z \in K} |F(s; z) - x(z)| \, ds + (1 + \delta) \sup_{z \in K} |x(z)| [\sup_{z \in K} |x(z)| + 1] \int_0^\infty s h(s) \, ds}{1 - (1 + \delta) \sup_{z \in K} |x(z)| \|h\|_{L^1}}. \quad (A.9)$$

Hence, we conclude that $\int_t^\infty \sup_{z \in K} |F(s; z) - x(z)| \, ds \to 0$ as $t \to \infty$, and so

$$\int_0^t \big(F(s, z) - x(z) \big) \, ds \to \int_0^\infty \big(F(s, z) - x(z) \big) \, ds, \quad (A.10)$$

as $t \to \infty$, locally uniformly in z for $\mathcal{R}(z) < \theta_c$. Hence, $\varphi(z)$ is well-defined and is analytic in $\mathcal{S}_{(-\infty, \theta_c)}$. By equations (2.17), (A.10) and the definitions of $\psi(z)$ and $\varphi(z)$ in (2.10) we have proved that locally uniformly in z for $\mathcal{R}(z) < \theta_c$,

$$e^{-t(\nu(x(z)-1))} \mathbb{E}\big[e^{zN_t} \big] = e^{\nu \int_0^t (F(s,z) - x(z)) \, ds} \to e^{\nu \varphi(z)} = \psi(z), \quad \text{as } t \to \infty.$$

To show the mod-ϕ convergence at speed $O(t^{-\nu})$, that is, for any compact set K, there exists some $C_K > 0$ such that $\sup_{z \in K} |e^{-t(\nu(x(z)-1))} \mathbb{E}[e^{zN_t}] - e^{\nu \varphi(z)}| \leq C_K t^{-\nu}$, it suffices to show that for any compact set $K \subset \{z \in \mathbb{C}; \mathcal{R}(z) < \|h\|_{L^1} - 1 - \log \|h\|_{L^1}\}$,

$$\int_0^\infty t^\nu \sup_{z \in K} |F(t; z) - x(z)| \, dt < \infty. \quad (A.11)$$

To see this, notice that

$$\sup_{z \in K} |e^{-t(\nu(x(z)-1))} \mathbb{E}\big[e^{zN_t} \big] - e^{\nu \varphi(z)}| = \sup_{z \in K} |e^{\nu \varphi(z)} \big(e^{\nu \int_t^\infty (F(s;z) - x(z)) \, ds} - 1 \big)|$$

$$\leq \sup_{z \in K} e^{\nu |\varphi(z)|} \big(e^{\nu \int_t^\infty \sup_{z \in K} |F(s;z) - x(z)| \, ds} - 1 \big).$$

Thus it suffices to show that

$$\int_t^\infty \sup_{z \in K} |F(s; z) - x(z)| \, ds \leq c_K t^{-\nu}, \quad (A.12)$$

for some $c_K > 0$. For any $T > 0$, by integration by parts,

$$\int_0^T t^v \sup_{z \in K} \left| F(t; z) - x(z) \right| dt$$

$$= -\int_0^T t^v d \left(\int_t^\infty \sup_{z \in K} \left| F(s; z) - x(z) \right| ds \right)$$

$$= -T^v \int_T^\infty \sup_{z \in K} \left| F(s; z) - x(z) \right| ds + v \int_0^T t^{v-1} \int_t^\infty \sup_{z \in K} \left| F(s; z) - x(z) \right| ds \, dt. \quad \text{(A.13)}$$

On the other hand, by Fubini's theorem and (A.11), we have

$$v \int_0^\infty t^{v-1} \int_t^\infty \sup_{z \in K} \left| F(s; z) - x(z) \right| ds \, dt = v \int_0^\infty \int_0^s t^{v-1} \sup_{z \in K} \left| F(s; z) - x(z) \right| dt \, ds$$

$$= \int_0^\infty s^v \sup_{z \in K} \left| F(s; z) - x(z) \right| dt \, ds < \infty. \quad \text{(A.14)}$$

Hence, we get $\lim_{T \to \infty} T^v \int_T^\infty \sup_{z \in K} |F(s; z) - x(z)| \, ds = 0$ by letting $T \to \infty$ in (A.13) and applying (A.14). This implies (A.12).

Next, let us prove (A.11). Notice that there exists $M > 0$ so that for any $s \geq M$,

$$\sup_{z \in \mathbb{Z}} \left| F(s; z) - x(z) \right| \leq (1 + \delta) \sup_{z \in K} \left| x(z) \right| \int_0^s \sup_{z \in K} \left| F(s - u; z) - x(z) \right| h(u) \, du$$

$$+ (1 + \delta) \sup_{z \in K} \left| x(z) \right| \sup_{z \in K} \left| x(z) - 1 \right| \int_s^\infty h(u) \, du.$$

Therefore, for every $s \geq 0$,

$$\sup_{z \in K} \left| F(s; z) - x(z) \right| \leq (1 + \delta) \sup_{z \in K} \left| x(z) \right| \int_0^s \sup_{z \in K} \left| F(s - u; z) - x(z) \right| h(u) \, du + g(s), \quad \text{(A.15)}$$

where

$$g(s) := C_1 \int_s^\infty h(u) \, du + C_2 1_{\{s \leq M\}}, \quad \text{(A.16)}$$

where

$$C_1 := (1 + \delta) \sup_{z \in K} \left| x(z) \right| \sup_{z \in K} \left| x(z) - 1 \right|, \qquad C_2 := \sup_{0 \leq s \leq M} \sup_{z \in K} \left| F(s; z) - x(z) \right|. \quad \text{(A.17)}$$

Let $p(s) := \sup_{z \in K} |F(s; z) - x(z)|$ and $q(s) := (1 + \delta) \sup_{z \in K} |x(z)| h(s)$ for every $s \geq 0$. Then, (A.15) can be re-written as $p(s) \leq \int_0^s q(s - u) p(u) \, du + g(s)$, $s \geq 0$, which implies

$$p(s) \leq \int_0^s q(s - u) \left(\int_0^u q(u - v) p(v) \, dv + g(u) \right) du + g(s)$$

$$= \int_0^s q^{*2}(s - u) p(u) \, du + \int_0^s q^{*1}(s - u) g(u) \, du + g(s), \quad s \geq 0.$$

By iterating and following the same argument as in the proof of generalized Gronwall's inequality in [11], we have

$$p(s) \leq g(s) + \int_0^s \sum_{k=1}^{\infty} q^{*k}(s-u)g(u)\,du,$$

which is equivalent to

$$\sup_{z \in K}\left|F(s; z) - x(z)\right| \leq g(s) + \int_0^s G(s-u)g(u)\,du, \tag{A.18}$$

where

$$G(t) := \sum_{k=1}^{\infty}\left((1+\delta)\sup_{z \in K}\left|x(z)\right|\right)^k h^{*k}(t), \tag{A.19}$$

where $h^{*k}(t)$ is the k-th order convolution of h, that is, $h^{*0}(t) := h(t)$ and $h^{*m}(t) := \int_0^t h^{*(m-1)}(t-s)h(s)\,ds$ for every $m \in \mathbb{N}$. Hence, it remains to show that

$$\int_0^{\infty} s^v g(s)\,ds < \infty, \tag{A.20}$$

$$\int_0^{\infty} s^v \int_0^s G(s-u)g(u)\,du\,ds < \infty. \tag{A.21}$$

Let us first prove (A.20). Note that

$$\int_0^{\infty} s^v g(s)\,ds = C_1 \int_0^{\infty} s^v \int_s^{\infty} h(u)\,du\,ds + C_2 \int_0^M s^v\,dv, \tag{A.22}$$

and by our assumption,

$$\int_0^{\infty} s^v \int_s^{\infty} h(u)\,du\,ds = \int_0^{\infty}\int_0^u s^v h(u)\,ds\,du = \frac{1}{v+1}\int_0^{\infty} u^{v+1}h(u)\,du < \infty, \tag{A.23}$$

thus (A.20) follows.

