MATHEMATICAL COMMUNICATIONS
Math. Commun. **18**(2013), 393–407

On the Bartlett spectrum of randomized Hawkes processes

Mark Kelbert[1,*], Nikolai Leonenko[2] and Vladimir Belitsky[3]

[1] *Department of Mathematics, University of Wales-Swansea, Singleton Park, Swansea, SA2 8PP, UK*
[2] *Cardiff School of Mathematics, Cardiff University, Senghennydd Road, Cardiff CF24 4AG, UK*
[3] *IME, University of Sao Paulo, Sao Paulo, Brazil*

Received January 30, 2013; accepted June 13, 2013

Abstract. We study the Bartlett spectrum of the randomized Hawkes process and demonstrate that it behaves very differently from the case of a classical Hawkes process. In particular, the Bartlett spectrum could have a singularity near the origin which indicates a long-range dependence property.

AMS subject classifications: 60G12, 60G55
Key words: point process, Hawkes process, randomization, Bartlett spectrum, long-range dependence, Mittag-Lefler function

1. Introduction

In this note we introduce randomized Hawkes processes and study their Bartlett spectra with a goal to provide an example of a point process with the singular Bartlett spectrum or long-range dependence. Long-range dependence or long-memory phenomena in stochastic processes was discovered in various applied models, see [3, 4, 20, 25, 29] for recent development. This is a rapidly developing subject in probability and statistics; a summary is given by [6, 17] (see also references therein). Book [17] contains an outstanding survey of the field, in particular, it discusses different definitions of long-range dependence of stationary processes in terms of the autocorrelation function (the integral of the correlation function diverges) or the spectrum (the spectral density has a singularity at zero). In the point processes context, the definition of long-range dependence has to be reconsidered by transforming the spectrum into the Bartlett spectrum [5].

For point processes the long-range dependence phenomena from different viewpoints were discussed in [2, 9, 12, 14, 15, 22, 28, 33, 34]. Here we present a new version of point processes with the singular Bartlett spectrum originated from the so-called Hawkes self-exciting processes (see [18, 19] and [13], p. 309) by 'randomizing' a parameter of the exciting function (cf. [5] or [13], p. 303–312). We show similarities and differences of the singular property of the Bartlett spectrum with that of the Bochner-Khintchine spectrum (as indication of long-range dependence) for stationary processes used in different models with randomization of parameters

*Corresponding author. *Email addresses:* `m.kelbert@swansea.ac.uk` (M. Kelbert), `LeonenkoN@cardiff.ac.uk` (N. Leonenko), `belitsky@ime.usp.br` (V. Belitsky)

before (see [3, 4, 9, 23, 32] and the references therein).

2. The Hawkes self-exciting process

Hawkes [18] (see also [19] or [13], p. 309) introduced a class of point processes which he called self-exciting processes (SEP). Nowadays, they are called linear SEP; generalized Hawkes processes with a non-linear rate function are studied in [9] and [11]. SEP are widely used in applications to seismology, cosmology, neurophysiology and DNA modeling, epidemiology and reliability (see, e.g. [10, 11, 30, 31] and [13], p. 183 and p. 309).

We define an SEP as a cluster process with cluster centers $\mathcal{P} = (t_n, n \in \mathbf{Z})$ produced by a stationary Poisson process of rate $\lambda > 0$. Each center t_n independently generates a cluster C_n with the following branching structure: the center t_n is said to be of generation zero. Given generations $0, 1, \ldots, k-1$ in C_n each point $\tau \in C_n$ of $(k-1)$th generation produces a finite Poisson process of offspring with a non-negative exciting function $h(. - \tau)$ such that

$$0 < \nu = \int_0^\infty h(u)\mathrm{d}u < 1 \tag{1}$$

the union of these offspring families form the generation k. Consider the associated point measure

$$\mathcal{T}(C) = \sum_{n \in \mathbf{Z}} \mathbf{1}_C(\tau_n), \quad C \in \mathcal{B},$$

where $\mathcal{B}$ is the Borel $\sigma-$field of $\mathbf{R}$, then the process

$$\mathcal{T}(t) = \mathcal{T}((0, t])$$

is called a Hawkes self-exciting point process (SEP). So, $\mathcal{T} = (\tau_n, n \in \mathbf{Z})$ is defined as a cluster process with a stationary Poisson process $\mathcal{P}$ of rate $\lambda > 0$ for cluster centers. The clusters form independent realizations of a finite branching process with the age-dependent probability of birth $h(.)$. Moreover, there exits a stationary version of SEP [13] which is considered hereafter.

Equivalently, SEP $\mathcal{T}$ can be described in terms of shot noise or filtering point processes, or in terms of point processes with a random intensity function of the form

$$\Lambda(t) = \Lambda(t, \mathcal{F}_t) = \lambda + \sum_{i \in \mathbf{Z}} h(t - \tau_i)\mathbf{1}_{\tau_i < t} = \lambda + \int_{-\infty}^t h(t - s)\mathrm{d}\mathcal{T}(s), \tag{2}$$

where $\mathcal{F}_t = \sigma(\tau_n \in \mathcal{T} : \ \tau_n \leq t)$ is the history of a point process $\mathcal{T}$. The interpretation of $\mathcal{F}_t-$stochastic intensity is as follows:

$$\mathbf{E}[\mathcal{T}(t, t + \mathrm{d}t) \mid \mathcal{F}_t] = \Lambda(t)\mathrm{d}t,$$

that is $\Lambda(t)$ is the (random) rate of the point process at time t, given the observation history $\mathcal{F}_t$.

