HAL open science

Tail asymptotics and precise large deviations for some Poisson cluster processes

Fabien Baeriswyl, Valérie Chavez-Demoulin, Olivier Wintenberger

To cite this version:

Fabien Baeriswyl, Valérie Chavez-Demoulin, Olivier Wintenberger. Tail asymptotics and precise large deviations for some Poisson cluster processes. 2024. hal-04071286v3

HAL Id: hal-04071286

https://hal.science/hal-04071286v3

Preprint submitted on 4 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Applied Probability Trust (4 February 2024)

TAIL ASYMPTOTICS AND PRECISE LARGE DEVIATIONS FOR SOME POISSON CLUSTER PROCESSES

FABIEN BAERISWYL,* *Université de Lausanne and Sorbonne Université*

VALÉRIE CHAVEZ-DEMOULIN,** *Université de Lausanne*

OLIVIER WINTENBERGER,*** *Sorbonne Université*

Abstract

We study the tail asymptotics of two functionals (the maximum and the sum of the marks) of a generic cluster in two sub-models of the marked Poisson cluster process, namely the renewal Poisson cluster process and the Hawkes process. Under the hypothesis that the governing components of the processes are regularly varying, we extend results due to [19] and [6] notably, relying on Karamata's Tauberian Theorem to do so. We use these asymptotics to derive precise large deviation results in the fashion of [32] for the above-mentioned processes.

Keywords: Renewal Poisson cluster process; Hawkes process; random maxima; random sums.

2020 Mathematics Subject Classification: Primary 60G70

Secondary 60G55; 60F10

1. Introduction

In this paper, we study the asymptotic properties of processes exhibiting clustering behaviour. Such processes are common in applications: for instance, earthquakes in seismology, where a main shock has

* Postal address: Département des Opérations, Anthropole, CH-1015 Lausanne, Suisse

* Email address: fabien.baeriswyl@unil.ch, fabien.baeriswyl@sorbonne-universite.fr

** Postal address: Département des Opérations, Anthropole, CH-1015 Lausanne, Suisse

*** Postal address: Laboratoire de Probabilités, Statistique et Modélisation, Sorbonne Université, Campus Pierre et Marie Curie, 4 place Jussieu, 75005 Paris, France

the ability to trigger a series of secondary shocks in a specific spatio-temporal neighbourhood; but also accidents giving rise to a series of subsequent claims in non-life insurance or heavy rainfall in meteorology to name a few. We will focus on two different processes that have effectively been used in these fields. The Hawkes process has been introduced in the pioneer works of [60] and [47], and has found applications in earthquake modeling (see e.g. [41]), in finance (see e.g. [9], [22]), in genome analysis (see [52]) or in insurance (see [56]). The renewal Poisson cluster process is a tool of choice in an insurance context for modelling series of claims arising from a single event (see e.g. [38] for a reference textbook), as well as in teletraffic modelling (see [19]) and in meteorology and weather forecast (see e.g. [20] or [49]).

The above processes, described heuristically and in specific contexts above, are part of the class of the so-called point processes: for a comprehensive overview, see the monographs of [13] and [14] or, more recently, and with connection to martingale theory, see [8]. Point process theory is an elegant framework describing the properties of random points occurring in general spaces. In both cases the temporal marked point process N possesses a representation as an infinite sum of Dirac measures (recall that the Dirac measure ε on $\mathcal{A}$ satisfies for every $A \in \mathcal{A}$ that $\varepsilon_x(A) = 1$ if $x \in A$ and $\varepsilon_x(A) = 0$ otherwise):

$$N(\cdot) = \sum_{i=1}^{\infty} \varepsilon_{T_i, A_i}(\cdot)$$

where T_i is the (random) time of occurrence of the ith event and A_i is its associated mark. The specific temporal marked point processes that we are interested in are cluster point processes. More specifically, we will assume that there exists an immigration process, under which independent points arise at a Poissonian rate; then, each of these immigrant events has the ability to trigger new points, called first generation offspring events. We will then look at two submodels. One is the renewal Poisson cluster process. It is complete with the immigrant events and their first generation offsprings. The term "renewal" comes from the fact that the times of the events form a renewal sequence. The other submodel is the Hawkes process in which every point of the first generation has the ability to generate new points, acting as an immigrant event, potentially generating therefore a whole cascade of points. Each immigrant event and its associated offspring events (whether direct children or indirect) form a generic cluster.

We will study the tail asymptotics of the partial maxima and sums of a transformation $X = f(A)$, for some nonnegative real valued function f, of the mark A of any event of N. Determining the behaviour of the maximum and the sum at the level of the cluster decomposition of a process is crucial to obtain limit theorems for partial maxima and sums of the whole process over finite intervals, see e.g. [55], [29] or [6]. Thus, we describe first a generic cluster from each of the above-mentioned processes.

For the renewal Poisson cluster process, we will consider a distributional representation of the maximum

of the marks in the generic cluster, denoted H^R,

$$H^R \stackrel{\mathrm{D}}{=} X \vee \bigvee_{j=1}^{K_A} X_j$$

where X is a transformation $f(A)$ of the mark A of the immigrant event, and X_j is the transformed mark of the jth first-generation offspring event. The number of offspring events, K_A, is random and possibly dependent on X. In particular, we will let the vector (X, K_A) be heavy-tailed, and assess whether the heavy-tailedness transfers to H^R. Details are relegated to Section 2. Note that under the hypothesis that X and K_A are independent the above distributional equation has received early consideration, e.g. in [59] or [28], where it is shown that H^R and X belong to the same maximum domain of attraction of some extreme value distribution (MDA for short - see [50], [15] or [17] for references on extreme value theory). A more recent advance in the case where X and K_A are dependent is to be found in [5], where a similar conclusion is reached about the MDA. Our emphasis is on the Fréchet MDA, which allows a certain refinement on the characterisation of the tail asymptotics.

We will also consider tail asymptotics for the sum functional, which for the very same renewal Poisson cluster process, and for a generic cluster, possesses the distributional representation

$$D^R \stackrel{\mathrm{D}}{=} X + \sum_{j=1}^{K_A} X_j$$

supposing again that (X, K_A) is heavy-tailed, We will also assess whether the heavy-tailedness of (X, K_A) transfers to D^R. This equation has received consideration under the hypothesis that X and K_A are independent, see [19]. We will retrieve their results in our framework. More recently, in the case of arbitrary dependence between X and K_A, similar asymptotics have been derived in [48].

We will then derive the very same kind of tail asymptotics, for the very same functionals of a generic cluster in the context of the Hawkes process. The distributional representation associated with the maximum of the marks in a generic cluster, denoted H^H, is given by

$$H^H \stackrel{\mathrm{D}}{=} X \vee \bigvee_{j=1}^{L_A} H_j^H$$

where L_A is the number of first-generation offspring events A of the event acting as immigrant, and H_j is the maximum of the marks of the offsprings of the jth offspring of the immigrant event considered, itself acting as immigrant for further subranches of the cluster, emphasising once again the cascade structure of the Hawkes process. The equation for H^H above is a special case of the higher-order Lindley equation, see [31]. Note that $X = f(A)$ and L_A are dependent through A. Letting L_A be Poisson distributed with

parameter κ_A, and (X, κ_A) be heavy-tailed we assess whether this transfers to H^H. This functional has received attention in the recent work of [5], where it was shown that H^H has the same MDA as that of X.

The distributional representation associated with the sum of the marks in a generic cluster in the Hawkes process, denoted D^H, is given by

$$D^H \stackrel{\mathrm{D}}{=} X + \sum_{j=1}^{L_A} D_j^H.$$

We will again let (X, κ_A) be heavy-tailed, and assess whether this transfers to D^H. This distributional equation, with cascade structure, has been extensively studied: see e.g. [4]; but also, as a main stochastic modelling approach to Google's PageRank algorithm, see [34], [27], [61], [10], [11] and, even more closely related to our results, [48]; in the context of random networks, see [37] or [36]; for a recent, theoretical advance as well as application to queuing systems, see [1] or [18].

The way we will deal with heavy-tailedness is through the classical notion of regular variation, introduced by J. Karamata in the 20th century (see e.g. [30]), which specifies that the functions of interest behave, in a neighbourhood of infinity, like power-law functions. For a thorough, textbook treatment of the topic in univariate settings, see [7]; we rely on [50], [51], [15] and [40] for the multivariate case.

The flexibility offered by our approach to the way we specify the regular variation of the governing components of our processes allows us, in the sequel, to extend results due to [53], [19], [26] or [16], that all studied the asymptotics of the tail of distributional quantities such as H and D in the above examples, but under various assumptions on the relations of the tails of X and K_A for the renewal Poisson cluster process, respectively X and L_A for the Hawkes process.

Finally, we use the results on the tails of H and D to derive (precise) large deviation principles for our processes of interest, in the flavour of [44], [39]. The "precise" terminology comes from the fact that we have exact asymptotic equivalence instead of logarithmic ones when assuming Cramér's condition. Early results on precise large deviations in the case of non-random maxima and sums can be found in [42], [43], [24] and [12]. The case of random maxima and sums of extended regularly varying random variables (a class containing regularly varying random variables) is to be found in [32], and we will rely on their results to derive our very own precise large deviation results. Contributions in this area for another subclass of subexponential distributions, namely the class of consistently varying random variables, can be found in [58] or [46]; for precise large deviations results on (negatively) dependent sequences, see [57] or [35].

The organisation of the paper is as follows: in Section 2, we describe the main processes of interest, that are part of the Poisson cluster process family; in Section 3, we recall some important notions and characterisations of (multivariate) regular variation; in Section 4, we derive the tail asymptotics for the

maximum of the marks in a generic cluster in the renewal Poisson cluster process; in Section 5, we derive the tail asymptotics for the sum of the marks in a generic cluster in the renewal Poisson cluster process; in Section 6, we derive the tail asymptotics for the maximum of the marks in a generic cluster in the Hawkes process; in Section 7, we derive the tail asymptotics for the sum of the marks in a generic cluster in the Hawkes process; in Section 8, we use the results from Section 4 to Section 7 to derive (precise) large deviations results for our processes of interest.

Notation

Vectors are usually in boldface. By "i.i.d." we classically mean independent and identically distributed and, consistently, "i.d." means identically distributed. We let $\lceil\cdot\rceil$ denote the upper integer part, $\lfloor\cdot\rfloor$ the lower integer part. For two functions $f(\cdot)$ and $g(\cdot)$, and $c \in \{0, \infty\}$, we note $f(x) = \mathcal{O}(g(x))$, as $x \to c$ whenever $\limsup_{x\to c}|f(x)/g(x)| \leqslant M$, for some finite $M > 0$; $f(x) = o(g(x))$, as $x \to c$ whenever $\lim_{x\to c}|f(x)/g(x)| = 0$; $f(x) \sim g(x)$, as $x \to c$ whenever $\lim_{x\to c} f(x)/g(x) = 1$. The product of two measures μ and ν is written as the tensor product $\mu \otimes \nu$.

2. Random functionals of clusters

We formally introduce the general Poisson cluster process, a class which includes the processes discussed in Section 1, keeping the spirit of the presentation and (most) notations from [6]. As hinted in Section 1, this process is made up of two components: an immigration process and an offspring process.

The immigration process, say N_0, is a marked homogeneous Poisson process (or marked PRM in short, for marked Poisson random measure) with representation given by:

$$N_0(\cdot) := \sum_{i=1}^{\infty} \varepsilon_{\Gamma_i, A_{i0}}(\cdot).$$

This point process has mean measure $\nu\text{Leb}\otimes F$, for $\nu > 0$, on the space $[0, \infty)\times\mathbb{A}$, where Leb is the Lebesgue measure, F is the common distribution function to all marks $(A_{i0})_{i\in\mathbb{N}}$, which take values on a measurable space $(\mathbb{A}, \mathcal{A})$, and where $\mathcal{A}$ corresponds to the Borel σ-field on $\mathbb{A}$. In particular, this means that the sequence of times $(\Gamma_i)_{i\in\mathbb{N}}$, corresponding to the arrivals of immigrant events, is a homogeneous Poisson process with rate given by νLeb. Since the space $\mathbb{A}$ can be quite general, applying a transformation $f(\cdot) : \mathbb{A} \to \mathbb{R}_+$ is natural, especially in practical applications. Note that we will also assume this transformation of the marks, i.e. we only consider nonnegative transformed marks in our models. For example, in a non-life insurance context, supposing that A_{i0} represents the characteristics of the ith accident, $f(A_{i0})$ could

represent the claim size pertaining to this accident. In subsequent sections, and to ease the notation, we shall denote $X_{i0} := f(A_{i0})$.

Conditioning on observing an immigration event at time Γ_i, the marked PRM N_0 is supplemented with an additional point process in $M_p([0,\infty) \times \mathbb{A})$ (the space of locally finite point measures on $[0,\infty) \times \mathbb{A}$) that we denote by $G_{A_{i0}}$. The cluster of points $G_{A_{i0}}$, occurring after time Γ_i, augments N_0 with triggered, offspring points or events.

The offspring cluster process, conditioned on observing an immigrant event (Γ_i, A_{i0}), admits the representation

$$G_{A_{i0}}(\cdot) := \sum_{j=1}^{K_{A_{i0}}} \varepsilon_{T_{ij}, A_{ij}}(\cdot)$$

where $(T_{ij})_{1 \leqslant j \leqslant K_{A_{i0}}}$ forms a sequence of nonnegative random variables indicating, for a fixed j, the random time from the immigrant event occurring at time Γ_i and the jth event of the cluster, and where $K_{A_{i0}}$ is a random variable with values in $\mathbb{N}_0$, corresponding to the number of events in the ith cluster. These events are the offspring of the immigrant event identified by (Γ_i, A_{i0}). A complete representation of the general Poisson cluster process is given by

$$N(\cdot) := \sum_{i=1}^{\infty} \sum_{j=0}^{K_{A_{i0}}} \varepsilon_{\Gamma_i + T_{ij}, A_{ij}}(\cdot)$$

providing we set $T_{i0} = 0$ for all $i \in \mathbb{N}$.

The first functional of interest is the maximum of the marks in the ith cluster, defined by

$$H_i := \bigvee_{j=0}^{K_{A_{i0}}} X_{ij}. \tag{1}$$

Above, for ease of notation, we have defined $X_{i0} = f(A_{i0})$; accordingly, we let $X_{ij} = f(A_{ij})$ for the transformation $f(\cdot) : \mathbb{A} \to \mathbb{R}_+$. The point process associated with the ith cluster is defined by

$$C_i(\cdot) := \varepsilon_{0, A_{i0}}(\cdot) + G_{A_{i0}}(\cdot).$$

It allows us to define the second functional of interest in this paper, namely the sum of all marks in the ith cluster, by

$$D_i := \int_{[0,\infty) \times \mathbb{A}} f(a) C_i(\mathrm{d}t, \mathrm{d}a). \tag{2}$$

In Section 8, we will look at the whole process on a subset of the temporal axis: at the level of the point process N, the sum of all marks in the finite time interval $[0, T]$, for $T > 0$, is given by

$$S_T := \int_{[0,T] \times \mathbb{A}} f(a) N(\mathrm{d}t, \mathrm{d}a). \tag{3}$$

From Section 4 to Section 7, we propose tail asymptotics for H_i and D_i in the settings of mainly two different submodels of the general Poisson cluster process, briefly described in the introduction, that we formally discuss next, keeping the presentation in [6], but fully described in Example 6.3 of [13]. However, we refer to the former reference for a complete description. In our work, we also assume that the sequence of marks (X_{ij}) is i.i.d.

2.1. Mixed binomial Poisson cluster process

In this model, the assumptions on N_0 are kept unchanged and the ith cluster has a representation of the form

$$C_i(\cdot) = \varepsilon_{0,A_{i0}}(\cdot) + G_{A_{i0}}(\cdot) = \varepsilon_{0,A_{i0}}(\cdot) + \sum_{j=1}^{K_{A_{i0}}} \varepsilon_{W_{ij},A_{ij}}(\cdot)$$

where $(K_{A_{i0}}, (W_{ij})_{j\geqslant 1}, (A_{ij})_{j\geqslant 0})_{i\geqslant 0}$ is an i.i.d. sequence, the sequence $(A_{ij})_{j\geqslant 0}$ is also i.i.d. for any fixed $i = 1, 2, \ldots$ and, finally, $(A_{ij})_{j\geqslant 1}$ is independent of both $K_{A_{i0}}$ and $(W_{ij})_{j\geqslant 1}$ for any $i = 1, 2, \ldots$. Note that this latter statement does not exclude dependence between A_{i0} and $K_{A_{i0}}$ (respectively $(W_{ij})_{j\geqslant 1}$). Additionally, it is assumed that $\mathbb{E}[K_A] < \infty$, where K_A denotes a generic random quantity distributed as $K_{A_{i0}}$.