Next, let us prove (A.21). Note that

$$\int_0^{\infty} s^v \int_0^s G(s-u)g(u)\,du\,ds$$

$$= C_1 \int_0^{\infty} s^v \int_0^s G(s-u)\int_u^{\infty} h(x)\,dx\,du\,ds + C_2 \int_0^{\infty} s^v \int_0^s G(s-u)1_{\{u \leq M\}}\,du\,ds,$$

and it is easy to check that

$$\int_0^{\infty} s^v \int_0^s G(s-u)1_{\{u \leq M\}}\,du\,ds = \int_0^{\infty}\int_u^{\infty} s^v G(s-u)\,ds\,1_{\{u \leq M\}}\,du$$

$$= \int_0^M \int_0^{\infty}(s+u)^v G(s)\,ds\,du$$

$$\leq 2^{v-1}M\int_0^{\infty}\left(s^v + M^v\right)G(s)\,ds,$$

where we used the inequality $(a + b)^v \leq 2^{v-1}(a^v + b^v)$ for any $a, b > 0$ and $v \geq 1$. We can compute that

$$\int_0^\infty G(s)\,ds = \sum_{k=1}^\infty \left((1 + \delta) \sup_{z \in K} |x(z)|\right)^k \int_0^\infty h^{*k}(s)\,ds = \sum_{k=1}^\infty \left((1 + \delta) \sup_{z \in K} |x(z)|\right)^k \|h\|_{L^1}^k,$$

which is finite since $(1 + \delta) \sup_{z \in K} |x(z)| \|h\|_{L^1} < 1$. Next, let us show that

$$\int_0^\infty s^v G(s)\,ds < \infty. \tag{A.24}$$

Notice that

$$\int_0^\infty s^v G(s)\,ds = \sum_{k=1}^\infty \left((1 + \delta) \sup_{z \in K} |x(z)|\right)^k \int_0^\infty s^v h^{*k}(s)\,ds$$

$$= \sum_{k=1}^\infty \left((1 + \delta) \sup_{z \in K} |x(z)|\right)^k \int_0^\infty s^v \int_0^s h(s - u)h^{*(k-1)}(u)\,du\,ds$$

$$= \sum_{k=1}^\infty \left((1 + \delta) \sup_{z \in K} |x(z)|\right)^k \int_0^\infty \int_u^\infty s^v h(s - u)ds\,h^{*(k-1)}(u)\,du$$

$$= \sum_{k=1}^\infty \left((1 + \delta) \sup_{z \in K} |x(z)|\right)^k \int_0^\infty \int_0^\infty (s + u)^v h(s)ds\,h^{*(k-1)}(u)\,du.$$

Note that for any $\delta' > 0$, there exists some $C(\delta') > 0$ such that for any $s, u \geq 0$,

$$(s + u)^v \leq C(\delta')s^v + (1 + \delta')u^v. \tag{A.25}$$

Therefore,

$$\int_0^\infty s^v h^{*k}(s)\,ds$$

$$\leq C(\delta') \int_0^\infty s^v h(s)\,ds \int_0^\infty h^{*(k-1)}(u)\,du + (1 + \delta') \int_0^\infty h(s)\,ds \int_0^\infty u^v h^{*(k-1)}(u)\,du$$

$$= C(\delta') \int_0^\infty s^v h(s)\,ds \|h\|_{L^1}^{k-1} + (1 + \delta')\|h\|_{L^1} \int_0^\infty u^v h^{*(k-1)}(u)\,du.$$

Let us define $A_k := \int_0^\infty s^v h^{*k}(s)\,ds$. Then, we have:

$$A_k \leq C(\delta')\|h\|_{L^1}^{k-1} A_1 + (1 + \delta')\|h\|_{L^1} A_{k-1}. \tag{A.26}$$

It follows that

$$A_k \leq C(\delta')\|h\|_{L^1}^{k-1}\left[1 + (1 + \delta')\|h\|_{L^1} + \left((1 + \delta')\|h\|_{L^1}\right)^2 + \cdots + \left((1 + \delta')\|h\|_{L^1}\right)^{k-2}\right]A_1$$
$$+ \left((1 + \delta')\|h\|_{L^1}\right)^{k-1} A_1.$$

Choose $\delta' > 0$ to be sufficiently small so that $(1 + \delta')\|h\|_{L^1} < 1$. Then, we have

$$A_k \le \frac{C(\delta')}{1 - (1 + \delta')\|h\|_{L^1}} \|h\|_{L^1}^{k-1} A_1 + (1 + \delta')^{k-1} \|h\|_{L^1}^{k-1} A_1$$

$$\le \left(\frac{C(\delta')}{1 - (1 + \delta')\|h\|_{L^1}} + 1 \right)(1 + \delta')^{k-1} \|h\|_{L^1}^{k-1} A_1.$$

Choose $\delta' > 0$ to be sufficiently small so that $(1 + \delta)(1 + \delta') \sup_{z \in K} |x(z)| \|h\|_{L^1} < 1$. Hence, we conclude that

$$\int_0^\infty s^v G(s)\,ds$$

$$= \sum_{k=1}^\infty \left((1 + \delta) \sup_{z \in K} |x(z)| \right)^k A_k$$

$$\le (1 + \delta) \sup_{z \in K} |x(z)| A_1 \left(\frac{C(\delta')}{1 - (1 + \delta')\|h\|_{L^1}} + 1 \right) \sum_{k=1}^\infty \left((1 + \delta) \sup_{z \in K} |x(z)| (1 + \delta')\|h\|_{L^1} \right)^{k-1}$$

$$= \frac{(1 + \delta) \sup_{z \in K} |x(z)| A_1 \left(\frac{C(\delta')}{1 - (1 + \delta')\|h\|_{L^1}} + 1 \right)}{1 - (1 + \delta)(1 + \delta') \sup_{z \in K} |x(z)| \|h\|_{L^1}} < \infty.$$

Finally, we can compute that

$$\int_0^\infty s^v \int_0^s G(s - u) \int_u^\infty h(x)\,dx\,du\,ds$$

$$= \int_0^\infty \int_u^\infty s^v G(s - u)\,ds \int_u^\infty h(x)\,dx\,du$$

$$= \int_0^\infty \int_0^\infty (s + u)^v G(s)\,ds \int_u^\infty h(x)\,dx\,du$$

$$\le 2^{v-1} \int_0^\infty \int_0^\infty (s^v + u^v) G(s)\,ds \int_u^\infty h(x)\,dx\,du$$

$$= 2^{v-1} \int_0^\infty s^v G(s)\,ds \int_0^\infty s h(s)\,ds + 2^{v-1} \int_0^\infty G(s)\,ds \cdot \frac{1}{v + 1} \int_0^\infty s^{v+1} h(s)\,ds < \infty.$$