If $\lambda > 0$ and $\nu < 1$, there exists a unique stationary cluster process $\mathcal{T}$ with rate $\tilde{\lambda} = \lambda/(1-\nu)$ satisfying (2) (see [8, 19]), where

$$\tilde{\lambda} = \frac{\mathbf{E}[\mathcal{T}(t, t + \mathrm{d}t)]}{\mathrm{d}t}.$$

The Bartlett spectral measure Γ of a (general) stationary point process $\mathcal{T} = (\tau_n, n \in \mathbf{Z})$ is defined, when it exists, by the relation

$$\mathbf{V}ar\Big(\int_{\mathbf{R}} \phi(t)\mathcal{T}(\mathrm{d}t)\Big) = \int_{\mathbf{R}} \mid \tilde{\phi}(\omega) \mid^2 \Gamma(\mathrm{d}\omega) \tag{3}$$

for any function $\phi \in \mathrm{L}^1 \cap \mathrm{L}^2 = \{\phi : \int_{\mathbf{R}} |\phi(t)|\mathrm{d}t < \infty, \int_{\mathbf{R}} |\phi(t)|^2\mathrm{d}t < \infty\}$, where

$$\tilde{\phi}(\omega) = \int_{\mathbf{R}} e^{i\omega t}\phi(t)\mathrm{d}t$$

is the Fourier transform of ϕ.

Formally, the Bartlett spectrum of a (general) stationary point process $\mathcal{T} = (\tau_n, n \in \mathbf{Z})$ is the Bochner-Khintchine spectrum of a generalized process

$$\tilde{\mathcal{T}}(t) = \sum_{n\in\mathbf{Z}} \delta(. - \tau_n)\mathbf{1}_{\tau_n < t}, \tag{4}$$

where $\delta(.)$ is the Dirac (generalized) function. This fact establishes a connection between the Bartlett spectrum Γ (see (3)) and the Bochner-Khintchine spectrum M of a stationary process (4) via the following formula

$$\mathbf{V}ar\Big(\int_{\mathbf{R}} \phi(t)\tilde{\mathcal{T}}(\mathrm{d}t)\Big) = \int_{\mathbf{R}} \mid \tilde{\phi}(\omega) \mid^2 M(\mathrm{d}\omega). \tag{5}$$

The Bartlett spectrum is well-defined if $\mathbf{E}\mathcal{T}^2(C) < \infty$ for all bounded Borel $C \in \mathbf{R}$. The Bartlett spectral density of the Hawkes process can be written as follows (see [18] or [13], p. 309)

$$\gamma(\omega) = \frac{\lambda}{2\pi(1-\nu)} \frac{1}{\mid 1 - \tilde{h}(\omega) \mid^2}, \quad \omega \in \mathbf{R}, \nu \in (0, 1). \tag{6}$$

Thus, the spectral density (6) tends to a finite constant as $\omega \to 0+$

$$\gamma(0) = \frac{\lambda}{2\pi(1-\nu)^3}.$$

This may be viewed as an indicator of short-range dependence. In this case, the variance grows linearly as $t \to \infty$:

$$\mathbf{V}ar\Big(\mathcal{T}(t)\Big) = 2\pi t\gamma(0)(1 + o(1)).$$

Moreover, under condition $\int_0^\infty uh(u)\mathrm{d}u < \infty$, the Central limit theorem holds:

$$\frac{[\mathcal{T}(t) - \lambda t/(1-\nu)]}{[\lambda t/(1-\nu)^3]^{1/2}} \Rightarrow^D N(0, 1)$$

as $t \to \infty$ (see [19]). Here $\Rightarrow^D$ stands for convergence in distributions, and $N(0, 1)$ is the standard normal law.

3. Randomized Hawkes processes

Here we propose a version of the Hawkes process with long-range dependence constructed via randomization of the exciting function, in contrast to [9], where the existence of critical SEP with the long-range dependence property was established under a heavy-tail condition on the exciting function, and investigate the Bartlett spectrum of such randomized SEP.

A randomized self-exciting process (RSEP) is defined via a randomized exciting function h_ζ depending on a random parameter ζ. The simplest example of an RSEP $\mathcal{T}_\zeta = (\tau_n^\zeta, n \in \mathbf{Z})$, or equivalently, $\mathcal{T}_\zeta(t) = \sum_{n\in\mathbf{Z}} \mathbf{1}_{(0,t)}(\tau_n^\zeta)$, with singular Bartlett spectra is as follows. Let ζ be a random variable and conditional on $\{\zeta = \nu\}$, $\mathcal{T}_\zeta$ is a classical Hawkes processes with the exciting function (9), i.e.

$$\Lambda_\nu(t) = \lambda + \int_{-\infty}^{t} h_\nu(t-s)\mathrm{d}\mathcal{T}(s) = \lambda + \sum_{n\in\mathbf{Z}} h_\nu(t-\tau_n^\nu)\mathbf{1}_{\tau_n^\nu < t}.$$

Alternatively, we can define the randomized Hawkes process $\mathcal{T}_\zeta$ by the following expression for its generating functional

$$\begin{aligned} G[\phi, \mathcal{T}_\zeta] &= \mathbf{E}\Big(\prod_{\tau_i^\zeta \in \mathcal{T}_\zeta} \phi(\tau_i^\zeta)\Big) = \mathbf{E}\exp\big[\int \log\phi(t)\mathrm{d}\mathcal{T}_\zeta(t)\big] \\ &= \mathbf{E}_\zeta \exp\big[\lambda \int_{\mathbf{R}} (F_\zeta(\phi(t+.)) - 1)\mathrm{d}t\big], \quad \phi \in S \end{aligned} \tag{7}$$

for a suitable space of functions S (see [13], p. 152–153). Here $F_\zeta(\phi(.))$ is the generating functional of a cluster generated by a single particle at the origin depending on the random parameter ζ, while $\phi(t+.)$ is simply the translation of $\phi(.)$.