2.2. Renewal Poisson cluster process

In this model, the ith cluster has the representation

$$C_i(\cdot) = \varepsilon_{0,A_{i0}}(\cdot) + G_{A_{i0}}(\cdot) = \varepsilon_{0,A_{i0}}(\cdot) + \sum_{j=1}^{K_{A_{i0}}} \varepsilon_{T_{ij},A_{ij}}(\cdot) \tag{4}$$

where all the assumptions from Section 2.1 hold, except that now, we denote the occurrence time sequence of the offspring events by $(T_{ij})_{j\geqslant 1}$ to emphasise that this forms a renewal sequence, that is, for any fixed $i = 1, 2, \ldots$, $T_{ij} = W_{i1} + \cdots + W_{ij}$. Note that this process is such that every Poisson immigrant has only $K_{A_{i0}}$ first generation offspring events. These points cannot generate further generations themselves, in contrast with the Hawkes process that we will introduce next.

Applying the transformation f on the marks of the events, we will, in Section 4 and Section 5, derive tail the asymptotics of generic versions of Equation (1) and Equation (2), given by:

1. for the maximum,

$$H^R \stackrel{\mathrm{D}}{=} X \vee \bigvee_{j=1}^{K_A} X_j; \tag{5}$$

2. for the sum,

$$D^R \stackrel{\mathrm{D}}{=} X + \sum_{j=1}^{K_A} X_j. \tag{6}$$

We isolate $X := f(A)$ from the rest of the transformed claims $(X_j) := \big(f(A_j)\big)$, to emphasise the possible dependence between X and K_A.

Remark 2.1. These two processes have been considered in the monograph [38]. The mixed binomial Poisson cluster process and the renewal Poisson cluster process are very similar in their description, and because their sole difference is the placement of the points along the time axis, we focus - in what follows - on the renewal Poisson cluster process. The results of Section 4 and Section 5 are directly applicable to the mixed binomial Poisson cluster process; the results of Section 8 also apply, upon the use of an alternative justification regarding the left-over effects to be discussed in that section. We refer to [6] and [5] for justifications.

2.3. Hawkes process

The specificity of the Hawkes process is that the clusters have a recursive pattern, in the sense that each point, whether immigrant or offspring, has the ability to act as an immigrant and generate a new cluster. To obtain the representation of the ith cluster G_{A_i}, one typically introduces a time shift operator θ_t, as in [6]. Let $m(\cdot) = \sum_{j=1}^{\infty} \varepsilon_{t_j,a_j}(\cdot)$ be a point measure: then, the time-shift operator is defined by

$$\theta_t m(\cdot) = \sum_{j=1}^{\infty} \varepsilon_{t_j+t,a_j}(\cdot)$$

for all $t \geqslant 0$. Then, the (recursive) representation of the ith cluster, conditioning on observing an immigration event (Γ_i, A_{i0}), is given by

$$C_i(\cdot) = \varepsilon_{0,A_{i0}}(\cdot) + G_{A_{i0}}(\cdot) = \varepsilon_{0,A_{i0}}(\cdot) + \sum_{j=1}^{L_{A_{i0}}} \Big(\varepsilon_{\tau^1_{ij},A^1_{ij}}(\cdot) + \theta_{\tau^1_{ij}} G_{A^1_{ij}}(\cdot)\Big)$$

where, given A_{i0}, the first-generation offspring process $N_{A_{i0}}(\cdot) := \sum_{j=1}^{L_{A_{i0}}} \varepsilon_{\tau^1_{ij},A^1_{ij}}(\cdot)$ is again a Poisson process, this time with (random) mean measure $\int h(s, A_{i0})\,\mathrm{d}s \otimes F$, and where the sequence $(G_{A^1_{ij}})_{j\geqslant 1}$ is i.i.d. and independent of the first-generation offspring process $N_{A_{i0}}$. Note that the sequence of times in the cluster representation $G_{A_{i0}}$, hereby denoted as (τ_{ij}), is the sequence of times of the first-generation offspring events. The function $h(\cdot)$ is referred to as the fertility function and controls both the displacement and the expected number of offspring(s) of a specific event. Hence, by definition, the number of first generation offspring events is Poisson and depends on the mark of the event acting as an immigrant to the stream

of points considered. Note that the above representation also emphasises the independence between the subclusters considered at any point, from the immigrant perspective. There is a connection with Galton-Watson theory that was historically used to show that the Hawkes process is a general Poisson cluster process (see [23]); we define it as part of this family, but the Hawkes process is classically introduced from the self-excitation perspective, that is, from the specification of the function $h(\cdot)$ (see e.g. [21]).

We propose in Section 6 and Section 7 tail asymptotics for the generic versions of Equation (2) and Equation (1), which satisfy, in the settings of the Hawkes process, fixed-point distributional equations of the form:

1. for the maximum,
$$H^H \overset{\mathrm{D}}{=} X \vee \bigvee_{j=1}^{L_A} H_j^H; \tag{7}$$
2. for the sum,
$$D^H \overset{\mathrm{D}}{=} X + \sum_{j=1}^{L_A} D_j^H; \tag{8}$$

where $L_A|A \sim \text{Poisson}(\kappa_A)$ and $\kappa_A = \int_{(0,\infty)} h(t,A)\,\mathrm{d}t$ and where (H_j^H) and (D_j^H) are i.i.d. copies of H^H and D^H, respectively. In this work, we always assume the subcriticality condition (in the terminology of branching processes) $\mathbb{E}[\kappa_A] < 1$, in order for clusters to be almost surely finite. This also implies that the expected total number of points in a cluster is given by $\frac{1}{1-\mathbb{E}[\kappa_A]}$, using a geometric series argument (see Chapter 12 in [8]). As pointed out in [1] and references therein, the combination of the subcriticality assumption, the fact that the random quantities involved in Equation (8) are nonnegative and the assumption that $\mathbb{E}[X] < \infty$ (to be made through the index of regular variation of X in further sections) yields the existence and uniqueness of a nonnegative solution to this distributional equation; for Equation (7), a discussion about the existence of potentially multiple solutions to the higher-order Lindley equation can be found in [2]. Lastly, note that Equation (7) and Equation (8) emphasise the cascade structure of the Hawkes process.

3. A word on regular variation

Throughout this paper, we will assume that the governing random components of our processes of interest are regularly varying, that is, roughly speaking, exhibit heavy tails. More specifically, we will assume that the random vector $\mathbf{X}$ is regularly varying. For the renewal Poisson cluster process, this amounts to assume that $\mathbf{X} = (X, K_A)$ is regularly varying, where X and K_A are defined as in Section 2.2;

for the Hawkes process, this amounts to assume that $\mathbf{X} = (X, \kappa_A)$ is regularly varying, where X and κ_A are defined in Section 2.3. The exact definition of regular variation varies in the literature depending on the context (see e.g. [50], [51], [15], [25], [54]). Hence, we first recall the definition of regular variation we use in this text in full generality, borrowing notations from [40]. We let $\mathbb{R}^d_{\mathbf{0}} = \mathbb{R}^d \backslash \{\mathbf{0}\}$ with $\mathbf{0} = (0, 0, \ldots, 0)$. We let $|\cdot|$ be any norm on $\mathbb{R}^d$ (by their equivalence). Note that, in subsequent sections, our framework is restricted to the case where $d = 2$.

Definition 3.1. Let $\mathbf{X}$ be a random vector with values in $\mathbb{R}^d$. Suppose that $|\mathbf{X}|$ is regularly varying with index $\alpha > 0$. Let (a_n) be a real sequence satisfying $n\mathbb{P}(|\mathbf{X}| > a_n) \to 1$, as $n \to \infty$. The random vector $\mathbf{X}$ (and its distribution) are said to be regularly varying if there exists a non-null Radon measure μ on the Borel σ-field of $\mathbb{R}^d_{\mathbf{0}}$ such that, for every μ-continuity set A, it holds that

$$\mu_n(A) := n\mathbb{P}\left(a_n^{-1}\mathbf{X} \in A\right) \to \mu(A), \text{ as } n \to \infty.$$

In the above definition, two remarks are in order:

1. the regular variation of $|\mathbf{X}|$ is univariate; standard definition applies, namely that the distribution of $|\mathbf{X}|$ has power-law tails, that is, $\mathbb{P}(|\mathbf{X}| > x) = x^{-\alpha}L(x)$ for $x > 0$, where $L(\cdot)$ is a slowly varying function;

2. the kind of convergence that takes place is vague convergence. The limiting measure possesses various nice properties, among which one can cite homogeneity: for any Borel set $B \subset \mathbb{R}^d_{\mathbf{0}}$ and $t > 0$, it holds that $\mu(tB) = t^{-\alpha}\mu(B)$.

Rather than using the sequential form as in Definition (3.1), it is possible to use an alternative continuous form. Additionally, a distinguished characterisation in the literature is through a limiting decomposition into "spectral" and 'radial" parts, see [51].

Proposition 3.1. (Theorem 6.1 in [51].) *A random vector $\mathbf{X}$ with values in $\mathbb{R}^d$ is regularly varying with index $\alpha > 0$ and non-null Radon measure μ on $\mathbb{R}^d_{\mathbf{0}}$ if and only if one of the following relations holds:*

1. *(Continuous form): The random variable $|\mathbf{X}|$ is regularly varying with index $\alpha > 0$ and*

$$\frac{\mathbb{P}\left(x^{-1}\mathbf{X} \in \cdot\right)}{\mathbb{P}(|\mathbf{X}| > x)} \xrightarrow{\text{v}} \mu(\cdot), \text{ as } x \to \infty.$$

2. *(Weak convergence to independent radial/spectral decomposition): the following limit holds*

$$\mathbb{P}\left(\left(\frac{\mathbf{X}}{x}, \frac{\mathbf{X}}{|\mathbf{X}|}\right) \in \cdot\right) \xrightarrow{\text{w}} \mathbb{P}\left((Y, \mathbf{\Theta}) \in \cdot\right), \text{ as } x \to \infty$$

where $Y \sim$ Pareto(α) with $\alpha > 0$ and is independent of $\mathbf{\Theta}$, which takes values on the unit sphere $\mathbb{S}^{d-1}$ defined by $\mathbb{S}^{d-1} = \{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| = 1\}$.

In Proposition (3.1), the notation $\xrightarrow{\text{v}}$ refers to vague convergence: we say that a sequence of measures (μ_n) (with $\mu_n \in M_+(E)$, the space of nonnegative Radon measure on $(E, \mathcal{E})$) converges vaguely to a measure $\mu \in M_+(E)$ if for all functions $f \in \mathbb{C}_K^+(E)$, we have $\int_E f(x)\mu_n(\mathrm{d}x) \to \int_E f(x)\mu(\mathrm{d}x)$, where $C_K^+(E)$ denotes the set of functions $f : E \to \mathbb{R}_+$ being continuous with compact support. For more details about vague convergence, see e.g. Chapter 3 in [51]. The notation $\xrightarrow{\text{w}}$ refers to the standard notion of weak convergence. The above characterisations have various consequences. The first property is a continuous mapping theorem, first proved in [25] in the framework of metric spaces. We use a simplified version fitting our settings, which we partially reproduce, from [40]. See also Proposition 4.3 and Corollary 4.2 in [33].

Proposition 3.2. (Theorem 2.2.30 in [40], Proposition 4.3 and Corollary 4.2 in [33].) *Let $\mathbf{X}$ be a random vector in $\mathbb{R}^d$ and suppose it is regularly varying with index $\alpha > 0$ and non-null Radon measure μ on $\mathbb{R}^d_{\mathbf{0}}$. Let $g(\cdot) : \mathbb{R}^d \to \mathbb{R}$ be a non-zero, continuous and positively homogeneous map of order γ, i.e. for every $\mathbf{x} \in \mathbb{R}^d$, $g(t\mathbf{x}) = t^\gamma g(\mathbf{x})$ for some $\gamma > 0$. Then, the following limit relation holds*

$$\frac{\mathbb{P}\left(x^{-1}g(\mathbf{X}) \in \cdot\right)}{\mathbb{P}\left(|\mathbf{X}|^\gamma > x\right)} \xrightarrow{\text{v}} \mu(g^{-1}(\cdot)), \text{ as } x \to \infty.$$

Note that for every $\epsilon > 0$, $\mu(g^{-1}(\{x \in \mathbb{R} : |x| > \epsilon\})) < \infty$. Moreover, if $\mu(g^{-1}(\cdot))$ is not the null measure on $\mathbb{R}_0$, then $g(\mathbf{X})$ is regularly varying with index α/γ and with non-null Radon measure

$$\frac{\mu(g^{-1}(\cdot))}{\mu(g^{-1}(\{x \in \mathbb{R} : |x| > 1\}))}.$$

Example 3.1. It is easily seen that the map defined by the projection on any coordinate of $\mathbf{X}$ is a continuous mapping satisfying the assumptions of Proposition (3.2) with $\gamma = 1$. If $d = 2$, $\mathbf{X} = (X_1, X_2)$ and $g(\mathbf{X}) := X_1$, then by the homogeneity property of the limiting Radon measure μ, as long as

$$\mu(\{(x_1, x_2) \in \mathbb{R}^2_{\mathbf{0}} : x_1 > 1\}) > 0$$

one obtains regular variation of X_1 with index $\alpha > 0$.

A second useful result, due to [54] again in the setting of metric spaces that we simplify here, shows that one can actually replace the norm $|\cdot|$ by any modulus. A modulus, as defined in Definition 2.2 of [54], is a function $\rho : \mathbb{R}^d \to [0, \infty)$ such that $\rho(\cdot)$ is non-zero, continuous and positively homogeneous of order 1. Proposition 3.1 in [54] then ensures the following.

Proposition 3.3. (Proposition 3.1 in [54].) *A random vector* $\mathbf{X}$ *with values in* $\mathbb{R}^d$ *is regularly varying with index* $\alpha > 0$ *and non-null Radon measure* μ *on* $\mathbb{R}^d_{\mathbf{0}}$ *if and only if there exists a modulus* ρ *such that* $\rho(\mathbf{X})$ *is regularly varying with index* $\alpha > 0$*, and a random vector* $\mathbf{\Theta}$ *taking values on* $\mathbb{S}^{d-1} := \{\mathbf{x} \in \mathbb{R}^d : \rho(\mathbf{x}) = 1\}$ *such that*

$$\mathbb{P}\left(\frac{\mathbf{X}}{\rho(\mathbf{X})} \in \cdot \;\middle|\; \rho(\mathbf{X}) > x\right) \xrightarrow{\mathrm{w}} \mathbb{P}\left(\mathbf{\Theta} \in \cdot\right), \text{ as } x \to \infty.$$

Finally, in subsequent sections, we shall also use an other characterisation via the regular variation of linear combinations, proven by [3]. We denote the inner product in $\mathbb{R}^d$ by $\langle \cdot, \cdot \rangle$.

Proposition 3.4. (Proposition 1.1 in [3].) *A random vector* $\mathbf{X}$ *with values in* $\mathbb{R}^d$ *is regularly varying with noninteger index* $\alpha > 0$ *if and only if there exists a slowly varying function* $L(\cdot)$ *such that, for all* $\mathbf{t} \in \mathbb{R}^d$*,*

$$\lim_{x \to \infty} \frac{\mathbb{P}\left(\langle \mathbf{t}, \mathbf{X} \rangle > x\right)}{x^{-\alpha} L(x)} = w(\mathbf{t}) \text{ exists,}$$

for some function $w(\cdot)$ *and there exists one* $\mathbf{t}_0 \neq 0$ *such that* $w(\mathbf{t}_0) > 0$.

The above result states that a random vector $\mathbf{X}$ is regularly varying with index $\alpha > 0$ if and only if all linear combinations of its components are regularly varying with the same index $\alpha > 0$. Note that it is not necessary for α in Proposition (3.4) to be noninteger for the above equivalence to hold; however, when this is not the case, there are some caveats that we avoid considering in our the results of upcoming sections (e.g., with α noninteger, we do not have to consider $t \in \mathbb{R}^d$ but rather $t \in \mathbb{R}^d_+$), see [3].

Finally, the last result of great importance in showing the transfer of regular variation in the subsequent sections is Karamata's Theorem, which can be found as Theorem 8.1.6 in [7]. Let X be a random variable, denote its associated Laplace-Stieltjes transform by $\varphi_X(s) := \mathbb{E}\left[e^{-sX}\right]$ for $s > 0$, and its n-th derivative by $\varphi_X^{(n)}(s) = \mathbb{E}\left[(-X)^n e^{-sX}\right]$. Let $\Gamma(\cdot)$ define the Gamma function.

Theorem 3.1. (Karamata's Tauberian Theorem, Theorem 8.1.6 in [7].) *The following statements are equivalent:*

1. *X is regularly varying with noninteger index $\alpha > 0$ and slowly varying function $L_X(\cdot)$, i.e.*

$$\mathbb{P}\left(X > x\right) \sim x^{-\alpha} L_X(x), \text{ as } x \to \infty.$$

2. *For a noninteger index $\alpha > 0$,*

$$\varphi_X^{(\lceil \alpha \rceil)}(s) \sim C_\alpha s^{\alpha - \lceil \alpha \rceil} L_X(1/s), \text{ as } s \to 0^+,$$

for $L_X(\cdot)$ a slowly varying function, where $C_\alpha := -\Gamma(\alpha+1)\Gamma(1-\alpha)/\Gamma(\alpha - \lfloor \alpha \rfloor)$.