This completes the proof.   □

Proof of Theorem 2. By Lemma 6, and Lemma 7, we have established the mod-ϕ convergence. Hence, by Theorem 3.2.2. in [15], for any $x > 0$, and $tx \in \mathbb{N}$,

$$\mathbb{P}(N_t = tx) = \frac{e^{-tI(x)}}{\sqrt{2\pi t \eta''(\theta^*)}} \left(\psi(\theta^*) + \frac{a_1}{t} + \frac{a_2}{t^2} + \cdots + \frac{a_{v-1}}{t^{v-1}} + O\left(\frac{1}{t^v}\right) \right), \qquad (A.27)$$

and for any $x > \eta'(0) = \frac{v}{1 - \|h\|_{L^1}}$ and $tx \in \mathbb{N}$,

$$\mathbb{P}(N_t \ge tx) = \frac{e^{-tI(x)}}{\sqrt{2\pi t \eta''(\theta^*)}} \frac{1}{1 - e^{-\theta^*}} \left(\psi(\theta^*) + \frac{b_1}{t} + \frac{b_2}{t^2} + \cdots + \frac{b_{v-1}}{t^{v-1}} + O\left(\frac{1}{t^v}\right) \right), \qquad (A.28)$$

where $I(x)$ is defined in (1.4) and $\eta'(\theta^*) = x$. Note that $\eta(\theta) = \nu(x(\theta) - 1)$, where $x(\theta) = e^{\theta + \|h\|_{L^1}(x(\theta)-1)}$. Thus, $\eta'(\theta) = \nu x'(\theta)$, and

$$x'(\theta) = \left(1 + \|h\|_{L^1} x'(\theta)\right)e^{\theta + \|h\|_{L^1}(x(\theta)-1)} = \left(1 + \|h\|_{L^1} x'(\theta)\right)x(\theta), \tag{A.29}$$

which implies that

$$x(\theta) = \frac{x'(\theta)}{1 + \|h\|_{L^1} x'(\theta)}. \tag{A.30}$$

Notice that $\eta'(\theta^*) = x$, and thus $x'(\theta^*) = \frac{\eta'(\theta^*)}{\nu} = \frac{x}{\nu}$, and $x(\theta^*) = \frac{x'(\theta^*)}{1+\|h\|_{L^1}x'(\theta^*)} \frac{x}{\nu+\|h\|_{L^1}x}$. Therefore,

$$\frac{x}{\nu} = x'\left(\theta^*\right) = \left(1 + \|h\|_{L^1} x'\left(\theta^*\right)\right)e^{\theta^* + \|h\|_{L^1}(x(\theta^*)-1)} = \left(1 + \frac{\|h\|_{L^1}x}{\nu}\right)e^{\theta^* + \frac{\|h\|_{L^1}x}{\nu+\|h\|_{L^1}x} - \|h\|_{L^1}},$$

which implies that $\theta^* = \log(\frac{x}{\nu+\|h\|_{L^1}x}) - \frac{\|h\|_{L^1}x}{\nu+\|h\|_{L^1}x} + \|h\|_{L^1}$. Hence, we have verified that

$$I(x) = \theta^* x - \eta\left(\theta^*\right) = x \log\left(\frac{x}{\nu + x\|h\|_{L^1}}\right) - x + x\|h\|_{L^1} + \nu. \tag{A.31}$$

Moreover, by the property of Legendre transform, $\eta''(\theta^*) = \frac{1}{I''(x)}$, which together with

$$I'(x) = \theta^* = \log\left(\frac{x}{\nu + \|h\|_{L^1}x}\right) - \frac{\|h\|_{L^1}x}{\nu + \|h\|_{L^1}x} + \|h\|_{L^1}, \tag{A.32}$$

implies that

$$I''(x) = \frac{\nu^2}{x(\nu + \|h\|_{L^1}x)^2}. \tag{A.33}$$

$\square$

A.2. Proofs of the results in Section 2.2

Proof of Theorem 8. Since we have established mod-ϕ convergence, and N_t is lattice distributed, the result follows from Corollary 3.3.5. in [15], which is restated in our (2.5).

Next, let us show that (2.26) holds. Recall from (A.33) that $I''(x) = \frac{\nu^2}{x(\nu+\|h\|_{L^1}x)^2}$. Therefore, for any $i \geq 2$, by Leibniz formula,

$$\begin{aligned}
I^{(i)}(x) &= \nu^2 \frac{d^{i-2}}{dx^{i-2}} x^{-1} \cdot \left(\nu + \|h\|_{L^1}x\right)^{-2} \\
&= \nu^2 \sum_{j=0}^{i-2} \frac{(i-2)!}{(i-2-j)!j!}(-1)^{i-2}(i-2-j)!\|h\|_{L^1}^j(j+1)!x^{-1-(i-2-j)}\left(\nu + \|h\|_{L^1}x\right)^{-j-2} \\
&= \nu^2(i-2)!(-1)^{i-2}\frac{x^{1-i}}{(\nu + \|h\|_{L^1}x)^2}\sum_{j=0}^{i-2}(j+1)\left(\frac{\|h\|_{L^1}x}{\nu + \|h\|_{L^1}x}\right)^j
\end{aligned}$$

$$= v^2(i-2)!(-1)^{i-2}\frac{x^{1-i}}{(v+\|h\|_{L^1}x)^2}\frac{d}{dy}\sum_{j=0}^{i-2}y^{j+1}\bigg|_{y=\frac{\|h\|_{L^1}x}{v+\|h\|_{L^1}x}}$$

$$= v^2(i-2)!(-1)^{i-2}\frac{x^{1-i}}{(v+\|h\|_{L^1}x)^2}\frac{(i-1)y^i - iy^{i-1}+1}{(y-1)^2}\bigg|_{y=\frac{\|h\|_{L^1}x}{v+\|h\|_{L^1}x}},$$

which implies (2.26).

Finally, let us compute $\eta'(0)$, $\eta''(0)$ for Theorem 8 as well as $\eta'''(0)$ for Remark 11. Let us recall that $\eta(\theta) = v(x(\theta)-1)$, where $x(\theta) = e^{\theta + \|h\|_{L^1}(x(\theta)-1)}$. Thus, we can compute $x'(\theta) = (1 + \|h\|_{L^1}x'(\theta))x(\theta)$ and $x''(\theta) = [\|h\|_{L^1}x''(\theta) + (1 + \|h\|_{L^1}x'(\theta))^2]x(\theta)$ and

$$x'''(\theta) = \big[\|h\|_{L^1}x'''(\theta) + 2\big(1 + \|h\|_{L^1}x'(\theta)\big)\|h\|_{L^1}x''(\theta)\big]x(\theta)$$

$$+ \big[\|h\|_{L^1}x''(\theta) + \big(1 + \|h\|_{L^1}x'(\theta)\big)^2\big]\big(1 + \|h\|_{L^1}x'(\theta)\big)x(\theta).$$

Note that $x(0) = 1$. By letting $\theta = 0$, we get the formulas for $\eta'(0)$, $\eta''(0)$ and $\eta'''(0)$. $\qquad\square$

A.3. Proof of Proposition 1(i): Derivations of $(a_k)_{k=1}^\infty$

Before we proceed, let us first introduce the Faà di Bruno's formula that will be used repeatedly in our proofs.