Note (see [19]) that $F_\nu(\phi(.))$ for a fixed $\zeta = \nu$ satisfies the functional equation

$$F_\nu(\phi(.)) = \phi(0)\exp\big[\int_{\mathbf{R}} (F_\zeta(\phi(t+.)) - 1)\mu_\nu(t)\mathrm{d}t\big]. \tag{8}$$

The behaviour of variance and the Bartlett spectrum of RSEP may differ drastically from the classical case. The simplest example of RSEP with a singular spectrum is as follows. Let ζ be a random variable with support $(0,1)$ which has Beta distribution (11). Suppose that, conditional on $\{\zeta = \nu\}$, $\mathcal{T}_\zeta$ has an exciting function

$$h(u) = h_\nu(u) = \nu\alpha e^{-\alpha u}, \quad u > 0, \alpha > 0, \nu \in (0,1) \tag{9}$$

with a randomized parameter $\zeta = \nu \in (0,1)$. Note that

$$\tilde{h}_\nu(\omega) = \frac{\nu\alpha}{\alpha - i\omega}, \quad \mid 1 - \tilde{h}_\nu(\omega) \mid^{-2} = \frac{\alpha^2 + \omega^2}{\alpha^2(1-\nu)^2 + \omega^2}, \quad \omega \in \mathbf{R}. \tag{10}$$

Let us assume that ν has Beta $B(a,b)$- distribution of the form

$$f(u) = \frac{1}{B(a,b)} u^{a-1}(1-u)^{b-1}, \quad 0 < u < 1, a > 0, b > 0. \tag{11}$$

Using the law of total variance $\mathbf{V}ar[Y] = \mathbf{E}\big(\mathbf{V}ar[Y|X]\big) + \mathbf{V}ar[\mathbf{E}(Y|X)]$, the expression for the Bartlett spectrum for the classical Hawkes process (6), and Parseval-Plancherel isometry formula it is easy to see that for all functions $\phi \in \mathrm{L}^1 \cap \mathrm{L}^2$ we have

$$\mathbf{V}ar\Big(\int_{\mathbf{R}} \phi(t)\mathcal{T}_\zeta(\mathrm{d}t)\Big) = \int_{\mathbf{R}} \mathrm{d}\omega \mid \tilde{\phi}(\omega) \mid^2 \int_0^1 \gamma_u(\omega) f(u)\mathrm{d}u + \left(\lambda \int_{\mathbf{R}} \phi(t)\mathrm{d}t\right)^2 \mathbf{V}ar\Big[\frac{1}{1-\nu}\Big], \tag{12}$$

where

$$\gamma_u(\omega) = \frac{\lambda}{2\pi(1-u)} \frac{\alpha^2+\omega^2}{\alpha^2(1-u)^2+\omega^2}, \quad \omega \in \mathbf{R}. \tag{13}$$

Theorem 1. *Let $\mathcal{T}_\nu$ be a Hawkes process with the random coefficient ν, which follows the Beta PDF as in (11). Then for $a > 0, b > 2$*

(i) $\mathbf{E}[\frac{1}{1-\nu}]^2 < \infty$.

(ii) The Bartlett spectral density of a process $\mathcal{T}_\nu$ is of the form

$$\gamma_\nu(\omega) = L_\nu(\omega) + M\delta(\omega), \quad \omega \in \mathbf{R} \tag{14}$$

where $\delta(\omega)$ is a unit mass at $\omega = 0$,

$$L_\nu(\omega) = \frac{\lambda(\alpha^2+\omega^2)}{2\pi B(a,b)} \int_0^1 u^{a-1}(1-u)^{b-2} \frac{\mathrm{d}u}{\alpha^2(1-u)^2+\omega^2}, \quad \omega \in \mathbf{R}, \tag{15}$$

and

$$M = \frac{\lambda^2}{B(a,b)}\left[B(a,b-2) - \frac{B(a,b-1)^2}{B(a,b)}\right]. \tag{16}$$

(iii) For $a = 1, b \in (2,3)$ the absolutely continuous component $L_\nu(\omega)$ of Bartlett spectral density (14) has the following singular property as $\omega \to 0+$

$$L_\nu(\omega) \approx \frac{C_0}{\omega^{3-b}}, \tag{17}$$

where

$$C_0 = C_0(\lambda,\alpha,b) = \frac{\lambda\alpha^2}{2\pi B(1,b)} \int_0^\infty \frac{y^{b-2}\mathrm{d}y}{(\alpha^2y^2+1)}.$$

(iv) For $a = 1, b \in (2,3)$ the first two moments of the Hawkes process $\mathcal{T}_\nu$ have the form

$$\mathbf{E}\mathcal{T}_\nu(t) = t\lambda \frac{B(1,b-1)}{B(1,b)}, \tag{18}$$

$$\mathbf{V}ar[\mathcal{T}_\nu(t)] = S_{sg}(t) + t^2 \frac{\lambda^2}{B(1,b)}\left[B(1,b-2) - \frac{B(1,b-1)^2}{B(1,b)}\right] \tag{19}$$

and, in the limit $t \to \infty$,

$$S_{sg}(t) \approx t^{4-b} C_0(\lambda, \alpha, b) \frac{\Gamma(b-2) \sin\left(\frac{\pi(3-b)}{2}\right)}{(3-b)(4-b)}. \tag{20}$$

The proof of Theorem 1 is given in Section 4.