Remark 3.1. Note that when X is regularly varying with index $\alpha \in (n, n+1)$, the $(n+1)$-th moment does not exist. Observe that the above trivially implies that, when $\alpha \in (n, n+1)$, $\varphi_X^{(n+1)}(s) = \varphi_X^{(\lceil \alpha \rceil)}(s) \to \infty$, as $s \to 0^+$, a property we will use repeatedly in subsequent sections.

4. Tail asymptotics of maximum functional in renewal Poisson cluster process

We now prove a single big-jump principle for the tail asymptotics of the distribution of the maximum functional of a generic cluster in the settings of the renewal Poisson cluster process. As mentioned in Remark (2.1), the conclusions reached for this process are of course valid for the mixed binomial Poisson cluster process.

Proposition 4.1. *Suppose the vector* (X, K_A) *in Equation* (5) *is regularly varying with index* $\alpha > 1$ *and non-null Radon measure* μ*. Then,*

$$\mathbb{P}\left(H^R > x\right) \sim (1 + \mathbb{E}\left[K_A\right])\mathbb{P}\left(X > x\right), \text{ as } x \to \infty.$$

Moreover, if $\mu(\{(x_1, x_2) \in \mathbb{R}^2_{+,\mathbf{0}} : x_1 > 1\}) > 0$*, then* H^R *is regularly varying with index* $\alpha > 1$*.*

Proof of Proposition (4.1). The proof can be found in Appendix (A). It uses a classical approach via conditioning on K_A and Taylor expansions and is given for completeness. □

Remark 4.1. In the proof of Proposition (4.1), one only needs X to be regularly varying for H^H to be regularly varying. However, to keep the same settings in terms of regular variation as for the upcoming results, we make the assumption that (X, K_A) is regularly varying and regular variation of X follows by considering the consequences of this assumption contained in Example (3.1). The case where $\mathbb{P}\left(X > x\right) = o(\mathbb{P}\left(K_A > x\right))$, $x \to \infty$, X regularly varying and K_A a stopping time with respect to $(A_j)_{j \geq 0}$ is treated in Proposition 3.1 and Corollary 4.2 of [5]. It is proved that H^R is also regularly varying but, more generally, that H^R falls in the same MDA than X. What we propose in Proposition (4.1) is merely a refinement for the Fréchet MDA, describing explicitly the tail of H^R when $\mathbb{P}\left(K_A > x\right) = \mathcal{O}(\mathbb{P}\left(X > x\right))$, $x \to \infty$, and K_A depending only on X_0.

5. Tail asymptotics of the sum functional in renewal Poisson cluster process

We now prove a result concerning the sum functional of a generic cluster in the settings of the renewal Poisson cluster process. Again, this extends easily to the mixed binomial Poisson cluster process.

Proposition 5.1. *Suppose the vector* (X, K_A) *in Equation* (6) *is regularly varying with noninteger index* $\alpha > 1$. *Then,* D^R *is regularly varying with the same index* α. *More specifically,*

$$\mathbb{P}\left(D^R > x\right) \sim \mathbb{P}\left(X + \mathbb{E}\left[X\right] K_A > x\right) + \mathbb{E}\left[K_A\right] \mathbb{P}\left(X > x\right), \text{ as } x \to \infty.$$

Proof of Proposition (5.1). First, note that the Laplace-Stieltjes transform of D^R in Equation (6) is given by

$$\begin{aligned}\varphi_{D^R}(s) := \mathbb{E}\left[e^{-sX - s\sum_{j=1}^{K_A} X_j}\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{-sX} e^{-s\sum_{j=1}^{K_A} X_j} \mid A\right]\right] \\ &= \mathbb{E}\left[e^{-sX} e^{K_A \log \mathbb{E}\left[e^{-sX}\right]}\right] =: \mathbb{E}\left[e^{-sX + K_A \log \varphi_X(s)}\right]\end{aligned}$$

upon recalling that $X := f(A)$ and K_A are independent conditionally on the ancestral mark A, and that $(X_j)_{j\geqslant 1}$ are i.i.d. and independent of A. We first show that, for any noninteger $\alpha \in (n, n+1)$, $n \in \mathbb{N}$,

$$\varphi_{D^R}^{(n+1)}(s) \sim \varphi_{X+\mathbb{E}[X]K_A}^{(n+1)}(s) + \mathbb{E}\left[K_A\right] \varphi_X^{(n+1)}(s), \text{ as } s \to 0^+,$$

where $\varphi_{X+\mathbb{E}[X]K_A}(s) := \mathbb{E}\left[e^{-sX - s\mathbb{E}[X]K_A}\right]$, and where $\varphi_X^{(n)}$ is the nth derivative of the Laplace-Stieltjes transform of a random variable X.

We have to consider the following expression:

$$\begin{aligned}\left|\varphi_{D^R}^{(n+1)}(s) - \left(\varphi_{X+\mathbb{E}[X]K_A}^{(n+1)}(s) + \mathbb{E}\left[K_A\right] \varphi_X^{(n+1)}(s)\right)\right| &= \Bigg|\mathbb{E}\left[\left(-X + K_A \frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n+1} e^{-sX + K_A \log \varphi_X(s)}\right] \\ &\quad + \mathbb{E}\left[K_A \frac{\varphi_X^{(n+1)}(s)}{\varphi_X(s)} e^{-sX + K_A \log \varphi_X(s)}\right] \\ &\quad - \mathbb{E}\left[\left(-X - \mathbb{E}\left[X\right] K_A\right)^{n+1} e^{-sX - s\mathbb{E}[X]K_A}\right] \\ &\quad - \mathbb{E}\left[K_A\right] \mathbb{E}\left[\left(-X\right)^{n+1} e^{-sX}\right] + C_{n+1}\Bigg| \\ &=: \left|B_1 + B_2 - B_3 - B_4 + C_{n+1}\right|. \end{aligned} \tag{9}$$

Consider first the difference $\left|B_1 - B_3\right|$. The following set of inequalities, directly due to the convexity of the function $\log \varphi_X(\cdot)$, will prove useful in controlling the above difference: for $s > 0$, we have

$$-s\mathbb{E}\left[X\right] K \leqslant K \log \varphi_X(s) \leqslant sK \frac{\varphi_X^{(1)}(s)}{\varphi_X(s)} \leqslant 0 \leqslant -sK \frac{\varphi_X^{(1)}(s)}{\varphi_X(s)} \leqslant -K \log \varphi_X(s) \leqslant s\mathbb{E}\left[X\right] K. \tag{10}$$

Using the basic decomposition $(a^{n+1}-b^{n+1}) = (a-b)\sum_{k=0}^{n} a^{n-k}b^k$ as well as Equation (10) yields

$$\begin{aligned}
\big|B_1 - B_3\big| &\leqslant \Bigg|\left(\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)} + \mathbb{E}\left[X\right]\right) \\
&\quad \cdot \mathbb{E}\left[K_A\left(\sum_{k=0}^{n}\left(-X + K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-k}\left(-X - \mathbb{E}\left[X\right]K_A\right)^k\right)e^{-sX+K_A\log\varphi_X(s)}\right]\Bigg| \\
&\leqslant \Bigg|\left(\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)} + \mathbb{E}\left[X\right]\right)\Bigg(\mathbb{E}\left[K_A\left(-X + K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^n e^{-sX+K_A\log\varphi_X(s)}\right] \\
&\quad + \mathbb{E}\left[K_A\left(-X - \mathbb{E}\left[X\right]K_A\right)^n e^{-sX+K_A\log\varphi_X(s)}\right] \\
&\quad + \mathbb{E}\left[K_A\left(\sum_{k=1}^{n-1}\left(-X + K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-k}\left(-X - \mathbb{E}\left[X\right]K_A\right)^k\right)e^{-sX+K_A\log\varphi_X(s)}\right]\Bigg)\Bigg| \\
&=: \big|G(B_{11}+B_{12}+B_{13})\big|.
\end{aligned}$$

where $G := \frac{\varphi_X^{(1)}(s)}{\varphi_X(s)} + \mathbb{E}\left[X\right]$.

We then treat each term separately. First, consider B_{11}. Using the binomial theorem, we have that

$$\left(-X + K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^n = \sum_{j=0}^{n}\binom{n}{j}\left(-X\right)^j\left(K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-j}.$$

Using the linearity of expectations, we separate the cases. Let $j=0$. Because $G>0$, $K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)} < 0$, using Equation (10) and the basic inequality $xe^{-x} \leqslant e^{-1}$, we get:

$$\begin{aligned}
\left|G\mathbb{E}\left[K_A\left(K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^n e^{-sX+K_A\log\varphi_X(s)}\right]\right| &\leqslant \frac{G}{s}\mathbb{E}\left[K_A^{n-1}\left|\left(\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-1}\right|\left(-K_A\log\varphi_X(s)\right)e^{K_A\log\varphi_X(s)}\right] \\
&\leqslant \frac{G}{s}\mathbb{E}\left[K_A^{n-1}\left|\left(\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-1}\right|e^{-1}\right]. \qquad (11)
\end{aligned}$$

In order to control the upper bound, we need to control G/s, and we have to distinguish two cases:

Case $\alpha \in (1,2)$: We have the identities

$$\frac{G}{s} = \frac{\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)} + \mathbb{E}\left[X\right]}{s} = \frac{\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)} - \varphi_X^{(1)}(s) + \varphi_X^{(1)}(s) + \mathbb{E}\left[X\right]}{s} = \varphi_X^{(1)}(s)\left(\frac{\frac{1}{\varphi_X(s)}-1}{s}\right) + \frac{\varphi_X^{(1)}(s) + \mathbb{E}\left[X\right]}{s}.$$

The limit as $s \to 0^+$ of $\frac{\frac{1}{\varphi_X(s)}-1}{s}$ is the derivative of $1/\varphi_X(s)$ at $s=0$ and hence is finite; it follows that

$$\varphi_X^{(1)}(s)\left(\frac{\frac{1}{\varphi_X(s)}-1}{s}\right) = \mathcal{O}\big(\varphi_X^{(1)}(s)\big), \text{ as } s \to 0^+.$$

Now note that, for the second term, if first X has negligible tails with respect to $X+\mathbb{E}[X]K_A$, by Lemma (C.2), it follows that

$$\frac{\varphi_X^{(1)}(s)+\mathbb{E}[X]}{s}=o\big(\varphi^{(2)}_{X+\mathbb{E}[X]K_A}(s)+\mathbb{E}[K_A]\,\varphi_X^{(2)}(s)\big),\ \text{as } s\to 0^+.$$

If X is regularly varying with the same index as $X+\mathbb{E}[X]K_A$, then clearly, by adapting the proof of Lemma (C.2), it follows that

$$\frac{\varphi_X^{(1)}(s)+\mathbb{E}[X]}{s}=\mathcal{O}\big(\varphi^{(2)}_{X+\mathbb{E}[X]K_A}(s)+\mathbb{E}[K_A]\,\varphi_X^{(2)}(s)\big),\ \text{as } s\to 0^+.$$

By a dominated convergence argument, the upper bound in Equation (11) is such that

$$\mathbb{E}\left[K_A^{n-1}\left|\left(\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-1}\right|\Big(-K_A\log\varphi_X(s)\Big)e^{K_A\log\varphi_X(s)}\right]=o(1),\ \text{as } s\to 0^+,$$

and combining with the arguments above, this proves that, no matter if X is lighter or as heavy as the modulus $X+\mathbb{E}[X]K_A$,

$$B_{11}=o\big(\varphi^{(2)}_{X+\mathbb{E}[X]K_A}(s)+\mathbb{E}[K_A]\,\varphi_X^{(2)}(s)\big),\ \text{as } s\to 0^+.$$

Case $\alpha\in(n,n+1)$, with $n\in\mathbb{N}\backslash\{1\}$: Using the definition of the derivative, as $s\to 0^+$,

$$\lim_{s\to 0^+}\frac{G}{s}=\lim_{s\to 0^+}\frac{\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}+\mathbb{E}[X]}{s}\ \text{and}\ \lim_{s\to 0^+}\frac{\frac{\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}+\mathbb{E}[X]}{s}}{\frac{\varphi_X^{(2)}(s)}{\varphi_X(s)}-\frac{(\varphi_X^{(1)}(s))^2}{(\varphi_X(s))^2}}=1,$$

and, for this range of $\alpha\in(n,n+1)$ with $n\in\mathbb{N}\backslash\{1\}$

$$\frac{\varphi_X^{(2)}(s)}{\varphi_X(s)}-\frac{(\varphi_X^{(1)}(s))^2}{(\varphi_X(s))^2}<\infty$$

which is finite since $\alpha\in(n,n+1)$ for $n\geqslant 2$. Because $\varphi_X^{(1)}(s)$ is finite, Equation (11) is finite. Upon applying Theorem (3.1), it follows that, as $s\to 0^+$,

$$\left|G\mathbb{E}\left[K_A\left(K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^n e^{-sX+K_A\log\varphi_X(s)}\right]\right|=o\big(\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}(s)+\mathbb{E}[K_A]\,\varphi_X^{(n+1)}(s)\big).$$

The treatment of terms where $j>0$ is easier: it is sufficient to note that, whenever X appears in the product, one can always "lose a power": suppose without loss of generality that $j=1$ in the decomposition due to the binomial theorem above; we are left to consider the following term

$$\left|G\mathbb{E}\left[K_A\left\{\binom{n}{1}\Big(-X\Big)^1\left(K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-1}\right\}e^{-sX+K_A\log\varphi_X(s)}\right]\right|.$$

This is smaller than

$$\frac{G}{s}\mathbb{E}\left[\binom{n}{1}K_A^n\left|\left(\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-1}\right|(sX)e^{-sX}\right] \leqslant \frac{G}{s}\mathbb{E}\left[\binom{n}{1}K_A^n\left|\left(\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-1}\right|e^{-1}\right],$$

and by similar reasoning as above, the expectation as well as the whole of the upper bound is finite. All in all, this shows that, as $s \to 0^+$,

$$\left|G\mathbb{E}\left[K_A\left\{\binom{n}{1}\Big(-X\Big)^1\left(K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-1}\right\}e^{-sX+K_A\log\varphi_X(s)}\right]\right| \\ = o\big(\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}(s) + \mathbb{E}\left[K_A\right]\varphi_X^{(n+1)}(s)\big)$$

Upon applying the same arguments on all terms making up B_{11}, using at times Hölder's inequality to justify that expectations of the form $\mathbb{E}\left[K_A\big(-X\big)^{j-1}\big(K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\big)^{n-j}\right]$ for $2 \leqslant j \leqslant n-1$ are finite, and one X is factorised as in the reasoning above, this is sufficient to show that

$$\left|GB_{11}\right| = o\big(\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}(s) + \mathbb{E}\left[K_A\right]\varphi_X^{(n+1)}(s)\big), \text{ as } s \to 0^+.$$

A completely analogous approach - omitted for brevity - shows that

$$\left|GB_{12}\right| = o\big(\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}(s) + \mathbb{E}\left[K_A\right]\varphi_X^{(n+1)}(s)\big), \text{ as } s \to 0^+$$

replacing only the appeal to Equation (10) by the fact that we can always find $s > 0$ small enough such that $s\mathbb{E}\left[X\right] \leqslant -2\log\varphi_X(s)$, which holds because of the following reasoning: since $\varphi_X(s)$ is differentiable at 0, by the integrability of X, one obtains

$$\lim_{s\to 0^+} -\frac{\log\varphi_X(s)}{s} = -\frac{\varphi_X^{(1)}(0)}{\varphi_X(0)} = -\mathbb{E}\left[-X\right] \iff \lim_{s\to 0^+} -\frac{\log\varphi_X(s)}{s} = \mathbb{E}\left[X\right].$$

By a similar argument, $-\frac{2\log\varphi_X(s)}{s} \to 2\mathbb{E}\left[X\right]$, as $s \to 0^+$. Hence, there exists $s > 0$ small enough such that

$$s\mathbb{E}\left[X\right] \leqslant -2\log\varphi_X(s).$$

Finally, consider $\left|GB_{13}\right|$. The sum given can be factorised as

$$\sum_{k=1}^{n-1}\left(-X + K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-k}\left(-X - \mathbb{E}\left[X\right]K_A\right)^k \\ = \left(-X + K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)\sum_{k=1}^{n-1}\left(-X + K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\right)^{n-1-k}\left(-X - \mathbb{E}\left[X\right]K_A\right)^k$$