Lemma 13 (Faà di Bruno's formula).

$$\frac{d^n}{dx^n}f\big(g(x)\big) = \sum_{\mathcal{S}_n}\frac{n!\,f^{(m_1+\cdots+m_n)}(g(x))}{m_1!1!^{m_1}m_2!2!^{m_2}\cdots m_n!n!^{m_n}}\cdot\prod_{j=1}^n\big(g^{(j)}(x)\big)^{m_j}, \tag{A.34}$$

where the sum is over the set $\mathcal{S}_n$ consisting of all the n-tuples of non-negative integers $(m_1,\ldots,m_n)$ satisfying the constraint $1\cdot m_1 + 2\cdot m_2 + 3\cdot m_3 + \cdots + n\cdot m_n = n$.

We recall from the proof of Theorem 3.2.2. and Remark 3.2.5. in [15] that for $\mathbb{Z}$-valued random variables X_n,

$$\mathbb{P}(X_n = t_n x) = \frac{e^{-t_n F(x)}}{\sqrt{2\pi t_n\eta''(\theta^*)}}\left(\psi(\theta^*) + \frac{a_1}{t_n} + \cdots + \frac{a_{v-1}}{t_n^{v-1}} + O\big(t_n^{-v}\big)\right), \tag{A.35}$$

as $n\to\infty$, where $a_k = \int_{-\infty}^\infty \alpha_{2k}(w)\frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\,dw$, where α_k's are defined via the expansion:

$$\sum_{k=0}^{2v-1}\frac{\alpha_k}{t_n^{k/2}} + O\big(t_n^{-v}\big) = \sum_{k=0}^{2v-1}\frac{\psi^{(k)}(\theta^*)}{k!}\left(\frac{iw}{\sqrt{t_n\eta''(\theta^*)}}\right)^k e^{-\frac{w^2}{\eta''(\theta^*)}\sum_{k=1}^{2v-1}\frac{\eta^{(k+2)}(\theta^*)}{(k+2)!}(\frac{iw}{\sqrt{t_n\eta''(\theta^*)}})^k}. \tag{A.36}$$

Let us define $f(x) = e^{-x}$ and

$$g(x) = \frac{w^2}{\eta''(\theta^*)}\sum_{k=1}^\infty\frac{\eta^{(k+2)}(\theta^*)}{(k+2)!}x^k. \tag{A.37}$$

Then we can compute that $g(0) = 0$ and for every $j \in \mathbb{N}$,

$$g^{(j)}(0) = \frac{w^2}{\eta''(\theta^*)} \frac{\eta^{(j+2)}(\theta^*)}{(j+2)(j+1)}. \tag{A.38}$$

By Faà di Bruno's formula, it follows that

$$f\big(g(x)\big) = \sum_{n=0}^{\infty} \sum \frac{n!(-1)^{m_1+\cdots+m_n}}{m_1!1!^{m_1} m_2!2!^{m_2} \cdots m_n!n!^{m_n}} \cdot \prod_{j=1}^{n} \left(\frac{w^2}{\eta''(\theta^*)} \frac{\eta^{(j+2)}(\theta^*)}{(j+2)(j+1)} \right)^{m_j} \frac{x^n}{n!}. \tag{A.39}$$

It follows that

$$\frac{\alpha_k}{t_n^{k/2}} = \sum_{\ell=0}^{k} \frac{\psi^{(k-\ell)}(\theta^*)}{(k-\ell)!} \sum_{\mathcal{S}_\ell} \frac{(-1)^{m_1+\cdots+m_\ell}}{m_1!1!^{m_1} m_2!2!^{m_2} \cdots m_\ell!\ell!^{m_\ell}}$$

$$\cdot \prod_{j=1}^{\ell} \left(\frac{w^2}{\eta''(\theta^*)} \frac{\eta^{(j+2)}(\theta^*)}{(j+2)(j+1)} \right)^{m_j} \frac{(iw)^k}{(\eta''(\theta^*))^{k/2}},$$

which implies that

$$\alpha_k(w) = \sum_{\ell=0}^{k} \frac{\psi^{(k-\ell)}(\theta^*)}{(k-\ell)!} \sum_{\mathcal{S}_\ell} \frac{(-1)^{m_1+\cdots+m_\ell}}{m_1!1!^{m_1} m_2!2!^{m_2} \cdots m_\ell!\ell!^{m_\ell}}$$

$$\cdot \prod_{j=1}^{\ell} \left(\frac{1}{\eta''(\theta^*)} \frac{\eta^{(j+2)}(\theta^*)}{(j+2)(j+1)} \right)^{m_j} \frac{(i)^k w^{k+2(m_1+\cdots+m_\ell)}}{(\eta''(\theta^*))^{k/2}}.$$

By the property of standard normal random variable,

$$\int_{-\infty}^{\infty} w^m \cdot \frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}} \, dw = (m-1)!!, \tag{A.40}$$

if m is even and 0 if m is odd. Hence, the formula for a_k follows.

A.4. Proof of Proposition 1(ii): Derivations of $(b_k)_{k=1}^{\infty}$

We recall from the proof of Theorem 3.2.2. and Remark 3.2.5. in [15] that for integer-valued random variables X_n,

$$\mathbb{P}(X_n \geq t_n x) = \frac{e^{-t_n F(x)}}{\sqrt{2\pi t_n \eta''(\theta^*)}} \frac{1}{1-e^{-\theta^*}} \left(\psi\big(\theta^*\big) + \frac{b_1}{t_n} + \frac{b_2}{t_n^2} + \cdots + \frac{b_{v-1}}{t_n^{v-1}} + O\left(\frac{1}{t_n^v}\right) \right), \tag{A.41}$$

where $b_k = \int_{\mathbb{R}} \beta_{2k}(w) \frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}} \, dw$, where $\beta_j(w)$ are given via the expansion:

$$g_n(w) = \sum_{k=0}^{2v-1} \frac{\beta_k(w)}{(t_n)^{k/2}} + O\big(t_n^{-v}\big), \tag{A.42}$$

where

$$g_n(w) := \sum_{k=0}^{\infty} e^{-k\theta^*} e^{-k\frac{iw}{\sqrt{t_n \eta''(\theta^*)}}} \cdot \psi\left(\theta^* + \frac{iw}{\sqrt{t_n \eta''(\theta^*)}}\right)$$

$$\cdot e^{t_n(\eta(\theta^* + \frac{iw}{\sqrt{t_n \eta''(\theta^*)}}) - \eta(\theta^*) - \eta'(\theta^*)\frac{iw}{\sqrt{t_n \eta''(\theta^*)}} + \frac{w^2}{2t_n})}$$

$$= \frac{\psi\left(\theta^* + \frac{iw}{\sqrt{t_n \eta''(\theta^*)}}\right)}{1 - e^{-\theta^* - \frac{iw}{\sqrt{t_n \eta''(\theta^*)}}}} \cdot e^{t_n(\eta(\theta^* + \frac{iw}{\sqrt{t_n \eta''(\theta^*)}}) - \eta(\theta^*) - \eta'(\theta^*)\frac{iw}{\sqrt{t_n \eta''(\theta^*)}} + \frac{w^2}{2t_n})}.$$