Remark 1. *Similar results can be obtained for the Erlang fertility rate*

$$h(u) = \nu \alpha^2 u e^{-\alpha u},$$

for $u > 0$ and $0 < \nu < 1$, instead of exponential (9). Indeed, in this case $1 - \tilde{h}_\nu(\omega) = 1 - \frac{\nu\alpha^2}{(\alpha - i\omega)^2}$, and

$$\gamma_\nu(\omega) = L_\nu(\omega) + M\delta(\omega), \quad \omega \in \mathbf{R}$$

where the constant M is given by (16), while the first term

$$L_\nu(\omega) = \frac{\lambda}{2\pi B(a,b)} \int_0^1 u^{a-1}(1-u)^{b-2} \frac{(\alpha^2 + \omega^2)^2 \mathrm{d}u}{[\alpha^2(1-u)^2 - \omega^2]^2 + 4\alpha^2\omega^2}, \quad \omega \in \mathbf{R}.$$

Thus for $a = 1, 2 < b < 3$ we have as $\omega \to 0+$

$$\gamma_\nu(\omega) \approx \frac{C_1}{\omega^{3-b}},$$

where now

$$C_1 = C_1(\lambda, \alpha, b) = \frac{\lambda\alpha^4}{2\pi B(1,b)} \int_0^\infty \frac{y^{b-2} \mathrm{d}y}{\alpha^4 y^2 + 4\alpha^2}.$$

Remark 2. *In this example $\mathbf{V}ar[\mathcal{T}_\nu(t)]$ increases asymptotically as t^2, while the first term $S_{sg}(t)$ behaves as $t^\beta, 1 < \beta = 4 - b < 2$ and contains information about the singularity parameter β. On the other hand, the singular property of the Bartlett spectrum (17) indicates the long-range dependence property of RSEP. In contrast to the singular property of the Bochner-Khintchine spectrum of ordinary stationary processes (cf. [17] or [3, 4, 23, 32]), the singular property of the Bartlett spectrum is more sophisticated and contains two terms (see (14)). Moreover, the variance in (17) is represented as the sum of two terms and only $S_{sg}(t)$ contains information about the singularity parameter, which is important for statistical inference (see again [6] or [17]). It is interesting that for the point processes we cannot use the randomization procedure proposed by [23, 32], since in this case the limiting process cannot be a point process in view of the Central limit theorem. On the other hand, the approach of [3, 4] is also not applicable to point processes, since it is based on the Lévy-Khintchine canonical representation of the characteristic function. For these reasons, our approach is a natural way for introducing long-range dependence in point processes, and is different from the existing in the literature (see again [2, 9, 12, 14, 15, 22, 28, 33, 34] and the references therein).*

4. Proof of Theorem 3.1

(i) This follows from the special form of Beta distribution.

(ii) This follows from (12) and observation that the second term in (12) has the form

$$\lambda^2 \mid \tilde{\phi}(0) \mid^2 \mathbf{V}ar\Big[\frac{1}{1-\nu}\Big].$$

Again, due to the special form of Beta distribution one gets the constant M in (16).

(iii) So, we concentrate on the absolutely continuous term. It is easy to compute $\tilde{\mu}_\nu(\omega) = \frac{\nu\alpha}{\alpha - i\omega}$,

$$\mid 1 - \tilde{\mu}_\nu(\omega) \mid^{-2} = \frac{\alpha^2+\omega^2}{\alpha^2(1-\nu)^2+\omega^2},$$

and the singular part of the Bartlett spectrum

$$L_\nu(\omega) = \frac{\lambda(\alpha^2+\omega^2)}{2\pi}\int_0^1 f_\nu(u)\frac{1}{(1-u)}\frac{\mathrm{d}u}{(\alpha^2(1-u)^2+\omega^2)}. \tag{21}$$

For the Beta-distribution with $a=1$, (21) takes the form

$$L_\nu(\omega) = \frac{\lambda(\alpha^2+\omega^2)}{2\pi B(1,b)} I, \quad I = \int_0^1 (1-u)^{b-2}\frac{\mathrm{d}u}{\alpha^2(1-u)^2+\omega^2}.$$

Observe that for $\omega \to 0+$

$$\int_0^1 \frac{x^{b-2}\omega^{3-b}}{\alpha^2x^2+\omega^2}\mathrm{d}x = \int_0^{1/\omega} \frac{y^{b-2}}{\alpha^2y^2+1}\mathrm{d}y \ \to \ \int_0^\infty \frac{y^{b-2}}{\alpha^2y^2+1}\mathrm{d}y.$$

Hence,

$$I \approx \frac{C_1(\alpha)}{\omega^{3-b}}, C_1(\alpha) = \int_0^\infty \frac{y^{b-2}\mathrm{d}y}{\alpha^2y^2+1}.$$

(iv) To compute $\mathbf{V}ar[\mathcal{T}_\nu(t)]$ we consider the test function $\phi(s) = \mathbf{1}_{(0,t]}(s)$, and obtain

$$\mathbf{V}ar[\mathcal{T}_\nu(t)] = S_{sg}(t) + t^2\frac{\lambda^2}{B(1,b)}\left[B(1,b-2) - \frac{B(1,b-1)^2}{B(1,b)}\right].$$

Introduce the function $g(\omega)$ which is continuous in a neighbourhood of zero, bounded on $[0,\infty)$:

$$g(\omega) = \frac{\lambda(\alpha^2+\omega^2)}{2\pi B(1,b)}\int_0^1 \frac{x^{b-2}\omega^{3-b}}{\alpha^2x^2+\omega^2}\mathrm{d}x,$$

and $g(\omega) \to g(0) = C_0(\lambda,\alpha,b)$ as $\omega \to 0+$. Then

$$S_{sg}(t) = t^2\int_0^\infty \frac{\sin^2\frac{\omega t}{2}}{\left(\frac{\omega t}{2}\right)^2}\omega^{b-3}g(\omega)\mathrm{d}\omega.$$

By using standard Abelian-like arguments we obtain for $\gamma = b - 3$ and $K(z) = \sin^2 \frac{z}{2} / \left(\frac{z}{2}\right)^2$

$$S(t) = \int_0^\infty K(\omega t)\omega^\gamma g(\omega)\mathrm{d}\omega \approx t^{-\gamma-1} g(0) \int_0^\infty K(z) z^\gamma \mathrm{d}z \tag{22}$$

as $t \to \infty$. In fact, (22) holds for γ such that

$$\int_0^\infty K(z)\,|z|^\gamma\,\mathrm{d}z < \infty. \tag{23}$$

Indeed, we observe that

$$t^{-\gamma-1} \int_0^\infty K(z) z^\gamma \mathrm{d}z = \int_0^\infty K(tz) z^\gamma \mathrm{d}z.$$