Now this yields, upon using Equation (10) and the basic inequality $xe^{-x} \leqslant e^{-1}$ in the last step,

$$\begin{aligned}
|GB_{13}| &= \frac{G}{s}\mathbb{E}\Bigg[K_A\Bigg|\Bigg(-sX + K_A s\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\Bigg)\Bigg|e^{-\left(sX - K_A\log\varphi_X(s)\right)} \\
&\qquad \cdot \sum_{k=1}^{n-1}\Bigg|\Bigg(-X + K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\Bigg)^{n-k-1}\Bigg(-X - \mathbb{E}\left[X\right]K_A\Bigg)^k\Bigg|\Bigg] \\
&\leqslant \frac{G}{s}\mathbb{E}\Bigg[K_A\Bigg(sX - K_A\log\varphi_X(s)\Bigg)e^{-\left(sX - K_A\log\varphi_X(s)\right)} \\
&\qquad \cdot \sum_{k=1}^{n-1}\Bigg|\Bigg(-X + K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\Bigg)^{n-k-1}\Bigg(-X - \mathbb{E}\left[X\right]K_A\Bigg)^k\Bigg|\Bigg] \\
&\leqslant \frac{G}{s}\mathbb{E}\Bigg[K_A e^{-1}\sum_{k=1}^{n-1}\Bigg|\Bigg(-X + K_A\frac{\varphi_X^{(1)}(s)}{\varphi_X(s)}\Bigg)^{n-k-1}\Bigg(-X - \mathbb{E}\left[X\right]K_A\Bigg)^k\Bigg|\Bigg].
\end{aligned}$$

The highest order of the product of the summands above is of power n: again, since $\alpha \in (n, n+1)$, using Hölder's inequality, the expectation is finite. Overall, this shows once again that

$$\left|GB_{13}\right| = o\big(\varphi_{X+\mathbb{E}[X]K_A}^{(n+1)}(s) + \mathbb{E}\left[K_A\right]\varphi_X^{(n+1)}(s)\big), \text{ as } s \to 0^+.$$

Collecting all of the above bounds, this shows that

$$\left|B_1 - B_3\right| = o\big(\varphi_{X+\mathbb{E}[X]K_A}^{(n+1)}(s) + \mathbb{E}\left[K_A\right]\varphi_X^{(n+1)}(s)\big), \text{ as } s \to 0^+.$$

Consider the difference $\left|B_2 - B_4\right|$, and note that one can write it as

$$\begin{aligned}
\left|B_2 - B_4\right| &= \left|\frac{\varphi_X^{(n+1)}(s)}{\varphi_X(s)}\mathbb{E}\left[K_A\big(1 - e^{-sX + K_A\log\varphi_X(s)}\big)\right] + \varphi_X^{(n+1)}(s)\Bigg(\frac{1}{\varphi_X(s)} - 1\Bigg)\mathbb{E}\left[K_A\right]\right| \\
&=: \left|B_{21} + B_{22}\right|.
\end{aligned}$$

Now, by a dominated convergence argument as before, one has that $\mathbb{E}\left[K_A\big(1 - e^{-sX + K_A\log\varphi_X(s)}\big)\right] = o(1)$, as $s \to 0^+$, and hence, that

$$\left|B_{21}\right| = o\big(\varphi_X^{(n+1)}(s)\big) = o\big(\varphi_{X+\mathbb{E}[X]K_A}^{(n+1)}(s) + \mathbb{E}\left[K_A\right]\varphi_X^{(n+1)}(s)\big), \text{ as } s \to 0^+.$$

Similarly, by the integrability of K_A,

$$\left|B_{22}\right| = o\big(\varphi_X^{(n+1)}(s)\big) = o\big(\varphi_{X+\mathbb{E}[X]K_A}^{(n+1)}(s) + \mathbb{E}\left[K_A\right]\varphi_X^{(n+1)}(s)\big), \text{ as } s \to 0^+.$$

Collecting the above, this implies that

$$\left|B_2 - B_4\right| = o\big(\varphi_{X+\mathbb{E}[X]K_A}^{(n+1)}(s) + \mathbb{E}\left[K_A\right]\varphi_X^{(n+1)}(s)\big), \text{ as } s \to 0^+.$$

Lastly, the terms making up C_{n+1} when $\alpha \in (n, n+1)$ are all the terms (and cross-products) of order strictly lower than $n+1$ and, consequently, are finite. It follows by Theorem (3.1) that

$$C_{n+1} = o\big(\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}(s) + \mathbb{E}\,[K_A]\,\varphi^{(n+1)}_X(s)\big), \text{ as } s \to 0^+.$$

All in all, this essentially shows that, as $s \to 0^+$,

$$\left|\varphi^{(n+1)}_{D^R}(s) - \big(\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}(s) + \mathbb{E}\,[K_A]\,\varphi^{(n+1)}_X(s)\big)\right| = o\big(\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}(s) + \mathbb{E}\,[K_A]\,\varphi^{(n+1)}_X(s)\big),$$

and hence that

$$\varphi^{(n+1)}_{D^R}(s) \sim \varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}(s) + \mathbb{E}\,[K_A]\,\varphi^{(n+1)}_X(s), \text{ as } s \to 0^+,$$

and this equivalence holds for any $\alpha \in (n, n+1)$, $n \in \mathbb{N}$.

Because the modulus $X + \mathbb{E}\,[X]\,K_A$ is regularly varying whenever (X, K_A) is - see Remark (5.1) - Karamata's Theorem (3.1) implies that

$$\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}(s) \sim C_\alpha s^{\alpha-\lceil\alpha\rceil} L_{X+\mathbb{E}[X]K_A}(1/s), \text{ as } s \to 0^+$$

for some slowly varying function $L_{X+\mathbb{E}[X]K_A}(\cdot)$. Then, suppose first that X is not regularly varying and has negligible tails with respect to the modulus $X + \mathbb{E}\,[X]\,K_A$. Then Lemma (C.1) yields that

$$\varphi^{(n+1)}_X(s) = o\big(\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}(s)\big), \text{ as } s \to 0^+$$

and hence, this implies that

$$\varphi^{(n+1)}_{D^R}(s) \sim C_\alpha s^{\alpha-\lceil\alpha\rceil} L_{X+\mathbb{E}[X]K_A}(1/s)\big(1 + o(1)\big), \text{ as } s \to 0^+$$

which yields by re-applying Karamata's Tauberian Theorem (3.1), that

$$\mathbb{P}\left(D^R > x\right) \sim x^{-\alpha} L_{D^R}(x) \sim x^{-\alpha} L_{X+\mathbb{E}[X]K_A}(x)\big(1 + o(1)\big), \text{ as } x \to \infty.$$

In the case where X is regularly varying, by Example (3.1), and because X has the same index $\alpha > 1$ as the modulus $X + \mathbb{E}\,[X]\,K_A$, if the limiting Radon measure is non-null on the correct subspace, Karamata's Tauberian Theorem (3.1) yields,

$$\varphi^{(n+1)}_X(s) \sim C_\alpha s^{\alpha-\lceil\alpha\rceil} L_X(1/s), \text{ as } s \to 0^+.$$

Then, for each $n \in \mathbb{N}$

$$\varphi^{(n+1)}_{D^R}(s) \sim 2C_\alpha s^{\alpha-\lceil\alpha\rceil}\big(L_{X+\mathbb{E}[X]K_A}(1/s) + \mathbb{E}\,[K_A]\,L_X(1/s)\big), \text{ as } s \to 0^+,$$

and because the sum of two slowly varying function is still a slowly varying function, $L_D^R(\cdot) := L_{X+\mathbb{E}[X]K_A}(\cdot) + \mathbb{E}[K_A] L_X(\cdot)$ is slowly varying. Applying again Karamata's Tauberian Theorem (3.1) in the other direction, yields

$$\mathbb{P}\left(D^R > x\right) \sim x^{-\alpha} L_{D^R}(x) \sim x^{-\alpha}\left(L_{X+\mathbb{E}[X]K_A}(x) + \mathbb{E}[K_A] L_X(x)\right), \text{ as } x \to \infty$$

which yields the desired result and the proof is complete. □

Remark 5.1. Note that the assumption that the random vector (X, K_A) is regularly varying with index $\alpha > 1$ ensures, by Proposition (3.3), that $X + \mathbb{E}[X] K_A$ is regularly varying with the same index $\alpha > 1$. Indeed, it can be easily seen that $\rho(X, K_A) := X + \mathbb{E}[X] K_A$ is a modulus (in the sense made precise in Section 3), provided that $\mathbb{E}[X] \neq 0$, which is a natural assumption to make, since X is taken to be nonnegative.

Remark 5.2. Note that the findings of Proposition (5.1) are consistent with the findings of [19]: in particular, if X and K_A are independent - which is the setting in the aforementioned paper - or even if X and K_A are asymptotically independent (i.e. if $\mathbb{P}(X > x, \mathbb{E}[X] K_A > x) = o\big(\mathbb{P}(X > x)\mathbb{P}(\mathbb{E}[X] K_A > x)\big)$, as $x \to \infty$) then the proposed asymptotics of Proposition (5.1) encompass three cases, depending on the relation between X and K_A:

1. when $\mathbb{P}(K_A > x) = o(\mathbb{P}(X > x))$, as $x \to \infty$, then $\mathbb{P}(X + \mathbb{E}[X] K_A > x) \sim \mathbb{P}(X > x)$, as $x \to \infty$. From Proposition (5.1), this means that

$$\mathbb{P}\left(D^R > x\right) \sim \mathbb{P}(X > x) + \mathbb{E}[K_A]\mathbb{P}(X > x) \sim (\mathbb{E}[K_A] + 1)\mathbb{P}(X > x), \text{ as } x \to \infty$$

   which is equivalent to Proposition 4.1 in [19];

2. when $\mathbb{P}(X > x) = o(\mathbb{P}(K_A > x))$, as $x \to \infty$, then $\mathbb{P}(X + \mathbb{E}[X] K_A > x) \sim (\mathbb{E}[X])^\alpha \mathbb{P}(K_A > x)$, as $x \to \infty$. From Proposition (5.1), this means that

$$\mathbb{P}\left(D^R > x\right) \sim (\mathbb{E}[X])^\alpha \mathbb{P}(K_A > x), \text{ as } x \to \infty$$

   which is equivalent to Proposition 4.3 in [19];

3. lastly, when $\mathbb{P}(K_A > x) \sim c\mathbb{P}(X > x)$, as $x \to \infty$, for $c > 0$, then $\mathbb{P}(X + \mathbb{E}[X] K_A > x) \sim \mathbb{P}(X > x) + c(\mathbb{E}[X])^\alpha \mathbb{P}(X > x)$, as $x \to \infty$. From Proposition (5.1), this means that

$$\mathbb{P}\left(D^R > x\right) \sim \left(\mathbb{E}[K_A] + 1 + c(\mathbb{E}[X])^{-\alpha}\right)\mathbb{P}(X > x), \text{ as } x \to \infty$$

   which is equivalent to Lemma 4.7 in [19].

Our approach offers a more flexible framework for dependence between the governing components of the clusters, namely X and K_A. Yet, in this latter direction, and more closely related to our results, [48] shows in a recent contribution that

$$\mathbb{P}\left(D^R > x\right) \sim 1_{\{\mathbb{E}[K_A]<\infty\}}\mathbb{E}\left[K_A\right]\mathbb{P}\left(X > x\right) + \mathbb{P}\left(\mathbb{E}\left[X\right]K_A > x\right), \text{ as } x \to \infty,$$

in the regime where (X, K_A) are arbitrarily dependent and either K_A is intermediate regularly varying and $\mathbb{P}\left(X > x\right) = o(\mathbb{P}\left(K_A > x\right)),$ as $x \to \infty$ (Theorem 6.10 in [48]) or X is intermediate regularly varying and $\mathbb{P}\left(K_A > x\right) = o(\mathbb{P}\left(X > x\right)),$ as $x \to \infty$ (Theorem 6.11 in [48]). The novelty in this paper is to propose similar asymptotics in the case where K_A and X are effectively tail equivalent.

Note that the content of Proposition (5.1) is a kind of "double" big-jump principle: the heavy-tailedness introduced by letting the vector (X, K_A) be regularly varying implies that there is two ways for the sum D^R to be large; either through a combination of the dependent variables X and K_A or through the classical single big-jump coming from the additional term $\mathbb{E}\left[K_A\right]\mathbb{P}\left(X > x\right)$ consisting of the offspring events.

6. Tail asymptotics of the maximum functional in the Hawkes process

We now propose a single big-jump principle concerning the maximum functional of a generic cluster in the settings of the Hawkes process. Recall that $\mathbb{E}\left[L_A\right] = \mathbb{E}\left[\kappa_A\right] = 1$.

Proposition 6.1. *Suppose the vector (X, κ_A) in Equation (7) is regularly varying with index $\alpha > 1$ and non-null Radon measure μ. Then,*

$$\mathbb{P}\left(H^H > x\right) \sim \frac{1}{1 - \mathbb{E}\left[\kappa_A\right]}\mathbb{P}\left(X > x\right), \text{ as } x \to \infty.$$

Moreover, if $\mu(\{(x_1, x_2) \in \mathbb{R}^2_{+,\mathbf{0}} : x_1 > 1\}) > 0$, then H^H is regularly varying with index $\alpha > 1$.

Proof of Proposition (6.1). The proof can be found in Appendix (B) and follows the same approach as the proof of Proposition (4.1). □

Remark 6.1. As hinted in Section 1, a closely related work concerning the maxima of the marks in a generic cluster of the Hawkes process can be found in [5]. Under the assumption that K_A is a stopping time with respect to a filtration including the information about (X_{ij}), it is shown in their Lemma 4.1 that H^H falls in the same MDA as X. What we propose in Proposition (6.1) is merely a refinement for the Fréchet MDA, describing explicitly the tail of H^H.

7. Tail asymptotics of the sum functional in Hawkes process

We now propose another "double" big-jump principle concerning the sum functional of a generic cluster in the setting of the Hawkes process. The tail approximation obtained in Proposition (7.1) below is in fact very similar to the one in Proposition (5.1), where both a single big-jump principle and a combination of the effects of the dependent variables X and κ_A yield large values for D^H.

Proposition 7.1. *Assume that (X, κ_A) in Equation (8) has a regularly varying distribution with noninteger index $\alpha > 1$. Then, (X, L_A) is regularly varying with the same index α. Further, D^H is regularly varying with index α. In fact,*

$$\mathbb{P}\left(D^H > x\right) \sim \frac{1}{1-\mathbb{E}\left[\kappa_A\right]}\mathbb{P}\left(X + \left(\frac{\mathbb{E}\left[X\right]}{1-\mathbb{E}\left[\kappa_A\right]}\right)\kappa_A > x\right), \text{ as } x \to \infty.$$

Proof of Proposition (7.1). Recall that the assumption that (X, κ_A) is regularly varying with index $\alpha > 1$ is equivalent to the regular variation of the linear combinations $t_1 X + t_2 \kappa_A$ for all $t_1, t_2 \in \mathbb{R}_+$ by Proposition (3.4). Similarly as in the proof of Proposition (5.1), if we can show, at any order $(n+1)$ for $n \in \mathbb{N}$, and for any $t_1, t_2 \in \mathbb{R}_+$, that the behaviour of $\varphi^{(n+1)}_{t_1 X + t_2 \kappa_A}(s) := \frac{\partial^{n+1}}{\partial s^{n+1}}\left(\mathbb{E}\left[e^{-s(t_1 X + t_2 \kappa_A)}\right]\right)$, and that of $\varphi^{(n+1)}_{t_1 X + t_2 L_A}(s) := \frac{\partial^{n+1}}{\partial s^{n+1}}\left(\mathbb{E}\left[e^{-s(t_1 X + t_2 L_A)}\right]\right)$, as $s \to 0^+$ are comparable, i.e. if

$$\varphi^{(n+1)}_{t_1 X + t_2 \kappa_A}(s) \sim \varphi^{(n+1)}_{t_1 X + t_2 L_A}(s), \text{ as } s \to 0^+,$$

then by Karamata's Theorem (3.1), we have

$$\mathbb{P}\left(t_1 X + t_2 \kappa_A > x\right) \sim \mathbb{P}\left(t_1 X + t_2 L_A > x\right), \text{ as } x \to \infty.$$

But this essentially means, reapplying Proposition (3.4), that (X, L_A) is regularly varying.