Note that

$$\sum_{k=0}^{2\nu-1} \frac{\beta_k(w)}{(t_n)^{k/2}} + O(t_n^{-\nu}) = \frac{1}{1 - e^{-\theta^* - \frac{iw}{\sqrt{t_n \eta''(\theta^*)}}}} \sum_{k=0}^{2\nu-1} \frac{\alpha_k(w)}{(t_n)^{k/2}}. \tag{A.43}$$

We define $f(x) = \frac{1}{1-e^{-\theta^*}x}$, and $g(x) = e^{-x}$. Then,

$$f(g(x)) = \sum_{n=0}^{\infty} \sum_{n=0}^{\infty} \frac{d^n}{dx^n} f(g(0)) \frac{x^n}{n!}, \tag{A.44}$$

where we can compute that $f^{(k)}(x) = e^{-k\theta^*} k! (1 - e^{-\theta^*}x)^{-k-1}$, and $g^{(k)}(x) = (-1)^k e^{-x}$, and by Faà di Bruno's formula, we get

$$\frac{d^n}{dx^n} f(g(0)) = \sum_{\mathcal{S}_n} \frac{n! e^{-\sum_{j=1}^n m_j \theta^*} (\sum_{j=1}^n m_j)! (1 - e^{-\theta^*})^{-\sum_{j=1}^n m_j - 1}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_n! n!^{m_n}} \cdot \prod_{j=1}^n (-1)^{j \cdot m_j}. \tag{A.45}$$

By the formula for $\alpha_k(w)$, we have

$$\beta_k(w) = \sum_{n=0}^k \sum_{\mathcal{S}_n} \frac{e^{-\sum_{j=1}^n m_j \theta^*} (\sum_{j=1}^n m_j)! (1 - e^{-\theta^*})^{-\sum_{j=1}^n m_j - 1}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_n! n!^{m_n}} \cdot \prod_{j=1}^n (-1)^{j \cdot m_j} \tag{A.46}$$

$$\cdot \sum_{\ell=0}^{k-n} \frac{\psi^{(k-n-\ell)}(\theta^*)}{(k-n-\ell)!} \sum_{\mathcal{S}_\ell} \frac{(-1)^{m_1+\cdots+m_\ell}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_\ell! \ell!^{m_\ell}}$$

$$\cdot \prod_{j=1}^\ell \left(\frac{1}{\eta''(\theta^*)} \frac{\eta^{(j+2)}(\theta^*)}{(j+2)(j+1)}\right)^{m_j} \frac{(i)^k w^{k+2(m_1+\cdots+m_\ell)}}{(\eta''(\theta^*))^{k/2}}.$$

Hence, the formula for b_k follows.

A.5. Proof of Proposition 3(i): Computations of $\eta^{(k)}(\theta^*)$

We recall that $\eta(\theta) = \nu(x(\theta) - 1)$ so that

$$\eta^{(k)}(\theta^*) = \nu x^{(k)}(\theta^*), \quad k \in \mathbb{N}. \tag{A.47}$$

Note that

$$x(\theta) = e^{\theta + \|h\|_{L^1}(x(\theta)-1)} = e^{\theta - \|h\|_{L^1}} e^{\|h\|_{L^1} x(\theta)}. \tag{A.48}$$

Therefore, by Leibniz formula,

$$x^{(k)}(\theta^*) = \sum_{\ell=0}^{k} \binom{k}{\ell} e^{\theta^* - \|h\|_{L^1}} \frac{d^\ell}{d\theta^\ell} e^{\|h\|_{L^1} x(\theta)} \bigg|_{\theta=\theta^*}. \tag{A.49}$$

By Faà di Bruno's formula, we get

$$\frac{d^\ell}{d\theta^\ell} e^{\|h\|_{L^1} x(\theta)} \bigg|_{\theta=\theta^*} = \sum_{\mathcal{S}_\ell} \frac{\ell! \cdot \|h\|_{L^1}^{m_1+\cdots+m_\ell} \cdot e^{\|h\|_{L^1} x(\theta^*)}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_\ell! \ell!^{m_\ell}} \cdot \prod_{j=1}^{\ell} \left(x^{(j)}(\theta^*)\right)^{m_j}. \tag{A.50}$$

Hence,

$$x^{(k)}(\theta^*) = e^{\theta^* - \|h\|_{L^1}} \frac{d^k}{d\theta^k} e^{\|h\|_{L^1} x(\theta)} \bigg|_{\theta=\theta^*} + \sum_{\ell=0}^{k-1} \binom{k}{\ell} e^{\theta^* - \|h\|_{L^1}} \frac{d^\ell}{d\theta^\ell} e^{\|h\|_{L^1} x(\theta)} \bigg|_{\theta=\theta^*} \tag{A.51}$$

$$= e^{\theta^* - \|h\|_{L^1}} \|h\|_{L^1} \cdot e^{\|h\|_{L^1} x(\theta^*)} \cdot x^{(k)}(\theta^*)$$

$$+ e^{\theta^* - \|h\|_{L^1}} \sum_{\mathcal{T}_k} \frac{k! \cdot \|h\|_{L^1}^{m_1+\cdots+m_{k-1}} \cdot e^{\|h\|_{L^1} x(\theta^*)}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_{k-1}! (k-1)!^{m_{k-1}}} \cdot \prod_{j=1}^{k-1} \left(x^{(j)}(\theta^*)\right)^{m_j}$$

$$+ \sum_{\ell=0}^{k-1} \binom{k}{\ell} e^{\theta^* - \|h\|_{L^1}} \sum_{\mathcal{S}_\ell} \frac{\ell! \cdot \|h\|_{L^1}^{m_1+\cdots+m_\ell} \cdot e^{\|h\|_{L^1} x(\theta^*)}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_\ell! \ell!^{m_\ell}} \cdot \prod_{j=1}^{\ell} \left(x^{(j)}(\theta^*)\right)^{m_j},$$

where $\mathcal{T}_k$ denotes the set of $(k-1)$-tuples of non-negative integers $(m_1, \ldots, m_{k-1})$ satisfying the constraint $1 \cdot m_1 + 2 \cdot m_2 + 3 \cdot m_3 + \cdots + (k-1) \cdot m_{k-1} = k$. Note that $x(\theta^*) = e^{\theta + \|h\|_{L^1}(x(\theta^*)-1)}$. Also that $\theta^* = \arg\max_{\theta \geq 0} \{\theta x - \nu(x(\theta)-1)\}$. Thus, $x = \nu x'(\theta^*)$, which gives $x'(\theta^*) = \frac{x}{\nu}$ and $x(\theta^*) = \frac{x}{\nu + \|h\|_{L^1} x}$.