Choose $n(t) \to 0$, $tn(t) \to \infty$, as $t \to \infty$, and note that

$$S(t) = t^{-\gamma-1}[g(0) \int_0^\infty K(z) z^\gamma \mathrm{d}z + F(t)],$$

where

$$F(t) = t^{\gamma+1} \int_0^\infty K(\omega t)\omega^\gamma [g(\omega) - g(0)]\mathrm{d}\omega.$$

Next, we use an estimate

$$\begin{aligned} |F(t)| &\le t^{\gamma+1} \int_0^{n(t)} |K(t\omega)|\,|g(\omega) - g(0)|\,\omega^\gamma \mathrm{d}\omega + t^{\gamma+1} \int_{n(t)}^\infty |K(t\omega)|\,|g(\omega) - g(0)|\,\omega^\gamma \mathrm{d}\omega \\ &\le \sup_{0 \le \omega \le n(t)} |g(\omega) - g(0)| \int_0^\infty K(z)\,|z|^\gamma\,\mathrm{d}z + 2 \sup_{0 \le \omega < \infty} |g(\omega)| \int_{tn(t)}^\infty K(z)\,|z|^\gamma\,\mathrm{d}z. \end{aligned}$$

Due to continuity of $g(\omega)$ at $\omega = 0$, $F(t) = o(1)$ as $t \to \infty$. Finally, we use the following well-known integral identity (see, e.g. [1], formula (2.1))

$$J(s) = \int_{\mathbf{R}} \frac{\sin^2 \frac{\omega t}{2}}{\left(\frac{\omega t}{2}\right)^2} |\omega|^s \mathrm{d}\omega = \frac{4\Gamma(1+s)\sin\left(\frac{\pi s}{2}\right)}{s(1-s)}, \quad -1 < s < 0. \tag{24}$$

Note, that $J(0) = 2\pi$. Hence, (24) implies that as $t \to \infty$

$$S_{sg}(t) \approx t^{4-b} g(0) \frac{2\Gamma(b-2)\sin\left(\pi \frac{(3-b)}{2}\right)}{(3-b)(4-b)}.$$

5. Mittag-Leffler (ML) exciting function

Consider the random variable ξ with Mittag-Leffler distribution (MLD), see, i.e., [21]

$$\begin{aligned} &F(x) = \mathbf{P}(\xi \le x) = 1 - E_{p,1}(-\alpha x^p), \quad 0 < p \le 1, \alpha > 0, x > 0, \text{ and} \\ &F(x) = 0, \quad x \le 0, \end{aligned} \tag{25}$$

which is infinitely divisible and has the density

$$f(x) = \frac{\mathrm{d}}{\mathrm{d}x} F(x) = \alpha x^{p-1} E_{p,p}(-\alpha x^p). \tag{26}$$

It is defined in terms of two parameters Mittag-Leffler function:

$$E_{p,q}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(q+pk)}, \quad z \in \mathbf{C}, \mathrm{Re}(p) > 0, \mathrm{Re}(q) > 0. \tag{27}$$

The Mittag-Leffler distribution represents a completely skewed geometrical stable distribution (see, e.g. [26, 27]). Note that for $p = 1$ MLD becomes an exponential distribution.

It is known [21, 26] that for $0 < p < 1, q > 0$ the following asymptotic expansion is valid in a sector around the negative real axis as $z \to -\infty$:

$$E_{p,q}(z) = -\sum_{k=1}^{N} \frac{z^{-k}}{\Gamma(q-pk)} + O\big(|z|^{-N-1}\big). \tag{28}$$

Thus, the density function of Mittag-Leffler distribution

$$f(x) \sim \frac{C}{x^{p+1}}, \quad x \to \infty, \tag{29}$$

where

$$C = -\frac{1}{\alpha \Gamma(-p)} > 0$$

for $0 < p < 1$. The characteristic function of the Mittag-Leffler distribution (see [27], pp. 220–221) is given by

$$\mathbf{E} e^{i\omega\xi} = \frac{1}{1 + \alpha^{-p}|\omega|^p (1 - i\mathrm{sign}(\omega)\tan\frac{\pi p}{2})}, \quad \omega \in \mathbf{R}, 0 < p < 1. \tag{30}$$

Consider now the Hawkes process with the exciting function

$$h(u) = h_\nu(u) = \nu\alpha u^{p-1} E_{p,p}(-\alpha u^p), \quad u > 0, 0 < p < 1 \tag{31}$$

and the random parameter $\zeta = \nu \in (0,1)$. Note that this exiting function has a power-law asymptotic, see (29). Then we obtain

$$\frac{1}{|1 - \tilde{h}_\nu(\omega)|^2} = \frac{\big(1 + \alpha^{-p}|\omega|^p\big)^2 + \alpha^{-2p}|\omega|^{2p}\tan^2\frac{\pi p}{2}}{\big(1 - \nu + \alpha^{-p}|\omega|^p\big)^2 + \alpha^{-2p}|\omega|^{2p}\tan^2\frac{\pi p}{2}}. \tag{32}$$

Assuming again that ν has Beta(a,b)-distribution with density function (11), we obtain

$$\begin{aligned} \mathbf{V}ar\Big(\int_{\mathbf{R}} \phi(t)\mathcal{T}_\zeta(\mathrm{d}t)\Big) &= \int_{\mathbf{R}} \mathrm{d}\omega \mid \tilde{\phi}(\omega) \mid^2 \int_0^1 \hat{\gamma}_u(\omega) f(u)\mathrm{d}u \\ &\quad + \left(\lambda \int_{\mathbf{R}} \phi(t)\mathrm{d}t\right)^2 \mathbf{V}ar\Big[\frac{1}{1-\nu}\Big], \end{aligned} \tag{33}$$

where now

$$\hat{\gamma}_u(\omega) = \frac{\lambda}{2\pi(1-u)} \frac{\left(1+\alpha^{-p}|\omega|^p\right)^2 + \alpha^{-2p}|\omega|^{2p}\tan^2\frac{\pi p}{2}}{\left(1-u+\alpha^{-p}|\omega|^p\right)^2 + \alpha^{-2p}|\omega|^{2p}\tan^2\frac{\pi p}{2}}, \quad \omega \in \mathbf{R} \tag{34}$$

Theorem 2. *Let $\mathcal{T}_\nu$ be a Hawkes process with the Mittag-Leffler exciting function and the random coefficient ν, which follows the Beta PDF as in (11). Then for $a > 0$, $b > 2$*

(i) $\mathbf{E}[\frac{1}{1-\nu}]^2 < \infty$.