The following bounds will be useful:

1. By a Taylor expansion, as $s \to 0^+$,

$$s - (1 - e^{-s}) \leqslant s^2/2. \tag{12}$$

2. For $s > 0$ small enough,

$$-(1 - e^{-s}) \leqslant -s/2. \tag{13}$$

First, note that it is possible to write $\varphi_{t_1 X + t_2 L_A}(\cdot)$ as a function of κ_A instead of L_A. Using the Tower property and recalling that $L_A | A \sim \text{Poisson}(\kappa_A)$ yields

$$\varphi_{t_1 X + t_2 L_A}(s) = \mathbb{E}\left[\mathbb{E}\left[e^{-s t_1 X - s t_2 L_A} \mid A\right]\right] = \mathbb{E}\left[e^{-s t_1 X - (1 - e^{-s t_2})\kappa_A}\right].$$

From this, and letting $\alpha \in (n, n+1)$, $n \geqslant 1$, simple derivations and collection of terms lead us to consider the difference given by

$$\begin{aligned}
\left|\varphi^{(n+1)}_{t_1X+t_2L_A}(s) - \varphi^{(n+1)}_{t_1X+t_2\kappa_A}(s)\right| = \Bigg|&\mathbb{E}\left[(-t_1X)^{n+1}\big(e^{-st_1X-(1-e^{-st_2})\kappa_A} - e^{-st_1X-st_2\kappa_A}\big)\right] \\
&+ I_1\mathbb{E}\left[(-t_1X)^{n}(-t_2\kappa_A)\big(e^{-st_1X-(1-e^{-st_2}\kappa_A)\kappa_A-I_2st_2} - e^{-st_1X-st_2\kappa_A}\big)\right] \\
&+ \ldots \\
&+ I_j\mathbb{E}\left[(-t_1X)(-t_2\kappa_A)^{n}\big(e^{-st_1X-(1-e^{-st_2}\kappa_A)\kappa_A-I_kst_2} - e^{-st_1X-st_2\kappa_A}\big)\right] \\
&+ \mathbb{E}\left[(-t_2\kappa_A)^{n+1}\big(e^{-st_1X-(1-e^{-st_2})\kappa_A-(n+1)st_2} - e^{-st_1X-st_2\kappa_A}\big)\right] + C_{n+1}\Bigg| \\
=: \left|B_1 + B_{21} + \ldots + B_{2j} + B_3 + C_{n+1}\right|&
\end{aligned}$$

where the constants of product terms $(B_{21}, \ldots, B_{2j})$ $I_1, I_2, \ldots, I_j, I_k \in \mathbb{N}$ depend on n.

Consider term B_1. Using Equation (12) and Equation (13) and the basic inequality $xe^{-x} \leqslant e^{-1}$, one can show that

$$\begin{aligned}
|B_1| &= \mathbb{E}\left[(t_1X)^{n+1}e^{-st_1X}\big(e^{-(1-e^{-st_2})\kappa_A} - e^{-st_2\kappa_A}\big)\right] \\
&\leqslant \mathbb{E}\left[(t_1X)^{n+1}e^{-st_1X}e^{-(1-e^{-st_2})\kappa_A}\kappa_A(st_2 - 1 + e^{-st_2})\right] \\
&\leqslant \mathbb{E}\left[(t_1X)^{n+1}e^{-st_1X}e^{-st_2\kappa_A/2}\kappa_A(st_2)^2/2\right] \\
&\leqslant \mathbb{E}\left[(st_1Xe^{-st_1X})\left(\frac{st_2\kappa_A}{2}e^{-st_2\kappa_A/2}\right)t_2(t_1X)^n\right] \\
&\leqslant \mathbb{E}\left[e^{-2}t_2(t_1X)^n\right].
\end{aligned}$$

and by the finiteness of the nth moment of X when $\alpha \in (n, n+1)$, the above expectation is finite. Hence, it follows, using Karamata's Theorem (3.1) and Remark (3.1), that

$$B_1 = o\big(\varphi^{(n+1)}_{t_1X+t_2\kappa_A}(s)\big), \text{ as } s \to 0^+.$$

Consider one representative for the cross-product terms, say, without loss of generality, B_{21}. Then, proceeding as before for term B_1, using Equation (12) and Equation (13) and the basic inequality

$xe^{-x} \leqslant e^{-1}$, yields

$$\begin{aligned}
|B_{21}| &= I_1 \mathbb{E}\left[(t_1X)^n (t_2\kappa_A) e^{-st_1X}\big((e^{-(1-e^{-st_2})\kappa_A - st_2} - e^{-st_2\kappa_A - st_1}) - (e^{-st_2\kappa_A} - e^{-st_2\kappa_A - st_1})\big)\right] \\
&\leqslant I_1 \mathbb{E}\left[(t_1X)^n e^{-st_1X} (t_2\kappa_A^2) e^{-(1-e^{-st_2})\kappa_A}(st_2 - 1 + e^{-st_2})\right] \\
&\leqslant I_1 \mathbb{E}\left[(t_1X)^n e^{-st_1X} (t_2\kappa_A^2) e^{-st_2\kappa_A/2}\frac{(st_2)^2}{2}\right] \\
&\leqslant I_1 \mathbb{E}\left[(st_1Xe^{-st_1X})\left(\frac{st_2\kappa_A}{2}e^{-st_2\kappa_A/2}\right)(t_1X)^{n-1}t_2^2\kappa_A\right] \\
&\leqslant I_1 \mathbb{E}\left[e^{-2}(t_1X)^{n-1}t_2^2\kappa_A\right].
\end{aligned}$$

Using Hölder's inequality, because the order of the product of X^{n-1} and κ_A is n, one obtains that the above expectation is finite. It follows from Karamata's Theorem (3.1) and Remark (3.1), that

$$B_{21} = o\big(\varphi^{(n+1)}_{t_1X+t_2\kappa_A}(s)\big), \text{ as } s \to 0^+,$$

and similarly for each cross product term $B_{22}, \ldots, B_{2j}$.

Consider now B_3. With similar tools as before, using Equation (12) and Equation (13) and the basic inequality $x^2e^{-x} \leqslant 4e^{-2}$, yields

$$\begin{aligned}
|B_3| &= \mathbb{E}\left[(t_2\kappa_A)^{n+1} e^{-st_1X}\big((e^{-(1-e^{-st_2})\kappa_A - 2st_2} - e^{st_2\kappa_A - 2st_2}) - (e^{-st_2\kappa_A} - e^{-st_2\kappa_A - 2st_2})\big)\right] \\
&\leqslant \mathbb{E}\left[(t_2\kappa_A)^{n+1} e^{-st_1X}(e^{-(1-e^{-st_2})\kappa_A - 2st_2} - e^{st_2\kappa_A - 2st_2})\right] \\
&\leqslant \mathbb{E}\left[(t_2\kappa_A)^{n+1} e^{-(1-e^{-st_2})\kappa_A}(st_2\kappa_A - (1-e^{-st_2})\kappa_A)\right] \\
&\leqslant 2\mathbb{E}\left[t_2^{n+1}\kappa_A^{n+2}(st_2/2)^2 e^{-st_2\kappa_A/2}\right] \\
&\leqslant 2\mathbb{E}\left[(t_2\kappa_A)^n 4e^{-2}\right],
\end{aligned}$$

which essentially shows once again, using Karamata's Theorem (3.1) and Remark (3.1), that

$$B_3 = o\big(\varphi^{(n+1)}_{t_1X+t_2\kappa_A}(s)\big), \text{ as } s \to 0^+.$$

Lastly, making up the remainder C_{n+1} are terms of strictly smaller order than $n+1$. These are finite and trivially, using Karamata's Theorem (3.1) and Remark (3.1),

$$C_{n+1} = o\big(\varphi^{(n+1)}_{t_1X+t_2\kappa_A}(s)\big), \text{ as } s \to 0^+.$$

Collecting all of the above results, it follows that

$$\big|\varphi^{(n+1)}_{t_1X+t_2L_A}(s) - \varphi^{(n+1)}_{t_1X+t_2\kappa_A}(s)\big| = o(\varphi^{(n+1)}_{t_1X+t_2\kappa_A}(s)), \text{ as } s \to 0^+$$

which essentially means that, for all $n \in \mathbb{N}$

$$\varphi^{(n+1)}_{t_1 X + t_2 L_A}(s) \sim \varphi^{(n+1)}_{t_1 X + t_2 \kappa_A}(s), \text{ as } s \to 0^+.$$

Now, by Karamata's Theorem (3.1), this means that, for all $t_1, t_2 \in \mathbb{R}_+$

$$\mathbb{P}\left(t_1 X + t_2 L_A > x\right) \sim \mathbb{P}\left(t_1 X + t_2 \kappa_A > x\right), \text{ as } x \to \infty,$$

and using Proposition (3.4), this means that (X, L_A) is regularly varying with index $\alpha > 1$. We conclude by applying Theorem 1 in [1] which yields the desired result. □

Remark 7.1. Proposition (7.1) is essentially about showing that if (X, κ_A) is regularly varying, then (X, L_A) is also regularly varying, furthermore with the same index $\alpha > 1$. The equivalence between the regularly varying property of κ_A and that of L_A is easy to prove and is to be found, for example, in [34]. The crucial step to obtain the tail asymptotic of D^H and its regularly varying property in Proposition (7.1) relies on Theorem 1 in [1]. In their even more general setting, the distribution of $X + cL_A$ is intermediate regularly varying, for all $c \in (\mathbb{E}\left[D^H\right] - \epsilon, \mathbb{E}\left[D^H\right] + \epsilon)$ for some $\epsilon > 0$: this assumption encompasses the case where (X, L_A) is regularly varying, but also the cases where X (respectively L_A) is intermediate regularly varying and L_A (respectively X) is lighter, in the sense that $\mathbb{P}\left(L_A > x\right) = o(\mathbb{P}\left(X > x\right))$, as $x \to \infty$ (respectively $\mathbb{P}\left(X > x\right) = o(\mathbb{P}\left(L_A > x\right))$, as $x \to \infty$).

Proposition (7.1) extends Lemma 5.2 in [6] by letting (X, κ_A) be regularly varying, while it is shown in the aforementioned paper that D^H is regularly varying in the case X is itself regularly varying and with noninteger $\alpha \in (0, 2)$. In the aforementioned paper, three cases are distinguished, with various assumptions on the relation between X and L_A. Note that we do not cover the case $\alpha \in (0, 1)$ in Proposition (7.1), which is studied in [6].

In a recent contribution concerning PageRank, Theorem 4.2 in [48] provides similar asymptotics as in Theorem 4.2, that can be specialised to our case when X and K_A are allowed to have any form of dependence but one has a negligible tail with respect to the other. The aforementioned theorem also applies to intermediate regularly varying X and K_A. The main connection and specialisation is the following one:

1. if K_A is regularly varying with index $\alpha > 1$ and $\mathbb{E}\left[X^{\alpha+\epsilon}\right] < \infty$ for some $\epsilon > 0$, and if $\mathbb{P}\left(X > x\right) = o(\mathbb{P}\left(K_A > x\right))$, as $x \to \infty$, then

$$\mathbb{P}\left(D^H > x\right) \sim \mathbb{E}\left[K_A\right] \mathbb{P}\left(X > x\right) + \mathbb{P}\left(\mathbb{E}\left[X\right] K_A > x\right), \text{ as } x \to \infty;$$

2. if X is regularly varying with index $\alpha > 1$ and $\mathbb{E}\left[K_A^{\alpha+\epsilon}\right] < \infty$ for some $\epsilon > 0$, and if $\mathbb{P}\left(K_A > x\right) = o(\mathbb{P}\left(X > x\right))$, as $x \to \infty$, then

$$\mathbb{P}\left(D^H > x\right) \sim (1 + \mathbb{E}\left[K_A\right])\mathbb{P}\left(X > x\right), \text{ as } x \to \infty.$$

Hence, our result essentially extends the above, allowing for tail equivalence between X and K_A.

8. Precise large deviations of cluster process functionals

In this section, we make use of the cluster asymptotics from Section 4 to Section 7 to derive (precise) large deviation results for the renewal Poisson cluster process as well as for the Hawkes process.

Notation wise, we let

$$N_T = \left|\{(i,j) : 0 \leqslant \Gamma_i \leqslant T, 0 \leqslant \Gamma_i + T_{ij} \leqslant T\}\right|$$

represent the number of events occurring in the time interval $[0,T]$, for $T > 0$, and we let

$$J_T = \left|\{(i,j) : 0 \leqslant \Gamma_i \leqslant T, T \leqslant \Gamma_i + T_{ij}\}\right|$$

represent the number of (ordered) events coming from clusters that started in the time interval $[0,T]$, but occurring after time $T > 0$. We will also need the following decomposition of the maximum: for $x > 0$

$$\left\{\max_{1\leqslant i\leqslant C_T} H_i - \max_{1\leqslant j\leqslant J_T} X_j > x\right\} \subseteq \left\{\max_{1\leqslant i\leqslant N_T} X_i > x\right\} \subseteq \left\{\max_{1\leqslant i\leqslant C_T} H_i > x\right\} \tag{14}$$

where $C_T \sim \text{Poisson}(\nu T)$ is the number of clusters starting in the interval $[0,T]$, for $T > 0$, and H_i is as in Equation (1). This is due to the fact that the immigration process is the classical homogeneous Poisson process with parameter $\nu > 0$, see Section 2. The upper bounding set in decompositions (14) overshoots by taking the maximum over all the events belonging to clusters initiated before time $T > 0$, i.e. this includes events occurring after time $T > 0$. This is convenient, since C_T and H are independent.

The precise large deviation results for the sum will necessitate another decomposition. Notation wise, rewriting Equation (3) using N_T yields:

$$S_T := \sum_{j=1}^{N_T} X_j,$$

and we let μ_{S_T} denote the expectation of S_T. Then we can decompose the deviation as:

$$\begin{aligned} S_T - \mu_{S_T} &= \sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] - \left(\sum_{j=1}^{J_T} f(A_j) - \mathbb{E}\left[\sum_{j=1}^{J_T} f(A_j)\right]\right) \\ &=: \sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] - \left(\varepsilon_T - \mathbb{E}\left[\varepsilon_T\right]\right). \end{aligned} \tag{15}$$

As in decomposition (14), the first difference overshoots by summing marks of all events belonging to clusters started before $T > 0$, and removing the left-over effect of events occurring after time $T > 0$ in a second step, denoted by ε_T. Again, note that C_T and D are independent.

Furthermore, regarding the left-over effect, the following properties hold:

1. (Property 1) in [5], for both the renewal Poisson cluster process and the Hawkes process, that $\mathbb{E}[J_T] = o(T)$, as $T \to \infty$;

2. (Property 2) in [6], for both the renewal Poisson cluster process and the Hawkes process, that $\mathbb{E}[\varepsilon_T] = o(\sqrt{T})$, as $T \to \infty$; and hence, in our settings, the condition $\mathbb{E}[\varepsilon_T] = o(T)$, as $T \to \infty$ holds as well.

8.1. Large deviations of maxima over an interval $[0, T]$

We now illustrate how the asymptotics of Proposition (4.1) and Proposition (6.1) help to determine the asymptotic behaviour of the whole processes on an interval. In what follows, we let H denote a generic maximum, i.e. it can either be H^R or H^D from Section 4 and Section 6. At the end of the section, we present some related work.

Proposition 8.1. *Suppose that the conditions of either Proposition* (4.1) *or those of Proposition* (6.1) *hold. Then, as $T \to \infty$, and for any $\gamma > 0$*

$$\lim_{T\to\infty} \sup_{x \geqslant \gamma\nu T} \left| \frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant N_T} X_i > x\right)}{\mathbb{E}[N_T]\,\mathbb{P}(X > x)} - 1 \right| = 0.$$

Proof of Proposition (8.1). Using decomposition (14)

$$\frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i - \max_{1\leqslant j\leqslant J_T} X_i > x\right)}{\mathbb{E}[N_T]\,\mathbb{P}(X > x)} \leqslant \frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant N_T} X_i > x\right)}{\mathbb{E}[N_T]\,\mathbb{P}(X > x)} \leqslant \frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i > x\right)}{\mathbb{E}[N_T]\,\mathbb{P}(X > x)}.$$

Upper bound: By the remark following Theorem 3.1 in [32] for any $\gamma > 0$,

$$\lim_{T\to\infty} \sup_{x \geqslant \gamma\nu T} \left| \frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i > x\right)}{\mathbb{E}[C_T]\,\mathbb{P}(H > x)} - 1 \right| = 0, \text{ as } T \to \infty.$$

Using the asymptotics of Proposition (4.1) and of Proposition (6.1),

$$\mathbb{E}[C_T]\,\mathbb{P}(H > x) \sim \mathbb{E}[N_T]\,\mathbb{P}(X > x), \text{ as } T \to \infty$$

for the x-values considered, i.e. when $x \geqslant \gamma\nu T$ for any $\gamma > 0$.