For $k \geq 1$, $x^{(k)}(\theta^*)$ can be computed recursively as:

$$x^{(k)}(\theta^*) = \frac{x(\theta^*)}{1 - \|h\|_{L^1} x(\theta^*)} \sum_{\mathcal{T}_k} \frac{k! \cdot \|h\|_{L^1}^{m_1+\cdots+m_{k-1}}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_{k-1}! (k-1)!^{m_{k-1}}} \cdot \prod_{j=1}^{k-1} \left(x^{(j)}(\theta^*)\right)^{m_j}$$

$$+ \frac{x(\theta^*)}{1 - \|h\|_{L^1} x(\theta^*)} \sum_{\ell=0}^{k-1} \binom{k}{\ell} \sum_{\mathcal{S}_\ell} \frac{\ell! \cdot \|h\|_{L^1}^{m_1+\cdots+m_\ell}}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_\ell! \ell!^{m_\ell}} \cdot \prod_{j=1}^{\ell} \left(x^{(j)}(\theta^*)\right)^{m_j}.$$

A.6. Proof of Proposition 3(ii): Computations of $\psi^{(k)}(\theta^*)$

Let us recall that

$$\psi(\theta) = e^{\nu \int_0^\infty [F(s;\theta) - x(\theta)] ds}. \tag{A.52}$$

where $F(\cdot;\theta)$ satisfies: $F(t;\theta) = e^{\theta + \int_0^t (F(t-s;\theta)-1)h(s)\,ds}$. Let $F^{(k)}(\cdot;\theta)$ denote the k-th partial derivative of $F(\cdot;\theta)$ w.r.t. θ. By Faà di Bruno's formula, we have

$$\psi^{(k)}(\theta^*) = \sum_{\mathcal{S}_k} \frac{k! \cdot v^{m_1+\cdots+m_k} \cdot \psi(\theta^*)}{m_1! 1!^{m_1} m_2! 2!^{m_2} \cdots m_k! k!^{m_k}} \cdot \prod_{j=1}^{k} \left(\int_0^\infty \left[F^{(j)}(s;\theta^*) - x^{(j)}(\theta^*) \right] ds \right)^{m_j}. \quad \text{(A.53)}$$

Moreover, by Leibniz formula, we can compute that

$$F^{(k)}(t;\theta^*) = \sum_{\ell=0}^{k} \binom{k}{\ell} e^{\theta^* - \int_0^t h(s)\,ds} \frac{d^\ell}{d\theta^\ell} e^{\int_0^t F(t-s;\theta)h(s)\,ds} \bigg|_{\theta=\theta^*}. \quad \text{(A.54)}$$

By Faà di Bruno's formula, similar to deriving the formulas for $x^{(k)}(\theta^*)$, we get (2.16).

A.7. Computations of a_1 and b_1 for Theorem 2

By Proposition 1, we can compute that

$$a_1 = -\frac{1}{2} \frac{\psi''(\theta^*)}{\eta''(\theta^*)} + \frac{1}{24} \frac{\psi(\theta^*)\eta^{(4)}(\theta^*) + 4\psi'(\theta^*)\eta^{(3)}(\theta^*)}{(\eta''(\theta^*))^2} - \frac{15}{72} \frac{\psi(\theta^*)(\eta^{(3)}(\theta^*))^2}{(\eta''(\theta^*))^3}, \quad \text{(A.55)}$$

$$b_1 = -\psi(\theta^*) \frac{1}{2} \frac{e^{-\theta^*} + e^{-2\theta^*}}{(1-e^{-\theta^*})^2} \frac{1}{\eta''(\theta^*)} - \frac{1}{2}\psi''(\theta^*) \frac{1}{\eta''(\theta^*)}$$

$$+ \psi(\theta^*) \left[\eta^{(4)}(\theta^*) \frac{1}{24} \frac{3}{(\eta''(\theta^*))^2} - \frac{1}{2}(\eta^{(3)}(\theta^*))^2 \frac{1}{36} \frac{15}{(\eta''(\theta^*))^3} \right]$$

$$+ \frac{e^{-\theta^*}}{1-e^{-\theta^*}} \frac{\psi'(\theta^*)}{\eta''(\theta^*)} + \psi'(\theta^*)\eta^{(3)}(\theta^*) \frac{1}{2(\eta''(\theta^*))^2} - \frac{e^{-\theta^*}}{1-e^{-\theta^*}} \eta^{(3)}(\theta^*) \frac{1}{2(\eta''(\theta^*))^2}. \quad \text{(A.56)}$$

Thus to obtain a_1 and b_1, we need to compute $\psi(\theta^*)$, $\psi'(\theta^*)$, $\psi''(\theta^*)$, $\eta''(\theta^*)$, $\eta^{(3)}(\theta^*)$ and $\eta^{(4)}(\theta^*)$ as provided in Proposition 3. We recall that $\psi(\theta) = e^{v \int_0^\infty [F(s;\theta)-x(\theta)]\,ds}$. Thus,

$$\psi'(\theta^*) = v \int_0^\infty \left[F^{(1)}(s;\theta^*) - x'(\theta^*) \right] ds \cdot \psi(\theta^*), \quad \text{(A.57)}$$

$$\psi''(\theta^*) = \left[\left(v \int_0^\infty \left[F^{(1)}(s;\theta^*) - x'(\theta^*) \right] ds \right)^2 \right.$$

$$\left. + v \int_0^\infty \left[F^{(2)}(s;\theta^*) - x''(\theta^*) \right] ds \right] \cdot \psi(\theta^*). \quad \text{(A.58)}$$

We first recall that:

$$F(t;\theta^*) = e^{\theta^* + \int_0^t (F(t-s;\theta^*)-1)h(s)\,ds}, \quad t \geq 0. \quad \text{(A.59)}$$

By differentiating (A.59) w.r.t. θ^* and applying (A.59) again, we obtain:

$$F^{(1)}(t;\theta^*) = \left[1 + \int_0^t F^{(1)}(t-s;\theta^*)h(s)\,ds \right] F(t;\theta^*), \quad t \geq 0. \quad \text{(A.60)}$$

Finally by differentiating (A.60) w.r.t. θ^* and applying (A.60) again, we obtain:

$$F^{(2)}(t;\theta^*) = \left[\int_0^t F^{(2)}(t-s;\theta^*)h(s)\,ds + \left(1 + \int_0^t F^{(1)}(t-s;\theta^*)h(s)\,ds \right)^2 \right] F(t;\theta^*), \quad \text{(A.61)}$$

for any $t \geq 0$. Let us recall that $\eta(\theta^*) = \nu(x(\theta^*) - 1)$ and

$$\eta''(\theta^*) = \nu x''(\theta^*), \qquad \eta^{(3)}(\theta^*) = \nu x^{(3)}(\theta^*), \qquad \eta^{(4)}(\theta^*) = \nu x^{(4)}(\theta^*). \quad \text{(A.62)}$$