(ii) The Bartlett spectral density of a process $\mathcal{T}_\nu$ is of the form

$$\gamma_\nu(\omega) = L_\nu(\omega) + M\delta(\omega), \quad \omega \in \mathbf{R}, \tag{35}$$

where $M\delta(\omega)$ is defined in (16), and

$$L_\nu(\omega) = \frac{\lambda}{2\pi B(a,b)} \left((1+\alpha^{-p}|\omega|^p)^2 + \alpha^{-2p}|\omega|^{2p}\tan^2\frac{\pi p}{2}\right) \int_0^1 \mathrm{d}u \cdot u^{a-1}(1-u)^{b-2}$$
$$\times \left[\left(1-u+\alpha^{-p}|\omega|^p\right)^2 + \alpha^{-2p}|\omega|^{2p}\tan^2\frac{\pi p}{2}\right]^{-1}, \quad \omega \in \mathbf{R}. \tag{36}$$

(iii) For $a = 1, b \in (2,3), 0 < p < 1$, the absolutely continuous component $L_\nu(\omega)$ of Bartlett spectral density (35) has the following singular property as $\omega \to 0+$

$$L_\nu(\omega) \approx \frac{\bar{C}_0}{\omega^{p(3-b)}}, \tag{37}$$

where

$$\bar{C}_0 = \bar{C}_0(\lambda, \alpha, b, p) = \frac{\lambda}{2\pi B(1,b)} \int_0^\infty \frac{y^{b-2}\mathrm{d}y}{y^2 + 2\alpha^{-p}y + \alpha^{-2p}(1+\tan^2\frac{\pi p}{2})}.$$

(iv) For $a = 1, b \in (2,3)$ the first two moments of the Hawkes process $\mathcal{T}_\nu$ have the form

$$\mathbf{E}\mathcal{T}_\nu(t) = t\lambda \frac{B(1,b-1)}{B(1,b)}, \tag{38}$$

$$\mathbf{V}ar[\mathcal{T}_\nu(t)] = Mt^2 + t^2 \int_0^\infty \frac{\sin^2\frac{\omega t}{2}}{\left(\frac{\omega t}{2}\right)^2} L_\nu(\omega)\mathrm{d}\omega$$
$$= S_{sg}(t) + t^2 \frac{\lambda^2}{B(1,b)} \left[B(1,b-2) - \frac{B(1,b-1)^2}{B(1,b)}\right], \tag{39}$$

and, $t \to \infty$,

$$S_{sg}(t) \approx t^{1-p(b-3)} \bar{C}_0 \frac{\Gamma(b-2)\sin\left(\pi\frac{(3-b)}{2}\right)}{(3-b)(4-b)}. \tag{40}$$

The proof of Theorem 2 follows the main ideas of the proof of Theorem 1. So we present the sketch of the proof only. For Beta distribution with $a = 1$, we have

$$L_\nu(\omega) = \frac{\lambda}{2\pi B(1,b)}\Big[\big(1+\alpha^{-p}|\omega|^p\big) + \alpha^{-2p}|\omega|^{2p}\tan^2\frac{\pi p}{2}\Big]I,$$

where

$$\begin{aligned} I &= \int_0^1 x^{b-2}\mathrm{d}x\Big[\big(x+\alpha^{-p}\omega^p\big)^2 + \alpha^{-2p}\omega^{2p}\tan^2\frac{\pi p}{2}\Big]^{-1} \\ &= \omega^{p(b-3)}\int_0^{\omega^{-p}} y^{b-2}\mathrm{d}y\Big[y^2 + 2\alpha^{-p}y + \alpha^{-2p}(1+\tan^2\frac{\pi p}{2})\Big]^{-1}. \end{aligned}$$

When $\omega \to 0+$, $0 < p < 1$

$$\begin{aligned} &\frac{\lambda}{2\pi B(1,b)}\Big[\big(1+\alpha^{-p}|\omega|^p\big) + \alpha^{-2p}|\omega|^{2p}\tan^2\frac{\pi p}{2}\Big] \\ &\qquad\times\int_0^{\omega^{-p}} y^{b-2}\mathrm{d}y\Big[y^2 + 2\alpha^{-p}y + \alpha^{-2p}(1+\tan^2\frac{\pi p}{2})\Big]^{-1} \to \bar{C}_0. \end{aligned}$$

Now

$$S_{sg}(t) = t^2\int_0^\infty \frac{\sin^2(\omega t/2)}{(\omega t/2)^2}\omega^{p(b-3)}\bar{f}(\omega)\mathrm{d}\omega,$$

where

$$\begin{aligned} \bar{f}(\omega) =& \frac{\lambda}{2\pi B(1,b)}\Big[\big(1+\alpha^{-p}|\omega|^p\big) + \alpha^{-2p}|\omega|^{2p}\tan^2\frac{\pi p}{2}\Big] \\ &\times\int_0^1 x^{b-2}\mathrm{d}x\Big[\big(x+\alpha^{-p}\omega^p\big)^2 + \alpha^{-2p}\omega^{2p}\tan^2\frac{\pi p}{2})\Big]^{-1} \\ \to \bar{f}(0) =& \frac{\lambda}{2\pi B(1,b)}\int_0^\infty y^{b-2}\mathrm{d}y\Big[y^2 + 2\alpha^{-p}y + \alpha^{-2p}(1+\tan^2\frac{\pi p}{2})\Big]^{-1}. \end{aligned}$$

The Abelian-like arguments are similar to the proof of Theorem 1.