Lower bound:

$$\begin{aligned}\frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i - \max_{1\leqslant j\leqslant J_T} X_j > x\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)} &= \frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i - \max_{1\leqslant j\leqslant J_T} X_j > x, \max_{1\leqslant j\leqslant J_T} X_j \leqslant x\varepsilon\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)} \\ &\quad + \frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i - \max_{1\leqslant j\leqslant J_T} X_j > x, \max_{1\leqslant j\leqslant J_T} X_j > x\varepsilon\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)} \\ &\geqslant \frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i > x(1+\varepsilon), \max_{1\leqslant j\leqslant J_T} X_j \leqslant x\varepsilon\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)} \\ &\geqslant \frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i > x(1+\varepsilon)\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)} \\ &\quad - \frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i > x(1+\varepsilon), \max_{1\leqslant j\leqslant J_T} X_j > x\varepsilon\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)}.\end{aligned}$$

The very last term in the lower bound is bounded above by

$$\frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i > x(1+\varepsilon), \max_{1\leqslant j\leqslant J_T} X_j > x\varepsilon\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)} \leqslant \frac{\mathbb{P}\left(\max_{1\leqslant j\leqslant J_T} X_j > x\varepsilon\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)}.$$

Conditioning on the values of J_T, using a union bound and the fact that the X_js are independent,

$$\begin{aligned}\mathbb{P}\left(\max_{1\leqslant j\leqslant J_T} X_j > x\varepsilon\right) &= \sum_{k=1}^{\infty}\mathbb{P}\left(\max_{1\leqslant j\leqslant k} X_j > x\varepsilon\right)\mathbb{P}\left(J_T=k\right) \\ &\leqslant \sum_{k=1}^{\infty}\sum_{j=1}^{k}\mathbb{P}\left(X_j > x\varepsilon\right)\mathbb{P}\left(J_T=k\right) \\ &\leqslant \sum_{k=1}^{\infty} k\mathbb{P}\left(X > x\varepsilon\right)\mathbb{P}\left(J_T=k\right) \\ &\leqslant \mathbb{E}\left[J_T\right]\mathbb{P}\left(X>x\varepsilon\right).\end{aligned}$$

Using Property (1) above, and Remark (8.1), which essentially says that $\mathbb{E}\left[N_T\right] = \mathcal{O}(T)$, as $T\to\infty$, and under the assumption that $x \geqslant \gamma\nu T$ for every $\gamma > 0$, it holds that $T\mathbb{P}\left(X>x\right) \to 0$ as $T\to\infty$, and it follows that, for any fixed $\epsilon > 0$,

$$\frac{\mathbb{P}\left(\max_{1\leqslant j\leqslant J_T} X_j > x\varepsilon\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)} = o(1), \text{ as } T\to\infty.$$

This implies that

$$\frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i - \max_{1\leqslant j\leqslant J_T} X_j > x\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)} \geqslant \frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i > x(1+\varepsilon)\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X>x\right)}.$$

Using again the remark following Theorem 3.1 in [32], it follows, for any $x \geqslant \gamma\nu T$, that

$$\lim_{T\to\infty}\sup_{x\geqslant\gamma\nu T}\left|\frac{\mathbb{P}\left(\max_{1\leqslant i\leqslant C_T} H_i > x(1+\varepsilon)\right)}{\mathbb{E}\left[C_T\right]\mathbb{P}\left(H > x(1+\varepsilon)\right)} - 1\right| = 0, \text{ as } T\to\infty.$$

Because H is regularly varying with index $\alpha > 1$, it follows that

$$\mathbb{E}\left[C_T\right]\mathbb{P}\left(H > x(1+\varepsilon)\right) = (1+\varepsilon)^{-\alpha}\mathbb{E}\left[C_T\right]\mathbb{P}\left(H > x\right), \text{ as } x \to \infty,$$

and using the asymptotics of Proposition (4.1) and of Proposition (6.1),

$$\mathbb{E}\left[C_T\right]\mathbb{P}\left(H > x(1+\varepsilon)\right) \sim (1+\varepsilon)^{-\alpha}\mathbb{E}\left[N_T\right]\mathbb{P}\left(X > x\right), \text{ as } T \to \infty.$$

Letting $\epsilon \to 0$, collecting the upper and lower bounds yields the desired result. □

Remark 8.1. Note that, by the independence of the clusters, we have:

1. for the renewal Poisson cluster process, $\mathbb{E}\left[N_T\right] = (\mathbb{E}\left[K_A\right]+1)\nu T$;
2. for the Hawkes process, $\mathbb{E}\left[N_T\right] = \frac{\nu T}{1-\mathbb{E}[\kappa_A]}$ (see e.g. Section 12.1 in [8]).

8.2. Large deviations of sums over an interval $[0, T]$

We finally illustrate how the results of Proposition (5.1) and Proposition (7.1) help to derive results for the renewal Poisson cluster process as well as for the Hawkes on an interval $[0, T]$. Note that D denotes a generic sum of the marks, i.e. it can either be D^H or D^R until specified.

Proposition 8.2. *1. Suppose the conditions of Proposition* (5.1) *hold. Then, as $T \to \infty$, for all $\gamma > 0$,*

$$\lim_{T\to\infty} \sup_{x \geqslant \gamma\nu T} \left| \frac{\mathbb{P}\left(S_T - \mu_{S_T} > x\right)}{\nu T\big(\mathbb{P}\left(X + \mathbb{E}\left[X\right]K_A > x\right) + \mathbb{E}\left[K_A\right]\mathbb{P}\left(X > x\right)\big)} - 1 \right| = 0.$$

2. Suppose the conditions of Proposition (7.1) *hold. Then, as $T \to \infty$, for all $\gamma > 0$,*

$$\lim_{T\to\infty} \sup_{x \geqslant \gamma\nu T} \left| \frac{\mathbb{P}\left(S_T - \mu_{S_T} > x\right)}{\mathbb{E}\left[N_T\right]\mathbb{P}\left(X + \left(\frac{\mathbb{E}[X]}{1-\mathbb{E}[\kappa_A]}\right)\kappa_A > x\right)} - 1 \right| = 0.$$

Proof of Proposition (8.2). We use decomposition (15), i.e.

$$\mathbb{P}\left(S_T - \mu_{S_T} > x\right) = \mathbb{P}\left(\sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] - \big(\varepsilon_T - \mathbb{E}\left[\varepsilon_T\right]\big) > x\right).$$

Upper bound: Note that

$$\mathbb{P}\left(S_T - \mu_{S_T} > x\right) \leqslant \mathbb{P}\left(\sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] > x - \mathbb{E}\left[\varepsilon_T\right]\right).$$

As $T \to \infty$, we can rewrite $x \geqslant \gamma\nu T$ as $x \geqslant \gamma^{'}\nu T + \mathbb{E}[\varepsilon_T]$, for some $0 < \gamma^{'} < \gamma$. Hence, under the assumption that $x \geqslant \gamma^{'}\nu T + \mathbb{E}[\varepsilon_T]$, then $x - \mathbb{E}[\varepsilon_T] \geqslant \gamma^{'}\nu T$, and since $C_T \sim \text{Poisson}(\nu T)$ is independent of D, using Lemma 2.1 and Theorem 3.1 in [32] yields

$$\lim_{T\to\infty} \sup_{x \geqslant \gamma^{'}\nu T} \left| \frac{\mathbb{P}\left(\sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] > x - \mathbb{E}[\varepsilon_T]\right)}{\nu T \mathbb{P}(D > x - \mathbb{E}[\varepsilon_T])} - 1 \right| = 0.$$

Recall that D is regularly varying with index $\alpha > 1$. Using Property (2) above, we can write $x - \mathbb{E}[\varepsilon_T] = x - o(T)$ as $T \to \infty$. Using the Potter bounds (see Theorem 1.5.6 in [7]), for all $I > 1$, $\eta > 0$, there exists X such that, for all $x - o(T) \geqslant X$,

$$\frac{\mathbb{P}(D > x - o(T))}{\mathbb{P}(D > x)} \leqslant I \max\left\{ \left(1 - \frac{o(T)}{x}\right)^{-\alpha+\eta}, \left(1 - \frac{o(T)}{x}\right)^{-\alpha+\eta} \right\}.$$

Because $x \geqslant \gamma\nu T + \mathbb{E}[\varepsilon_T]$, the above upper bound becomes uniformly close to 1, as $T \to \infty$. In combination with the above, it follows that, as $T \to \infty$, uniformly for $x \geqslant \gamma^{'}\nu T + \mathbb{E}[\varepsilon_T]$,

$$\mathbb{P}(S_T - \mu_{S_T} > x) \leqslant \nu T \mathbb{P}(D > x).$$

Lower bound: Let $\delta > 0$, and note that

$$\begin{aligned}
\mathbb{P}(S_T - \mu_{S_T} > x) &= \mathbb{P}\left(\sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] - \left(\varepsilon_T - \mathbb{E}[\varepsilon_T]\right) > x,\ \varepsilon_T - \mathbb{E}[\varepsilon_T] \leqslant x\delta\right) \\
&\quad + \mathbb{P}\left(\sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] - \left(\varepsilon_T - \mathbb{E}[\varepsilon_T]\right) > x,\ \varepsilon_T - \mathbb{E}[\varepsilon_T] > x\delta\right) \\
&\geqslant \mathbb{P}\left(\sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] - \left(\varepsilon_T - \mathbb{E}[\varepsilon_T]\right) > x,\ \varepsilon_T - \mathbb{E}[\varepsilon_T] \leqslant x\delta\right) \\
&\geqslant \mathbb{P}\left(\sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] > x(1+\delta)\right) \\
&\quad - \mathbb{P}\left(\sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] - \left(\varepsilon_T - \mathbb{E}[\varepsilon_T]\right) > x,\ \varepsilon_T - \mathbb{E}[\varepsilon_T] > x\delta\right) \\
&\geqslant \mathbb{P}\left(\sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] > x(1+\delta)\right) - \mathbb{P}(\varepsilon_T - \mathbb{E}[\varepsilon_T] > x\delta).
\end{aligned}$$

By taking $x \geqslant \gamma^{'}\nu T + \mathbb{E}[\varepsilon_T]$, it follows that $x\delta \geqslant \gamma^{'}(\nu T)\delta + \mathbb{E}[\varepsilon_T]\delta$, and hence, using Markov's inequality

$$\mathbb{P}(\varepsilon_T - \mathbb{E}[\varepsilon_T] > x\delta) \leqslant \mathbb{P}\left(\varepsilon_T \geqslant \gamma^{'}\nu T\delta + (\delta+1)\mathbb{E}[\varepsilon_T]\right) \leqslant \frac{\mathbb{E}[\varepsilon_T]}{\gamma^{'}\nu T\delta + (\delta+1)\mathbb{E}[\varepsilon_T]}.$$

Using Property (2) above, letting $T \to \infty$, for a fixed $\delta > 0$, the denominator in the upper bound goes to infinity, and this implies that

$$\lim_{T\to\infty} \mathbb{P}\left(\varepsilon_T - \mathbb{E}\left[\varepsilon_T\right] > x\delta\right) = 0 \text{ for } x \geqslant \gamma^{'}\nu T + \mathbb{E}\left[\varepsilon_T\right].$$

Since $x \geqslant \gamma^{'}\nu T + \mathbb{E}\left[\varepsilon_T\right] \geqslant \gamma\nu T$, using again Theorem 3.1 in [32], it follows that

$$\lim_{T\to\infty} \sup_{x\geqslant\gamma\nu T} \left| \frac{\mathbb{P}\left(\sum_{i=1}^{C_T} D_i - \mathbb{E}\left[\sum_{i=1}^{C_T} D_i\right] > x(1+\delta)\right)}{\nu T \mathbb{P}\left(D > x(1+\delta)\right)} - 1 \right| = 0.$$

Since D is regularly varying with index $\alpha > 1$, letting $\delta \to 0$ yields

$$\nu T \mathbb{P}\left(D > x(1+\delta)\right) \sim \nu T (1+\delta)^{-\alpha} \mathbb{P}\left(D > x\right) \sim \nu T \mathbb{P}\left(D > x\right).$$

It follows that, uniformly for $x \geqslant \gamma\nu T$, and as $T \to \infty$,

$$\nu T \mathbb{P}\left(D > x\right) \leqslant \mathbb{P}\left(S_T - \mu_{S_T} > x\right).$$

Collecting the above upper and lower bounds,

1. the asymptotics of Proposition (5.1) yields

$$\lim_{T\to\infty} \sup_{x\geqslant\gamma\nu T} \left| \frac{\mathbb{P}\left(S_T - \mu_{S_T} > x\right)}{\nu T\big(\mathbb{P}\left(X + \mathbb{E}\left[X\right] K_A > x\right) + \mathbb{E}\left[K_A\right] \mathbb{P}\left(X > x\right)\big)} - 1 \right| = 0.$$

2. the asymptotics of Proposition (7.1) yields

$$\lim_{T\to\infty} \sup_{x\geqslant\gamma\nu T} \left| \frac{\mathbb{P}\left(S_T - \mu_{S_T} > x\right)}{\mathbb{E}\left[N_T\right] \mathbb{P}\left(X + \left(\frac{\mathbb{E}[X]}{1-\mathbb{E}[\kappa_A]}\right)\kappa_A > x\right)} - 1 \right| = 0,$$

and recalling that $\frac{\nu T}{1-\mathbb{E}[\kappa_A]} = \mathbb{E}\left[N_T\right]$, this concludes the proof.

□

Remark 8.2. Early contributions to the (non-uniform) precise large deviations results for non-random sums of i.i.d. regularly varying random variables can be found in [42], [43], [24], or [45].

The proofs of Proposition (8.1) and Proposition (8.2) heavily rely on the work of [32], in which the authors show that, under the assumption that the process of integer-valued non-negative random variables $(N_T)_{T>0}$ is such that

1. $N_T/\lambda_T \xrightarrow{\mathrm{P}} 1$, as $\lambda_T \to \infty$, where $\lambda_T = \mathbb{E}\left[N_T\right]$;

2. the following limit holds:

$$\sum_{k>(1+\delta)\lambda_T} \mathbb{P}\left(N_T > k\right)(1+\epsilon)^k \to 0, \text{ as } \lambda_T \to \infty.$$

Furthermore, if the process (N_T) is independent of the sequence (X_j), by their Theorem 3.1, if the distribution of X is extended regularly varying, for any $\gamma > 0$,

$$\lim_{T\to\infty} \sup_{x\geqslant\gamma\lambda_T} \left| \frac{\mathbb{P}\left(S_T - \mu_{S_T} > x\right)}{\lambda_T \mathbb{P}\left(X > x\right)} - 1 \right| = 0, \text{ and } \lim_{T\to\infty} \sup_{x\geqslant\gamma\lambda_T} \left| \frac{\mathbb{P}\left(\max_{1\leqslant j\leqslant N_T} X_j > x\right)}{\lambda_T \mathbb{P}\left(X > x\right)} - 1 \right| = 0$$

where $S_T = \sum_{j=1}^{N_T} X_j$. Note that the authors show that the Poisson process C_T satisfies the assumptions above, but the second condition is difficult to show for more complicated processes. Hence, the trick is to bound the processes at hand in this work by a process governed by an independent variable, in our context C_T which is Poisson distributed and satisfies the settings of [32].

Note that the work in [32] extends the precise large deviation principles already studied in [12] (in the case of non-random sums) to the case of random sums.

In [58], the authors relax the two assumptions used in [32] and mentioned above, and reduce them into the single condition that

$$\mathbb{E}\left[N_T^{\beta+\epsilon} 1_{\{N_T>(1+\delta)\lambda_T\}}\right] = \mathcal{O}(\lambda_T), \text{ as } T \to \infty,$$

for fixed $\epsilon, \delta > 0$ small and β the (upper) index of extended regular variation, and prove similar precise large deviation results as [32]. In [46], the authors study another subclass of the subexponential family, namely the consistently varying random variables, and prove similar precise large deviations under the same conditions as [58].

Under the assumption that the sequence (X_j) exhibits negative dependence, i.e.

$$\mathbb{P}\left(\bigcap_{j=1}^{n}\{X_j \leqslant x_j\}\right) \leqslant M \prod_{j=1}^{n} \mathbb{P}\left(X_j \leqslant x_j\right) \text{ and } \mathbb{P}\left(\bigcap_{j=1}^{n}\{X_j > x_j\}\right) \leqslant M \prod_{j=1}^{n} \mathbb{P}\left(X_j > x_j\right)$$

for some $M > 0$, all $x_1, \ldots, x_n \in \mathbb{R}$, more recent literature such as [57] or [35] propose extensions and similar results to those of [46] under the same consistently varying random variables.

While our framework is more restrictive on the aspect that our sequence $(X_j)_{1\leqslant j\leqslant N_T}$ has elements that are regularly varying, which is a subclass of the extended regularly varying distributions, and that furthermore the elements of the sequence are independent, knowledge of the tail asymptotics of the cluster functionals allowed us to derive expressions that resemble known precise large deviations principles for random maxima and sums of independent random variables, even though, clearly, N_T and (X_j) are

dependent over a time window $[0,T]$. This comes at the cost of an extra term, for the sums the marks over a finite time interval, of an extra left-over effect $\mathbb{E}\left[\varepsilon_T\right]$ that vanishes as T becomes large.

Appendix A.