We can compute from the formulas in Proposition 3 that

$$x'(\theta) = \frac{x(\theta)}{1 - \|h\|_{L^1} x(\theta)}, \qquad x''(\theta) = \frac{(1 + \|h\|_{L^1} x'(\theta))^2 x(\theta)}{1 - \|h\|_{L^1} x(\theta)}, \quad \text{(A.63)}$$

$$x^{(3)}(\theta) = \frac{(3(1 + \|h\|_{L^1} x'(\theta))\|h\|_{L^1} x''(\theta) + (1 + \|h\|_{L^1} x'(\theta))^3)x(\theta)}{1 - \|h\|_{L^1} x(\theta)}, \quad \text{(A.64)}$$

$$x^{(4)}(\theta) = \frac{(3(\|h\|_{L^1} x''(\theta))^2 + 3(1 + \|h\|_{L^1} x'(\theta))\|h\|_{L^1} x^{(3)}(\theta))x(\theta)}{1 - \|h\|_{L^1} x(\theta)}$$

$$+ \frac{(6\|h\|_{L^1}(1 + \|h\|_{L^1} x'(\theta))^2 x''(\theta) + (1 + \|h\|_{L^1} x'(\theta))^4)x(\theta)}{1 - \|h\|_{L^1} x(\theta)}. \quad \text{(A.65)}$$

Thus $x'(\theta^*)$, $x''(\theta^*)$, $x^{(3)}(\theta^*)$ and $x^{(4)}(\theta^*)$ can be computed from $x(\theta^*)$. Note that $\theta^* = \arg\max_{\theta \geq 0}\{\theta x - \nu(x(\theta) - 1)\}$. Thus, $x = \nu x'(\theta^*)$, which gives $x'(\theta^*) = \frac{x}{\nu}$ and $x(\theta^*) = \frac{x}{\nu + \|h\|_{L^1} x}$. Hence, we get

$$x'(\theta^*) = \frac{x}{\nu}, \qquad x''(\theta^*) = \left(1 + \frac{\|h\|_{L^1} x}{\nu} \right)^2 \frac{x}{\nu}, \quad \text{(A.66)}$$

$$x^{(3)}(\theta^*) = \left(1 + \frac{\|h\|_{L^1} x}{\nu} \right)^3 \left(1 + \frac{3\|h\|_{L^1} x}{\nu} \right) \frac{x}{\nu}, \quad \text{(A.67)}$$

$$x^{(4)}(\theta^*) = 3\left(1 + \frac{\|h\|_{L^1} x}{\nu} \right)^4 \left(\left(\frac{\|h\|_{L^1} x}{\nu} \right)^2 + \left(1 + \frac{3\|h\|_{L^1} x}{\nu} \right) \frac{\|h\|_{L^1} x}{\nu} \right) \frac{x}{\nu}$$

$$+ \left(1 + \frac{\|h\|_{L^1} x}{\nu} \right)^4 \left(1 + \frac{6\|h\|_{L^1} x}{\nu} \right) \frac{x}{\nu}. \quad \text{(A.68)}$$

A.8. Proof of Proposition 12

Proof of Proposition 12. By change of measure for counting processes, see, for example, [34]

$$\mathbb{E}[1_{\{N_t \geq xt\}}] = \hat{\mathbb{E}}\left[1_{\{N_t \geq xt\}} \frac{d\mathbb{P}}{d\hat{\mathbb{P}}}\bigg|_{\mathcal{F}_t} \right] \quad \text{(A.69)}$$

$$= \hat{\mathbb{E}}\left[1_{\{N_t \geq xt\}} \cdot e^{\int_0^t \log(\frac{\lambda_s}{\hat{\lambda}_s})\,dN_s - \int_0^t (\lambda_s - \hat{\lambda}_s)\,ds} \right]$$

$$= \hat{\mathbb{E}}\left[1_{\{N_t \geq xt\}} \cdot e^{-(\log \gamma)N_t + (\gamma - 1)\int_0^t \lambda_s\,ds} \right]$$

$$= \hat{\mathbb{E}}\Big[1_{\{N_t \geq xt\}} \cdot e^{-(\log\gamma)N_t + (\gamma-1)vt + (\gamma-1)\int_0^t \int_0^{s^-} h(s-u)\,dN_u\,ds}\Big]$$

$$= \hat{\mathbb{E}}\Big[1_{\{N_t \geq xt\}} \cdot e^{-(\log\gamma)N_t + (\gamma-1)vt + (\gamma-1)\int_0^{t^-} [\int_u^t h(s-u)\,ds]\,dN_u}\Big]$$

$$= \hat{\mathbb{E}}\Big[1_{\{N_t \geq xt\}} \cdot e^{-(\log\gamma)N_t + (\gamma-1)vt + (\gamma-1)\|h\|_{L^1}N_t - (\gamma-1)\int_0^t H(t-u)\,dN_u}\Big]$$

$$= e^{-tI(x)}\hat{\mathbb{E}}\Big[1_{\{N_t \geq xt\}} \cdot e^{((\gamma-1)\|h\|_{L^1} - \log\gamma)(N_t - xt) - (\gamma-1)\int_0^t H(t-u)\,dN_u}\Big],$$

where we recall that $H(t) = \int_t^\infty h(s)\,ds$ denotes the right tail of the exciting function. $\qquad\square$

Acknowledgements

References

[1] Abergel, F. and Jedidi, A. (2015). Long-time behavior of a Hawkes process-based limit order book. *SIAM J. Financial Math.* **6** 1026–1043. MR3418223 https://doi.org/10.1137/15M1011469

[2] Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R.J.A. (2015). Modeling financial contagion using mutually exciting jump processes. *J. Financ. Econ.* **117** 585–606.

[3] Aït-Sahalia, Y., Laeven, R.J.A. and Pelizzon, L. (2014). Mutual excitation in Eurozone sovereign CDS. *J. Econometrics* **183** 151–167. MR3276385 https://doi.org/10.1016/j.jeconom.2014.05.006

[4] Alfonsi, A. and Blanc, P. (2016). Dynamic optimal execution in a mixed-market-impact Hawkes price model. *Finance Stoch.* **20** 183–218. MR3441291 https://doi.org/10.1007/s00780-015-0282-y

[5] Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J.F. (2013). Modelling microstructure noise with mutually exciting point processes. *Quant. Finance* **13** 65–77. MR3005350 https://doi.org/10.1080/14697688.2011.647054

[6] Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J.F. (2013). Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Process. Appl.* **123** 2475–2499. MR3054533 https://doi.org/10.1016/j.spa.2013.04.007

[7] Bauwens, L. and Hautsch, N. (2009). *Handbook of Financial Time Series*: *Modelling Financial High Frequency Data Using Point Processes*. Berlin: Springer.