6. Appendix. An alternative derivation of the LDP rate function

The explicit form of the Large deviation (LD) rate function of linear Hawkes processes was found in [7]. In papers [35] and [37], the rate function of a non-linear Hawkes process is derived. The latter presents an alternative proof in the linear case as well. Here we include one more alternative representation of the rate function in terms of the Lambert function W (see below). Denote the (random) number of points in a cluster generated by a single particle at 0 by S and its radius by L. Define the set

$$\mathcal{D}_S = \{\theta \in \mathbf{R} : \mathbf{E}e^{\theta S} < \infty\}, \tag{41}$$

and its interior part by $\mathcal{D}_S^0$. In [7], the following result is proved

Theorem 3. *Assume that the function $\theta \to \mathbf{E}e^{\theta S}$ is essentially smooth (cf. [16]) and $0 \in \mathcal{D}_S^0$. Moreover, we assume that $\mathbf{E}[Le^{\theta S}] < \infty$ for all $\theta \in \mathcal{D}_S^0$. Then $\mathcal{T}((0,t])/t$ satisfies a LDP on $\mathbf{R}$ with speed t and a good rate function*

$$\Lambda^*(x) = \sup_{\theta \in \mathbf{R}} (\theta x - \Lambda(\theta)), \tag{42}$$

where $\Lambda(\theta) = \lambda(\mathbf{E}[e^{\theta S}] - 1)$.

Remind that the rate function $\Lambda^*(x)$ is good iff the sets $\{\Lambda^*(x) \le a\}$ are compact. In order to find an explicit representation for the rate function, we proceed as follows. Consider the power series $a(x), f(x) \in \mathcal{Q}[[x]]$, the set of formal power series with zero constant term

$$a(x) = \sum_{n \ge 1} a_n x^n, \; f(x) = \sum_{n \ge 1} f_n x^n.$$

Suppose

$$f(a(x)) = x \quad \text{and} \quad a(f(x)) = x.$$

Then the Lagrange inversion formula gives a formula for the coefficients of $a(x)$

$$a(x)\Big|_{x^n} = \frac{1}{n}\left(\frac{x}{f(x)}\right)^n\Big|_{x^{n-1}}. \tag{43}$$

Now take $f(x) = xe^{-x}$, then the defining property of $a(x)$ becomes

$$a(x)e^{-a(x)} = x.$$

Using Lagrange inversion, we get

$$a_n = a(x)\Big|_{x^n} = \frac{1}{n}\left(\frac{x}{f(x)}\right)^n\Big|_{x^{n-1}} = \frac{1}{n}e^{nx}\Big|_{x^{n-1}} = \frac{1}{n}\frac{n^{n-1}}{(n-1)!} = \frac{n^{n-1}}{n!}.$$

This implies the following identity

$$\sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} x^k = a(x), \tag{44}$$

where $a(x)$ is the solution of equation $x = a(x)e^{-a(x)}$ or $a(x) = -W(-x)$, where W is the Lambert function defined in (50).

As follows from [24], Theorem 2.1.12, for the Hawkes processes

$$\mathbf{P}(S = k) = \frac{(k\nu)^{k-1}}{k!} e^{-k\nu}. \tag{45}$$

In view of (44), this means

$$\mathbf{E}\left[e^{\theta S}\right] = \nu^{-1} a(\nu e^{\theta - \nu}), \tag{46}$$

i.e. $\mathbf{E}\left[e^{\theta S}\right] = -\nu^{-1} W_0(-\nu e^{\theta-\nu})$, here we select the principal branch $W_0(x)$. This implies $-e^{-1} \le -\nu e^{\theta-\nu} \le 0$ or $\theta \in \Theta = (-\infty, \nu - \ln \nu - 1)$.

Referring to (42), one obtains

$$\Lambda^*(x) = x\theta_x + \lambda - \frac{\lambda}{\nu} a(\nu e^{\theta_x - \nu}],$$

where

$$a'(\nu e^{\theta_x - \nu}) e^{\theta_x - \nu} = \frac{x}{\lambda}. \tag{47}$$

Differentiation in (47) with respect to θ immediately implies

$$a(\nu e^{\theta_x - \nu}) = \frac{\nu x}{\lambda + \nu x}. \tag{48}$$

So, we obtain the Large Deviation rate function

$$\Lambda^*(x) = x\theta_x + \lambda - \frac{\nu x}{\lambda + \nu x}, \quad x > 0, \tag{49}$$

where θ_x is the solution of (48). It is straightforward that this result agrees with that of [35, 36, 37].

For convenience, we present below basic facts about the Lambert function. In 1758, Lambert introduced the function $W(x)$ satisfying $W(x)e^{W(x)} = x$. If x is real, then for $-\frac{1}{e} \le x < 0$ there are two possible real values of $W(x)$, we use the principle branch $W_0(x)$ which is uniquely specified by conditions $W_0\big(-\frac{1}{e}\big) = -1, W_0(0-) = 0$. In 1877, L. Euler considered a series expansion, while in 1959 E. M. Wright used complex branches of $W(z), z \in C$. The principal branch of W is analytic at 0:

$$W_0(z) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} z^n. \tag{50}$$

Clearly, $W_0\big(-\frac{1}{e}\big) = -1$ but its derivative has a singularity at $z = -\frac{1}{e}$. Thus the radius of convergence cannot exceed $\frac{1}{e}$ and, in fact, it equals $\frac{1}{e}$, as easily seen from the ratio test. Let us compute the derivatives of $W(x)$:

$$\frac{d^n W(x)}{dx^n} = \frac{e^{-nW(x)} p_n(W(x))}{(1 + W(x))^{2n-1}}, \quad n \ge 1, \tag{51}$$

where $p_1(w) = 1$ and

$$p_{n+1}(w) = -(nw + 3n - 1)p_n(w) + (1 + w)p_n'(w), \quad n \ge 1.$$

Acknowledgement

This work has been conducted under Grant 2011/20133-0 provided by the FAPESP. MK expresses his gratitude to NUMEC and IME, Universidade de São Paulo, Brazil, for their warm hospitality. The authors are indebted to the anonymous referee for several stimulating suggestions.