Proof of Proposition (4.1)

Proof of Proposition (4.1). By conditioning and using the independence of X and X_j, $j \geq 1$, and that of K_A and X_j, $j \geq 1$, we obtain

$$\begin{aligned}\mathbb{P}\left(H^R > x\right) &= 1 - \sum_{k=0}^{\infty} \mathbb{P}\big(X \leqslant x \mid K_A = k\big)\big(\mathbb{P}\left(X \leqslant x\right)\big)^k \mathbb{P}\left(K = k\right) \\ &= 1 - \sum_{k=0}^{\infty} \mathbb{P}\big(X \leqslant x \mid K_A = k\big) \exp\big(k \log\big(1 - \mathbb{P}\left(X > x\right)\big)\big)\mathbb{P}\left(K = k\right). \end{aligned} \tag{16}$$

A Taylor expansion on the exponential term, as $x \to \infty$ (and hence, as $\mathbb{P}\left(X > x\right) \to 0$ by the integrability of X), gives

$$\begin{aligned}\exp\big(k \log\big(1 - \mathbb{P}\left(X > x\right)\big)\big) &= \exp\big(-k\mathbb{P}\left(X > x\right) - o\big(k\mathbb{P}\left(X > x\right)\big)\big) \\ &= \big(1 - k\mathbb{P}\left(X > x\right) + o\big(k\mathbb{P}\left(X > x\right)\big)\big) \exp\big(-o\big(k\mathbb{P}\left(X > x\right)\big)\big)\end{aligned}$$

where the last equality follows by another Taylor expansion of the first exponential term in the second equality, as $x \to \infty$.

Plugging the above expansion in Equation (16) yields

$$\begin{aligned}\mathbb{P}\left(H^R > x\right) &= 1 - \sum_{k=0}^{\infty} \mathbb{P}\left(X \leqslant x, K_A = k\right) \exp\big(-o\big(k\mathbb{P}\left(X > x\right)\big)\big) \\ &\quad + \mathbb{P}\left(X > x\right) \sum_{k=0}^{\infty} k\mathbb{P}\left(X \leqslant x, K_A = k\right) \exp\big(-o\big(k\mathbb{P}\left(X > x\right)\big)\big) \\ &\quad - o(\mathbb{P}\left(X > x\right)) \sum_{k=0}^{\infty} \mathbb{P}\left(X \leqslant x, K_A = k\right) \exp\big(-o\big(k\mathbb{P}\left(X > x\right)\big)\big) \\ &=: 1 - B_1 + B_2 - B_3.\end{aligned}$$

We treat each term separately. For term $1 - B_1$, remarking that $1 = \mathbb{P}\left(X \leqslant x\right) + \mathbb{P}\left(X > x\right)$, we obtain

$$1 - B_1 = \mathbb{P}\left(X > x\right) + \sum_{k=0}^{\infty} \mathbb{P}\left(X \leqslant x, K_A = k\right)\big(1 - \exp\big(-o(k\mathbb{P}\left(X > x\right))\big)\big)$$

Using the basic inequality $1-e^{-x}\leqslant x$, term $1-B_1$ is bounded by

$$0\leqslant 1-B_1\leqslant \mathbb{P}\left(X>x\right)+o(\mathbb{E}\left[K_A\right]\mathbb{P}\left(X>x\right)),\ \text{as } x\to\infty.$$

For term B_2, we can write

$$\begin{aligned}B_2&=\mathbb{P}\left(X>x\right)\sum_{k=0}^{\infty}k\big(\mathbb{P}\left(X\leqslant x,K_A=k\right)\exp\big(-o(k\mathbb{P}\left(X>x\right))\big)+\mathbb{P}\left(K=k\right)-\mathbb{P}\left(K=k\right)\big)\\&=\mathbb{P}\left(X>x\right)\mathbb{E}\left[K_A\right]-\mathbb{P}\left(X>x\right)\sum_{k=0}^{\infty}k\mathbb{P}\left(X>x,K_A=k\right)\\&\quad+\mathbb{P}\left(X>x\right)\sum_{k=0}^{\infty}k\mathbb{P}\left(X\leqslant x,K_A=k\right)\big(\exp\big(-o(k\mathbb{P}\left(X>x\right))\big)-1\big)\\&=:B_{21}-B_{22}+B_{23}.\end{aligned}$$

Note that B_{22} is bounded above by $B_{22}\leqslant\mathbb{E}\left[K_A\right]\mathbb{P}\left(X>x\right)$, and hence, by a dominated convergence argument and the integrability of X, we have that

$$\mathbb{P}\left(X>x\right)\sum_{k=0}^{\infty}k\mathbb{P}\left(X>x,K_A=k\right)=o\big(\mathbb{P}\left(X>x\right)\big),\ \text{as } x\to\infty.$$

For term B_{23}, which is negative since for all $k\geqslant 0$, $0\leqslant e^{-o(k\mathbb{P}(X>x))}\leqslant 1$, we bound it below by

$$\begin{aligned}&-\mathbb{P}\left(X>x\right)\sum_{k=0}^{\infty}k\mathbb{P}\left(X\leqslant x,K_A=k\right)\\&\qquad\leqslant\mathbb{P}\left(X>x\right)\sum_{k=0}^{\infty}k\mathbb{P}\left(X\leqslant x,K_A=k\right)\big(\exp\big(-o(k\mathbb{P}\left(X>x\right))\big)-1\big)\\&\qquad\leqslant 0,\end{aligned}$$

and hence, by a dominated convergence argument, we obtain that, as $x\to\infty$,

$$\mathbb{P}\left(X>x\right)\sum_{k=0}^{\infty}k\mathbb{P}\left(X\leqslant x,K_A=k\right)\big(\exp\big(-o(k\mathbb{P}\left(X>x\right))\big)-1\big)=o\big(\mathbb{P}\left(X>x\right)\big).$$

Collecting the above results, we see that, essentially,

$$B_2=\mathbb{P}\left(X>x\right)\mathbb{E}\left[K_A\right]+o\big(\mathbb{P}\left(X>x\right)\big),\ \text{as } x\to\infty.$$

Finally, by very similar arguments to those employed for B_2 and omitted for brevity,

$$B_3=o\big(\mathbb{P}\left(X>x\right)\big),\ \text{as } x\to\infty.$$

Collecting the above, it essentially follows that

$$P(H^R > x) = \mathbb{P}\left(X > x\right) + \mathbb{P}\left(X > x\right)\mathbb{E}\left[K_A\right] + o\big(\mathbb{P}\left(X > x\right)\big), \text{ as } x \to \infty.$$

The desired result follows at once by taking the limit, as $x \to \infty$, and upon using the assumption that the limiting Radon measure is non-null on the subspace $\{(x_1, x_2) \in \mathbb{R}^2_{+,\mathbf{0}} : x_1 > 1\}$, which implies by means of Example (3.1) that H^R is regularly varying with index $\alpha > 1$. □

Appendix B.

Proof of Proposition (6.1)

Proof of Proposition (6.1). By conditioning and using the independence of X and H^H, and that of L_A and H^H, we obtain as in the proof of Proposition (4.1)

$$\mathbb{P}\left(H^H > x\right) = 1 - \sum_{k=0}^{\infty} \mathbb{P}\big(X \leqslant x \mid L_A = k\big) \exp\left(k \log\left(1 - \mathbb{P}\left(H^H > x\right)\right)\right)\mathbb{P}\left(L_A = k\right).$$

A Taylor expansion on the exponential term, as $x \to \infty$ (and hence, as $\mathbb{P}\left(H^H > x\right) \to 0$ by the integrability of H^H), yields, as $x \to \infty$,

$$\exp\left(k \log\left(1 - \mathbb{P}\left(H^H > x\right)\right)\right)$$
$$= \big(1 - k\mathbb{P}\left(H^H > x\right) + o\big(k\mathbb{P}\left(H^H > x\right)\big)\big) \exp\big(-o\big(k\mathbb{P}\left(H^H > x\right)\big)\big).$$

From here on, the proof follows the same lines as that of Proposition (5.1), except that the tail of H^H appears here rather than the tail of X. The proof is omitted for brevity, but we retrieve

$$P(H^H > x) = \mathbb{P}\left(X > x\right) + \mathbb{E}\left[L_A\right]\mathbb{P}\left(H^H > x\right) + o\big(\mathbb{P}\left(H^H > x\right)\big), \text{ as } x \to \infty$$

which yields the desired result. □

Appendix C.

Proof of Proposition (5.1)

We need the following Lemma in order to prove Lemma (C.2) used in the proof of Proposition (5.1):

Lemma C.1. *Suppose (X, K_A) is regularly varying with index $\alpha \in (n, n+1)$, for $n \in \mathbb{N}$. Additionally, suppose that X has negligible tails with respect to $X + \mathbb{E}\left[X\right]K_A$, i.e. $\mathbb{P}\left(X > x\right) = o(\mathbb{P}\left(X + \mathbb{E}\left[X\right]K_A > x\right))$, as $x \to \infty$. Then,*

$$\varphi_X^{(n+1)}(s) = o\big(\varphi_{X+\mathbb{E}[X]K_A}^{(n+1)}(s)\big), \text{ as } s \to 0^+.$$

Proof. Note that

$$\varphi_X^{(n+1)}(s) = \mathbb{E}\left[(-X)^{n+1}e^{-sX}\right] = \int_0^\infty x^{n+1}e^{-sx}\,\mathrm{d}(-\mathbb{P}\left(X \geqslant x\right))$$
$$= \left[-x^{n+1}e^{-sx}\mathbb{P}\left(X \geqslant x\right)\right]_0^\infty + \int_0^\infty \left((n+1)x^n e^{-sx} - sx^{n+1}e^{-sx}\right)\mathbb{P}\left(X \geqslant x\right)\mathrm{d}x.$$

The first term above vanishes; upon substituting, the second term yields

$$\int_0^\infty \left((n+1)x^n e^{-sx} - sx^{n+1}e^{-sx}\right)\mathbb{P}\left(X \geqslant x\right)\mathrm{d}x$$
$$= \int_0^\infty \left((n+1)(y/s)^n e^{-y} - s(y/s)^{n+1}e^{-y}\right)\mathbb{P}\left(X \geqslant y/s\right)\frac{\mathrm{d}y}{s}$$
$$= s^{-(n+1)}\int_0^\infty \left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\mathbb{P}\left(X \geqslant y/s\right)\mathrm{d}y.$$

Fix $\varepsilon > 0$ small and split the above integral into

$$s^{-(n+1)}\int_0^\infty \left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\mathbb{P}\left(X \geqslant y/s\right)\mathrm{d}y$$
$$= s^{-(n+1)}\Bigg(\int_0^\varepsilon \left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\mathbb{P}\left(X \geqslant y/s\right)\mathrm{d}y$$
$$+ \int_\varepsilon^\infty \left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\mathbb{P}\left(X \geqslant y/s\right)\mathrm{d}y\Bigg) =: I_1 + I_2.$$

Consider integral I_2 first. For some values $y \in [\varepsilon, \infty)$ the expression $(n+1)y^n e^{-y} - y^{n+1}e^{-y}$ might be negative, so bound I_2 above by its absolute value. Additionally, upon using the hypothesis of negligibility of the tail of X with respect to the tail of $X + \mathbb{E}\left[X\right]K_A$, it follows that, for any $\delta > 0$, for any fixed $\varepsilon > 0$ and $y > \varepsilon$, there is s_0 such that for all $s \leqslant s_0$, $\mathbb{P}\left(X \geqslant y/s\right) \leqslant \delta_\varepsilon \mathbb{P}\left(X + \mathbb{E}\left[X\right]K_A > y/s\right)$. All in all, because $X + \mathbb{E}\left[X\right]K_A$ is regularly varying with index $\alpha \in (n, n+1)$, this yields as an upper bound

$$|I_2| \leqslant s^{-(n+1)}\int_\varepsilon^\infty \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right|\delta_\varepsilon \mathbb{P}\left(X + \mathbb{E}\left[X\right]K_A > y/s\right)\mathrm{d}y$$
$$\leqslant s^{-(n+1)}\int_\varepsilon^\infty \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right|\delta_\varepsilon (y/s)^{-\alpha} L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y$$
$$\leqslant s^{\alpha-(n+1)}\delta_\varepsilon \int_\varepsilon^\infty \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| y^{-\alpha} L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y.$$

Because $\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)y^{-\alpha}$ is integrable over $[0, \infty)$, it follows from Proposition 4.1.2 (b) in [7] that, as $s \to 0^+$,

$$s^{\alpha-(n+1)}\delta_\varepsilon \int_\varepsilon^\infty \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| y^{-\alpha} L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y$$
$$\sim s^{\alpha-(n+1)} L_{X+\mathbb{E}[X]K_A}(1/s)\delta_\varepsilon \int_\varepsilon^\infty \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| y^{-\alpha}\,\mathrm{d}y.$$

For each fixed value of $\varepsilon > 0$, and as $s \to 0^+$, it is possible to take $\delta_\varepsilon > 0$ as small as needed so that to guarantee that

$$\delta_\varepsilon \int_\varepsilon^\infty \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| y^{-\alpha}\,\mathrm{d}y = o(1) \text{ as } s \to 0^+.$$

This implies that, for a fixed $\epsilon > 0$, as $s \to 0^+$,

$$\begin{aligned}\frac{|I_2|}{s^{\alpha-(n+1)}L_{X+\mathbb{E}[X]K_A}(1/s)} &\leqslant \frac{s^{\alpha-(n+1)}L_{X+\mathbb{E}[X]K_A}(1/s)\delta_\varepsilon \int_\varepsilon^\infty \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| y^{-\alpha}\,\mathrm{d}y}{s^{\alpha-(n+1)}L_{X+\mathbb{E}[X]K_A}(1/s)}\\ &\leqslant o(1).\end{aligned}$$

Consider now integral I_1. Because X is stochastically dominated by $X+\mathbb{E}\left[X\right]K_A$, and using the regular variation of the latter quantity, this yields

$$\begin{aligned}|I_1| &\leqslant s^{-(n+1)} \int_0^\varepsilon \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| \mathbb{P}\left(X + \mathbb{E}\left[X\right]K_A > y/s\right)\mathrm{d}y\\ &\leqslant s^{\alpha-(n+1)} \int_0^\varepsilon \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| y^{-\alpha} L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y.\end{aligned}$$

Because the function $\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)y^{-\alpha}$ is integrable over $[0,\varepsilon)$, it follows by Proposition 4.1.2 (a) in [7] that, as $s \to 0^+$,

$$\begin{aligned}&s^{\alpha-(n+1)} \int_0^\varepsilon \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| y^{-\alpha} L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y\\ &\qquad\qquad \sim s^{\alpha-(n+1)}L_{X+\mathbb{E}[X]K_A}(1/s) \int_0^\varepsilon \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| y^{-\alpha}\,\mathrm{d}y.\end{aligned}$$

It follows that

$$\begin{aligned}\frac{|I_1|}{s^{\alpha-(n+1)}L_{X+\mathbb{E}[X]K_A}(1/s)} &\leqslant \frac{s^{\alpha-(n+1)}L_{X+\mathbb{E}[X]K_A}(1/s)\int_0^\varepsilon \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| y^{-\alpha}\,\mathrm{d}y}{s^{\alpha-(n+1)}L_{X+\mathbb{E}[X]K_A}(1/s)}\\ &\leqslant \int_0^\varepsilon \left|\left((n+1)y^n e^{-y} - y^{n+1}e^{-y}\right)\right| y^{-\alpha}\,\mathrm{d}y\end{aligned}$$

and, because one can take $\varepsilon > 0$ as small as needed, this shows that

$$I_1 + I_2 = o\big(s^{\alpha-(n+1)}L_{X+\mathbb{E}[X]K_A}(1/s)\big), \text{ as } s \to 0^+.$$

Finally, because $n+1 = \lceil\alpha\rceil$, and using Karamata's Tauberian Theorem (3.1), which implies that $\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A} \sim C_\alpha s^{\alpha-\lceil\alpha\rceil}L_{X+\mathbb{E}[X]K_A}(1/s)$, as $s \to 0^+$, this shows that

$$\varphi^{(n+1)}_X(s) = o(\varphi^{(n+1)}_{X+\mathbb{E}[X]K_A}), \text{ as } s \to 0^+.$$

□

Lemma C.2. *Suppose (X, K_A) is regularly varying with index $\alpha \in (1,2)$ and slowly varying function $L_{X+\mathbb{E}[X]K_A}(\cdot)$ and X has a negligible tail compared to the modulus $X + \mathbb{E}[X]K_A$, i.e. $\mathbb{P}(X > x) = o(\mathbb{P}(X + \mathbb{E}[X]K_A > x))$, as $x \to \infty$. Then,*

$$\frac{\varphi_X^{(1)}(s) + \mathbb{E}[X]}{s} = o\big(\varphi^{(2)}_{X+\mathbb{E}[X]K_A}(s) + \mathbb{E}[K_A]\,\varphi_X^{(2)}(s)\big), \text{ as } s \to 0^+.$$

Proof. Let $\alpha \in (1,2)$. We assess

$$\begin{aligned}\frac{\varphi_X^{(1)}(s) + \mathbb{E}[X]}{s} &= \mathbb{E}\left[\frac{X\big(1-e^{-sX}\big)}{s}\right]\\ &= \int_0^\infty \frac{x(1-e^{-sx})}{s}\,\mathrm{d}(-\mathbb{P}(X \geqslant x))\\ &= \left[-\frac{x(1-e^{-sx})}{s}\mathbb{P}(X \geqslant x)\right]_0^\infty + \int_0^\infty \left(\frac{\big(1-e^{-sx}\big)}{s} + xe^{-sx}\right)\mathbb{P}(X \geqslant x)\,\mathrm{d}x.\end{aligned}$$