[8] Bordenave, C. and Torrisi, G.L. (2007). Large deviations of Poisson cluster processes. *Stoch. Models* **23** 593–625. MR2362700 https://doi.org/10.1080/15326340701645959

[9] Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Probab.* **24** 1563–1588. MR1411506 https://doi.org/10.1214/aop/1065725193

[10] Chavez-Demoulin, V., Davison, A.C. and McNeil, A.J. (2005). Estimating value-at-risk: A point process approach. *Quant. Finance* **5** 227–234. MR2240248 https://doi.org/10.1080/14697680500039613

[11] Chu, S.C. and Metcalf, F.T. (1967). On Gronwall's inequality. *Proc. Amer. Math. Soc.* **18** 439–440. MR0212529 https://doi.org/10.2307/2035473

[12] Crane, R. and Sornette, D. (2008). Robust dynamic classes revealed by measuring the response function of a social system. *Proc. Natl. Acad. Sci. USA* **105** 15649–15653. https://doi.org/10.1073/pnas.0803685105

[13] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. New York: Springer. MR1619036 https://doi.org/10.1007/978-1-4612-5320-4

[14] Errais, E., Giesecke, K. and Goldberg, L.R. (2010). Affine point processes and portfolio credit risk. *SIAM J. Financial Math.* **1** 642–665. MR2719785 https://doi.org/10.1137/090771272

[15] Féray, V., Méliot, P.-L. and Nikeghbali, A. (2016). *Mod-ϕ Convergence*: *Normality Zones and Precise Deviations*. *SpringerBriefs in Probability and Mathematical Statistics*. Cham: Springer. MR3585777 https://doi.org/10.1007/978-3-319-46822-8

[16] Gao, X., Zhou, X. and Zhu, L. (2018). Transform analysis for Hawkes processes with applications in dark pool trading. *Quant*. *Finance* **18** 265–282. MR3750735 https://doi.org/10.1080/14697688.2017.1403151

[17] Gao, X. and Zhu, L. (2018). Functional central limit theorems for stationary Hawkes processes and application to infinite-server queues. *Queueing Syst*. **90** 161–206. MR3850052 https://doi.org/10.1007/s11134-018-9570-5

[18] Gao, X. and Zhu, L. (2018). Limit theorems for Markovian Hawkes processes with a large initial intensity. *Stochastic Process*. *Appl*. **128** 3807–3839. MR3860011 https://doi.org/10.1016/j.spa.2017.12.001

[19] Gao, X. and Zhu, L. (2018). Large deviations and applications for Markovian Hawkes processes with a large initial intensity. *Bernoulli* **24** 2875–2905. MR3779705 https://doi.org/10.3150/17-BEJ948

[20] Giot, P. (2005). Market risk models for intraday data. *Eur*. *J*. *Finance* **11** 309–324.

[21] Gusto, G. and Schbath, S. (2005). FADO: A statistical method to detect favored or avoided distances between occurrences of motifs using the Hawkes' model. *Stat*. *Appl*. *Genet*. *Mol*. *Biol*. **4** Art. 24, 28. MR2170440 https://doi.org/10.2202/1544-6115.1119

[22] Hainaut, D. (2016). A bivariate Hawkes process for interest rates modelling. *Econ*. *Model*. **57** 180–196.

[23] Hawkes, A.G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58** 83–90. MR0278410 https://doi.org/10.1093/biomet/58.1.83

[24] Hawkes, A.G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *J*. *Appl*. *Probab*. **11** 493–503. MR0378093 https://doi.org/10.2307/3212693

[25] Jaisson, T. and Rosenbaum, M. (2015). Limit theorems for nearly unstable Hawkes processes. *Ann*. *Appl*. *Probab*. **25** 600–631. MR3313750 https://doi.org/10.1214/14-AAP1005

[26] Jaisson, T. and Rosenbaum, M. (2016). Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *Ann*. *Appl*. *Probab*. **26** 2860–2882. MR3563196 https://doi.org/10.1214/15-AAP1164

[27] Karabash, D. and Zhu, L. (2015). Limit theorems for marked Hawkes processes with application to a risk model. *Stoch*. *Models* **31** 433–451. MR3395721 https://doi.org/10.1080/15326349.2015.1024868

[28] Mohler, G.O., Short, M.B., Brantingham, P.J., Schoenberg, F.P. and Tita, G.E. (2011). Self-exciting point process modeling of crime. *J*. *Amer*. *Statist*. *Assoc*. **106** 100–108. MR2816705 https://doi.org/10.1198/jasa.2011.ap09546

[29] Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *J*. *Amer*. *Statist*. *Assoc*. **83** 9–27.

[30] Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Ann*. *Inst*. *Statist*. *Math*. **50** 379–402.

[31] Pernice, V., Staude, B., Cardanobile, S. and Rotter, S. (2011). How structure determines correlations in neuronal networks. *PLoS Comput*. *Biol*. **7** e1002059, 14. MR2821638 https://doi.org/10.1371/journal.pcbi.1002059

[32] Porter, M.D. and White, G. (2012). Self-exciting hurdle models for terrorist activity. *Ann*. *Appl*. *Stat*. **6** 106–124. MR2951531 https://doi.org/10.1214/11-AOAS513

[33] Reynaud-Bouret, P. and Schbath, S. (2010). Adaptive estimation for Hawkes processes; application to genome analysis. *Ann*. *Statist*. **38** 2781–2822. MR2722456 https://doi.org/10.1214/10-AOS806

[34] Sokol, A. and Hansen, N.R. (2015). Exponential martingales and changes of measure for counting processes. *Stoch*. *Anal*. *Appl*. **33** 823–843. MR3378040 https://doi.org/10.1080/07362994.2015.1040890

[35] Stabile, G. and Torrisi, G.L. (2010). Risk processes with non-stationary Hawkes claims arrivals. *Methodol*. *Comput*. *Appl*. *Probab*. **12** 415–429. MR2665268 https://doi.org/10.1007/s11009-008-9110-6

[36] Torrisi, G.L. (2016). Gaussian approximation of nonlinear Hawkes processes. *Ann*. *Appl*. *Probab*. **26** 2106–2140. MR3543891 https://doi.org/10.1214/15-AAP1141

[37] Torrisi, G.L. (2017). Poisson approximation of point processes with stochastic intensity, and application to nonlinear Hawkes processes. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **53** 679–700. MR3634270 https://doi.org/10.1214/15-AIHP730

[38] Zhang, X., Blanchet, J., Giesecke, K. and Glynn, P.W. (2009). Rare event simulation for a generalized Hawkes process. In *Proceedings of the* 2009 *Winter Simulation Conference*.

[39] Zhang, X., Blanchet, J., Giesecke, K. and Glynn, P.W. (2015). Affine point processes: Approximation and efficient simulation. *Math*. *Oper*. *Res*. **40** 797–819. MR3423739 https://doi.org/10.1287/moor.2014.0696

[40] Zhao, Q., Erdogdu, M.A., He, H.Y., Rajaraman, A. and Jure, L. (2015). SEISMIC: A self-exciting point process model for predicting tweet popularity. In *Proceedings of the* 21*th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1513–1522.

[41] Zhu, L. (2013). Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. *Insurance Math*. *Econom*. **53** 544–550. MR3130449 https://doi.org/10.1016/j.insmatheco.2013.08.008

[42] Zhu, L. (2013). Moderate deviations for Hawkes processes. *Statist*. *Probab*. *Lett*. **83** 885–890. MR3040318 https://doi.org/10.1016/j.spl.2012.12.011

[43] Zhu, L. (2013). Central limit theorem for nonlinear Hawkes processes. *J*. *Appl*. *Probab*. **50** 760–771. MR3102513 https://doi.org/10.1239/jap/1378401234

[44] Zhu, L. (2014). Process-level large deviations for nonlinear Hawkes point processes. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **50** 845–871. MR3224291 https://doi.org/10.1214/12-AIHP532

[45] Zhu, L. (2015). Large deviations for Markovian nonlinear Hawkes processes. *Ann*. *Appl*. *Probab*. **25** 548–581. MR3313748 https://doi.org/10.1214/14-AAP1003