References

[1] K. M. Abadir, W. Distaso, L. Giraitis, *An $I(d)$ model with trend and cycles*, Journ. Econometrics **163**(2011), 186–199.

[2] P. Arby, P. Flandrin, *Point processes, long-range dependence and wavelets*, in: *Wavelets in Medicine and Biology*, (A. Albroubi, M. Unser, Eds.), CRC Press, Boca Raton, FL, 1996, 413–437.

[3] O. E. Barndorff-Nielsen, *Superposition of Ornstein-Uhlenbeck type processes*, Theory Probab. Appl. **45**(2001), 175–194.

[4] O. E. Barndorff-Nielsen, N. Leonenko, *Spectral properties of superposition of Ornstein-Uhlenbeck type processes*, Methodol. Comput. Appl. Probab. **7**(2005), 335–352.

[5] M. S. Bartlett, *The spectral analysis of point processes*, J. R. Stat. Soc. Ser. B **25**(1963), 245–254.

[6] J. Beran, *Statistics for Long Memory Processes*, Chapman and Hall, NY, 1994.

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

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

[9] P. Brémaud, L. Massoulié, *Hawkes branching point processes without ancestors*, J. Appl. Probab. **34**(2001), 205–222.

[10] P. Brémaud, L. Massoulié, *Power spectra of general shot noisees and Hawkes point processes with a random excitation*, Adv. in Appl. Probab. **38**(2002), 122–135.

[11] P. Brémaud, G. Nappo, G. L. Tottisi, *Rate of convergence to equilibrium of marked Hawkes process*, J. Appl. Probab. **39**(2002), 123–136.

[12] D. J. Daley, *The Hurst index of long range dependent processes*, Ann. Probab. **27**(1999), 2035–2041.

[13] D. J. Daley, D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Springer, NY, 2003.

[14] D. J. Daley, R. Versilo, *Long range dependence of point processes, with queueing examples*, Stochastic Process. Appl. **70**(1997), 265–282.

[15] D. J. Daley, T. Rolski, R. Versilo, *Long range dependent of point processes and their Palm-Khinchine distributions*, Adv. in Appl. Probab. **32**(2000), 1051–1063.

[16] A. Dembo, O. Zeitouni, *Large Deviation Technique and Applications*, Springer, NY, 1998.

[17] P. Doukhan, G. Oppenheim, M. S. Taqqu, *Theory and Applications of Long-Range Dependence*, Birkhäuser, Boston, 2003.

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

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

[20] C. C. Heyde, N. N. Leonenko, *Student processes*, Adv. in Appl. Probab. **37**(2005), 342–365.

[21] N. Haubold, A. M. Mathai, R. K. Saxena, *Mittag-Leffler functions and their applications*, J. Appl. Math. **2011**, Art. ID 298628, 51 pp.

[22] B. González-Arévalo, G. Samorodnitsky, *Buffer content of a leaky-bucket system with long range dependent input traffic*, J. Appl. Probab. **40**(2003), 581–601.

[23] C. W. J. Granger, *Long-memory relationships and the aggregation of dynamic models*, J. Econometrics **14**(1980), 227–238.

[24] P. Jagers, *Branching Processes with Biological Applications*, J. Wiley, London, 1975.

[25] M. Kelbert, N. Leonenko, M. D. Ruiz-Medina, *Fractional random fields associated with stochastic fractional heat equation*, Adv. in Appl. Probab. **37**(2005), 108–133.

[26] T. J. Kozubowski, M. M. Meershaert, *A bivariate infinitely divisible distribution with exponential and Mittag-Leffler marginals*, Statist. Probab. Lett. **79**(2009), 1596–1601.

[27] S. Kotz, T. J. Kozubowski, K. Podgoródski, *The Laplace Distributions and Generalizations*, Springer, NY, 2001.

[28] R. Kulik, R. Szekl, *Sufficient conditions for long-range dependence of stationary point processes on the real line*, J. Appl. Probab. **38**(2001), 570–581.

[29] N. N. Leonenko, M. D. Ruiz-Medina, M. S. Taqqu, *Fractional elliptic, hyperbolic and parabolic random fields*, Electron. J. Probab. **16**(2011), 1134–1172.

[30] J. Moller, J. G. Rasmussen, *Perfect simulation of Hawkes processes*, Adv. in Appl. Probab. **37**(2005), 629–646.

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

[32] G. Oppenheim, M.-C. Viano, *Aggregation of random parameters Ornstein-Uhlenbeck or AR processes: some convergence results*, J. Time Ser. Anal. **25**(1994), 335–350.

[33] F. Ramires-Peres, R. Serfing, *Short noise on cluster processes with cluster models, and studies of long range dependence*, Adv. in Appl. Probab. **33**(2001), 631–657.

[34] G. Samorodnitsky, *Extreme value theory, ergodic theory and the boundary between short memory and long memory for stationary stable processes*, Ann. Probab. **32**(2004), 1438–1468.

[35] L. Zhu, *Moderate deviations for Hawkes processes*, Statist. Probab. Lett. **83**(2013), 885–890.

[36] L. Zhu, *Large deviations for Markovian non-linear Hawkes processes*, arXiv:1108.2432 [math.Pr].

[37] L. Zhu, *Process level Large deviations for non-linear Hawkes point processes*, arXiv:1108.2431 [math.Pr].