Since X is integrable, one has that $x\mathbb{P}(X \geqslant x) = o(1)$, as $x \to \infty$, so that the first expression on the right-hand side above vanishes; for the second integral, fix $\varepsilon > 0$ small and write

$$\begin{aligned}\int_0^\infty \left(\frac{\big(1-e^{-sx}\big)}{s} + xe^{-sx}\right)\mathbb{P}(X \geqslant x)\,\mathrm{d}x &= \int_0^\infty s^{-2}(1-e^{-y}+ye^{-y})\mathbb{P}(X \geqslant y/s)\,\mathrm{d}y,\\ &= \int_0^\varepsilon s^{-2}(1-e^{-y}+ye^{-y})\mathbb{P}(X \geqslant y/s)\,\mathrm{d}y\\ &\quad + \int_\varepsilon^\infty s^{-2}(1-e^{-y}+ye^{-y})\mathbb{P}(X \geqslant y/s)\,\mathrm{d}y\\ &=: (I_1 + I_2).\end{aligned}$$

Consider integral I_2 first. A similar argument as in the proof of Lemma (C.1) for integral I_2 there yields the following upper bound

$$\begin{aligned}I_2 &\leqslant \int_\varepsilon^\infty s^{-2}(1-e^{-y}+ye^{-y})\delta_\varepsilon \mathbb{P}(X + \mathbb{E}[X]K_A > y/s)\,\mathrm{d}y\\ &\leqslant s^{\alpha-2}\delta_\varepsilon \int_\varepsilon^\infty (1-e^{-y}+ye^{-y})y^{-\alpha}L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y.\end{aligned}$$

As $\varepsilon \to 0$, the above integral diverges. But for a (small) fixed value of $\varepsilon > 0$, upon using Proposition 4.1.2 (b) in [7], as $s \to 0^+$,

$$\begin{aligned}&s^{\alpha-2}\int_\varepsilon^\infty (1-e^{-y}+ye^{-y})y^{-\alpha}\delta_\varepsilon L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y\\ &\qquad\qquad \sim s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)\delta_\varepsilon \int_\varepsilon^\infty (1-e^{-y}+ye^{-y})y^{-\alpha}\,\mathrm{d}y.\end{aligned}$$

As $s \to 0^+$, and as in the proof of Lemma (C.1), it is possible to take $\delta_\varepsilon > 0$ as small as needed in order to ensure that

$$\delta_\varepsilon \int_\varepsilon^\infty (1 - e^{-y} + ye^{-y})y^{-\alpha}\,\mathrm{d}y = o(1) \text{ as } s \to 0^+.$$

This implies that, as $s \to 0^+$,

$$\frac{I_2}{s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)} \leqslant \frac{s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)\delta_\varepsilon \int_\varepsilon^\infty (1 - e^{-y} + ye^{-y})y^{-\alpha}\,\mathrm{d}y}{s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)} \leqslant o(1).$$

Consider now integral I_1. Because X is stochastically dominated by $X + \mathbb{E}[X]K_A$, for any fixed $\varepsilon > 0$, we have

$$\begin{aligned} I_1 &\leqslant \int_0^\varepsilon s^{-2}(1 - e^{-y} + ye^{-y})\mathbb{P}(X + \mathbb{E}[X]K_A \geqslant y/s)\,\mathrm{d}y \\ &= \int_0^\varepsilon s^{\alpha-2}(1 - e^{-y} + ye^{-y})y^{-\alpha}L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y. \end{aligned}$$

A Taylor expansion on the function $f(y) = e^{-y} + ye^{-y}$ yields $1 = e^{-y} - ye^{-y} + 2ye^{-y} - y^2e^{-y} + o(-y)$, and we get that

$$\int_0^\varepsilon s^{\alpha-2}(1 - e^{-y} + ye^{-y})y^{-\alpha}L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y \approx s^{\alpha-2}\int_0^\varepsilon (2ye^{-y} - y^2e^{-y})y^{-\alpha}L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y.$$

Because the integral $\int_0^\varepsilon (2ye^{-y} - y^2e^{-y})y^{-\alpha}\,\mathrm{d}y < \infty$ for $\alpha \in (1,2)$ and $\varepsilon > 0$ small, even if it is potentially large for values of α close to 2, it follows from Proposition 4.1.2. (a) in [7] that, as $s \to 0^+$,

$$s^{\alpha-2}\int_0^\varepsilon (2ye^{-y} - y^2e^{-y})y^{-\alpha}L_{X+\mathbb{E}[X]K_A}(y/s)\,\mathrm{d}y \sim s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)\int_0^\varepsilon (2ye^{-y} - y^2e^{-y})y^{-\alpha}\,\mathrm{d}y.$$

Hence, as $s \to 0^+$

$$\begin{aligned} \frac{I_1}{s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)} &\leqslant \frac{s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)\int_0^\varepsilon (2ye^{-y} - y^2e^{-y})y^{-\alpha}\,\mathrm{d}y}{s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)} \\ &\leqslant \int_0^\varepsilon (2ye^{-y} - y^2e^{-y})y^{-\alpha}\,\mathrm{d}y \end{aligned}$$

and because one can take $\varepsilon > 0$ as small as needed, it essentially follows, all in all, that

$$(I_1 + I_2) = o(s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)), \text{ as } s \to 0^+.$$

Because $X + \mathbb{E}[X]K_A$ is regularly varying, by Karamata's Tauberian Theorem (3.1),

$$\frac{s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)}{\varphi^{(2)}_{X+\mathbb{E}[X]K_A}(s) + \mathbb{E}[K_A]\,\varphi^{(2)}_X(s)} \sim \frac{s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s)}{C_\alpha s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s) + \mathbb{E}[K_A]\,\varphi^{(2)}_X(s)}, \text{ as } s \to 0^+.$$

Applying the result of Lemma (C.1), this yields that

$$s^{\alpha-2}L_{X+\mathbb{E}[X]K_A}(1/s) = \mathcal{O}\big(\varphi^{(2)}_{X+\mathbb{E}[X]K_A}(s) + \mathbb{E}[K_A]\,\varphi^{(2)}_X(s)\big), \text{ as } s \to 0^+.$$

This yields the desired result. □

Acknowledgements

The authors would like to thank the two anonymous referees for their suggestions that helped to shorten the present paper, for pointing out unexplored relevant references, and for numerous helpful comments making it more readable. The authors would also like to acknowledge the French Agence Nationale de la Recherche (ANR) and the project with reference ANR-20-CE40-0025-01 (T-REX project), and more specifically the members of the T-REX project for organising the VALPRED3 and VALPRED4 workshops at the CNRS Centre Paul Langevin in Aussois, during which fruitful discussions led to great improvement of this article.

Funding information

The are no funding bodies to thank relating to the creation of this article.

Competing interests

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

References

[1] Asmussen, S. and Foss, S. (2018). Regular variation in a fixed-point problem for single-and multiclass branching processes and queues. *Advances in Applied Probability* **50,** 47–61. 4, 9, 25

[2] Basrak, B., Conroy, M., Olvera-Cravioto, M. and Palmowski, Z. (2022). Importance sampling for maxima on trees. *Stochastic Processes and Their Applications* **148,** 139–179. 9

[3] Basrak, B., Davis, R. A. and Mikosch, T. (2002). A characterization of multivariate regular variation. *Annals of Applied Probability* 908–920. 12

[4] Basrak, B., Kulik, R. and Palmowski, Z. (2013). Heavy-tailed branching process with immigration. *Stochastic Models* **29,** 413–434. 4

[5] Basrak, B., Milinčević, N. and Žugec, P. (2023). On extremes of random clusters and marked renewal cluster processes. *Journal of Applied Probability* **60,** 367–381. 3, 4, 8, 13, 21, 27

[6] Basrak, B., Wintenberger, O. and Žugec, P. (2019). On the total claim amount for marked Poisson cluster models. *Advances in Applied Probability* **51,** 541–569. 1, 2, 5, 7, 8, 25, 27

[7] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1989). *Regular Variation* vol. 27. Cambridge University Press. 4, 12, 30, 36, 37, 38, 39

[8] Brémaud, P. (2020). *Point Process Calculus in Time and Space: An Introduction with Applications* vol. 98. Springer Nature. 2, 9, 29

[9] Chavez-Demoulin, V., Davison, A. C. and McNeil, A. J. (2005). Estimating value-at-risk: a point process approach. *Quantitative Finance* **5,** 227–234. 2

[10] Chen, N., Litvak, N. and Olvera-Cravioto, M. (2014). Pagerank in scale-free random graphs. In *Algorithms and Models for the Web Graph: 11th International Workshop, WAW 2014, Beijing, China, December 17-18, 2014, Proceedings 11*. Springer. pp. 120–131. 4

[11] Chen, N., Litvak, N. and Olvera-Cravioto, M. (2017). Generalized pagerank on directed configuration networks. *Random Structures & Algorithms* **51,** 237–274. 4

[12] Cline, D. B. and Hsing, T. (1991). Large deviation probabilities for sums and maxima of random variables with heavy or subexponential tails. *Preprint, Texas A&M University* **501,**. 4, 32

[13] Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes. Vol. I. Probability and its applications*. Springer-Verlag, New York. 2, 7

[14] Daley, D. J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*. Springer New York. 2

[15] De Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction* vol. 21. Springer. 3, 4, 10

[16] Denisov, D., Foss, S. and Korshunov, D. (2010). Asymptotics of randomly stopped sums in the presence of heavy tails. 4

[17] Embrechts, P., Klüppelberg, C. and Mikosch, T. (2013). *Modelling Extremal Events: for Insurance and Finance* vol. 33. Springer Science & Business Media. 3

[18] Ernst, P. A., Asmussen, S. and Hasenbein, J. J. (2018). Stability and busy periods in a multiclass queue with state-dependent arrival rates. *Queueing Systems* **90,** 207–224. 4

[19] Faÿ, G., González-Arévalo, B., Mikosch, T. and Samorodnitsky, G. (2006). Modeling teletraffic arrivals by a poisson cluster process. *Queueing Systems* **54,** 121–140. 1, 2, 3, 4, 20

[20] Foufoula-Georgiou, E. and Lettenmaier, D. P. (1987). A Markov renewal model for rainfall occurrences. *Water resources research* **23,** 875–884. 2

[21] Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58,** 83–90. 9

[22] Hawkes, A. G. (2018). Hawkes Processes and Their Applications to Finance: a Review. *Quantitative Finance* **18,** 193–198. 2

[23] Hawkes, A. G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability* **11,** 493–503. 9

[24] Heyde, C. C. (1967). On large deviation problems for sums of random variables which are not attracted to the normal law. *The Annals of Mathematical Statistics* **38,** 1575–1578. 4, 31

[25] Hult, H. and Lindskog, F. (2006). Regular variation for measures on metric spaces. *Publications de l'Institut Mathématique* **80,** 121–140. 10, 11

[26] Hult, H. and Samorodnitsky, G. (2008). Tail probabilities for infinite series of regularly varying random vectors. *Bernoulli* **14,** 838 – 864. 4

[27] Jelenković, P. R. and Olvera-Cravioto, M. (2010). Information ranking and power laws on trees. *Advances in Applied Probability* **42,** 1057–1093. 4

[28] Jessen, H. A. and Mikosch, T. (2006). Regularly varying functions. *Publications de l'Institut Mathématique* **80,** 171–192. 3

[29] Karabash, D. and Zhu, L. (2015). Limit theorems for marked Hawkes processes with application to a risk model. *Stochastic Models* **31,** 433–451. 2

[30] Karamata, J. (1933). Sur un mode de croissance régulière. théorèmes fondamentaux. *Bulletin de la Société Mathématique de France* **61,** 55–62. 4

[31] Karpelevich, F., Kelbert, M. Y. and Suhov, Y. M. (1994). Higher-order lindley equations. *Stochastic Processes and their Applications* **53,** 65–96. 3

[32] Klüppelberg, C. and Mikosch, T. (1997). Large deviations of heavy-tailed random sums with applications in insurance and finance. *Journal of Applied Probability* **34,** 293–308. 1, 4, 27, 28, 30, 31, 32

[33] Lindskog, F., Resnick, S. I. and Roy, J. (2014). Regularly Varying Measures on Metric Spaces: Hidden Regular Variation and Hidden Jumps. *Probability Surveys* **11,** 270–314. 11

[34] Litvak, N., Scheinhardt, W. R. and Volkovich, Y. (2007). In-degree and Pagerank: why do they follow similar power laws? *Internet Mathematics* **4,** 175–198. 4, 25

[35] Liu, L. (2009). Precise large deviations for dependent random variables with heavy tails. *Statistics & Probability Letters* **79,** 1290–1298. 4, 32

[36] Markovich, N. (2023). Extremal properties of evolving networks: local dependence and heavy tails. *Annals of Operations Research* 1–32. 4

[37] Markovich, N. M. and Rodionov, I. V. (2020). Maxima and sums of non-stationary random length sequences. *Extremes* **23,** 451–464. 4

[38] Mikosch, T. (2009). *Non-life insurance mathematics: an introduction with the Poisson process.* Springer Science & Business Media. 2, 8

[39] Mikosch, T. and Nagaev, A. V. (1998). Large deviations of heavy-tailed sums with applications in insurance. *Extremes* **1,** 81–110. 4

[40] Mikosch, T. and Wintenberger, O. (2023+). *Extremes for Time Series.* To appear. 4, 10, 11

[41] Musmeci, F. and Vere-Jones, D. (1992). A space-time clustering model for historical earthquakes. *Annals of the Institute of Statistical Mathematics* **44,** 1–11. 2

[42] Nagaev, A. V. (1969). Integral limit theorems taking large deviations into account when Cramér's condition does not hold. I. *Theory of Probability & Its Applications* **14,** 51–64. 4, 31

[43] Nagaev, A. V. (1969). Integral limit theorems taking large deviations into account when Cramér's condition does not hold. II. *Theory of Probability & Its Applications* **14,** 193–208. 4, 31

[44] Nagaev, S. V. (1979). Large deviations of sums of independent random variables. *The Annals of Probability* 745–789. 4

[45] Nagaev, S. V. (1979). Large deviations of sums of independent random variables. *The Annals of Probability* 745–789. 31

[46] Ng, K. W., Tang, Q., Yan, J.-A. and Yang, H. (2004). Precise large deviations for sums of random variables with consistently varying tails. *Journal of Applied Probability* **41,** 93–107. 4, 32

[47] Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association* **83,** 9–27. 2

[48] Olvera-Cravioto, M. (2021). PageRank's behavior under degree correlations. *The Annals of Applied Probability* **31,** 1403–1442. 3, 4, 21, 25

[49] Onof, C., Chandler, R. E., Kakou, A., Northrop, P., Wheater, H. S. and Isham, V. (2000). Rainfall modelling using Poisson-cluster processes: a review of developments. *Stochastic Environmental Research and Risk Assessment* **14,** 384–411. 2

[50] Resnick, S. I. (1987). *Extremes Values, Regular Variation and Point Processes.* Springer-Verlag, New York. 3, 4, 10

[51] Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer Science & Business Media. 4, 10, 11

[52] Reynaud-Bouret, P. and Schbath, S. (2010). Adaptive estimation for Hawkes processes; application to genome analysis. *The Annals of Statistics* **38,** 2781–2822. 2

[53] Robert, C. Y. and Segers, J. (2008). Tails of random sums of a heavy-tailed number of light-tailed terms. *Insurance: Mathematics and Economics* **43,** 85–92. 4

[54] Segers, J., Zhao, Y. and Meinguet, T. (2016). Polar decomposition of regularly varying time series in star-shaped metric spaces. *arXiv preprint arXiv:1604.00241*. 10, 11, 12

[55] Stabile, G. and Torrisi, G. L. (2010). Risk processes with non-stationary hawkes claims arrivals. *Methodology and Computing in Applied Probability* **12,** 415–429. 2

[56] Swishchuk, A. (2018). Risk model based on compound Hawkes process. *Wilmott* **2018,** 50–57. 2

[57] Tang, Q. (2006). Insensitivity to negative dependence of the asymptotic behavior of precise large deviations. 4, 32

[58] Tang, Q., Su, C., Jiang, T. and Zhang, J. (2001). Large deviations for heavy-tailed random sums in compound renewal model. *Statistics & Probability Letters* **52,** 91–100. 4, 32

[59] Tillier, C. and Wintenberger, O. (2018). Regular variation of a random length sequence of random variables and application to risk assessment. *Extremes* **21,** 27–56. 3

[60] Vere-Jones, D. and Ozaki, T. (1982). Some examples of statistical estimation applied to earthquake data: I. Cyclic Poisson and self-exciting models. *Annals of the Institute of Statistical Mathematics* **34,** 189–207. 2

[61] Volkovich, Y. and Litvak, N. (2010). Asymptotic analysis for personalized web search. *Advances in applied probability* **42,** 577–604. 4